On the Forecasting of Financial Volatility
Using Ultra-High Frequency Data
On the forecasting of financial volatility using ultra-high frequency data

António A. F. Santos

Faculty of Economics, Monetary and Financial Research Group (GEMF), University of Coimbra, Portugal

Abstract. The measurement of the volatility is key in financial markets. This is true not only because decisions are made in an environment of uncertainty, but because sometimes the volatility element overpowers all the remaining aspects in the decision process. Huge movements in the prices of the assets (volatility) can lead to huge losses and also huge gains. There are models to establish the fair prices for certain kind of assets, in that the only parameter that is not directly observable is the parameter characterizing the volatility. However, it is well established in the literature that the evolution of the volatility can be forecasted. Several parametric models have been proposed for modeling the volatility evolution, for example, the Autoregressive Conditional Heteroscedastic (ARCH) and the Stochastic Volatility model (SV). Nowadays, we live in a “Big Data” world, and even for non-professionals of financial markets, it is possible to record data obtained at every second. Recently, measures of volatility have been developed using intraday data, for example, the measure of realized volatility. One of the main aspects to consider is that intraday data and measures of realized volatility are associated with unequal time-spaced observations. In this paper, we compare the forecasts of the volatility evolution using intradaily observations and daily observations, and by trying to conciliate both kind of forecasts, for the data obtained from US and European stock markets, we find out that the use of measures of realized volatility represent an important improvement in volatility forecasting, that can be added to the more well established models that are used in this context, ARCH and SV models.

Keywords: ARCH models, Big data, Intraday data, Realized volatility, Stochastic volatility

JEL Classification: C11, C15, C53, C55, G17

* Address: Av. Dias da Silva, 165; 3004-512 Coimbra, Portugal; email: aasantos@fe.uc.pt
1 Introduction

The volatility is a highly studied subject in financial econometrics and a measure essential to decision agents acting in financial markets. As a measure of risk is used in different situations and enters in many models helping decision agents to take more rational decisions. Since the 80s (Engle 1982, Bollerslev 1986, Taylor 1986), the amount of research in this area is overwhelming.

Econometric models need data, and recently this is an abundant resource. The models have to be redefined to keep up the pace with these new developments. After a more parametric analysis of the volatility, at the end of the 90s, new developments have appeared. Some are more model-free developments, and one of the most important in this area is the development of realized volatility measures (Andersen and Bollerslev 1997, Andersen et al. 1999, 2001, 2003, 2005, 2011, Gonçalves and Meddahi 2009, Hansen and Lunde 2006). New developments have been considered, some recent research has been trying to establish a link between the realized volatility measure and the stochastic volatility models (Takahashi et al. 2009, Koopman and Scharth 2012). Our aim is to verify if the computations of realized volatility are compatible with forecasted volatility through the stochastic volatility model.

An empirical fact is that the volatility presents some degree of forecastability. Most of the studies have considered the ARCH family of models. The first results indicated that ARCH-type models would provide a very poor forecast of the volatility. This was due to the interpretation of the GARCH model as an ARMA on the squared returns. A step forward is to assume that the GARCH models do not define forecasts for the squared returns, but for another measure of volatility, in this case the integrated volatility.

The integrated volatility and correspondent estimator the realized volatility are quantities defined through the interpretation of the evolution of asset prices in a continuous framework. In this setting, a more natural link is given through a discrete-time stochastic volatility model, which is considered as the discrete version of the continuous model that considers the usual diffusion process associated with prices coupled with a subordinated process for the evolution of the volatility.
The novelty of this paper is to use volatility forecasts obtained from the stochastic volatility model which are compared to the realized volatility measure. Comparisons were already made using ARCH-type of forecasts, but in this case the forecasts are based on particle filter algorithms associated with the SV models, and the algorithms are robust to outliers commonly found in financial time series. Ultra-high-frequency data with individual US and European stocks are used to make the necessary comparisons. The article is organized as follows. In Section 2 we review some main results associated with the evolution of stock prices as a continuous stochastic process and the link with measures of volatility and related models. In Section 3 we introduce our version of the particle filter algorithm used to forecast the volatility for daily returns. In Section 4 we consider an empirical demonstration where the algorithms and the measures associated with the volatility evolution are compared. In Section 5 we present some conclusions.

2 Prices and volatility evolution

A standard model establishes that $P(t)$, the price path, is the solution of the differential equation $dP(t) = \mu(t)P(t)dt + \sigma(t)P(t)dW(t)$, where $\mu(t)$ is a mean evolution function, the $\sigma^2(t)$ a variance evolution function and $W(t)$ the standard Brownian motion. From this representation a measure of interest is the integrated volatility

$$IV_t = \int_0^t \sigma^2(s)ds.$$  \hspace{1cm} (1)

If the evolution of the price respects the price diffusion equation, for a period of length $t$, the volatility associated with the return for the same period is given by (1). As this quantity is unobservable, estimators have to be defined to estimate it. The simplest and most known is the realized volatility.

For $y_t$, the return at $t$, that represents the return associated with one period (a day), consider the partition $0 < t_1 < t_2 < \ldots < t_n < 1$, intraday returns are given by $y_i = p_t - p_{t_{i-1}}, i = 1, \ldots, n$, where $p_t$ is the log-price at $t$. The realized volatility is given by the sum of
the squares of the intraday returns,

\[ RV_t = \sum_{i=1}^{n} y_{i,t}^2. \]  

(2)

This can be used as an estimator to the integrated volatility and research has been conducted trying to establish when \( RV_t \) is a consistent estimator to \( IV_t \). The statistical properties of the estimator \( RV_t \) to the quantity \( IV_t \) were thoroughly analyzed in Barndorff-Nielsen and Shephard (2002), Barndorff-Nielsen and Shephard (2004), Zhang et al. (2005) and Hansen and Lunde (2006).

2.1 Stochastic volatility

Considering the evolution of the volatility that also follows a stochastic process, this can be connected with the stochastic process for the prices

\[
\begin{align*}
    dP(t) &= \mu(t) P(t) dt + \sigma(t) P(t) dW(t) \\
    d\log(\sigma^2(t)) &= \beta \log(\sigma^2(t)) + \sigma_\eta dZ(t)
\end{align*}
\]

where \( Z(t) \) is a Brownian motion independent from \( W(t) \). For daily observations a discretization of this process assumes the form of a nonlinear state-space model characterized by

\[
\begin{align*}
    y_t &= \exp \left( \frac{\alpha_t}{2} \right) \varepsilon_t \\
    \alpha_{t+1} &= \mu + \phi(\alpha_t - \mu) + \sigma_\eta \eta_{t+1}
\end{align*}
\]

where \( (\varepsilon_t, \eta_{t+1}) \) follow a bivariate normal distribution \( N(0, I_2) \), and a set of parameters \( \theta = (\mu, \phi, \sigma_\eta) \). Another important component of the model is the vector of states \( \alpha = (\alpha_1, \ldots, \alpha_n) \). This kind of formulation first appeared in Taylor (1986). The states define the daily evolution of the volatility. They vary over time in a stochastic manner and are governed by an AR(1) process. With a Bayesian approach and MCMC techniques the distribution for the parameters and for states are approximated (Jacquier et al. 1994, Shephard and Pitt 1997, Kim et al. 1998, Chib et al. 2002, Jacquier et al. 2004, Omori et al. 2007).
Here the aim is to define volatility forecasts. It is assumed that the parameters are known. More realistically, the parameters are estimated using a sample prior to the period associated with the forecasts. The volatility forecasts are defined through the predictive distribution associated with the states, with density \( f(\alpha_{t+1}|y_{1:t}) \), where \( y_{1:t} = (y_1, \ldots, y_t) \) is a set of available data until \( t \). Assuming a quadratic loss function the forecasts are defined as

\[
\sigma_{t+1|t}^2 = E(\exp(\alpha_{t+1})|y_{1:t}).
\] (3)

As in this environment the posterior distributions of interest are not assuming an easy-to-calculate form, which would also allow the definition of the moments easily, numerical procedures have to be considered as a way of approximating the aforementioned distributions.

3 Particle filter and volatility forecasts

State-space models first have appeared in engineering, signal processing and object tracking, and lately have been introduced in other fields like biology, medicine (medical imaging), and also in economics. When the needed distributions have to be redefined with the new information that arrives, with distributions that are intractable in an analytical manner, their redefinition often becomes a very demanding computational task. Different techniques and algorithms have been considered. One approach used to approximate sequentially the filter distributions belongs to a broad family of algorithms known as Sequential Monte Carlo (SMC), where the importance sampling (IS) plays a central role.

3.1 Sequential Monte Carlo

Considering a state-space model, different from the linear gaussian state-space model, assuming that the parameters are known or have been estimated previously, and the model can be characterized by two equations, system and measurement equations, the aim is to approximate the difficult-to-compute density \( f(\alpha_{1:t}|y_{1:t}) \) for \( t = 1, \ldots, n \), where \( \alpha_{1:t} = (\alpha_1, \ldots, t) \) and \( y_{1:t} = (y_1, \ldots, t) \). The initial state is assumed having a given density \( f(\alpha_0) \). The dimension of the problem
(density) is increasing as new information becomes available through $y_t$ for $t = 1, \ldots, n$.

For $t = n$ the density $f(\alpha_1:n|y_1:n)$ is known as the smoothing density, $f(\alpha_n|y_1:n)$ is the filter density, and a fixed lag smoothing density for $n - l$, $f(\alpha_{n-l}|y_1:n)$, with $0 < l < n$ can also be of interest. As new information arrives, with nowadays computational resources, approximate successively $f(\alpha_1:t|y_1:t)$, for $t = 1, \ldots, n$ using standard Monte Carlo simulation procedures is feasible.

With the SMC, the main idea is to explore the Markovian structure of the system equation. For a given $t$ we can approximate $f(\alpha_1:t|y_1:t)$ by sampling directly from the random vector $\alpha_1:t|y_1:t$, or assuming that we have an approximation at $t - 1$ to $f(\alpha_1:t-1|y_1:t-1)$ given by a set of particles and their respective weights, conditional on those particles an approximation to $f(\alpha_1:t|y_1:t)$ is defined by sampling from a density of dimension one. Some references can be found in Gordon et al. (1993), Carpenter et al. (1999), Pitt and Shephard (1999), Doucet et al. (2001), Pitt and Shephard (2001), Fearnhead and Clifford (2003), Cappé et al. (2004), Douc et al. (2009), Briers et al. (2010), Fearnhead et al. (2010), Li et al. (2015).

### 3.2 Stochastic volatility with particle filter

Considering a SMC with IS approach jointly with a resampling step, in a state-space context with sequential sampling, the aim is to update sequentially the filter distribution, i.e., approximate the density $f(\alpha_t|y_1:t)$. Using importance sampling an importance density is given by $g(\alpha_t) \propto f(y_t|\alpha_t)f(\alpha_t|\alpha_{t-1}, y_t)$, which in this case would be the best choice. This leads to importance weights with variance equal to zero. Usually it is not possible to sample from this density and the importance density $g(\alpha_t)$ is used to approximate the optimal one. There are some differences, which are corrected by the weights. If particles with equal weights are desired a resampling step is performed.

A seminal article associated with the SV model is due to Jacquier et al. (1994), not in context of filtering but the estimation of the parameters in the model. Nonetheless the algorithms were not very efficient, this article had an enormous impact on the research associated with the analysis of volatility using the SV model. Still in terms of
estimation of the model, within a Bayesian framework with MCMC, another major breakthrough was considered in Shephard and Pitt (1997), who proposed gaussian approximations to the distributions of the states instead of gamma as in Jacquier et al. (1994). Further developments using these results, still in the field of parameters estimation are found in Kim et al. (1998), Chib et al. (2002), Jacquier et al. (2004), Omori et al. (2007). When the aim is to define approximations to filter densities, Pitt and Shephard (1999) have considered the same kind of approximations. However, they were based on a first order Taylor approximation, which was shown by Smith and Santos (2006), that they were not robust when information contained in more extreme observations need to be considered. Smith and Santos (2006) considered a second order Taylor approximation for the likelihood which combined with the predictive density for the states leads to improvements in the particle filter algorithm. In the spirit of the auxiliary particle filter of Pitt and Shephard (1999), avoids the blind proposals as in the bootstrap filter of Gordon et al. (1993), takes into account the information of $y_t$, and by constituting a robust approximation of the target density avoids in some sense the degeneracy of the weights.

Here we develop the aforementioned results by using an even more robust approximation for the optimal importance density $g(\alpha_t) \propto f(y_t|\alpha_t) f(\alpha_t|\alpha_{t-1}, y_t)$. Considering the logarithm of $f(y_t|\alpha_t) f(\alpha_t|\alpha_{t-1})$

$$\ell(\alpha_t) \propto -\frac{\alpha_t}{2} - \frac{y_t^2}{2\exp(\alpha_t)} - \frac{(\alpha_t - \mu - \phi(\alpha_{t-1} - \mu))^2}{2\sigma_y^2},$$

this function is concave on $\alpha_t$, and considering the first and second order conditions to maximize the function in order to $\alpha_t$, the first derivative is equal to zero, $\ell'(\alpha_t) = 0$, and solving in order to $\alpha_t$ the solution is

$$\alpha_t^* = W\left(\frac{y_t^2 \sigma_y^2 e^{-\gamma}}{2}\right) + \gamma; \text{ with } \gamma = \mu(1 - \phi) + \phi\alpha_{t-1} - \frac{\sigma_y^2}{2},$$

where $W(\cdot)$ is the Lambert function. The second derivative is given by

$$\ell''(\alpha_t) = -\frac{2e^{\alpha_t} + \sigma_y^2 y_t^2}{2\sigma_y^2 e^{\alpha_t}}.$$
which is strictly negative for all $\alpha_t$, so $\alpha^*_t$ maximizes the function $\ell(\alpha_t)$, which defines the global maximum. By defining a second order Taylor expansion of $\ell(\alpha_t)$ around $\alpha^*_t$ the log-kernel of a gaussian density is obtained with mean $m_t = \alpha^*_t$ and variance

$$s^2_t = \frac{2\sigma_\eta^2 e^{m_t}}{2e^{m_t} + \sigma_\eta^2 y^2_t}.$$  

This gaussian density will constitute the importance density used to update the filter densities.

In the procedures implemented the estimates of interest were approximated using only particles with equal weights, which means that resampling steps are performed. Assuming that at $t-1$ we have a set of $N$ particles $\alpha_{t-1}^N = \{\alpha_{t-1,1}, \ldots, \alpha_{t-1,N}\}$ with associated equal weights $1/N$, which approximate the density $f(\alpha_{t-1}|y_{1:t-1})$, the algorithm proceeds as follows

1. Resample $N$ particles from $\alpha_{t-1}^N$, obtaining $\alpha_t^N = \{\alpha_{t,1}, \ldots, \alpha_{t,N}\}$.
2. For each element of the set, $\alpha_{t,i}$, $i = 1, \ldots, N$, sample a value from a normal distribution with mean and variance defined by (4) and (5), obtaining the set $\{\alpha_{t,1}^*, \ldots, \alpha_{t,N}^*\}$
3. Calculate the weights,

$$w_i = \frac{f(y_t|\alpha_{t,i}^*)f(\alpha_{t,i}^*|\alpha_{t-1,i})}{g(\alpha_{t,i}^*|m_t, s^2_t)}; \quad \pi_i = \frac{w_i}{\sum_{i=1}^N w_i}$$

4. Resample from the set $\{\alpha_{t,1}^*, \ldots, \alpha_{t,N}^*\}$ with the set of weights $\{\pi_1, \ldots, \pi_N\}$ obtaining a sample $\{\alpha_{t|1:t,1}, \ldots, \alpha_{t|1:t,N}\}$ where to each particle a weight of $1/N$ is associated.

For the one step-ahead volatility forecast, having the approximation to the density $f(\alpha_t|y_{1:t})$, due to the structure of the system equation in the SV model, AR(1) with gaussian noise, it is easy to sample from $f(\alpha_{t+1}|y_{1:t})$ and approximate the necessary measures. The main question is, can the density $f(\alpha_t|y_{1:t})$ be well approximated by the algorithm proposed?

### 4 An empirical demonstration

In this section a demonstration is presented using intraday data recorded for stocks in US and European markets for a period span-
ning more than one year. The forecasts for the volatility are compared with the realized volatility measure calculated using the intraday dataset. The forecasts are obtained using the GARCH model and a simple-to-use method like the EWMA advocated by RiskMetrics, and the aforementioned SV model, for which we presented a particle filter algorithm. This particle filter algorithm approximates the distributions of the states used in the definition of the forecasts for the volatilities. This kind of comparisons were made already in other contexts using ARCH-type of models. The origin of such comparisons is related to the doubts on the ability of these to forecast the volatility evolution. The GARCH as an ARMA on the squared returns defines the forecasts through the mean, which can be compared with the observables through the Mean Squared Error (MSE) or an equivalent measure. The other way is to establish a linear regression on the observables (squared returns) and the volatility forecasts and measure the goodness of the fit, for example, using a standard measure as is the $R^2$ in the context of a linear regression.

4.1 Data description

Data associated with ten stocks were obtained for each market. Due to space restrictions only the results of two in each market are presented. The data have all transactions (time, price and volume) for the period of January 31, 2014 to April 29, 2015. There are some missing data (days) due to recording problems, but a total of 252 days are considered, which makes that for each stock considered, around one million observations are available for analysis. Daily data were obtained from public available dataset in the site of Yahoo! Inc.. To obtain intraday data a software package was build in Objective C and used to record the data from the publicly available web pages at the site BATS Global Markets, Inc..

Using data associated with stocks, we consider observations unequally spaced in time, and each observation corresponds to an effective transaction with time of the transaction, price and volume. To demonstrate the results obtained, two stocks from the US market, Cisco Systems Inc. (CSCO) and Exxon Mobil Corporation (XOM) and two from the european markets, Bayer AG (BAYN) and Electricite de France SA (EDF) were considered. The aim is to test the
volatility forecasts and the algorithm proposed with the particle filter, but also check if some differences in volatility structure between US and Europe markets are found. In Table 1 are presented some of the main characteristics of the intraday data used in this article. For the four stocks considered, four series of intraday data were defined with prices, time between successive transactions, and volume traded. Each series contains around one million observations. Common patterns can be highlighted. The mean value of the volume for each transaction is related to the mean of the price. On the other hand, the mean of the time span in seconds between each transaction is around 5 seconds. For each stock there is an important number of simultaneous transactions, meaning time differences in seconds equal to zero, which must be taken into account, if some kind of duration model is included in the analysis. The probability of zero seconds between transactions is very significative.

4.2 Volatility forecasts and realized volatility measures

The volatility forecasts are defined through three models, first, the SV model using the particle filter to approximate the predictive density of the states, which define the forecasts. The second is a GARCH(1,1) without mean coefficient, from which, after estimating the parameters, the forecasts are defined as an affine combination on the squared returns and the respective variance. Finally, also a simple exponential moving average popularized by RiskMetrics. The weight attributed to the previous forecast is $\lambda = 0.96$. These forecasts are compared with the realized volatility measures which are calculated using the dataset described in the previous section.

<table>
<thead>
<tr>
<th>Stock</th>
<th>MPrice</th>
<th>MVolume</th>
<th>MTDS</th>
<th>NObs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSCO</td>
<td>25.97</td>
<td>275.97</td>
<td>4.93</td>
<td>1,175,834</td>
</tr>
<tr>
<td>XOM</td>
<td>93.15</td>
<td>103.65</td>
<td>5.49</td>
<td>1,068,629</td>
</tr>
<tr>
<td>BAYN</td>
<td>112.04</td>
<td>74.48</td>
<td>6.03</td>
<td>1,251,314</td>
</tr>
<tr>
<td>EDF</td>
<td>24.62</td>
<td>172.05</td>
<td>8.73</td>
<td>865,322</td>
</tr>
</tbody>
</table>
Theoretically the main measure of interest is the integrated volatility, which is unobservable, and the forecasts are compared with the realized volatility. This measure can still be highly variable, which makes difficult to evaluate how close are the forecasted volatilities from the integrated volatility. The results obtained are inconclusive in the sense that no model used to forecast the volatilities dominates the others in all cases. There are cases where the trivial-to-implement EWMA dominates the forecasts provided by the more elaborated models. Maybe a wise way to proceed is to combine the volatility forecasts from the different models to approximate better the realized volatility.

Before defining the forecasts, the parameters had to be estimated. The parameters of the GARCH(1,1) model characterized by the equations \( y_t = \sigma_t \varepsilon_t \) and \( \sigma_t^2 = \omega + \gamma_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \) were estimated through the maximum likelihood method. The parameters for the SV model were estimated through the mean of the posterior distribution of the parameters using Bayesian methods and MCMC procedures in the approximation of the respective posterior distribution. In the estimation samples of daily returns from January 2, 2001 to December 31, 2013 were used, which constitutes around 3,300 daily observations. A summary of the estimation results are given in the Table 2.

**Table 2.** Estimation of the GARCH and SV model. Under the estimates for the GARCH is presented the Standard error and for the SV the standard deviation of the respective chain. For GARCH maximum likelihood estimation was used, and for SV, Bayesian estimation with MCMC.

<table>
<thead>
<tr>
<th>Estimation</th>
<th>( \omega )</th>
<th>( \gamma_1 )</th>
<th>( \beta_1 )</th>
<th>( \mu(1-\phi) )</th>
<th>( \phi )</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSCO</td>
<td>0.042</td>
<td>0.022</td>
<td>0.969</td>
<td>0.05</td>
<td>0.959</td>
<td>0.297</td>
</tr>
<tr>
<td></td>
<td>(0.0075)</td>
<td>(0.0027)</td>
<td>(0.0037)</td>
<td>(0.0110)</td>
<td>(0.0072)</td>
<td>(0.0252)</td>
</tr>
<tr>
<td>XOM</td>
<td>0.037</td>
<td>0.0798</td>
<td>0.902</td>
<td>0.009</td>
<td>0.978</td>
<td>0.166</td>
</tr>
<tr>
<td></td>
<td>(0.0085)</td>
<td>(0.0088)</td>
<td>(0.0109)</td>
<td>(0.0037)</td>
<td>(0.0044)</td>
<td>(0.0182)</td>
</tr>
<tr>
<td>BAYN</td>
<td>0.030</td>
<td>0.046</td>
<td>0.950</td>
<td>0.092</td>
<td>0.8979</td>
<td>0.544</td>
</tr>
<tr>
<td></td>
<td>(0.0076)</td>
<td>(0.0041)</td>
<td>(0.0045)</td>
<td>(0.0151)</td>
<td>(0.0119)</td>
<td>(0.0383)</td>
</tr>
<tr>
<td>EDF</td>
<td>0.068</td>
<td>0.0647</td>
<td>0.916</td>
<td>0.070</td>
<td>0.923</td>
<td>0.327</td>
</tr>
<tr>
<td></td>
<td>(0.0192)</td>
<td>(0.0099)</td>
<td>(0.0128)</td>
<td>(0.0202)</td>
<td>(0.0209)</td>
<td>(0.0523)</td>
</tr>
</tbody>
</table>

Using the estimated parameters in the respective models, one step-ahead forecasts for the variance of the returns are defined. Assuming that at \( t \) an estimate the variance is given by \( \hat{\sigma}_t^2 \), through the
information available until \( t \) a forecast is defined for \( t+1 \). With the EWMA the forecast is defined as \( \hat{\sigma}_{t+1|t}^2 = \lambda \hat{\sigma}_t + (1 - \lambda) y_t^2 \). For the GARCH the forecasts are defined as \( \hat{\sigma}_{t+1|t}^2 = \omega + \gamma_1 y_t^2 + \beta_1 \hat{\sigma}_t^2 \). For the SV model, assuming that a given approximation for the density \( f(\alpha_t|y_1:t) \) is available at \( t \), obtained using particle filter methods, simulated values are obtained from \( f(\alpha_{t+1}|y_1:t) \) and used to approximate \( E(\exp(\alpha_{t+1})|y_1:t) \). This constitutes the forecasted volatility through the SV model.

In Figure 1 is depicted the evolution of the realized volatility and respective forecasts obtained through the models considered in this article. Different paths for the realized volatility and respective forecasts are obtained, which results in different fits. It can be verified that if the aim is to forecast the realized volatility, no forecasting model dominates the others, and surprisingly, or maybe not, in certain cases the EWMA (no computational effort) gives very good results.
Table 3. Summary statistics for the fit of volatility forecasts to realized volatility. Three forecasting models were used, SV, EWMA and GARCH. The values for $R^2$ were obtained from the regression $\log RV_{t+1} = \beta_0 + \beta_1 \log \text{forecast}_{t+1} + \text{res}$. The Mean Square Error (MSE) on the log of the variables is also presented.

<table>
<thead>
<tr>
<th></th>
<th>CSCO</th>
<th>XOM</th>
<th>BAYN</th>
<th>EDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.199</td>
<td>0.334</td>
<td>0.108</td>
<td>0.108</td>
</tr>
<tr>
<td>MSE</td>
<td>0.590</td>
<td>0.400</td>
<td>0.423</td>
<td>0.345</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.098</td>
<td>0.356</td>
<td>0.131</td>
<td>0.131</td>
</tr>
<tr>
<td>MSE</td>
<td>0.361</td>
<td>0.331</td>
<td>0.239</td>
<td>0.340</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.044</td>
<td>0.410</td>
<td>0.150</td>
<td>0.150</td>
</tr>
<tr>
<td>MSE</td>
<td>0.777</td>
<td>0.348</td>
<td>0.249</td>
<td>0.338</td>
</tr>
</tbody>
</table>

Even using a model that presents great flexibility, as the SV model, and computational intensive methods to forecast the volatility, adequate forecasts for the volatility are not easy to define. In many cases more heuristic methods may present a reasonable power, namely when compared with the results of more elaborated models taking into account the computational complexity involving in calculating the forecasts with both kind of models. The values found for the $R^2$ are in line with what is found in the literature, but the main result is that with the different adjustments, which are illustrated in Figure 1 and Table 3, further research can be produced to combine forecasts of the different models to obtain a more clear picture for the evolution of the volatility.

5 Conclusion

Important information related to the evolution of the volatility of stock returns can be obtained using intraday data. One of the most used is the realized volatility measure. This measure is not so noisy when compared with the squared returns. Theoretically it constitutes an estimator of an unobservable measure of interest, the integrated volatility, but still it has associated an important level of noise, which makes the forecasting difficult to implement. Levels of fit commonly found in other settings when the aim is to forecast the evolution of the mean of a variable are not found in forecasting the volatilities. Even using a very flexible model like the SV model and computationally intensive methods, the forecast ability is still limited. More research is needed to define methods of combining the different sources of forecasts as a way of improving volatility forecasts.


<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-14</td>
<td>The Determinants of Entrepreneurship at the Country Level: A Panel Data Approach</td>
<td>Gonçalo Brás &amp; Elias Soukiazis</td>
</tr>
<tr>
<td>2015-13</td>
<td>Budget, Expenditures Composition and Political Manipulation: Evidence from Portugal</td>
<td>João T. Addison, Orgul D. Ozturk &amp; Si Wang</td>
</tr>
<tr>
<td>2015-12</td>
<td>The Occupational Feminization of Wages</td>
<td>John T. Addison, Orgul D. Ozturk &amp; Si Wang</td>
</tr>
<tr>
<td>2015-09</td>
<td>The Impact of the Chinese Exchange Policy on Foreign Trade with the European Union</td>
<td>Ana Cardoso &amp; António Portugal Duarte</td>
</tr>
<tr>
<td>2015-08</td>
<td>The links between the Companies’ Market Price Quality and that of its Management and Business Quality: A system panel data approach</td>
<td>Dinis Santos &amp; Elias Soukiazis</td>
</tr>
<tr>
<td>2015-07</td>
<td>Education and Software piracy in the European Union</td>
<td>Nicolas Dias Gomes, Pedro André Cerqueira &amp; Luís Alçada-Almeida</td>
</tr>
<tr>
<td>2015-06</td>
<td>A Monetary Analysis of the Liquidity Trap</td>
<td>João Braz Pinto &amp; João Sousa Andrade</td>
</tr>
<tr>
<td>2015-05</td>
<td>Efficient Skewness/Semivariance Portfolios</td>
<td>Rui Pedro Brito, Hélder Sebastião &amp; Pedro Godinho</td>
</tr>
<tr>
<td>2015-04</td>
<td>Size Distribution of Portuguese Firms between 2006 and 2012</td>
<td>Rui Pascoal, Mário Augusto &amp; Ana M. Monteiro</td>
</tr>
<tr>
<td>2015-03</td>
<td>Optimum Currency Areas, Real and Nominal Convergence in the European Union</td>
<td>João Sousa Andrade &amp; António Portugal Duarte</td>
</tr>
<tr>
<td>2015-02</td>
<td>Estimating State-Dependent Volatility of Investment Projects: A Simulation Approach</td>
<td>Pedro Godinho</td>
</tr>
<tr>
<td>2015-01</td>
<td>Is There a Trade-off between Exchange Rate and Interest Rate Volatility? Evidence from an M-GARCH Model</td>
<td>António Portugal Duarte, João Sousa Andrade &amp; Adelaide Duarte</td>
</tr>
<tr>
<td>2014-21</td>
<td>Estado Social, Quantis, Não-Linearidades e Desempenho Económico: Uma Avaliação Empírica</td>
<td>Adelaide Duarte, Marta Simões &amp; João Sousa Andrade</td>
</tr>
</tbody>
</table>
Assessing the Impact of the Welfare State on Economic Growth: A Survey of Recent Developments
- Marta Simões, Adelaide Duarte & João Sousa Andrade

Business Cycle Synchronization and Volatility Shifts
- Pedro André Cerqueira

The Public Finance and the Economic Growth in the First Portuguese Republic
- Nuno Ferraz Martins & António Portugal Duarte

On the Robustness of Minimum Wage Effects: Geographically-Disparate Trends and Job Growth Equations
- John T. Addison, McKinley L. Blackburn & Chad D. Cotti

Determinants of Subjective Well-Being in Portugal: A Micro-Data Study
- Sara Ramos & Elias Soukiazis

Changes in Bargaining Status and Intra-Plant Wage Dispersion in Germany. A Case of (Almost) Plus Ça Change?
- John T. Addison, Arnd Kölling & Paulino Teixeira

The Renewables Influence on Market Splitting: the Iberian Spot Electricity Market
- Nuno Carvalho Figueiredo, Patrícia Pereira da Silva & Pedro Cerqueira

Drivers for Household Electricity Prices in the EU: A System-GMM Panel Data Approach
- Patrícia Pereira da Silva & Pedro Cerqueira

Effectiveness of Intellectual Property Regimes: 2006-2011
- Noemí Pulido Pavón & Luis Palma Martos

Dealing with Technological Risk in a Regulatory Context: The Case of Smart Grids
- Paulo Moisés Costa, Nuno Bento & Vítor Marques

Stochastic Volatility Estimation with GPU Computing
- António Alberto Santos & João Andrade

The Impact of Expectations, Match Importance and Results in the Stock Prices of European Football Teams
- Pedro Godinho & Pedro Cerqueira

Is the Slovak Economy Doing Well? A Twin Deficit Growth Approach
- Elias Soukiazis, Eva Muchova & Pedro A. Cerqueira

The Role of Gender in Promotion and Pay over a Career
- John T. Addison, Orgul D. Ozturk & Si Wang

Output-gaps in the PIIGS Economies: An Ingredient of a Greek Tragedy
- João Sousa Andrade & António Portugal Duarte

Software Piracy: A Critical Survey of the Theoretical and Empirical Literature
- Nicolas Dias Gomes, Pedro André Cerqueira & Luís Alcada Almeida

Agriculture in Portugal: Linkages with Industry and Services
- João Gaspar, Gilson Pina & Marta C. N. Simões

Effects of Taxation on Software Piracy Across the European Union
- Nicolas Dias Gomes, Pedro André Cerqueira & Luís Alcada Almeida

A Crise Portuguesa é Anterior à Crise Internacional
- João Sousa Andrade

Collective Bargaining and Innovation in Germany: Cooperative Industrial Relations?
- John T. Addison, Paulino Teixeira, Katalin Evers & Lutz Bellmann

Market Efficiency, Roughness and Long Memory in the PSI20 Index Returns: Wavelet and Entropy Analysis
- Rui Pascoal & Ana Margarida Monteiro

Do Size, Age and Dividend Policy Provide Useful Measures of Financing Constraints? New Evidence from a Panel of Portuguese Firms
- Carlos Carreira & Filipe Silva

A Política Orçamental em Portugal entre Duas Intervenções do FMI: 1986-2010
- Carlos Fonseca Marinheiro

Distortions in the Neoclassical Growth Model: A Cross-Country Analysis
- Pedro Brinca
Learning, Exporting and Firm Productivity: Evidence from Portuguese Manufacturing and Services Firms
- Carlos Carreira

Equity Premia Predictability in the EuroZone
- Nuno Silva

Human Capital and Growth in a Services Economy: the Case of Portugal
- Marta Simões & Adelaide Duarte

Does Voter Turnout Affect the Votes for the Incumbent Government?
- Rodrigo Martins & Francisco José Veiga

Determinants of Worldwide Software Piracy Losses
- Nicolas Dias Gomes, Pedro André Cerqueira & Luís Alçada Almeida

Despesa Pública em Educação e Saúde e Crescimento Económico: Um Contributo para o Debate sobre as Funções Sociais do Estado
- João Sousa Andrade, Marta Simões & Adelaide P. S. Duarte

Duration dependence and change-points in the likelihood of credit booms ending
- Vítor Castro & Megumi Kubota

Job Promotion in Mid-Career: Gender, Recession and 'Crowding'
- John T. Addison, Orgul D. Ozturk & Si Wang

Mathematical Modeling of Consumer’s Preferences Using Partial Differential Equations
- Jorge Marques

The Effects of Internal and External Imbalances on Italy’s Economic Growth. A Balance of Payments Approach with Relative Prices No Neutral.
- Elías Soukiazis, Pedro André Cerqueira & Micaela Antunes

A Regional Perspective on Inequality and Growth in Portugal Using Panel Cointegration Analysis
- Marta Simões, João Sousa Andrade & Adelaide Duarte

Macroeconomic Determinants of the Credit Risk in the Banking System: The Case of the GIPSI
- Vítor Castro

Majority Vote on Educational Standards
- Robert Schwager

Productivity Growth and Convergence: Portugal in the EU 1986-2009
- Adelaide Duarte, Marta Simões & João Sousa Andrade

What Determines the Duration of a Fiscal Consolidation Program?
- Luca Agnello, Vítor Castro & Ricardo M. Sousa

Minimum Wage Increases in a Recessionary Environment
- John T. Addison, McKinley L. Blackburn & Chad D. Cotti

The International Monetary System in Flux: Overview and Prospects
- Pedro Bação, António Portugal Duarte & Mariana Simões

Are There Change-Points in the Likelihood of a Fiscal Consolidation Ending?
- Luca Agnello, Vítor Castro & Ricardo M. Sousa

The Dutch Disease in the Portuguese Economy
- João Sousa Andrade & António Portugal Duarte

Is There Duration Dependence in Portuguese Local Governments’ Tenure?
- Vítor Castro & Rodrigo Martins

Testing for Nonlinear Adjustment in the Portuguese Target Zone: Is there a Honeymoon Effect?
- António Portugal Duarte, João Soares da Fonseca & Adelaide Duarte

Portugal Before and After the European Union
- Fernando Alexandre & Pedro Bação

The International Integration of the Eastern Europe and two Middle East Stock Markets
- José Soares da Fonseca
2012-21 Are Small Firms More Dependent on the Local Environment than Larger Firms? Evidence from Portuguese Manufacturing Firms
- Carlos Carreira & Luís Lopes

2012-20 Macroeconomic Factors of Household Default. Is There Myopic Behaviour?
- Rui Pascoal

2012-19 Can German Unions Still Cut It?
- John Addison, Paulino Teixeira, Jens Stephani & Lutz Bellmann

2012-18 Financial Constraints: Do They Matter to R&D Subsidy Attribution?
- Filipe Silva & Carlos Carreira

2012-17 Worker Productivity and Wages: Evidence from Linked Employer-Employee Data
- Ana Sofia Lopes & Paulino Teixeira

2012-16 Slovak Economic Growth and the Consistency of the Balance-of-Payments Constraint Approach
- Elias Soukiazis & Eva Muchova

2012-15 The Importance of a Good Indicator for Global Excess Demand
- João Sousa Andrade & António Portugal Duarte

2012-14 Measuring Firms’ Financial Constraints: A Rough Guide
- Filipe Silva & Carlos Carreira

2012-13 Convergence and Growth: Portugal in the EU 1986-2010
- Marta Simões, João Sousa Andrade & Adelaide Duarte

2012-12 Where Are the Fragilities? The Relationship Between Firms’ Financial Constraints, Size and Age
- Carlos Carreira & Filipe Silva

2012-11 An European Distribution of Income Perspective on Portugal-EU Convergence
- João Sousa Andrade, Adelaide Duarte & Marta Simões

2012-10 Financial Crisis and Domino Effect
- Pedro Bação, João Maia Domingues & António Portugal Duarte

2012-09 Non-market Recreational Value of a National Forest: Survey Design and Results
- Paula Simões, Luís Cruz & Eduardo Barata

2012-08 Growth rates constrained by internal and external imbalances and the role of relative prices: Empirical evidence from Portugal
- Elias Soukiazis, Pedro André Cerqueira & Micaela Antunes

2012-07 Is the Erosion Thesis Overblown? Evidence from the Orientation of Uncovered Employers
- John Addison, Paulino Teixeira, Katalin Evers & Lutz Bellmann

2012-06 Explaining the interrelations between health, education and standards of living in Portugal. A simultaneous equation approach
- Ana Poças & Elias Soukiazis

2012-05 Turnout and the Modeling of Economic Conditions: Evidence from Portuguese Elections
- Rodrigo Martins & Francisco José Veiga

2012-04 The Relative Contemporaneous Information Response. A New Cointegration-Based Measure of Price Discovery
- Helder Sebastião

- Elias Soukiazis, Pedro Cerqueira & Micaela Antunes

2012-02 As Ações Portuguesas Seguem um Random Walk? Implicações para a Eficiência de Mercado e para a Definição de Estratégias de Transaçao
- Ana Rita Gonzaga & Helder Sebastião

2012-01 Consuming durable goods when stock markets jump: a strategic asset allocation approach
- João Amaro de Matos & Nuno Silva

2011-21 The Portuguese Public Finances and the Spanish Horse
- João Sousa Andrade & António Portugal Duarte

2011-20 Fitting Broadband Diffusion by Cable Modem in Portugal
- Rui Pascoal & Jorge Marques
<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011-19</td>
<td>A Poupança em Portugal</td>
<td>Fernando Alexandre, Luís Aguiar-Conraria, Pedro Bação &amp; Miguel Portela</td>
</tr>
<tr>
<td>2011-17</td>
<td>The Portuguese Stock Market Cycle: Chronology and Duration Dependence</td>
<td>Vítor Castro</td>
</tr>
<tr>
<td>2011-16</td>
<td>The Fundamentals of the Portuguese Crisis</td>
<td>João Sousa Andrade &amp; Adelaide Duarte</td>
</tr>
<tr>
<td>2011-15</td>
<td>The Structure of Collective Bargaining and Worker Representation: Change and Persistence in the German Model</td>
<td>John T. Addison, Paulino Teixeira, Alex Bryson &amp; André Pahnke</td>
</tr>
<tr>
<td>2011-14</td>
<td>Are health factors important for regional growth and convergence? An empirical analysis for the Portuguese districts</td>
<td>Ana Poças &amp; Elias Soukiazis</td>
</tr>
<tr>
<td>2011-13</td>
<td>Financial constraints and exports: An analysis of Portuguese firms during the European monetary integration</td>
<td>Filipe Silva &amp; Carlos Carreira</td>
</tr>
<tr>
<td>2011-12</td>
<td>Growth Rates Constrained by Internal and External Imbalances: a Demand Orientated Approach</td>
<td>Elias Soukiazis, Pedro Cerqueira &amp; Micaela Antunes</td>
</tr>
<tr>
<td>2011-11</td>
<td>Inequality and Growth in Portugal: a time series analysis</td>
<td>João Sousa Andrade, Adelaide Duarte &amp; Marta Simões</td>
</tr>
<tr>
<td>2011-10</td>
<td>Do financial Constraints Threat the Innovation Process? Evidence from Portuguese Firms</td>
<td>Filipe Silva &amp; Carlos Carreira</td>
</tr>
<tr>
<td>2011-08</td>
<td>From Goal Orientations to Employee Creativity and Performance: Evidence from Frontline Service Employees</td>
<td>Filipe Coelho &amp; Carlos Sousa</td>
</tr>
<tr>
<td>2011-07</td>
<td>The Portuguese Business Cycle: Chronology and Duration Dependence</td>
<td>Vítor Castro</td>
</tr>
<tr>
<td>2011-06</td>
<td>Growth Performance in Portugal Since the 1960’s: A Simultaneous Equation Approach with Cumulative Causation Characteristics</td>
<td>Elias Soukiazis &amp; Micaela Antunes</td>
</tr>
<tr>
<td>2011-05</td>
<td>Heteroskedasticity Testing Through Comparison of Wald-Type Statistics</td>
<td>José Murteira, Esmeralda Ramalho &amp; Joaquim Ramalho</td>
</tr>
<tr>
<td>2011-04</td>
<td>Accession to the European Union, Interest Rates and Indebtedness: Greece and Portugal</td>
<td>Pedro Bação &amp; António Portugal Duarte</td>
</tr>
<tr>
<td>2011-03</td>
<td>Economic Voting in Portuguese Municipal Elections</td>
<td>Rodrigo Martins &amp; Francisco José Veiga</td>
</tr>
<tr>
<td>2011-02</td>
<td>Application of a structural model to a wholesale electricity market: The Spanish market from January 1999 to June 2007</td>
<td>Vítor Marques, Adelino Fortunato &amp; Isabel Soares</td>
</tr>
<tr>
<td>2011-01</td>
<td>A Smoothed-Distribution Form of Nadaraya-Watson Estimation</td>
<td>Ralph W. Bailey &amp; John T. Addison</td>
</tr>
</tbody>
</table>

A série Estudos do GEMF foi iniciada em 1996.