UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

DEPARTAMENTO DE ESTATÍSTICA

AN IMPLEMENTATION OF THE LHAR-CJ MODEL WITH FUNCTIONAL COEFFICIENTS

LEONARDO GABRIEL DA PAZ

Porto Alegre
2018

LEONARDO GABRIEL DA PAZ

AN IMPLEMENTATION OF THE LHAR-CJ MODEL WITH FUNCTIONAL COEFFICIENTS

Trabalho de Conclusão de Curso submetido como requisito parcial para a obtenção do grau de bacharel

Orientador Metodológico
Prof. Dr. Eduardo de Oliveira Horta

Porto Alegre
2018

Instituto de Matemática e Estatística
Departamento de Estatística

An implementation of the LHAR-CJ model with functional coefficients

Leonardo Gabriel da Paz

Banca examinadora:

Prof. Dr. Eduardo de Oliveira Horta
UFRGS

Prof. Dr. Hudson da Silva Torrent
UFRGS

An implementation of the LHAR-CJ model with functional coefficients

Leonardo Gabriel da Paz[a,*], Eduardo Horta[a]

[a] *Universidade Federal do Rio Grande do Sul, Department of Statistics, 9500 Bento Gonçalves Av., 43–111, Porto Alegre, RS, Brazil 91509-900.*

Abstract

This article aims to compare the forecast performance of the LHAR-CJ model, proposed in Corsi and Renò (2012) and a LHAR-CJ model with functional coefficients for a Vale return series. This new model, instead of estimating fixed coefficients for each variable in the autoregressive model, estimates a functional coefficient that is state dependent, where the state is represented by the lagged realized volatility. In other words, the coefficients are functions of the states of the response variable. We found out that, for this data, the functional coefficients model has a better forecast performance with the right smoothness parameter.

Keywords: realized volatility, volatility forecasting, continuous volatility, functional coefficients, leverage effect, jumps

1. Introduction

The relevance of the financial market has increased over the years, but its relevance comes with its complexity. Many articles and new models has been published trying to supply the demand from the market, exploring new approaches and estimation methods to improve the forecasting performance of models in general.

*Leonardo Gabriel da Paz
Email address: leonardogpaz@gmail.com (Leonardo Gabriel da Paz)

The HAR-RV (Heterogeneous Auto-Regressive with Realized Volatility) model, published in Corsi (2009), was designed specifically to have a better forecast performance than the previous models. The HAR-RV model uses the realized volality incorporating intraday returns to calculate, estimate, and predict the variability of the data. Originally, this model estimated the coefficients with the sum of the intraday squared returns. However, a new approach to calculate the volatility continuously was presented in Corsi and Renò (2012).

This new approach, the LHAR-CJ (Leverage Heterogeneous Auto-Regressive with Continuous volatility and Jumps) model, not just estimates de volatility continuously, but also incorporates leverage effect and jumps to the model. It allowed a better forecast performance than the usual HAR-RV model for the tested data (see Corsi and Renò (2012) for more details). However, even with that improvement, some research professionals started to publish new approachs to these models.

One example is the model proposed by Chen et al. (2017), that incorporates time-dependent functional coefficients in the original HAR-RV model. In this article, it is also possible to see a forecast performance improvement using the functional coefficients. In view of the positive forecasting impact brought by incorporating functional coefficients in HAR-RV type models, we decided to apply them in the LHAR-CJ model. However, in our approach the functional relationship is state-dependent rather than time-dependent.

In the next section, we explain the theoretical background used to implement the new model (basically an overview of the LHAR-CJ model and functional coefficients). Following, we describe how we actually implemented the model and what kind of computational resources and functions we used. In the same section, we analyze the results and compare the new model with the previous one. Section 4 concludes.

2. Theoretical Review

In this section, we explain the theoretical background to implement the model.

2.1. The LHAR-CJ Model

We herein shortly present the LHAR-CJ model, following Corsi and Renò (2012) closely. The reader is referred to this text for details.

Before presenting the model, we need to define some expressions. To calculate the variables, such as the continuous volatility and the jumps, we need to apply some transformations to the data (prices and returns). The two-scale realized volatility estimator is given by

$$\text{TSRV} = \frac{\left(\sum_{j=1}^{n-K}(X_{j+K} - X_j)^2 - \sum_{j=1}^{n-1}(X_{j+1} - X_j)^2\right)}{K\left(1 - \frac{n-K+1}{nK}\right)}, \tag{2.1}$$

where X_j is the intraday logarithmic price with $j = 1, ...n$, where n is the number of prices each day. K is the length in ticks of the subsampling interval (we used $K = 10$ following the bibliography).

To calculate the continuous volatility and the jumps, we also need the threshold bipower variation estimator introduced in Corsi et al. (2010) and given by

$$\begin{aligned} \text{TBPV}_t = & \frac{\pi}{2}\frac{M}{M-2}\sum_{j=0}^{M-2}|\Delta_{t,j}X| \cdot |\Delta_{t,j+1}X| \\ & \times I_{\{(\Delta_{t,j}X)^2 \leq \vartheta_{j-1}\}} I_{\{(\Delta_{t,j+1}X)^2 \leq \vartheta_j\}}, \end{aligned} \tag{2.2}$$

where $\Delta_{t,j}X$ is the jth intraday return of day t, with $j = 1, ..., M$. $I_{(\cdot)}$ in an indicator function and ϑ_j is a threshold function (see Corsi and Renò (2012)).

Given these estimators, we can define the continuous volatility element as $\widehat{C}_t = \text{TBPV}_t$, the jump effect as $\hat{J}_t = \max(\text{TSRV}_t - \text{TBPV}_t, 0)$ and the sum of both estimators as $\widehat{V}_t = \text{TBPV}_t + \text{TSRV}_t$. Now, consider r_t as the daily return open-to-close. We can define the leverage effect as $r_t^- = min(r_t, 0)$.

Following, we define the aggregated effects of continuous volatility, jumps and leverage, which are defined by an integer number h of days as follows

$$\log \widehat{V}_t^{(h)} = \frac{1}{h}\sum_{j=1}^{h} \log \widehat{V}_{t-j+1}, \log \widehat{C}_t^{(h)} = \frac{1}{h}\sum_{j=1}^{h} \log \widehat{C}_{t-j+1},$$

$$r_t^{(h)} = \frac{1}{h}\sum_{j=1}^{h} r_{t-j+1}, \hat{J}_t^{(h)} = \sum_{j=1}^{h} \hat{J}_{t-j+1},$$

Pay attention to the fact that the logarithmics on the right sides of the equations are just a symbolic representation.

Then, defining $d = 1$, $w = 5$ and $m = 22$, the one-step LHAR-CJ considering daily, weekly and monthly periods consists in the following equation

$$\begin{aligned} \log V_{t+1} = c &+ \beta^{(d)} \log C_t^{(d)} + \beta^{(w)} \log C_t^{(w)} + \beta^{(m)} \log C_t^{(m)} \\ &+ \alpha^{(d)} \log(1 + J_t^{(d)}) + \alpha^{(w)} \log(1 + J_t^{(w)}) \\ &+ \alpha^{(m)} \log(1 + J_t^{(m)}) + \gamma^{(d)-} r_t^{(d)-} + \gamma^{(w)-} r_t^{(w)-} \\ &+ \gamma^{(m)-} r_t^{(m)-} + \epsilon_{t+1} \end{aligned} \quad (2.3)$$

2.2. Functional Coefficients

Basically, models with functional coefficients estimate not only a number for each coefficient in the regression, but a function. In the time series scenario, that function may be time-dependent or space-dependent. If time-dependent, each coefficient will have the form of $\beta_j(t)$, and you can see an example of that in Chen et al. (2017). If space-dependent, each coefficient has the form of $\beta_j(X_t)$, where X_t is the state of a time series in the time t. In our case, the time series used as reference is the V_t, in other words, our coefficients will have the form of $\beta_j(\widehat{V}_t)$.

3. Methodology

The data used in this paper is the financial returns of Vale, from 2009 to 2017. The returns were obtained from the intraday prices, calculated from open-to-close. The prices data were treated using the Economática method.

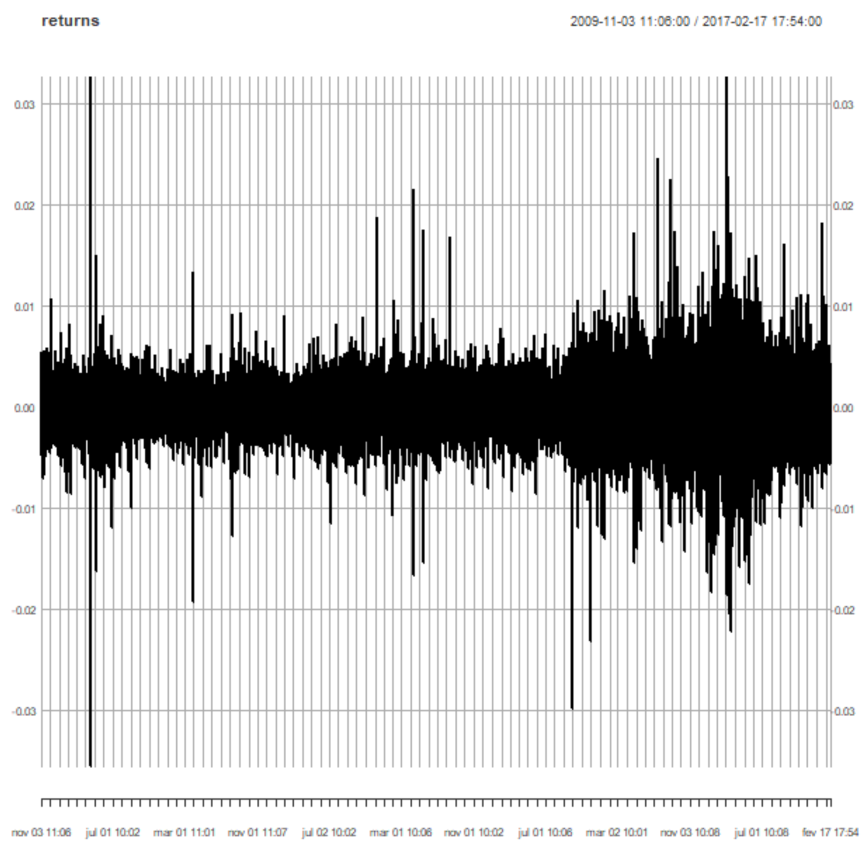

Figure 1: Graphic of the returns open-to-close of Vale

Beggining the analysis, we estimated all V_t, C_t, J_t and r_t^- components and fitted the LHAR-CJ model with and without functional coefficients. At this stage, we used all data available (1793 days and intraday returns calculated every minute). The numeric parameter to estimate the functional coefficient model was the lagged V_t (one-step), and the smoothness parameter chosen was $h = 3$. In Figures 2–4 below, we have the graphics, where the black curve represents the estimated functional coefficient and the red straight line represents the coefficients from the standard LHAR-CJ model.

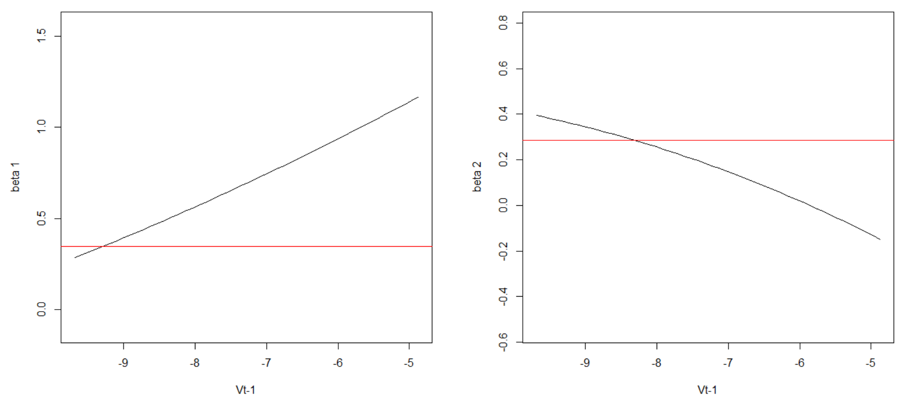

Figure 2: Graphic of β_1 and β_2

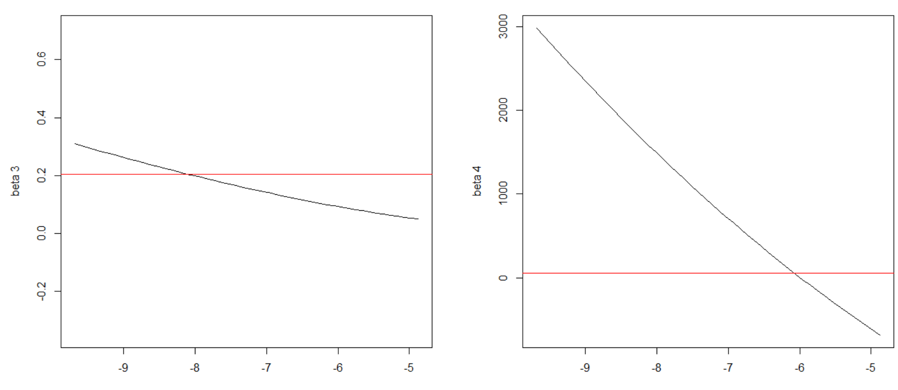

Figure 3: Graphic of β_3 and β_4

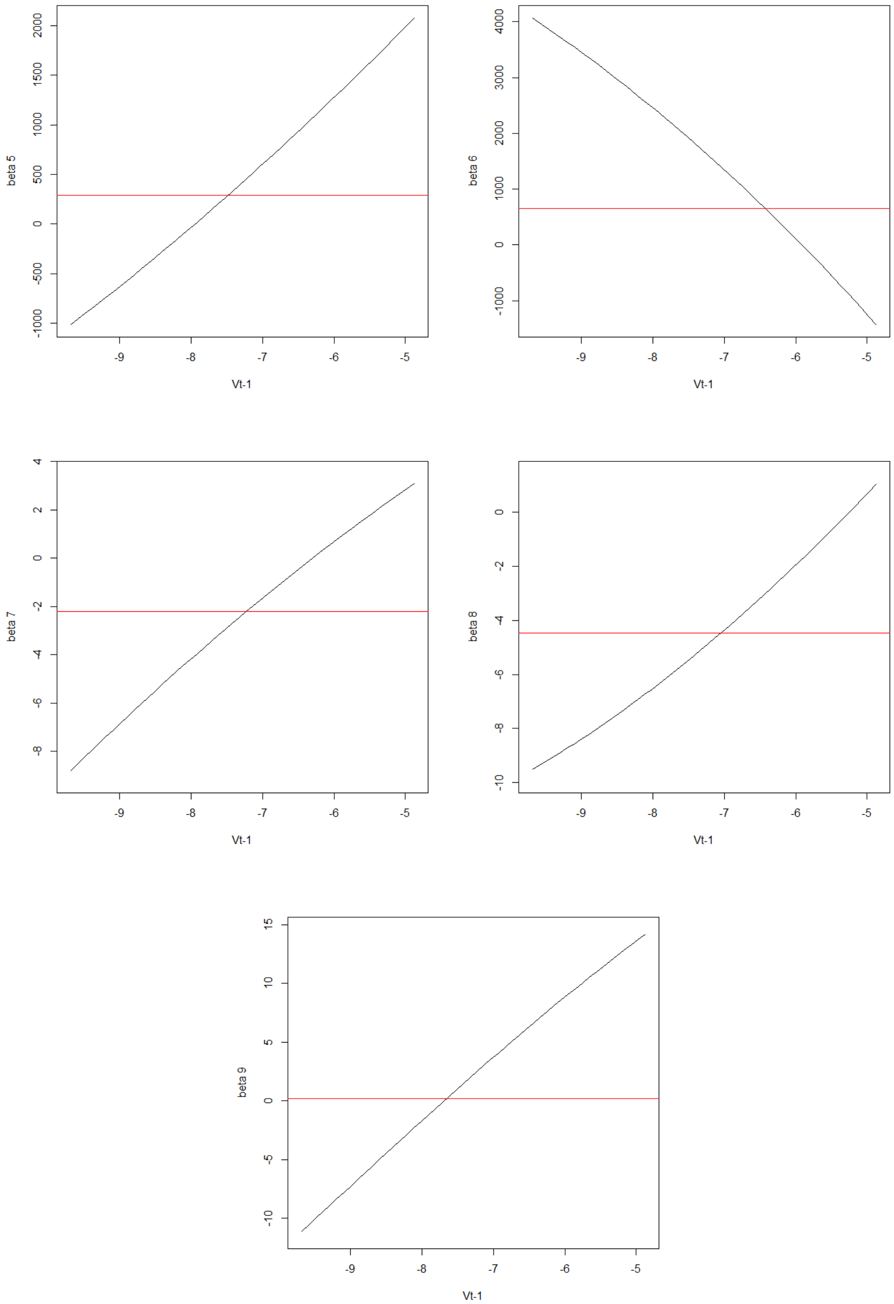

Figure 4: Graphic of β_5, β_6, β_7, β_8 and β_9.

As we can see, there is a big difference between the constant value and the functional coefficients, and that difference is reflected on the forecast results following next.

To do the forecast analysis, we generated a one-step forecast based on a expanding window going from the interval of days $1, ..., 1272$ to the interval of days $1, ..., 1769$. In other words, we estimated the model with the pre-set interval of days, then forecasted the next day. This was done for both LHAR-CJ models, with and without functional coefficients.

In the functional coefficients model, we set a grid for the smoothness parameter (h), going from 0.6 to 1.4 by 0.2, and then, realizing that the MSE (Mean Squared Error) was decreasing as the h increased, we expanded the grid, with a new part going from 3 to 15 by one. This grid may be refined in future work.

On the table below, we can se the MSE in all cases described above.

Table 1: Table of MSE

		LHAR-CJFC			
		h			
	LHAR-CJ	**0.6**	**0.8**	**1.0**	**1.2**
		0.1654552	0.1424039	0.1339300	0.1284922
		h			
		1.4	**3**	**4**	**5**
		0.1253136	0.1215540	0.1215564	0.1215987
		h			
		6	**7**	**8**	**9**
MSE	0.1288502	0.1216328	0.1216573	0.1216748	0.1216875
		h			
		10	**11**	**12**	**13**
		0.1216969	0.1217041	0.1217097	0.1217141

		h	
		14	**15**
		0.1217177	0.1217205

As we can see, with $h = 3$ we obtain the minimum MSE for the values of h tested, and its MSE is also less than the MSE of the model with constant coefficients.

Therefore, using the MSE criteria, we could obtain a better forecast performance with the model with functional coefficients for the presented data.

4. Conclusion

The aim of the study was to show that we can get a better forescast performance using functional coefficients and explore the volatility models area. Given that, we may say that the we achieved our goals, and the study may be useful to start other research projects about functional coefficients and volatility models.

5. Acknowledgements

The author Leonardo Gabriel da Paz wishes to thank Eduardo Horta for all his support and knowledge developing this paper.

References

Chen, X.B., Gao, J., Li, D., Silvapulle, P., 2017. Nonparametric estimation and forecasting for time-varying coefficient realized volatility models. Journal of Business & Economic Statistics , 1–13.

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

Corsi, F., Pirino, D., Reno, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. Journal of Econometrics 159, 276–288.

Corsi, F., Renò, R., 2012. Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. Journal of Business & Economic Statistics 30, 368–380.