The Conditional Heterocedasticity on the Argentine Inflation. An Analysis for the Period from 1943 to 2013

Juan Carlos Abril[1] and María De Las Mercedes Abril[2]

Universidad Nacional de Tucumán, Facultad de Ciencias Económicas and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Tucumán. Argentina.

[1]*jabril@herrera.unt.edu.ar, mabrilblanco@hotmail.com*[2]

Abstract. Numerous economic time series do not have a constant mean and in practical situations, we often see that the variance of observational error is subject to a substantial variability over time. This phenomenon is known as *volatility*. To take into account the presence of volatility in an economic series, it is necessary to resort to models known as *conditional heteroscedastic* models. In these models, the variance of a series at a given time point depends on past information and other data available up to that time point, so that a conditional variance must be defined, which is not constant and does not coincides with the overall variance of the observed series.

There is a very large variety of nonlinear models in the literature, which are useful for the analysis of any economic time series with volatility, but we will focus in analyzing our series of interest using *ARCH* type models introduced by Engle (1982) and their extensions . These models are non-linear in terms of variance.

Our objective will be the study of the monthly inflation data of Argentina for the period from January 1943 to December 2013. The data is officially published by the National Institute of Statistics and Censuses (or INDEC as it is known in Argentina). Although it is a very long period in which various changes and interventions took place, it can be seen that certain general patterns of behavior have persisted over time, which allows us to admit that the study can be appropriately based on available information.

Keywords: Inflation, heterocedasticity, volatility, time series.

Introduction

Numerous economic time series do not have a constant mean and in practical situations, we often see that the variance of observational error is subject to substantial variability over time. This phenomenon is known as *volatility*.

To take into account the presence of volatility in an economic series, it is necessary to resort to models known as *conditional heteroscedastic models*. In these models, the variance of a series at a given time point depends on past information and other data available up to that time point, so a conditional variance can be defined, which is not constant and does not coincide with the overall variance of the observed series.

An important feature of any economic time series is that they are not generally serially correlated but they are dependent. Thus linear models such as those belonging to the family of autoregressive moving average models or *ARMA* models may not be appropriate to describe these series.

The first inspection of series such as the one we will study suggests that they do not have a mean and a constant variance. A stochastic variable in which the variance is constant is said to be *homocedastic* as opposed to what would be a *heteroscedastic* variable. For those series where there is volatility, the unconditional variance can be constant even when the conditional variance in some periods is unusually large.

Our objective will be the study of the monthly inflation data of Argentina for the period from January 1943 to December 2013. The data is officially published by the National Institute of

Statistics and Censuses (INDEC). Although it is a very long period in which basic changes, basket changes and interventions took place even in the INDEC itself, it can be seen that certain general patterns of behavior have persisted over time, which allows us to admit that the study is adequately based on the information available.

Method

Research Model

Modeling volatility

Linear models of the *ARMA* type, for example, admit that the disturbances have zero mean and constant variance, usually equal to one (this is equivalent to saying that these disturbances are white noise). Under these conditions, the conditional variance given over the past history, that is given F_{t-1}, is constant over time.

There is a very large variety of nonlinear models available in the literature to address these situations, but we will focus on *ARCH* models or *autoregressive conditional heteroscedastic models* introduced by R. Engle (1982) and their extensions . These models are nonlinear in terms of variance.

In the analysis of non-linear model, *errors* (also called *innovations*, because they represent the new part of the series that cannot be predicted from the past) known as ε_t, are generally assumed IID and the model has the form

$$y_t = g(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) + \varepsilon_t h(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots)$$
$$= g_t + \varepsilon_t h_t = \mu_t + \varepsilon_t h_t, \qquad (1)$$

where $g(\bullet) = g_t = \mu_t$, represents the conditional mean and $h^2(\bullet) = h_t^2$ is the conditional variance.

***ARCH* type models**

The basic idea of this model is that a series y_t is not serially correlated but depends on past prices through a quadratic function.

An *ARCH(q)* model can be expressed as

$$y_t = \mu_t + \varepsilon_t h_t$$
$$\varepsilon_t \sim i.i.d\ D(0,1) \qquad (2)$$
$$\sigma^2 = h_t^2 = \omega + \sum_{i=1}^{q} \alpha\ z_{t-1}^2,$$

where $z_t = y_t - \mu_t$ and $D(\bullet)$ is a probability density function with zero mean and variance equal to one.

An *ARCH* model such as the one just described adequately describes volatility clustering. The conditional variance of y_t is an increasing function of the square of the shock that occurs at time $t-1$. Consequently, if y_t is sufficiently large in absolute value, σ_t^2 and thus y_t are expected to be also large in absolute value. It is necessary to take into account that even when the conditional variance in an *ARCH* type model varies with time, that is, $\sigma_t^2 = E(z_t^2|F_{t-1})$, the unconditional variance of z_t is constant and, given that $\omega > 0$ and $\sum_{i=1}^{q} a_i < 1,$ we have

$$\sigma_t^2 = E\{E(z_t^2|F_{t-1})\} = \frac{\omega}{1 - \sum_{i=1}^{q} a_i}. \qquad (3)$$

In most practical applications, excess kurtosis in an *ARCH* model, along with a normal distribution, is not enough to be able to explain what a dataset like ours. Therefore, we can make use of other distributions. For example, we can suppose that ε_t follows a *t* Student distribution with mean 0, variance equal to 1 and υ degrees of freedom, that is, ε_t is *ST (0,1, υ)*. In this case, the unconditional kurtosis for the *ARCH (1)* model is

$\lambda\left(1-\alpha_1^2\right)/\left(1-\lambda\alpha_1^2\right)$ where $\lambda = 3(\upsilon-2)/(\upsilon-4)$.

Due to the additional coefficient υ, the ARCH (1) model based on a *t* distribution will have heavier tails than one based on a normal distribution.

The calculation of σ_t^2 in (2) depends on past quadratic residuals, z_t^2, that are not observed for $t = 0, -1, \ldots, -q+1$. To initialize the process, unobserved quadratic residuals are set to a value equal to the sample mean.

GARCH type models

While Engle (1982) certainly made the greatest contribution to financial econometrics, *ARCH* type models are rarely used in practice because of their simplicity.

A good generalization of this model is found in the *GARCH* type models introduced by Bollerslev (1986). This model is also a weighted average of the past quadratic residuals. This model is more parsimonious than *ARCH* models, and even in its simplest form, has proven to be extremely successful in predicting conditional variances.

It should be noted that *GARCH* type models are not the only extension and there are at least twelve specifications related to them that will be the object of future research.

Generalized *ARCH* models (or *GARCH* models as they are also known) are based on an infinite *ARCH* specification and allow to reduce the number of parameters to be estimated by imposing nonlinear constraints on them. The *GARCH(p, q)* model is expressed as follows

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} a_i z_{t-1}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. \qquad (4)$$

As in the case of *ARCH* models, it is necessary to impose some restrictions on σ_t^2 to ensure that it is positive for all *t*. Bollerslev (1986) showed that ensuring that

$\omega > 0, a_i \geq 0$ (for $i = 1, \ldots, q$) and

$\beta_j \geq 0$ (for $i = 1, \ldots, p$) is sufficient to guarantee that the conditional variance is positive.

In terms of the estimation process, we can say that many authors have proposed using a Student *t* distribution in combination with a *GARCH* model to adequately model heavy tails on economic or financial time series whose data are of high frequency, which will be seen below.

Sample

First, we proceed to plot the series. Within the study of a series, graphic methods are an excellent way to begin an investigation.

After plotting it, we can immerse ourselves in a detailed study of the subject under consideration. Among the functions that comply with the graphs that we will present are the following:

1. They make the data under study more visible, systematized and synthesized.

2. They reveal their variations and their historical or spatial evolution.

3. They can show the relationships between the various elements of a system or a process and show indications of the future correlation between two or more variables.

In addition to this, the application of these methods suggests new research hypotheses and allows the subsequent implementation of statistical models ranging from the simplest to those that are much more refined, thus achieving a better analysis of the data and its fluctuations over time.

Section (a) of Figure 1 shows the monthly levels of the Consumer Price Index. In section (b) of the same figure we can see the first differences of the logarithm of the monthly IPC level. This is what is popularly known as *inflation* and will be the series object of our work.

After a careful inspection of this last section, we can see that there are periods where the volatility is low and can be confused with the presence of seasonality especially within the period from January 1943 until the end of 1974. Then, there is a clear a period of very important volatility until late 1977, which is repeated with similar characteristics between 1983 and the end of 1985. Subsequently, between 1987 and the end of 1992 we have a period of high volatility. From late 2001, the volatility in the series under study is almost null, and it almost disappears from 2009.

Figure 1: *(a) Monthly levels of the Consumer Price Index from January 1943 to December 2013. (b) First difference of the logarithm of the level of the Consumer Price Index from January 1943 to December 2013.*
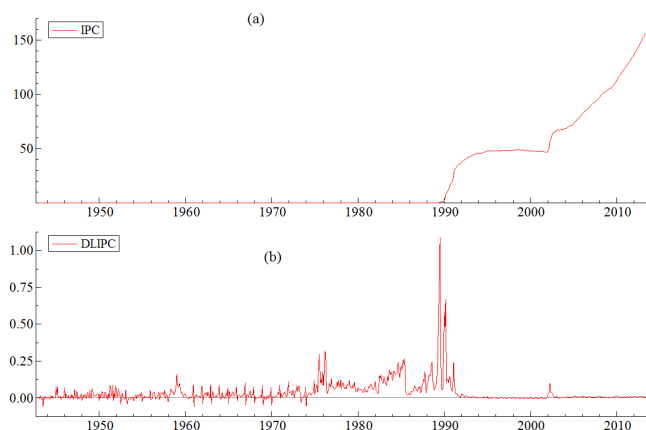

Figure 2 shows in section (a) the autocorrelation for the inflation series under study, also in section (b) we can see the partial autocorrelation function. Finally, in section (c), we have the estimated density function which is represented by a red line, compared to the normal density function which is represented by a green line.

From the study of this figure we see that the series is not stationary, although it has some seasonal components and its distribution is different from that of a normal variable, so it is possible that we must make their corresponding estimates using a *t* Student distribution.

Figure 2: *(a) Autocorrelation function of the inflation series for the period from January 1943 to December 2013. (b) Partial autocorrelation function of the same series under study. (c) Estimated density function compared to normal density (green line).*
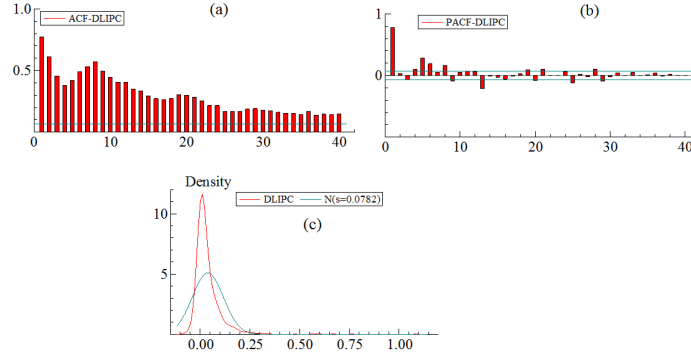
Data Analysis and Discussion

Different alternatives were tried with respect to the modeling of the inflation series in Argentina for the period between January of 1943 and December of 2013. After analyzing them and seeing the values of different goodness of fit statistics, like the Akaike Criterion, the Schwarz Criterion or the Hannan - Quinn Criterion, we are left with a model based on the equation (1), where

y_t is the series under study and its explicit specification is given by

$$y_t = \mu_t + \varepsilon_t h_t, \tag{5}$$

where ε_t has a t distribution with 2.21167 degrees of freedom. The conditional mean, μ_t is equal to a general mean μ a seasonal component and an *ARMA (1, 1)* process, which is explicitly

$$\mu_t = \mu + \gamma_t + \varphi\, y_{t-1} + v + \theta v_{t-1}, \tag{6}$$

where $v_t = \varepsilon_t h_t$ and γ_t is a seasonal component that satisfies the following condition

$$\gamma_t = -\gamma_{t-1} - \ldots - \gamma_{t-11}, \tag{7}$$

that is, its sum is equal to zero over the previous year. This is achieved with the introduction of adequate dummy variables, Besides that, the conditional variance in (5) is given by

$$h_t^2 = \sigma_t^2 = \omega + \alpha(y_{t-1} - \mu_{t-1})^2 + \beta\sigma_{t-1}^2, \tag{8}$$

that is, a *GARCH(1, 1)* with a constant given by ω. In the estimation process for our case, we could see that the constant was not significant different from zero, so we decided to remove it in the final formulation.

Figure 3: *Characteristics of the inflation series for the period between 1943 and 2013. (a) Inflation series. (b) Residues from the inflation series. (c) Quadratic residues. (d) Standardized residuals. (e) Conditional mean from the application of the volatility. (f) Conditional standard deviation. (g) Conditional variance. (h) Standardized residuals compared to a t distribution with 0, 1 and 2.21167 degrees of freedom.*
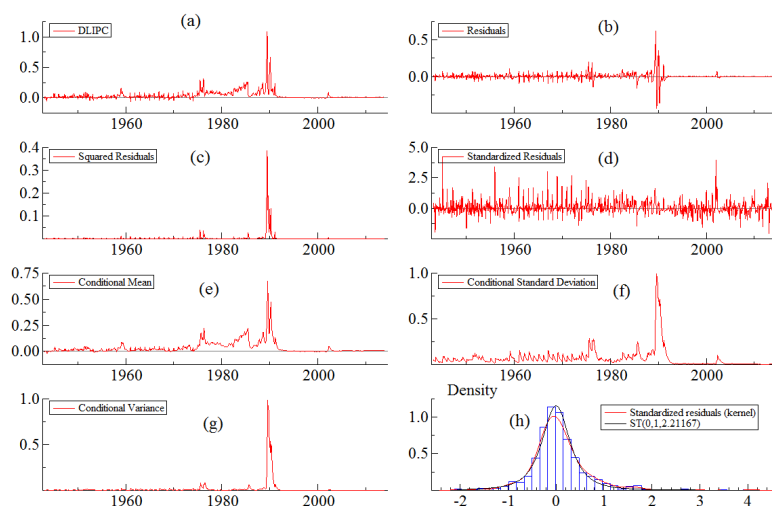

In the (a) the series of inflation again , in (b) we can see the residuals, in (c) we have the quadratic residuals while in (d) we have are the standardized residuals of the series. In section (e) of Figure 3 we show the estimated conditional mean for the proposed volatility model, in (f) we see the conditional standard deviation, while in (g) we observe the estimated conditional variance that arises from the application of the model to our series under study.

Looking closely at sections (a), (e), (f) and (g) of Figure 3, we can say that the model adequately explains the inflation series of our country for the period between 1943 and 2013. This same characteristic arises again in section (h) when comparing the distribution of the standardized residuals with a Student t distribution with 2.21167 degrees of freedom.

An interesting fact to note is that towards the end of the period under consideration, in particular from October or November 2004, the conditional variance or volatility of the series becomes practically insignificant.

Figure 4: *Prediction of the conditional mean versus the inflation series for the period between 1943 and 2013 (top). Conditional variance prediction for the model for the period between 1943 and 2013 (bottom).*
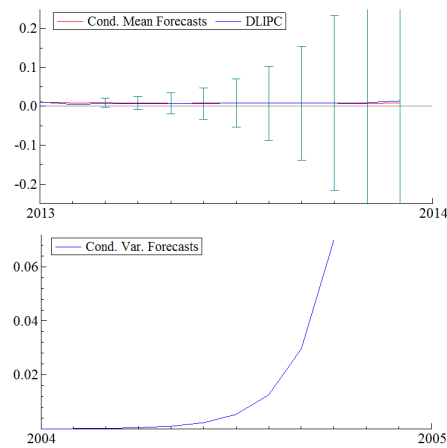

At the top of Figure 4 we can see the last ten observations of the series under study, highlighted in blue, and the corresponding predictions of the conditional mean. The vertical bars correspond to the 95% confidence interval that serve to compare the predicted value with the one we observed. In the lower graphic we tried to predict the conditional variance corresponding to the last ten observations of our series under study, however, this was not possible. As we have already expressed, the conditional variance, or volatility, as of October 2004 is equal to zero. This situation corroborates the fact that from that date no volatility predictions can be made for this series, which is in line with the beginning of a period of lack of confidence in official statistics. Statistically, this involved the prediction of values different from those that may have been in reality, which gave rise to an extremely smooth series that does not coincide with the rest of it nor with the reality lived.

Figure 5 is similar to Figure 4, but in the latter the bars corresponding to the confidence intervals were removed and the vertical axis scale was changed in order to have a better observation of the original series and the predicted conditional mean. One can clearly see a great coincidence between both Figures and the analysis that we can do is exactly the same for both cases.

Figure 5: *Prediction of the conditional mean versus the inflation series for the period between 1943 and 2013 (top).*
Conditional variance prediction for the model for the period between 1943 and 2013 (bottom).
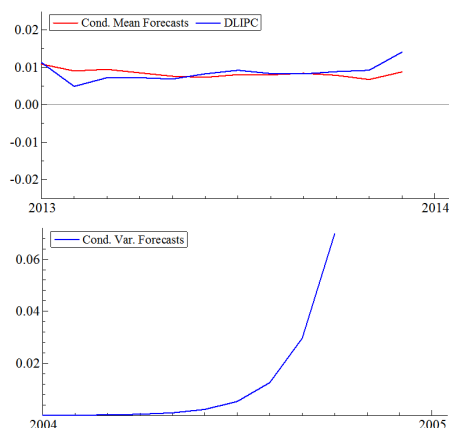
Final Remarks

In this initial stage of our investigation, we set out to analyze methods to treat a great variety of data with irregularities that happen in time series. Integrated autoregressive moving averages models (or *ARIMA* models) are often considered to provide the main basis for modeling in any time series. However, given the current state of research in time series research, there may be more attractive, and above all more efficient alternatives. Numerous economic time series do not have a constant mean and also in most cases phases are observed where relative calm reigns, followed by periods of important changes, that is, that the variability changes over time. Such behavior is what is called the *volatility*.

Among the models we have presented are the ones belonging to the *ARCH* family. The *ARCH* models or autoregressive models with conditional heteroscedasticity were first presented by Engle in 1982 with the aim of estimating the variance of inflation in Great Britain. The basic idea of this model is that y_t is not serially correlated but the conditional volatility or variance of the series depends on the past returns by means of a quadratic function. However this kind of models are rarely used in practice because of their simplicity. A good generalization of this model is found in the *GARCH* type models introduced by Bollerslev (1986). This model is also a weighted average of the last quadratic residuals, but it is more parsimonious than the *ARCH* type models and even in its simplest form has proven to be extremely successful in predicting conditional variances, so we decided to make use of them when working with our data.

Before applying any statistical method to the data under study it is fundamental to observe them graphically in order to be able to become familiar with them. This can have numerous benefits, as we have explained at the beginning of our analysis, as this process will serve as an indicator of ideas for a more detailed later study. This was the first step of our work in which we could see the main features of the series and served us to make an appropriate adjustment of it. Again it is necessary to emphasize that although this is a very long period to analyze; where there were numerous changes such as the base, and interventions in the INDEC, it is possible to make a very interesting study where the main characteristics of the series are appreciated.

We decided to fit an appropriate *GARCH* type model that captures the main characteristics of the data. We saw that it takes adequately to the volatility of the series but however, it presents some difficulties when making predictions. It remains for later work to analyze if we can use another specification that takes into account this fact that captures even more the characteristics of the series that can do a *GARCH* model like the one we have seen. With this analysis we have been able to estimate that from October 2004, volatility is equal to zero. This situation corroborates the fact that from that date no volatility predictions can be made for this series, which is consonance with the beginning of a period of lack of confidence in the official statistics. Statistically, this gave rise to an extremely smooth series that does not match the rest nor with the reality lived.

References

Abril, Juan Carlos. (1999), *Análisis de Series de Tiempo Basado en Modelos de Espacio de Estado*, EUDEBA: Buenos Aires.

Abril, Juan Carlos (2004). *Modelos para el Análisis de las Series de Tiempo*. Ediciones Cooperativas: Buenos Aires.

Abril, Maria de las Mercedes. (2014). *El Enfoque de Espacio de Estado de las Series de Tiempo para el Estudio de los Problemas de Volatilidad*. Tesis Doctoral. Universidad Nacional de Tucumán. Argentina.

Baillie, R. T. y Bollerslev, T. (1989). The Message in Daily Exchange Rates: A Conditional-Variance Tale. *Journal of Business and Economic Statistics*, **7**, 297-305.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, **31**, 307-27.

Bollerslev, T. (1987). A Conditionally Heteroscedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics*, **69**, 542-547.

Bollerslev, T., Chou, R. Y. y Kroner, K. F. (1992). *ARCH* Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, **52**, 5-59.

Bollerslev, T. y Wooldridge, J. M. (1992). Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances. *Econometric Reviews*, **11**, 143-172.

Box, G. E. P. y Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control* (Revised edition), Holden-Day Inc.: San Francisco.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of the United Kingdom Inflation. *Econometrical*, **50**, 987-1007.

Goldfarb, D. (1970). A Family of Variable Metric Updates Derived by Variational Means. *Mathematics of Computation*, **24 (109)**, 23-26.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press: Cambridge.

Hsieh, D. A. (1989). Modeling Heteroskedasticity in Daily Foreign Exchange Rates. *Journal of Business and Economic Statistics*, **7**, 307--317.

Pagan, A. (1996). The Econometrics of Financial Markets. *Journal of Empirical Finance*, **3**, 15-102.

Palm, F. C. (1996). *GARCH* Models of Volatility. En *Handbook of Statistics,* edited by G. Maddala, y C. Rao, 209-240. Elsevier Science, Amsterdam.

Palm, F. C. y Vlaar, P. J. G. (1997). Simple Diagnostics Procedures for Modeling Financial Time Series. *Allgemeines Statistisches Archiv*, **81**, 85-101