

Intraday-patterns in the Colombian Exchange Market Index and VaR: Evaluation of Different Approaches

Patrones del IGBC y valor en riesgo: evaluación del desempeño de diferentes metodologías para datos intra-día

JULIO CÉSAR ALONSO-CIFUENTES^a, MANUEL SERNA-CORTÉS^b

CIENFI - DEPARTAMENTO DE ECONOMÍA, FACULTAD DE CIENCIAS ADMINISTRATIVAS Y ECONÓMICAS, UNIVERSIDAD ICESI, CALI, COLOMBIA

Abstract

This paper evaluates the performance of 16 different parametric, non-parametric and one semi-parametric specifications to calculate the Value at Risk (VaR) for the Colombian Exchange Market Index (IGBC). Using high frequency data (10-minute returns), we model the variance of the returns using GARCH and TGARCH models, that take in account the leverage effect, the day-of-the-week effect, and the hour-of-the-day effect. We estimate those models under two assumptions regarding returns' behavior: Normal distribution and t distribution. This exercise is performed using two different ten-minute intraday samples: 2006-2007 and 2008-2009. For the first sample, we found that the best model is a TGARCH(1,1) without day-of-the week or hour-of-the-day effects. For the 2008-2009 sample, we found that the model with the correct conditional VaR coverage would be the GARCH(1,1) with the day-of-the-week effect, and the hour-of-the-day effect. Both methods perform better under the t distribution assumption.

Key words: Leverage, Finance, GARCH model, Risk estimation, Stock returns.

Resumen

El documento evalúa el desempeño de 16 métodos paramétricos, uno no paramétrico y uno semiparamétrico, para estimar el VaR (Valor en Riesgo) de un portafolio conformado por el Índice General de la Bolsa de Valores de Colombia (IGBC). El ejercicio se realiza analizando dos muestras de datos intra-día con una periodicidad de 10 minutos para los períodos 2006-2007 y 2008-2009. Los modelos paramétricos evaluados consideran la presencia o no

^aProfessor. E-mail: jcalonso@icesi.edu.co

^bResearch Assistant. E-mail: mserna.cortes@gmail.com

de patrones de comportamiento, tales como: el efecto “Leverage”, el efecto día de la semana, el efecto hora y el efecto día-hora. Nuestros resultados muestran que para la primera muestra el mejor modelo es un TGARCH(1,1) sin el efecto día de la semana ni la hora del día y bajo el supuesto de una distribución t . Para la segunda muestra, 2008-2009, el método que presenta el mejor comportamiento corresponde al modelo GARCH(1,1), que tiene en cuenta el efecto del día y la hora. Estos dos modelos presentan una correcta cobertura condicional y menor función de pérdida.

Palabras clave: apalancamiento, estimación de riesgo, finanzas, GARCH, rendimientos financieros,.

1. Introduction

On January 5, 2007 the Colombian Stock Market Index (IGBC, from the Spanish acronym) dropped 3.1% within the first ten minutes after the stock market opened. Such a drop in the stock market had never occurred before, and it would only happen again in the year 2009. At the time of closing that day, the IGBC had bounced back to such an extent that the overall index loss was 2.1% for that day. This means that the IGBC was down 334 points for the first ten minutes of trade, but then it recouped during the course of the day with a cumulative overall loss of 280.8 points at the end of the day. A financial analyst who only keeps track of information about the index at closing will probably come to the conclusion that it was a relatively ordinary day. That, however, was not just an ordinary day for traders. The kind of risk that materialized during the first ten minutes of trade that day would have gone unnoticed if an analyst had focused on a daily time horizon.

In fact, the behavior of the IGBC during the course of a day seems to follow a relatively stable pattern which can be taken into consideration to improve risk measures. This paper is aimed to illustrate how the involvement of previously documented behavior patterns can be used for improving the performance of risk measures.

It is a hard fact that making decisions in financial markets is exposed to different sources of risk. Hence, there is a need to acknowledge the importance of measuring risk and developing techniques that allow to make improved decisions considering the market circumstances. After the instability episodes and financial crises in the 1980s¹ and 1990s², measuring financial risk has become a routine everyday task in the “back office” of financial institutions (see Alonso & Berggrun 2008). It appears, moreover, that the financial crisis in 2008 led risk management to become the center of discussion again. The reliability of methods such as the Value-at-Risk (VaR) approach is part of the discussion where na reun-

¹For example, the external debt crisis in most Latin American countries in the 1980s or the collapse of the New York Stock Exchange in 1987.

²For example, the burst of the financial and real state bubbles in Japan in the 1990s and of the dot com companies in the late 1990s, the Mexican “Tequila” crisis in 1994, the financial crisis in Southeast Asia in 1997, the Russian crisis in 1997, and the Argentinean crisis in 1998.

derstanding of the limitations of this approach to detect risk in the latest financial crisis has raised great interest among academics, regulators, and the mass media.

VaR is a measure (an estimate) of the largest potential loss for a given time horizon and a given significance level under circumstances which are considered “normal” in the market. VaR is, without question, the most popular measure of financial risk among regulators, financial market stakeholders, and academics. Yet, despite its popularity, the recent financial crisis in 2008 made evident the limitations of this risk measurement tool. In fact, at the onset of the crisis, the collective conscience in the financial community seemed to concur that one of the major culprits of the financial crisis was, undoubtedly, excessive reliance on the VaR (Nocera 2009)³. Such reliance actually meant that the appropriation of the meaning and interpretation of the VaR disappeared at some point in time along with the bubble. The market players believed that mathematical and statistical models would be sufficient for managing risk and forgot that VaR was only one of the components of the analysis. Although it was a good way to measure risk, it still had limitations as to estimating it and incorporating other kinds of risks such as liquidity and systemic risk.

During the mortgage bubble, the easy earnings derived at a risk that had been transformed into mathematical certainty, which made people overlook the true meaning of the VaR. Agents forgot that the VaR was only intended to describe what occurred 99% of the times. A VaR of USD 25,000, for example, implied that this amount was not only the most one could lose 99% of the times, but also the least one could lose 1% of the times. It was precisely this 1% where analysts had to link other analyses to incorporate the quantification of liquidity risk or scenarios where the economy would go into a recession and portfolio diversification (systemic risk) had little importance. It was maybe losing sight of this 1% of the times what allowed the bubble to go on for such a long time. Financial entities focused on minimum-risk low-yield investments (VaR), but when they lost that 1% of the times, they did so in a disproportional matter.

After the storm associated with the financial crisis, it now seems clear to both academics and financial analysts that risk measures were not responsible for the crisis; what failed was judgment on the part of the individuals who interpreted these numbers (for a documented discussion of this issue, see Nocera (2009)). It is also evident that these measures, especially VaR, should not fall into oblivion, but should, on the contrary, be polished. Although VaR has some major limitations, incorporating these limitations to the analyses allows for more effective risk management 99% of the times-disregarding them would be going to the extreme of absolute risk aversion.

Consequently, the latest financial crisis is not the end of risk measurement. It is, on the contrary, a wake-up call to encourage reaching a deeper understanding of it, particularly of how it is calculated and interpreted. It is precisely at this point where our work is geared tow and illustrating how the incorporation of previously

³Nocera (2009) examines the issue of excessive reliance on the VaR as a result of a lack of an understanding of this measure as a supporting tool for analyzing risk.

documented behavior patterns can be used for improving the performance of risk measures.

It is important to acknowledge that the calculation of the VaR typically follows a simple and intuitive concept, but estimating it poses some practical difficulties. This kind of measure is difficult to estimate because it requires knowledge of the future value distribution of the asset or portfolio being reviewed. In most approaches, the distribution function is not directly estimated. A distribution is assumed for which parameters are calculated for the first moment (mean) and second moment around the mean (variance). Therefore, in practice, various approaches to its calculation take into account considerations that go from assuming normal distribution with constant variance or yields to assume other kinds of distribution and to allow the variance to be updated on a period-after-period basis.

Regardless of the kind of approach used for calculating VaR, a daily time horizon is the most common way to estimate it. This customary calculation of the VaR on a daily basis is partly due to the need to report this number to the regulatory agencies. Nevertheless, in recent years the calculation of the VaR for shorter periods of time has become increasingly popular because of two factors. Firstly, there is a need to have information regarding the risk associated with their business⁴ on the part of the stakeholders involved in the market on minute-to-minute basis. On the other hand, there is an increasing availability of intraday information and a widespread use of statistical methods and computer capabilities for processing this information.

The purpose of this work is to evaluate the performance of various approaches to estimate VaR for the following ten minutes. As far as the authors are aware, these kinds of exercises for the Colombian case have not been published in the past. To accomplish this objective, VaR is calculated using different approaches, including parametric, non-parametric, and semi-parametric approaches, and a portfolio that replicates the Colombian Stock Market General Index (IGBC, from its Spanish acronym). To this end, it is essential to acknowledge that the calculus of the VaR entails predicting the performance of the conditional distribution of the portfolio for the following period. A conditional distribution can be different from one day of the week to another (see Alonso & Romero 2009) and even within the same day (see Alonso & García 2009). Therefore, we use VaR models that capture both the weekday effect and the hour-effect as well as the most commonly discussed (Alonso & Arcos 2006) stylized facts such as, the volatility clustering and the heavy tails of the distribution of returns.

This paper is organized as follows: The first section provides a brief introduction and the second provides a brief discussion of the calculations and the evaluation of the Value at Risk in this exercise. The third section addresses the kind of data to be used and the necessary considerations for estimating models using intraday data. The fourth section summarizes the results obtained, and finally, the last section presents the final comments.

⁴See Giot (2000) for an extensive discussion of the reasons for the usefulness of calculating VaR for short time periods.

2. Calculation and Evaluation of Value at Risk

As mentioned above, the concept behind Value at Risk (VaR) is very straightforward and intuitive; these characteristics make this technique very popular. Nevertheless, despite its conceptual simplicity, its calculation poses a relatively sophisticated statistical problem. VaR is intuitively defined as the maximum loss expected from a portfolio with a certain confidence level in a given period of time (see, for example, Alonso & Berggrun 2008).

Formally speaking, the VaR for the following trading period $t + 1$ given the information available in the current period t ($VaR_{t+1|t}$) is defined as:

$$P(Z_{t+1} < VaR_{t+1|t}) = \alpha \quad (1)$$

whereas Z_{t+1} stands for future yield (in Colombian pesos) of the portfolio value for the following period and $(1 - \alpha)$ is the level of confidence of the VaR. Therefore, the calculation of the VaR depends on the assumptions regarding the function of distribution of potential losses or gains (absolute yield) from the portfolio Z_{t+1} .

It can be easily proved that if Z_{t+1} follows a distribution with its first two finite moments (such as a normal or a t-distribution), then the value at risk will be as follows:

$$VaR_{t+1|t} = F(\alpha) \cdot \sigma_{t+1} \quad (2)$$

whereas σ_{t+1} stands for the standard deviation of the distribution of Z_{t+1} , and $F(\alpha)$ represents the α percentile of the corresponding (standardized) distribution. Thus, the calculation of the VaR critically depends on two assumptions regarding the behavior of the distribution of Z_{t+1} : (i) its volatility (standard deviation σ) and (ii) its distribution $F(\cdot)$.

As noted earlier, there are several methodological approaches to estimate VaR. These approaches can be classified in three large groups: (i) historical simulation or the non-parametric approach, which does not assume a distribution and does not require estimating parameters; (ii) a parametric approach which involves assuming a distribution and estimating a set of parameters; and (iii) a semi-parametric approach which involves techniques that combine estimating parameters and using the non-parametric approach, such as, for example, filtered historical simulation.

In general, the results obtained after using the various kind of approaches are different and, in each case, the adequacy of these models must be evaluated on an individual case basis. On the other hand, backtesting the performance of an approach to the calculation of the VaR is not an easy task either. A description of the various types of approaches used here for estimating VaR is provided below, including a discussion of the methods to be used for evaluating the performance of the various approaches.

2.1. Some Approaches for the Estimation of VaR

The most common non-parametric approach is the historical simulation. This kind of approach involves determining the α percentile based on historical data.

In other words, this method assumes that past realizations of yield values from the portfolio represent the best approximation of the portfolio yield distribution for the following period. Therefore, $VaR_{t+1|t}$ will be equal to the α percentile of historical portfolio yield values. This approximation will be named as the specification 1.

On the other hand, any parametric approach involves assuming a given distribution function $F(\cdot)$ and the behavior of the parameter that characterizes it, e.g., σ_{t+1} . A well-documented fact of yields on assets is the presence of clustered variance (volatility clustering) (for a discussion of this stylized fact for the Colombian case, see, for example, Alonso & Arcos (2006)). This means, in other words, that volatility is not constant and, therefore, σ will be a function of time. Taking into account this stylized fact, the VaR of a portfolio can then be estimated using the following expression:

$$VaR_{t+1|t} = F(\alpha) \cdot \sigma_{t+1|t} \quad (3)$$

where $\sigma_{t+1|t}$ stands for the standard deviation for period $t+1$ subject to the information available in period t . Thus, it will be necessary to model the conditional variance in order to obtain a one-step-ahead forecast and to assume a distribution for calculating VaR.

Following Alonso & García (2009), we will use ten different approximations in this exercise to estimate the behavior of the variance⁵, and for each parametric approach, a model is estimated assuming a normal distribution and a t-distribution⁶.

In our case, we will consider eight parametric specifications of the GARCH model. Specification 2 reflects the GARCH(p,q) model proposed by Bollerslev (1986). We will particularly use the GARCH(1,1) model, which, as suggested by Brooks (2008), is usually sufficient to capture the clustered volatility phenomenon. Hence, specification 2 can be represented as follows:

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 z_t^2 \quad (4)$$

where z_t stands for the error in the mean equation, and α_0 , α_1 and α_2 are non-negative. Additionally, a necessary and sufficient condition for the variance generating process to be stationary is that $\alpha_1 + \alpha_2 < 1$.

Specification 3, which was proposed by Berument & Kiyamaz (2003), among others, incorporates dummy variables in the GARCH(1,1) model, capturing the effect of each day on the volatility of returns⁷. In this case, the variance has the following behavior:

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 z_t^2 + \sum_{i=1}^4 \beta_i D_{it} \quad (5)$$

⁵The mean is modeled using an autoregressive moving average (ARMA) process, particularly an ARMA(1,1) model that was selected based on the Akaike, Schwarz and Hannan-Quinn information criteria.

⁶A Student's t distribution is assumed in order to take into account a possible heavy tail behavior in the yield distribution, which is another stylized fact of yields from a portfolio or an asset (Alonso & Arcos 2006).

⁷These dummy variables are also incorporated into the mean equation.

where D_{1t} equals one, if day t is a Monday or, otherwise, zero. $D_{2t} = 1$ if t is a Tuesday or zero otherwise, and so forth, respectively. Summarizing, D_{it} are the dummy variables for the first four days of the week.

The day of the week effect (DOW) has been documented in several countries. The findings of recent studies such as Mittal & Jain (2009), for the case of India, and, Kamath & Chinpiao (2010) for the case of Turkey, have shown that this effect exists in emerging markets. For the Colombian case, Alonso & Romero (2009) and Rivera (2009) have shown that this effect is present in the volatility of the IGBC and the Colombian peso-US dollar exchange rate.

Following Giot (2000), specification 4 considers the hour of the day effect (HOD) in our GARCH(1,1) model for variance. There are a fairly good number of studies available in financial literature documenting a U-shaped behavior of volatility and returns within a day. Panas (2005) presented an extensive bibliographic review that documents the presence of this effect in various financial markets worldwide. This specification is aimed at capturing intraday behavior by means of dummy variables. Taking into account that the trading hours at the Colombian Stock Exchange run from 9:00 am to 1:00 pm in the periods being reviewed, this means that there are, in total, four hours for trading. Hence, the hour of the day effect is incorporated into the model using three dummy variables for time,

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1\sigma_t^2 + \alpha_2z_t^2 + \sum_{i=1}^3 \beta_i H_{it} \tag{6}$$

where H_{it} takes the value of one, if t equals trading hour i , or zero otherwise.

The fifth specification to be considered incorporates both the day of the week effect and the hour of the day effect into the GARCH(1,1) model. Dummy variables are included, which take the value of one taking into account time and day. All together, $(4 \times 5) - 1 = 19$ dichotomous variables are used. In this case, the variance is modeled as follows:

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1\sigma_t^2 + \alpha_2z_t^2 + \sum_{i=1}^5 \sum_{j=1}^4 \varphi_{ij} D_{it} H_{jt} - \varphi_{54} D_{5t} H_{4t} \tag{7}$$

The GARCH(1,1) specifications considered so far do not capture one of the common stylized facts observed for yields: The leverage effect. The leverage effect reflects the trend in volatility of having a greater increase when the price drops than when the price rises to the same extent. The TGARCH model (GARCH model with a threshold)⁸ is customarily used in financial literature for capturing the asymmetric behavior of volatility. Thus, specification 6 corresponds to a TGARCH model. This model matches the model proposed by Glosten, Jagannathan & Runkle (1993), which can be expressed as follows:

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1\sigma_t^2 + \alpha_2z_t^2 + \alpha_3d_tz_t^2 \tag{8}$$

where $d_t = 1$ if $z_t < 0$ and $d_t = 0$ if $z_t > 0$. It can also be expected that if the leverage effect is present, then $\alpha_3 > 0$. The condition for obtaining non-negative

⁸This model is also known as the GARCH GJR model.

variances will continue to be that α_0 , α_1 and α_2 must be non-negative values. Another condition is that $\alpha_1 + \alpha_3 \geq 0$. On the other hand, the model continues to be admissible if $\alpha_3 < 0$, provided that $\alpha_1 + \alpha_3 \geq 0$. A necessary and sufficient condition for the variance generating process to be stationary is $\alpha_1 + \alpha_2 < 1$.

Specifications 7, 8, and 9 incorporate the day of the week effect, the hour of the day effect, and the day of the week and hour of the day effect, respectively, into the TGARCH model.

TABLE 1: Summary of specifications of the models used in this exercise.

Specification	Acronym	Model
1	HS	Historical simulation
2	GARCH	$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 z_t^2$
3	GARCH + DOW	$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 z_t^2 + \sum_{i=1}^4 \beta_i D_{it}$
4	GARCH + HOD	$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 z_t^2 + \sum_{i=1}^3 \beta_i H_{it}$
5	GARCH + DOW + HOD	$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 z_t^2$ $+ \sum_{i=1}^5 \sum_{j=1}^4 \varphi_{ij} D_{it} H_{jt} - \varphi_{54} D_{5t} H_{4t}$
6	TGARCH	$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 z_t^2 + \alpha_3 d_t z_t^2$
7	TGARCH + DOW	$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 z_t^2 + \alpha_3 d_t z_t^2 + \sum_{i=1}^4 \beta_i D_{it}$
8	TGARCH + HOD	$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 z_t^2 + \alpha_3 d_t z_t^2 + \sum_{i=1}^3 \beta_i H_{it}$
9	TGARCH + DOW + HOD	$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 z_t^2 + \alpha_3 d_t z_t^2$ $+ \sum_{i=1}^5 \sum_{j=1}^4 \varphi_{ij} D_{it} H_{jt} - \varphi_{54} D_{5t} H_{4t}$
10	FHS	Filtered historical simulation

Note: DOW: Day of the week effect, HOD: Hour of the day effect.

d_t is defined as follows: $d_t = 1$ if $z_t < 0$ and $d_t = 0$ if $z_t > 0$.

H_{it} are the dummy variables for the first three hours of trading at the stock exchange.

D_{it} are the dummy variables for the first four days of the week.

Lastly, a semi-parametric approach, an average-filtered historical simulation is considered. This approach makes it possible to filter an autocorrelation of yields. Thus, yields are filtered by using an ARMA (p,q)⁹ process. The estimated residual will be used to perform a historical simulation as described above. This approach has an advantage over historical simulation since it provides an empirical function of density that is more “realistic” in capturing autocorrelation (see, for example,

⁹As described above, in this case an ARMA(1,1) model will be used.

Dowd 2005). A summary of the specifications described above is presented in Table 1.

2.2. Approaches to Assess the Estimated VaR Models

The fit of our models is assessed based on two backtesting or calibration tests and one loss-function criterion. The difficulty in assessing a calculation of VaR lies in the fact that the VaR cannot be observed directly. If one wants to conduct an assessment outside of the sample, in practice there is only information about the realization of yield for the following period, but information about the realization of VaR for that period will not be available.

The most commonly used test available in the literature is Kupiec's (1995) proportion of failures. The purpose of this test is to determine whether the observed proportion of losses that exceed the VaR (also known as proportion of failures) is consistent with the theoretical proportion of failures which provided the basis for constructing the VaR. In other words, the model must provide (non-conditional) coverage in constructing the VaR. In particular, according to the null hypothesis that our model has a "good fit", the n number of failures follows a binomial distribution¹⁰. In general, considering a total number of observations N and a theoretical level of proportion of failures equal to α (significance level), the probability of observing n losses is calculated as follows:

$$P(n|N, \alpha) = \binom{N}{n} \alpha^n (1 - \alpha)^{N-n} \quad (9)$$

To test the null hypothesis that the proportion of failures (ρ) is the same as the theoretically expected value (α) ($H_0 : \rho = \alpha$), Kupiec (1995) suggested the following statistic:

$$t_U = \frac{\hat{\rho} - \alpha}{\sqrt{\hat{\rho}(1 - \hat{\rho})/N}} \quad (10)$$

where $\hat{\rho}$ is the observed proportion of failures. Kupiec (1995) demonstrated that t_U follows a t distribution with $N - 1$ degrees of freedom.

Christoffersen (1998) suggested a test which considers that the calculation of VaR for $t + 1$ represents a forecast subject to the information available in period t . Then, VaR provides coverage subject to the information available at t and, therefore, the backtest should take this into consideration.

The idea underlying this test is that if the best model of VaR is being used, then using all information available at the time of predicting VaR, one should not be able to predict whether the VaR value was exceeded or not. This means that the observed number of failures must be random over time. Thus, a risk model will be said to have suitable non-conditional coverage if the probability of failure equals ρ , i.e. $P(PL_{t+1} > VaR^p_{t+1}) = \rho$.¹¹ A risk model will be said to have correct conditional coverage if $P_t(PL_{t+1} > VaR^p_{t+1}) = \rho$.

¹⁰A random variable is defined which takes the value of one if a loss is greater than the VaR or zero otherwise.

¹¹Where PL_{t+1} stands for the portfolio loss in period $t + 1$.

Therefore, having correct non-conditional coverage means that a model has failures with a probability of ρ on average as the days go by. Having correct conditional coverage, on the other hand, means that the model has failures with a probability of ρ every day, given all the information available on the previous day. It must be noted that correct non-conditional coverage is a necessary, yet not sufficient, condition for correct conditional coverage.

Christoffersen's (1998) idea entails separating specific forecasts being tested and then testing each individual forecast separately. The first is the equivalent of examining whether the model generates a correct proportion of failures, i.e., whether it provides correct non-conditional coverage. The latter implies to test that observed failures are statistically independent from each other. This means that failures should not cluster over time. Evidence of such clustering would mean that the model specification is not correct, even if the model meets the non-conditional coverage requirement.

Given that the theoretical probability of failures is α , Christoffersen (1998) suggests a test that can be expressed in terms of a likelihood ratio (LR) test. Under the null hypothesis of correct non-conditional coverage, the test statistic will be as follows:

$$LR_{uc} = -2 \ln[(1 - \alpha)^{N-n} \alpha^n] + 2 \ln[(1 - \rho)^{N-n} \rho^n] \quad (11)$$

This statistic follows an χ_1^2 distribution. Coming back to the independence test, let n_{kl} be the number of days on which status l occurs at t after status k occurred a $t - 1$, where the status refer to failures or non-failures. Besides, let π_{kl} be the probability that status l occurs for any t , given that the status at $t - 1$ was k . Under the null hypothesis of independence, the test statistic is as follows:

$$LR_{ind} = -2 \ln[(1 - \hat{\pi}_2)^{n_{00} + n_{11}} \hat{\pi}_2^{n_{01} + n_{11}}] + 2 \ln[(1 - \hat{\pi}_{01})^{n_{00}} \hat{\pi}_{01}^{n_{01}} (1 - \hat{\pi}_{11})^{n_{10}} \hat{\pi}_{11}^{n_{11}}] \quad (12)$$

This statistic also follows an χ_1^2 . Additionally, the estimated probabilities are defined as follows:

$$\hat{\pi}_{01} = \frac{n_{11}}{n_{00} + n_{01}}, \quad \hat{\pi}_{11} = \frac{n_{11}}{n_{10} + n_{11}}, \quad \hat{\pi}_2 = \frac{n_{01} + n_{11}}{n_{00} + n_{10} + n_{01} + n_{11}} \quad (13)$$

Overall, under the combined hypothesis of correct coverage and independence –i.e., the hypothesis of correct conditional coverage– the test statistic is as follows:

$$LR_{cc} = LR_{uc} + LR_{ind} \quad (14)$$

which follows an χ_2^2 distribution. Thus, Christoffersen's (1998) test allows testing the hypothesis of coverage and independence concurrently. It also tests these hypotheses separately, making possible to identify where the model is failing.

Meanwhile, López (1998) proposes a different approach to evaluate the behavior of VaR using a utility function for selecting the best model based on a set of models that meet the correct conditional coverage requirement. López's (1998)

loss function considers the number of failures and the magnitude of each failure in the following manner:

$$\Psi_{t+1} = \begin{cases} 1 + (PL_{t+1} - VaR_{t+1|t})^2 & \text{if } PL_{t+1} < VaR_{t+1|t} \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

where Ψ_{t+1} represents the loss function.

Thus, by penalizing the method with the largest failures, the intent is to find a model that minimizes:

$$\Psi = \sum_{t=1}^N \Psi_{t+1} \quad (16)$$

3. Description of Our Exercise

In order to evaluate the behavior of the ten approaches¹² above with $\alpha = 0.05$ a recursive window is used. The evaluation exercise involves the following steps:

1. Calculate the VaR for period $T + 1$ (next 10 minutes) using the first T observations;
2. Save the estimated $VaR_{T+1|T}$ and compare it against the observed loss or gain;
3. Update the sample by incorporating an additional observation;
4. Repeat 1,000 times steps 1 to 3 using the last 1,000 observations; and
5. Perform the tests described above.

Below a description is provided of the data used and some special considerations for using intraday data.

3.1. Data

We used ten-minute observations of the returns from the IGBC (General Colombian Stock Exchange Index). In order to achieve our objective of determining the behavior of the various VaR's specifications for a very short time horizon. This exercise was carried out with two samples that represent different environments in the international and macroeconomic markets.

The two samples were used to perform different comparisons of the effectiveness of the VaR's specifications in a relatively steady scenario (2006-2007) versus a scenario of increased uncertainty and volatility (2008-2009). The first sample began at 9:00 am on December 27, 2006, and ended at 1:00 pm on November 9, 2007, with a total of 5,088 observations. The second sample (2008-2009), which

¹²The exercise is carried out under the assumptions of either a normal distribution or a t-distribution.

corresponds to the financial crisis period, started at 9:00 am on June 3, 2008, and ended at 1:00 pm on March 17, 2009, with a total of 4,655 observations.

The IGBC series for the first period was obtained from the Bloomberg information system, while the series for the second period of analysis was obtained from Reuters¹³ financial information platform. Figures 1 and 2 show the IGBC series, both for 2006-2007 and 2008-2009, including the returns, the corresponding histograms, and probability charts for the normal, t with 3 and 4 degrees of freedom theoretical distributions.

Based on the charts it is possible to infer that the distribution of yields has relatively heavy tails in comparison to the normal distribution. The descriptive statistics for both samples are reported in Table 2. Jarque-Bera's normality test leads to the conclusion that there is no evidence in favor of the normal distribution of yields for either of the samples. This result is consistent with the stylized facts of yields as discussed by Alonso & Arcos (2006) for this same series.

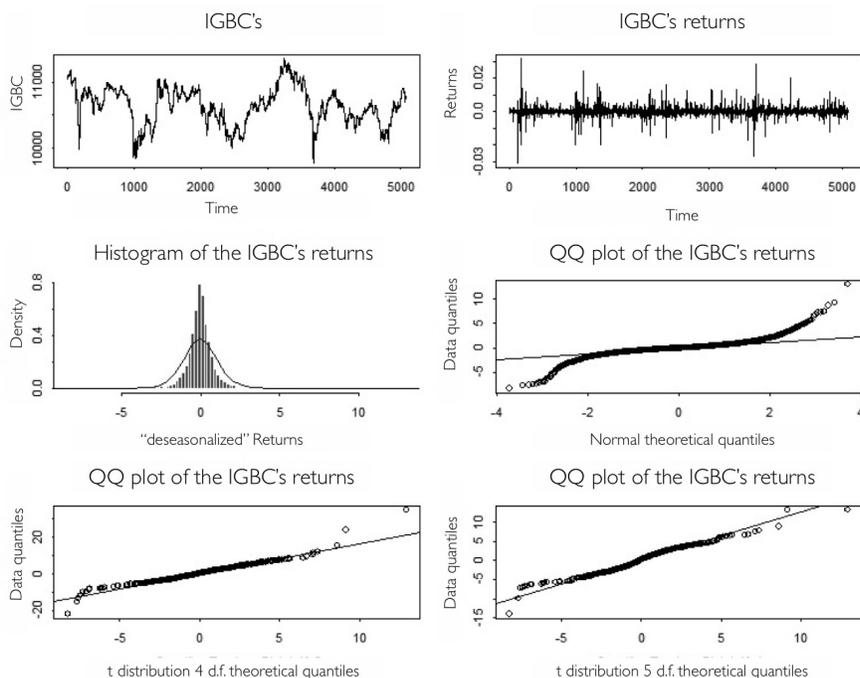


FIGURE 1: IGBC series and returns 2006-2007.

This means that the probability of obtaining extreme values is much greater in the empirical distribution of yields than the expected from a normal distribution. Consequently, in addition to the parametric estimation of VaR under an

¹³The use of different sources of information does not pose any issues to this exercise. Both sources obtain information from the registry system at the Colombian Stock Exchange, so data reported from both sources are identical. There was a change in the source of information because one of the information service providers charged a more convenient fee.

assumption of normality, the parametric VaR estimation was carried out using a Student's t distribution, which relatively adjusts better to the reality of data used¹⁴, as shown by the qq-plots of the t distribution with 3 and 4 degrees of freedom for the 2006-2007 sample and 4 and 5 degrees of freedom for the 2008-2009 sample. As can be seen, the theoretical percentiles from the aforementioned distribution not only adjust more closely to those observed, but also incorporate the stylized fact of heavy tails.

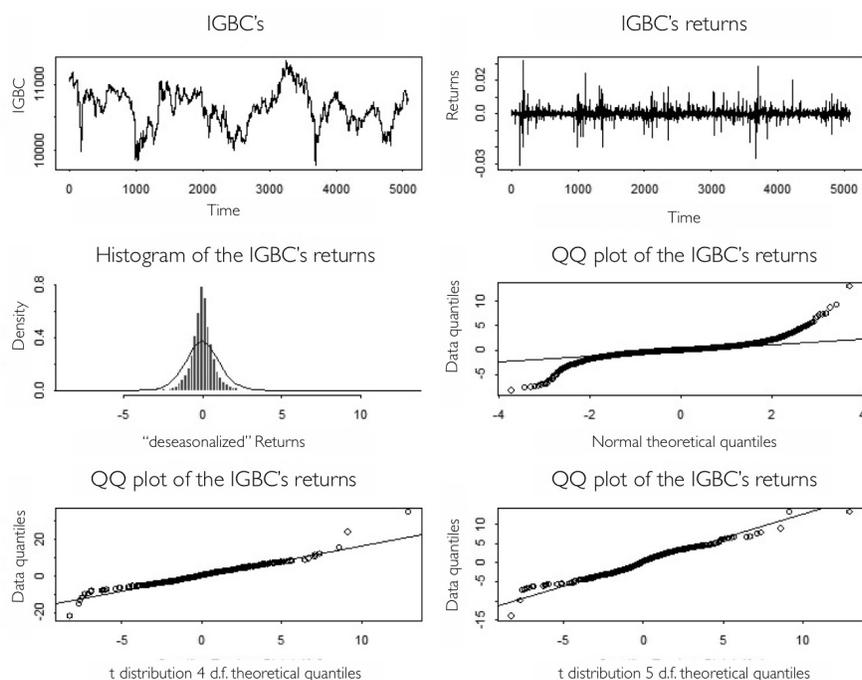


FIGURE 2: IGBC series and returns 2008-2009.

Table 2 shows the apparent symmetry that can be observed in each of the histograms. The kurtosis of both samples is relatively high, especially for the sample from the period that coincides with the financial crisis. On the other hand, the variance of the sample from the financial crisis period is 2.23 greater than that of the other sample. These two results confirm that the financial crisis period is more volatile. These characteristics of the samples, but particularly of the second sample, represent a challenge to the modeling of variance based on GARCH models.

¹⁴The degrees of freedom were estimated for each iteration based on the conditional variance of the returns assumed under GARCH models.

TABLE 2: Descriptive statistics of returns for every 10 minutes of the IGBC for both samples.

Sample	2006-2007	2008-2009
Mean	-4.60E-006	-0.5083E-004
Variance	4.89E-06	1.092E-05
Asymmetry coefficient	0.22	0.30
Kurtosis	44.56	196.23
Jarque-Bera	77671.30***	6764.49***

(***) The null hypothesis of normality is rejected with a 99 % level of confidence.

3.2. Special Considerations for Intraday Data

The nature of the data used for this research brings some methodological problems as mentioned by Andersen & Bollerslev (1997) and Giot (2005). By modeling the volatility of returns at high frequencies (i.e., every 5, 10 or 20 minutes), Andersen & Bollerslev (1997) show that the existence of bias is more likely to occur in GARCH and ARCH parameters with high-frequency data when a GARCH model is estimated. Particularly, the probability that the sum of the coefficients equals one increases¹⁵, and thus the probability of estimating non-stationary models for the variance process will also increase. This means that using higher frequency samples involves the risk of capturing the “noise” associated with intraday seasonality and, ultimately, the existence of bias in the estimation of parameters for the GARCH model.

As suggested in literature, there are several alternatives for preventing bias or noise associated with intraday seasonality. Andersen & Bollerslev (1997) propose the use of “deseasonalized” returns (r_t^*). Deseasonalization can be assumed to be deterministic, and when intraday observations are available at regular intervals (e.g., every 10 or 30 minutes), deseasonalized returns can be calculated using the formula:

$$r_t^* = \frac{r_t}{\sqrt{\phi(i_t)}} \quad (17)$$

where r_t stands for observed returns and $\phi(i_t)$ represents the deterministic component of intraday seasonality. To calculate this component, Giot (2005) proposes an average of all square returns that correspond to the same time and day of the week of the observed return r_t . Hence, for 10-minute periods, for each of the five days of the week, the same number of $\phi(i_t)$ is obtained as the number of 10-minute periods in a trading day¹⁶. Therefore, the specifications will be estimated using the “deseasonalized” series. Later, intraday seasonality will be incorporated in order to calculate the VaR.

¹⁵In this case, the probability that $\alpha_1 + \alpha_2 = 1$ increases, which implies that the variance process will explode.

¹⁶In the case of the IGBC index, there are four hours of trading and six 10-minute periods per hour. This means that there are 24 different $\phi(i_t)$.

4. Results

Table 3 shows the proportion of failures for each of the samples as well as for the non-parametric approach (specification 1), the semi-parametric approach (specification 10), and GARCH and TGARCH specifications under the assumption of a normal distribution. For the ten approaches, the hypothesis of correct unconditional coverage is rejected if the Kupiec’s (1995) test is used. In other words, the forecast proportion of failures using our models is different from the observed proportion of failures.

In fact, it can be observed in Table 3 that the proportions of failure are lower than the 5%, expected proportion of failures. This could be an indication that these specifications are fairly conservative in estimating the VaR. Table 4 on the other hand, reports the same results for GARCH and TGARCH specifications that use a t-distribution. The results are different. In the case of the first sample, the hypothesis of correct unconditional coverage cannot be rejected for specifications 2, 3, 6, and 7. This means that the observed proportion of failures for these specifications is the same as the theoretical proportion used for designing the VaR. For other specifications, the coverage is relatively lower than theoretically expected ($\alpha = 0.05$). For the second sample, specification 8 is the only specification using does not provide correct unconditional coverage.

Thus, if only unconditional coverage is considered, those specifications using a t-distribution exhibit a better behavior than those where a normal distribution is assumed.

TABLE 3: Proportion of failures and Kupiec’s (1995) test. Normal distribution.

Spec.	Sample 2006-2007		Sample 2008-2009	
	$\hat{\rho}$	t_U	$\hat{\rho}$	t_U
1	0.038	-1.985**	0.016	-8.569**
2	0.030	-3.708**	0.028	-4.217**
3	0.027	-4.487**	0.032	-3.234**
4	0.033	-3.009**	0.032	-2.234**
5	0.031	-3.467**	0.035	-2.581**
6	0.030	-3.708**	0.028	-4.217**
7	0.027	-4.487**	0.03	-3.708**
8	0.033	-3.009**	0.033	-3.009**
9	0.031	-3.467**	0.032	-3.234**
10	0.035	-2.367**	0.020	-7.234**

t_U = Kupiec’s t-statistic.

(**) Rejects the null hypothesis of non-conditional coverage ($\rho = 0.05$) with a 5% significance level.

Let us now consider Lopez’s magnitude loss function (see Table 5). Under the assumption of normality for parametric approaches, it is found that, for the first sample (2006-2007), the third specification is the one that minimizes the loss

TABLE 4: Proportion of failures and Kupiec's (1995) test. t-distribution.

Spec. t-distribution	Sample 2006-2007		Sample 2008-2009	
	$\hat{\rho}$	t_U	$\hat{\rho}$	t_U
2	0.04	-1.614	0.041	-1.435
3	0.042	-1.261	0.041	-1.435
4	0.038	-1.985**	0.042	-1.261
5	0.037	-2.178**	0.043	-1.091
6	0.039	-1.797	0.041	-1.435
7	0.040	-1.435	0.042	-1.261
8	0.037	-2.178**	0.033	-3.009**
9	0.037	-2.178**	0.044	-0.925

t_U = Kupiec's t-statistic.

(**) Rejects the null hypothesis of non-conditional coverage ($\rho = 0.05$) with a 5% significance level.

TABLE 5: Results of Lopez's (1998) loss function. Normal distribution.

Spec.	Sample 2006-2007	Sample 2008-2009
	Ψ	Ψ
1	2333046836358.57	2647046846407.870
2	1653110917303.23	451650358867.146
3	1578224677463.77*	453951659137.534
4	1684453701968.83	431233167551.233
5	1603126383687.97	435309131396.309
6	1657300659233.63	453207153145.048
7	1581276624589.99	452283235911.192
8	1689813156174.26	430361362409.036*
9	1613824157927.02	433582401277.377
10	2033048131561.07	2036136846407.654

(*) Lowest loss from Lopez's magnitude loss function.

The units of measure for this test are square Colombian pesos.

The initial portfolio value for each period equals 100 million Colombian pesos.

function. On the other hand, for the 2008-2009 sample, specification 8 is the one that exhibits the best behavior with regard to this criterion. None of these specifications, however, provides correct unconditional coverage.

In the case of parametric specifications where a t-distribution is assumed, we find that specification 6 is the one that minimizes Lopez's loss for the first sample, and specification 5 does the same for the second sample. Both specifications provide correct coverage for their corresponding samples.

This would mean that, based on these two criteria, the VaR calculated from a TGARCH model, without considering week day effect or day time effect and a t-distribution, and a GARCH model considering week day and day time effects

TABLE 6: Results of Lopez’s (1998) loss function. t-distribution.

Spec.	Sample 2006-2007	Sample 2008-2009
t-distribution	Ψ	Ψ
2	1850889691317.980	512915110922.186
3	1872451233890.160	504166038902.471
4	1858706617825.380	511117944321.063
5	1865178096796.600	502818217742.356*
6	1846764259665.510*	541384048077.450
7	1868859264549.290	550978138480.490
8	1852149851803.150	3540151295984.520
9	1858768144055.600	529705197041.650

(*) Lowest loss from Lopez’s magnitude loss function.

The units of measure for this test are square Colombian pesos. The initial portfolio value for each period equals 100 million Colombian pesos.

and a t-distribution would be the best specifications for estimating VaR for the first and second samples, respectively.

Finally, Table 7 shows the results for Christoffersen’s (1998) correct conditional coverage test for the estimated models under the assumption of normality. It can be observed that, for the 2006-2007 sample, there are four specifications that stand out, namely, 1, 4, 8, and 9, because there is not sufficient evidence to reject the null hypothesis of correct conditional coverage. Thus, for this sample, there are a GARCH model considering the day time effect (specification 4), a GARCH model with leverage and day time effects (specification 8), a historical simulation (specification 1), and a filtered historical simulation (specification 10). None of these approaches, however, provides the lowest Lopez’s loss function for that sample. The results differ from those of the 2008-2009 sample. In fact, based on Christoffersen’s (1998) test, none of the specifications provides correct conditional coverage.

If we consider the parametric models estimated under the assumption of a t-distribution (see Table 8), we find that, for the first sample, all models with the exception of models 5, 8, and 9, the hypothesis of correct conditional coverage cannot be rejected. Out of these specifications, specification 6 (TGARCH without considering the day and time effect) is the one that exhibits the lowest loss function. For the second sample, model 8 is the only one that rejects the hypothesis of correct conditional coverage. And in this case the specification 5 (GARCH model considering day and time effect) provides both correct conditional coverage and the lowest Lopez’s loss function.

López’s (1998) loss function allows to make a comparison of models that have correct conditional coverage estimated both under the assumption of normal distribution and the assumption of a t-distribution for each of the samples. For the 2006-2007 sample, specification 6 (TGARCH model without considering day and time effect), which was estimated under the assumption of a t-distribution, mini-

mizes Lopez's loss function and, therefore, has a better behavior than a historical simulation or a filtered historical simulation. For the second sample, the best model is model 5, which represents a GARCH(1,1) model considering week day and day time effect, estimated under the assumption of a t-distribution¹⁷.

TABLE 7: Christoffersen's (1998) coverage and independence test. Normal distribution.

Spec.	Sample 2006-2007			Sample 2008-2009		
	LR_{uc}	LR_{ind}	LR_{cc}	LR_{uc}	LR_{ind}	LR_{cc}
1	3.29	2.43	0.864	32.74**	0.01	32.74 ⁺⁺
2	9.77**	-0.51	9.26 ⁺⁺	12.04**	0.02	12.06 ⁺⁺
3	13.28**	0.02	13.29 ⁺⁺	7.78**	0.03	7.81 ⁺⁺
4	6.88**	-1.28	5.59	7.78**	0.035	7.81 ⁺⁺
5	8.74**	-0.77	7.96 ⁺⁺	5.27**	0.04	5.31
6	9.77**	-0.51	9.26 ⁺⁺	12.04**	0.02	12.06 ⁺⁺
7	13.28**	0.0206	13.299 ⁺⁺	9.769**	0.0285	9.797 ⁺⁺
8	6.88**	-1.2873	5.591	6.878**	0.0381	6.916 ⁺⁺
9	8.74**	-0.7769	7.962 ⁺⁺	7.777**	0.0347	7.811 ⁺⁺
10	4.93	2.93	1.99	35.84**	0.05	35.79 ⁺⁺

(**) Rejects the null hypothesis of non-conditional coverage at a 5% significance level.

(++) Rejects the null hypothesis of correct conditional coverage at a 5% significance level.

TABLE 8: Christoffersen's (1998) coverage and independence test. t-distribution.

Spec.	Sample 2006-2007			Sample 2008-2009		
	LR_{uc}	LR_{ind}	LR_{cc}	LR_{uc}	LR_{ind}	LR_{cc}
2	2.25	-2.84	-0.59	1.812	0.07	1.89
3	1.42	-3.23	-1.81	1.812	0.07	1.89
4	3.29	3.86 ^{oo}	7.15 ⁺⁺	1.42	0.079	1.50
5	3.89**	-0.74	3.15	1.08	0.0858	1.17
6	2.75	-2.639	0.11	1.81	0.074	1.89
7	1.81	-3.039	-1.23	1.42	0.079	1.50
8	3.89**	-0.74	3.15	6.88**	0.67	7.56 ⁺⁺
9	3.99**	-0.74	3.15	0.79	0.09	0.88

(**) Rejects the null hypothesis of non-conditional coverage at a 5% significance level.

(oo) Rejects the null hypothesis of independence at a 5% significance level.

(++) Rejects the null hypothesis of correct conditional coverage at a 5% significance level.

5. Final Remarks and Conclusions

In order to test our hypothesis that intraday behavior patterns could provide relevant information that could be used for improving risk measures such as the

¹⁷After completing all of the calculations reported above, an exercise was carried out in order to guarantee the robustness of our results both at the beginning and at the end of the two samples being reviewed. For this purposed, the same exercises were replicated, starting with one month, two months, and three months less of data. The results and conclusions remained unchanged. This exercise was also carried out omitting one month, two months, and three months of data at the end of both samples. The conclusions did not change substantially. For the purpose of saving space, these results are not reported here.

VaR, we evaluated the behavior for the next ten minutes of trading at the Colombian Stock Exchange using 18 different ways to estimate VaR for a portfolio with the same behavior as that of the Colombian Stock Exchange index. We considered a non-parametric approach, eight parametric models under the assumption of a normal distribution, and 8 models under the assumption of a t-distribution and one semi-parametric approach. These methods were applied to two different samples, one for a relatively steady period (2006-2007)¹⁸ and another sample for a scenario of increased uncertainty and volatility (2008-2009)¹⁹.

The parametric specifications include the day of the week effect and the hour of the day effect as well as different ways to forecast volatility for the following ten minutes (for a summary of specifications used, see Table 1).

In all cases, prior to the estimation of the models, data is deseasonalized following Giot's (2005) recommendations. The results obtained can be summarized as described below. Firstly, in the case of parametric VaRs with the assumption of normality, we find that there is no model that provides correct non-conditional coverage for the two samples being considered.

Secondly, for both samples, the estimated VaR models under the assumption of a t-distribution have, overall, a better performance than those under the assumption of a normal distribution. Thirdly, we found that, using Christoffersen's (1998) test and López's (1998) loss function to compare models that have correct conditional coverage, we found that the TGARCH(1,1), model without considering week-day and day-time effect and a t-distribution, is the best model for the 2006-2007 sample²⁰. For the second sample, the best model is GARCH(1,1), which considers week-day effect and day-time effect estimated under the assumption of a t-distribution. This result validates our hypothesis that intraday behavior patterns can provide relevant information for improving risk measures such as the VaR.

The normal probability charts, Jarque-Bera's normality test, and conditional coverage tests are useful for inferring that, in general, using the assumption of a t-distribution seems to be a better approach than using a normal distribution assumption, which supports our results.

Lastly, our results suggest that there is a need to study intraday behavior of stock portfolios in more detail and encourage a review of approaches that incorporate the dynamic of each of the assets that comprise the portfolio. In other words, it will be necessary to investigate the effect of modeling the multivariate conditional distribution of all assets involved. In order to be able to achieve this, the conditional matrix of variance and covariance will have to be estimated.

¹⁸This sample, consisting of 5,088 observations in total, runs from 9:00 am on December 27, 2006, to 1:00 pm on November 9, 2007.

¹⁹This sample, consisting of 4,655 observations in total, begins at 9:00 am on June 3, 2008 and ends at 1:00 pm on March 17, 2009.

²⁰It is worth mentioning that Alonso & García (2009) found that, using the first sample, the best model for forecasting the IGBC average for the next ten minutes is a model that did not consider the day or time effect. In other words, these authors showed that day and time are not important when it comes to forecasting the behavior (average) of the IGBC for the next ten minutes. Our results for this sample allow drawing a similar conclusion with regard to the behavior of the variance.

[Recibido: agosto de 2010 — Aceptado: enero de 2011]

References

- Alonso, J. C. & Arcos, M. A. (2006), ‘Cuatro hechos estilizados de las series de rendimientos: una ilustración para Colombia’, *Estudios Gerenciales* **22**(110).
- Alonso, J. C. & Berggrun, L. (2008), *Introducción al Análisis de Riesgo Financiero*, Colección Discernir. Serie Ciencias Administrativas y Económicas, Universidad ICESI, Cali, Colombia.
- Alonso, J. C. & García, J. C. (2009), ‘¿qué tan buenos son los patrones del IGBC para predecir su comportamiento?: una aplicación con datos de alta frecuencia’, *Estudios Gerenciales* **25**(112), 1–50.
- Alonso, J. C. & Romero, F. (2009), The day-of-the-week effect: The Colombian exchange rate and stock market case, in ‘Selected Abstracts and Papers. Latin American Research Consortium 2009’, pp. 112–120.
- Andersen, T. G. & Bollerslev, T. (1997), ‘Intraday periodicity and volatility persistence in financial markets’, *The Journal of Empirical Finance* **4**(2-3), 115–158.
- Berument, H. & Kiyamaz, H. (2003), ‘The day of the week effect on stock market volatility and volume: International evidence’, *Review of Financial Economics* **12**(3), 363–380.
- Bollerslev, T. (1986), ‘Generalized autoregressive conditional heteroskedasticity’, *Journal of Econometrics* **31**(3), 307–327.
- Brooks, C. (2008), *Introductory Econometrics for Finance*, Cambridge University Press, London.
- Christoffersen, P. (1998), ‘Evaluating interval forecasts’, *International Economic Review* **39**(4), 841–862.
- Dowd, K. (2005), *Measuring Market Risk*, 2 edn, John Wiley & Sons Ltd, England.
- Giot, P. (2000), Intraday value-at-risk, CORE Discussion Papers 2000045, Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE).
- Giot, P. (2005), ‘Market risk models for intraday data’, *European Journal of Finance* **11**, 309–324.
- Glosten, L., Jagannathan, R. & Runkle, D. E. (1993), ‘On the relation between the expected value and the volatility of the nominal excess return on stocks’, *Journal of Finance* **48**(5), 1779–1801.
- Kamath, R. & Chinpiao, L. (2010), ‘An investigation of the day-of-the-week effect on the Istanbul stock exchange of Turkey’, *Journal of International Business Research* **9**(1), 15–27.

- Kupiec, P. H. (1995), 'Techniques for verifying the accuracy of risk measurement models', *Journal of Derivatives* **3**(2), 73–84.
- López, J. A. (1998), 'Methods for evaluating value at risk estimates', *Economic Policy Review* **4**(3).
- Mittal, S. K. & Jain, S. (2009), 'Stock market behaviour: evidences from Indian market', *Vision* **13**(3), 19–29.
- Nocera, J. (2009), 'Risk mismanagement', *The New York Times* .
- Panas, E. (2005), 'Generalized Beta distributions for describing and analysing intraday stock market data: Testing the U-shape pattern', *Applied Economics* **37**(2), 191–199.
- Rivera, D. M. (2009), 'Modelación del efecto del día de la semana para los índices accionarios de Colombia mediante un modelo STAR GARCH', *Revista de Economía del Rosario* **12**(1), 1–24.