



REPORTE DE ESTABILIDAD FINANCIERA

Enero de 2016

Comparación de métodos para la estimación de la incertidumbre del valor en riesgo

Santiago Gamba Santamaría
Oscar Fernando Jaulín Méndez
Luis Fernando Melo Velandia
Carlos Andrés Quicazán Moreno

COMPARACIÓN DE MÉTODOS PARA LA ESTIMACIÓN DE LA INCERTIDUMBRE
DEL VALOR EN RIESGO*

SANTIAGO GAMBA SANTAMARÍA[‡]
OSCAR FERNANDO JAULÍN MÉNDEZ[†]
LUIS FERNANDO MELO VELANDIA[£]
CARLOS ANDRÉS QUICAZÁN MORENO[‡]

RESUMEN. El Valor en Riesgo (VaR) es una medida de riesgo de mercado ampliamente usada por administradores de riesgo y autoridades regulatorias. Sin embargo, a pesar de que existe una gran variedad de metodologías propuestas en la literatura para la estimación del VaR, pocas de ellas dicen algo acerca de su distribución o sus intervalos de confianza. Este artículo compara distintas metodologías para calcular esos intervalos. Se utilizaron métodos basados en normalidad asintótica, teoría del valor extremo y *bootstrap* de submuestra. Usando simulaciones de Monte Carlo, se encontró que estas aproximaciones son válidas sólo para cuantiles altos. Particularmente, en términos de porcentaje de cobertura, estas metodologías presentan un buen desempeño para el VaR(99%) y un bajo desempeño para el VaR(95%) y el VaR(90%). En general, estos resultados se confirman a través de un ejercicio empírico aplicado a los bonos de deuda pública colombiana.

Palabras clave: Valor en Riesgo, intervalos de confianza, *data tilting*, *bootstrap* de submuestra.

JEL Codes: C51, C52, C53, G32.

ABSTRACT. Value at Risk (VaR) is a market risk measure widely used by risk managers and market regulatory authorities. There is a variety of methodologies proposed in the literature for the estimation of VaR. However, few of them get to say something about its distribution or its confidence intervals. This paper compares different methodologies for computing such intervals. Several methods, based on asymptotic normality, extreme value theory and subsample bootstrap, are implemented. Using Monte Carlo simulations, it is found that these approaches are only valid for high quantiles. Particularly, there is a good performance at VaR(99%), in terms of coverage rates, and bad performance for VaR(95%) and VaR(90%). In general, these results are confirmed by conducting an empirical exercise using Colombian public debt bonds.

Keywords: Value at Risk, confidence intervals, data tilting, subsample bootstrap.

JEL Codes: C51, C52, C53, G32.

* Las opiniones expresadas aquí pertenecen a los autores y no necesariamente representan las opiniones del Banco de la República o su junta directiva, por lo tanto se eximen de toda responsabilidad. Los autores agradecen a Daniel Mariño y Esteban Gómez por sus comentarios y sugerencias.

[‡] Estudiante en práctica, Unidad de Econometría del Banco de la República. E-mail: gambas@javeriana.edu.co.

[†] Profesional, Departamento de Estabilidad Financiera del Banco de la República. E-mail: ojaulime@banrep.gov.co.

[£] Econometrista principal, Unidad de Econometría del Banco de la República. E-mail: lmelovel@banrep.gov.co.

[‡] Profesional especializado, Departamento de Estabilidad Financiera del Banco de la República. E-mail: equicamo@banrep.gov.co.

1. INTRODUCCIÓN

El Valor en Riesgo (VaR) es una medida de riesgo de mercado ampliamente usada por administradores de riesgo y autoridades regulatorias. Se define como la máxima pérdida en el valor de un activo para un nivel de probabilidad (α) y un horizonte de pronóstico dados; el VaR corresponde entonces al α -ésimo cuantil de la distribución de los retornos del activo. Existen diferentes métodos para obtener el estimador puntual del VaR, sin embargo, pocos de ellos dicen algo acerca de su incertidumbre. La estimación de intervalos de confianza resuelve este problema debido a que cuantifica dicha incertidumbre, en ese sentido un intervalo angosto implica una estimación más precisa mientras que un intervalo ancho muestra una alta varianza en la medida. En Colombia, el VaR se usa para determinar el nivel de capital de las instituciones financieras, por lo que se hace necesario tener precaución con el uso de esta medida cuando los intervalos de confianza resultan ser considerablemente anchos.

Las metodologías que permiten estimar intervalos de confianza para el VaR se basan en diferentes supuestos, algunas de ellas suponen normalidad asintótica. Por ejemplo, Moraux [2011] propone una primera aproximación asumiendo que los retornos son independientes e idénticamente distribuidos. Usando el método delta estima la varianza asintótica del VaR y con ella calcula los intervalos de confianza. Sin embargo, esta metodología no incorpora los hechos estilizados de las series de tiempo financieras (comportamiento heteroscedástico y de colas pesadas).

Por esta razón, otros métodos permiten que las series de tiempo sigan un proceso ARMA-GARCH. Algunos de ellos no asumen ninguna distribución en particular. Spierdijk [2014] menciona que algunos de estos métodos fallan cuando el supuesto de normalidad asintótica no se tiene o cuando el tamaño de muestra no es lo suficientemente grande. Chan et al. [2007], Gao y Song [2008] y Francq y Zakoïan [2015] emplean diferentes aproximaciones basadas en Quasi Máxima Verosimilitud (QML por sus siglas en inglés) con el objeto de estimar el cuantil de los residuales estandarizados sin hacer uso del supuesto de normalidad asintótica. Gao y Song [2008] usan simulación histórica filtrada, Francq y Zakoïan [2015] proponen una reparametrización de los errores estandarizados del GARCH, mientras Chan et al. [2007] se basan en teoría del valor extremo.

Otro método para estimar los intervalos de confianza del VaR se basa en técnicas de *bootstrap* sobre los errores estandarizados. Sin embargo, el *bootstrap* convencional no es consistente cuando se viola el supuesto de normalidad asintótica, y su inconsistencia se incrementa bajo la presencia de dinámicas GARCH con cuarto momento no acotado [Hall y Yao, 2003]. Por lo anterior, Spierdijk [2014] propone una metodología de *bootstrap* de submuestra para modelos ARMA-GARCH usando también estimación por QML. Esta metodología presenta un buen desempeño cuando no se confirma el supuesto de normalidad asintótica; y comparado con el *bootstrap* convencional, no requiere que el cuarto momento de los errores sea finito.

Las metodologías expuestas por Gao y Song [2008] y Francq y Zakoïan [2015] no toman en cuenta colas pesadas en las innovaciones GARCH¹, como sí lo hacen Chan et al. [2007]. Por esta razón,

¹Es importante tener en cuenta que no hay evidencia de que éstas sean robustas a ese hecho.

se evalúa el desempeño de los estimadores por *data tilting* y normalidad asintótica desarrollados por Chan et al. [2007] y la metodología expuesta por Spierdijk [2014]. Finalmente se comparan estos resultados con los de la aproximación normal de Moraux [2011].

El propósito de este artículo es comparar estas metodologías, con el objeto de identificar los escenarios bajo los cuales las técnicas de intervalos de confianza se desempeñan adecuadamente. Esto se realiza a través de la evaluación de los porcentajes de cobertura para cada uno de los métodos usando simulaciones de Monte Carlo. Adicionalmente, se incluye una aplicación empírica con el objetivo de analizar qué alternativa es la más adecuada para la serie de retornos de la deuda pública del sistema financiero colombiano.

El resto del artículo está organizado de la siguiente manera. La sección dos explica los métodos que son usados para calcular los intervalos de confianza del VaR. En la sección tres se muestra un ejercicio de simulación. La sección cuatro contiene una aplicación empírica para la deuda pública del sistema financiero colombiano. Finalmente, algunas conclusiones son presentadas en la sección cinco.

2. METODOLOGÍA

En esta sección, se consideran cuatro métodos para estimar los intervalos de confianza del VaR de unos retornos r_t .

2.1. Normalidad asintótica. Moraux [2011] considera el caso simple en el que el VaR es obtenido bajo el supuesto de normalidad, $r_t \stackrel{iid}{\sim} N(\mu, \sigma^2)$. Entonces, $\widehat{VaR}(t+1|t)_\alpha = \widehat{\mu} + \widehat{\sigma}q_\alpha$, donde q_α es el α -ésimo cuantil de la distribución normal estándar.

Dados los siguientes resultados asintóticos, $\sqrt{T}(\widehat{\mu} - \mu) \xrightarrow{d} N(0, \sigma^2)$ y $\sqrt{T}(\widehat{\sigma}^2 - \sigma^2) \xrightarrow{d} N(0, \sigma^4)$, y usando el método delta, él demuestra que $\sqrt{T}(\widehat{VaR} - VaR) \xrightarrow{d} N(0, \sigma^2(1 + 0,5q_\alpha^2))$.

Por lo tanto, el intervalo de confianza con nivel β del VaR es:

$$\widehat{VaR}(t+1|t)_\alpha \pm \frac{\widehat{\sigma}}{\sqrt{T}} \sqrt{1 + 0,5q_\alpha^2} \ q_{0,5(1+\beta)} \quad (1)$$

Aunque esta aproximación es muy simple, no tiene en cuenta los hechos estilizados de las series financieras. En particular, no considera las distribuciones con colas pesadas y los conglomerados de volatilidad, que sí son tenidos en cuenta por los métodos explicados a continuación.

2.2. Estimador de Hill asintótico. Este método usa el estimador puntual de un cuantil propuesto por Chan et al. [2007]. Usando teoría del valor extremo, ellos demuestran que este estimador tiene una distribución normal asintótica.

Suponga que los retornos siguen un modelo $AR(P) - GARCH(p, q)$,

$$r_t = \mu + \sum_{i=1}^P \phi_i r_{t-i} + \varepsilon_t \quad (2)$$

$$\sigma_t^2 = c + \sum_{i=1}^p b_i \varepsilon_{t-i}^2 + \sum_{i=1}^q a_i \sigma_{t-i}^2 \quad (3)$$

donde $\varepsilon_t = \sigma_t z_t$, $c < 0$, $b_i \geq 0$, $a_i \geq 0$ para cada i , $\sum_{i=1}^p b_i + \sum_{i=1}^q a_i < 1$, y $\{z_t\}$ es una secuencia de variables aleatorias independientes e idénticamente distribuidas, con media cero y varianza uno.

Por tanto, el $\text{VaR}(\alpha)$ un período adelante es de la siguiente forma:

$$\text{VaR}(t+1 | t)_\alpha = \mu + \sum_{i=1}^P \phi_i r_{t-i+1} + \sigma_{t+1|t} q_\alpha \quad (4)$$

donde q_α es el α -ésimo cuantil de z_{t+1} .

Bajo el supuesto de que las colas de la distribución de z_t decrecen a una tasa constante γ , la estimación del exceso sobre un umbral, usando la teoría de valor extremo, puede ser analizada en términos de la secuencia $\{\hat{z}_t\}$ de residuos estandarizados.

Si $\hat{z}_{m,1} \leq \hat{z}_{m,2} \leq \dots \leq \hat{z}_{m,m}$ son los residuos ordenados de los últimos m residuos estandarizados $\hat{z}_v, \dots, \hat{z}_T$, con $v = T - m + 1$, el estimador de Hill de γ es:

$$\hat{\gamma} = \left[\frac{1}{k} \sum_{i=1}^k \log \frac{\hat{z}_{m,m-i+1}}{\hat{z}_{m,m-k}} \right]^{-1} \quad (5)$$

donde k es el número de observaciones extremas, y $\hat{z}_{m,m-k}$ puede ser considerado como el umbral.

Hill [1975] propone el siguiente estimador puntual de q_α :

$$\hat{q}_\alpha = (1 - \alpha)^{-\frac{1}{\hat{\gamma}}} \left(\frac{k}{m} \right)^{\frac{1}{\hat{\gamma}}} \hat{z}_{m,m-k} \quad (6)$$

Y por lo tanto, el estimador puntual del $\text{VaR}(t+1 | t)_\alpha$ es:

$$\widehat{\text{VaR}}(t+1 | t)_\alpha = \hat{\mu} + \sum_{i=1}^P \hat{\phi}_i r_{t-i+1} + \hat{\sigma}_{t+1|t} \hat{q}_\alpha \quad (7)$$

Chan et al. [2007] muestran que este estimador es consistente y converge a la distribución normal estándar.

$$\frac{\hat{\gamma} \sqrt{k}}{\left| \log \left(\frac{k}{m(1-\alpha)} \right) \right|} \left[\frac{\widehat{\text{VaR}}(t+1 | t)_\alpha}{\text{VaR}(t+1 | t)_\alpha} - 1 \right] \xrightarrow{d} N(0, 1) \quad (8)$$

Por lo tanto, el intervalo del $\widehat{VaR}(t+1|t)_\alpha$, con un nivel de confianza β es:

$$\left(\widehat{VaR}(t+1|t)_\alpha \exp \left\{ q_{(1-\beta)/2} \left| \log \left(\frac{k}{m(1-\alpha)} \right) \right| / (\hat{\gamma} \sqrt{k}) \right\}, \widehat{VaR}(t+1|t)_\alpha \exp \left\{ q_{(1+\beta)/2} \left| \log \left(\frac{k}{m(1-\alpha)} \right) \right| / (\hat{\gamma} \sqrt{k}) \right\} \right) \quad (9)$$

2.3. Data tilting. Chan et al. [2007] proponen usar *data tilting* para estimar los intervalos de confianza del VaR. Este método es una aproximación no paramétrica que puede entenderse como una generalización de la metodología de verosimilitud empírica,² donde las observaciones son ponderadas para minimizar una función de distancia.³

Usando z_t y $z_{m,m-k}$ como se definieron en la sección 2.2 y $\delta_t = I(\hat{z}_t \geq \hat{z}_{m,m-k})$, esta aproximación se calcula de la siguiente manera.

Primero, para cada vector de ponderaciones fijas $w = (w_v, \dots, w_T)$ tal que $w_t \geq 0$ y $\sum_{t=v}^T w_t = 1$, se resuelve el siguiente problema de optimización:

$$(\hat{\gamma}(w), \hat{c}(w)) = \underset{\gamma, c}{argmax} \sum_{t=v}^T w_t \log \left((c\gamma \hat{z}_t^{-\gamma-1})^{\delta_t} (1-c \hat{z}_{m,m-k}^{-\gamma})^{1-\delta_t} \right) \quad (10)$$

Lo que da como resultado:

$$\hat{\gamma}(w) = \frac{\sum_{t=v}^T w_t \delta_t}{\sum_{t=v}^T w_t \delta_t (\log \hat{z}_t - \log \hat{z}_{m,m-k})} \quad (11)$$

$$\hat{c}(w) = \hat{z}_{m,m-k}^{\hat{\gamma}(w)} \sum_{t=v}^T w_t \delta_t \quad (12)$$

Es importante mencionar que tanto $\hat{\gamma}(w)$ como $\hat{c}(w)$ serán parte de las restricciones que se explicarán más adelante.

Definiendo la función de distancia⁴ como:

$$D_l(w) = \begin{cases} -m^{-1} \sum_{t=v}^T \log(mw_t), & \text{para } l = 0 \\ \sum_{t=v}^T w_t \log(mw_t), & \text{para } l = 1 \\ (l(1-l))^{-1} (1-m^{-1} \sum_{t=v}^T (mw_t)^l), & \text{para } l \neq 0, 1 \end{cases} \quad (13)$$

Se encuentran los pesos que minimizan la distancia,

$$(2m)^{-1} L(VaR(t+1|t)_\alpha) = \min_w D_l(w) \quad (14)$$

²Un completo análisis de este método se encuentra en Owen [1988], Owen [1990], y Owen [2001], entre otros.

³ Este método es usado para construir regiones de confianza y una de sus ventajas es que permite calcularlas con cierto nivel de asimetría.

⁴Esta función mide la distancia entre los pesos no restringidos y los pesos dado una distribución uniforme ($1/n$) y se basan en medidas de divergencia en potencia [Hall y Yao, 2003].

sujeto a:

$$\begin{aligned}
 w_t &\geq 0, \\
 \sum_{t=v}^T w_t &= 1, \\
 \hat{\gamma}(w) \log \left((VaR(t+1|t)_\alpha - \hat{\mu}_{t+1|t}) / (\hat{\sigma}_{t+1|t} \hat{z}_{m,m-k}) \right) &= \log \left(\sum_{t=v}^T w_t \delta_t / (1 - \alpha) \right)
 \end{aligned} \tag{15}$$

Usando el método de optimización de los multiplicadores de Lagrange para $D_1(w)$ ⁵:

$$w_t = \begin{cases} \frac{1}{m} e^{-1-\lambda_1}, & \text{si } \delta_t = 0 \\ \frac{1}{m} \exp \left\{ -1 - \lambda_1 + \lambda_2 \left(\frac{\log \left((VaR(t+1|t)_\alpha - \hat{\mu}_{t+1|t}) / (\hat{\sigma}_{t+1|t} \hat{z}_{m,m-k}) \right)}{A_2(\lambda_1)} \right) - \frac{1}{A_1(\lambda_1)} \right. \\ \left. - \frac{A_1(\lambda_1)}{A_2^2(\lambda_1)} \log \left(\hat{z}_t / \hat{z}_{m,m-k} \right) \log \left((VaR(t+1|t)_\alpha - \hat{\mu}_{t+1|t}) / (\hat{\sigma}_{t+1|t} \hat{z}_{m,m-k}) \right) \right\}, & \text{si } \delta_t = 1 \end{cases} \tag{16}$$

donde λ_1 y λ_2 satisfacen

$$\sum_{t=v}^T w_t = 1, \quad \hat{\gamma}(w) \log \left((VaR(t+1|t)_\alpha - \hat{\mu}_{t+1|t}) / (\hat{\sigma}_{t+1|t} \hat{z}_{m,m-k}) \right) = \log \left(\sum_{t=v}^T w_t \delta_t / (1 - \alpha) \right) \tag{17}$$

y

$$A_1(\lambda_1) = 1 - \frac{m-k}{m} e^{-1-\lambda_1}, \quad A_2(\lambda_1) = A_1(\lambda_1) \frac{\log \left((VaR(t+1|t)_\alpha - \hat{\mu}_{t+1|t}) / (\hat{\sigma}_{t+1|t} \hat{z}_{m,m-k}) \right)}{\log \left(A_1(\lambda_1) / (1 - \alpha) \right)} \tag{18}$$

Bajo ciertas condiciones Chan et al. [2007] muestran que:

$$L(VaR(t+1|t)_\alpha^0) \xrightarrow{d} \chi^2(1) \tag{19}$$

donde $VaR(t+1|t)_\alpha^0$ denota el valor verdadero del $VaR(t+1|t)_\alpha$.

Por lo tanto, basado en este resultado, un intervalo de confianza con un nivel β para $VaR(t+1|t)_\alpha^0$ es:

$$I'_\beta = \{VaR(t+1|t)_\alpha : L(VaR(t+1|t)_\alpha) \leq u_\beta\} \tag{20}$$

Donde u_β es el valor crítico al nivel β de una distribución $\chi^2(1)$.

⁵Chan et al. [2007] afirman que la función de distancia para $l = 1$ tiene buenas propiedades de robustez.

2.4. **Bootstrap de submuestra.** Spierdijk [2014] propone una metodología de *bootstrapping* submuestal sobre los errores para estimar los intervalos de confianza, asumiendo que los retornos siguen un modelo ARMA-GARCH. El procedimiento incluye los siguientes pasos:

1. Usar el método QML para estimar el modelo ARMA-GARCH a los retornos de la serie r_1, \dots, r_T .
2. Obtener una submuestra aleatoria (con tamaño l) sin reemplazamiento de los residuales estandarizados del paso anterior, $\tilde{z}_1, \dots, \tilde{z}_l$.
3. Recuperar los retornos de *bootstrap* $\tilde{r}_1, \dots, \tilde{r}_l$ usando los parámetros estimados en el paso 1 y $\tilde{z}_1, \dots, \tilde{z}_l$.
4. Volver a usar QML para estimar un ARMA-GARCH sobre $\tilde{r}_1, \dots, \tilde{r}_l$.
5. Calcular el \tilde{q}_z^α como el α -ésimo cuantil muestral de los residuos estandarizados obtenidos en el paso anterior.
6. Usando los parámetros estimados del paso 4 y r_1, \dots, r_T calcular $\tilde{\mu}_{t+1|t}$, $\tilde{\sigma}_{t+1|t}$ para $t = 1 \dots T$.
7. Calcular $\widetilde{VaR}(t+1|t)_\alpha = \tilde{\mu}_{t+1|t} + \tilde{\sigma}_{t+1|t} \tilde{q}_z^\alpha$.

Finalmente, el anterior procedimiento se repite B -veces, con lo que se obtiene $\widetilde{VaR}(t+1|t)_{\alpha,1}, \dots, \widetilde{VaR}(t+1|t)_{\alpha,B}$. Por lo tanto el intervalo de confianza a un nivel β es:

$$\left[\widetilde{VaR}(t+1|t)_{\alpha,1} + q_{VaR}^{(1-\beta)/2}, \widetilde{VaR}(t+1|t)_{\alpha,1} + q_{VaR}^{(1+\beta)/2} \right] \quad (21)$$

donde q_{VaR}^α es el cuantil empírico de $\left\{ \widetilde{VaR}(t+1|t)_{\alpha,1} - \widetilde{VaR}(t+1|t)_\alpha, \dots, \widetilde{VaR}(t+1|t)_{\alpha,B} - \widetilde{VaR}(t+1|t)_\alpha \right\}$ y $\widetilde{VaR}(t+1|t)_\alpha = \hat{\mu}_{t+1|t} + \hat{\sigma}_{t+1|t} \hat{q}_z^\alpha$.

3. EJERCICIO DE SIMULACIÓN

Para este ejercicio de simulación dos procesos son considerados. El primero corresponde a un proceso *i.i.d.* con valor esperado $\mu/(1-\phi_1)$ y varianza $c/(1-a_1-b_1)$. El segundo sigue un proceso $AR(1) - GARCH(1,1)$. Los parámetros, descritos en las ecuaciones (2) y (3), son: $\mu = 1$, $\phi_1 = 0,1$, $c = 0,1$, $b_1 = 0,05$, $a_1 = 0,92$. Los dos procesos son simulados con tres diferentes distribuciones, normal y *t* de Student con 3 y 10 grados de libertad, respectivamente. Tanto el número de repeticiones como el tamaño de muestra son de 1000.

El objetivo de estas simulaciones es evaluar el comportamiento de las metodologías, calculando el intervalo de confianza al 90% con un horizonte de predicción de un día del VaR(99%) explicado en la sección 2. Para este propósito, se estiman las razones de cobertura para cada método: normalidad asintótica, normalidad asintótica del estimador de Hill, *data tilting* y *bootstrap* de submuestra.⁶ Las tasas de cobertura se definen como la proporción de repeticiones en las que el intervalo contiene el VaR real simulado.

⁶ Siguiendo a Chan et al. [2007] y Spierdijk [2014], k es igual a $1,5(\log T)^2$ para el estimador asintótico de Hill y B es definido como $l_3 = (3,5T^{1/2} + 2T^{2/3})$ para el *bootstrap* de submuestra, respectivamente.

Los Cuadros 1, 2 y 3 muestran las razones de cobertura para diferentes simulaciones. La primera columna de estas tablas especifica los parámetros de la simulación mientras que los parámetros que no son mencionados toman los valores que se describieron anteriormente. La segunda columna indica el método evaluado y el resto de columnas contienen los indicadores de cobertura para cada PGD (*iid* y GARCH) y las tres distribuciones usadas (normal, t_3 , t_{10}).

El Cuadro 1 muestra las simulaciones para un conjunto de valores de los parámetros de la ecuación de media $\{(\mu; \phi_1)\} = \{(0;0), (1;0,1), (1;0,4), (1;0,9)\}$. En el Cuadro 2, el ejercicio se realiza para un conjunto de valores de parámetros de la ecuación de volatilidad $\{(c; b_1; a_1)\} = \{(0,1;0,05;0,92), (0,1;0,005;0,99), (0,1;0,3;0,4)\}$. Finalmente, el Cuadro 3 reporta las razones de cobertura para diferentes probabilidades del VaR ($\alpha = \{0,90;0,95;0,99\}$) y varios niveles de confianza del intervalo ($\beta = \{0,90;0,95;0,99\}$).

En el caso *iid* del Cuadro 1, puede decirse que todas las metodologías tienen un buen comportamiento, ya que todas las razones de cobertura son cercanas al nivel del intervalo de confianza, fijado al 90%. En particular, los métodos de *data tilting* y *bootstrap* de submuestra presentan las razones de cobertura más cercanas al valor esperado. Sin embargo, su desempeño se deteriora cuando aumenta la persistencia en la ecuación de media ($\phi_1 = 0,4$ y $\phi_1 = 0,9$) en el caso de efectos GARCH. En el Cuadro 2, sucede algo similar cuando hay baja persistencia en la ecuación de volatilidad, $(c, b_1, a_1) = (0,1, 0,3, 0,4)$.

El Cuadro 3 deja ver que las razones de cobertura para la mayoría de metodologías tienen buenos resultados para cuantiles altos en el nivel del VaR, $\alpha = 0,99$. En particular, los intervalos calculados con los métodos del estimador asintótico de Hill y *data tilting* obtienen resultados pobres con un $VaR(95\%)$ y $VaR(90\%)$.

CUADRO 1. Tasas de cobertura simuladas para diferentes parámetros AR

AR Parámetros	Método	<i>iid</i>			GARCH		
		Normal	t_3	t_{10}	Normal	t_3	t_{10}
$\mu = 0$ $\phi_1 = 0$	Normal asintótico	0.90	0.49	0.83	0.90	0.51	0.87
	Hill asintótico	0.94	0.87	0.94	0.88	0.78	0.88
	<i>Data Tilting</i>	0.92	0.89	0.93	0.84	0.82	0.89
	<i>Bootstrap</i> de submuestra	0.91	0.93	0.88	0.86	0.87	0.85
$\mu = 1$ $\phi_1 = 0,1$	Normal asintótico	0.88	0.49	0.85	0.88	0.54	0.82
	Hill asintótico	0.99	0.93	0.97	0.95	0.88	0.94
	<i>Data Tilting</i>	0.90	0.90	0.92	0.87	0.86	0.88
	<i>Bootstrap</i> de submuestra	0.90	0.92	0.91	0.92	0.88	0.94
$\mu = 1$ $\phi_1 = 0,4$	Normal asintótico	0.89	0.48	0.85	0.62	0.56	0.62
	Hill asintótico	0.99	0.95	0.98	0.97	0.90	0.95
	<i>Data Tilting</i>	0.89	0.91	0.91	0.86	0.82	0.88
	<i>Bootstrap</i> de submuestra	0.90	0.92	0.91	0.10	0.16	0.05
$\mu = 1$ $\phi_1 = 0,9$	Normal asintótico	0.89	0.46	0.84	0.26	0.28	0.25
	Hill asintótico	1.00	0.99	1.00	1.00	0.98	0.99
	<i>Data Tilting</i>	0.90	0.89	0.91	0.85	0.81	0.87
	<i>Bootstrap</i> de submuestra	0.92	0.94	0.88	0.12	0.13	0.14

Tasas de cobertura simuladas para los intervalos de confianza al 90% del VaR(99%) un periodo adelante. El proceso generador de datos (PGD) asociado a las columnas 3 hasta la 5 (*iid*) es $r_t \stackrel{iid}{\sim} F$ con media $\mu/(1-\phi_1)$ y varianza $c/(1-b_1-a_1)$, con $c=0,1$, $b_1=0,05$ y $a_1=0,92$, donde F es normal, t de Student con 3 grados de libertad o t de Student con 10 grados de libertad. Mientras que el PGD asociado a las últimas tres columnas (GARCH) es $r_t = \mu + \phi_1 r_{t-1} + \varepsilon_t$ y $\sigma_t^2 = c + b_1 \varepsilon_{t-1}^2 + a_1 \sigma_{t-1}^2$ con $c=0,1$, $b_1=0,05$ y $a_1=0,92$. Los residuos estandarizados del modelo GARCH se distribuyen normal, t de Student con 3 grados de libertad o t de Student con 10 grados de libertad.

Como se esperaba, el método de normalidad asintótica muestra malos resultados usando una distribución t de Student con 3 grados de libertad. También presenta bajos niveles de cobertura cuando se viola el supuesto *iid* de correlación serial, i.e. escenarios GARCH con $\phi_1 > 0,1$ y $b_1 = 0,3$. En el caso del estimador de Hill, esta metodología tiende a sobre estimar los intervalos de confianza en la mayoría de los casos. Adicionalmente, este método subestima la cobertura para simulaciones que calculan un VaR(90%) y el VaR(95%) como se muestra en el Cuadro 3. Por otro lado, cuando hay un alto nivel de persistencia en la ecuación de volatilidad, que corresponde al grupo de simulaciones del Cuadro 2, el estimador asintótico de Hill muestra cobertura más baja, acercándose al valor esperado.

En general, para el método de *data tilting* se obtienen porcentajes de cobertura cercanos a los esperados, aunque se encuentran levemente por debajo de dicho valor bajo la presencia de efectos GARCH en la serie, en especial cuando se tiene alta persistencia en la media.

CUADRO 2. Tasas de cobertura simuladas para diferentes parámetros GARCH

GARCH Parámetros	Método	<i>iid</i>			GARCH		
		Normal	t_3	t_{10}	Normal	t_3	t_{10}
$c = 0,1$ $b_1 = 0,05$ $a_1 = 0,92$	Normal asintótico	0.88	0.50	0.85	0.86	0.51	0.81
	Hill asintótico	0.98	0.93	0.97	0.95	0.87	0.94
	<i>Data Tilting</i>	0.90	0.90	0.92	0.87	0.86	0.88
	<i>Bootstrap</i> de submuestra	0.90	0.92	0.91	0.92	0.88	0.94
$c = 0,1$ $b_1 = 0,005$ $a_1 = 0,99$	Normal asintótico	0.89	0.50	0.82	0.87	0.51	0.88
	Hill asintótico	0.96	0.89	0.94	0.95	0.86	0.93
	<i>Data Tilting</i>	0.89	0.90	0.93	0.89	0.85	0.85
	<i>Bootstrap</i> de submuestra	0.91	0.92	0.89	0.94	0.91	0.92
$c = 0,1$ $b_1 = 0,3$ $a_1 = 0,5$	Normal asintótico	0.89	0.49	0.84	0.55	0.39	0.53
	Hill asintótico	1.00	0.98	1.00	0.99	0.94	0.97
	<i>Data Tilting</i>	0.88	0.90	0.91	0.84	0.83	0.86
	<i>Bootstrap</i> de submuestra	0.86	0.93	0.89	0.48	0.66	0.63

Tasas de cobertura simuladas para los intervalos de confianza al 90% del VaR(99%) un periodo adelante. El PGD asociado a las columnas 3 hasta la 5 (*iid*) es $r_t \stackrel{iid}{\sim} F$ con media $\mu/(1-\phi_1)$ y varianza $c/(1-b_1-a_1)$, con $\mu = 1$ y $\phi_1 = 0,1$, donde F es normal, t de Student con 3 grados de libertad o t de Student con 10 grados de libertad. Mientras que el PGD asociado a las últimas tres columnas (GARCH) es $r_t = \mu + \phi_1 r_{t-1} + \varepsilon_t$ y $\sigma_t^2 = c + b_1 \varepsilon_{t-1}^2 + a_1 \sigma_{t-1}^2$ con $\mu = 1$ y $\phi_1 = 0,1$. Los residuos estandarizados del modelo GARCH se distribuyen normal, t de Student con 3 grados de libertad o t de Student con 10 grados de libertad.

El *bootstrap* de submuestra presenta los indicadores de cobertura más cercanos al valor esperado de los mismos; sin embargo, los niveles más bajos se obtienen cuando hay una alta correlación serial, i.e., cuando $\phi_1 = 0,9$, y cuando hay una baja persistencia en la varianza (tercer grupo de simulaciones del Cuadro 2).

Es importante mencionar que los cuatro métodos de intervalos de confianza evaluados son sensibles a diferentes escenarios. El grado de persistencia en media afecta negativamente el desempeño de las cuatro metodologías, mientras que el grado de persistencia en varianza resulta en tasas de cobertura más cerca de los valores esperados. Por otro lado, algunas distribuciones de cola pesada pueden afectar la razón de cobertura de los estimadores realizados por el método de normalidad asintótica, como lo muestran los resultados para la distribución t de Student con 3 grados de libertad en los Cuadros 1, 2 y 3. Finalmente, los porcentajes de cobertura de los métodos del estimador asintótico de Hill y *data tilting* disminuyen cuando se consideran los efectos GARCH en lugar de una distribución *iid*, al igual que con la existencia de colas pesadas.

CUADRO 3. Tasas de cobertura simuladas para diferentes niveles de confianza y diferentes niveles del VaR

$VaR(\alpha)$ $C.I.(\beta)$	Método	<i>iid</i>			GARCH		
		Normal	t_3	t_{10}	Normal	t_3	t_{10}
$\alpha = 0,99$ $\beta = 0,90$	Normal asintótico	0.90	0.47	0.83	0.90	0.54	0.85
	Hill asintótico	0.98	0.91	0.97	0.96	0.87	0.94
	<i>Data Tilting</i>	0.90	0.90	0.92	0.87	0.86	0.88
	<i>Bootstrap</i> de submuestra	0.90	0.92	0.91	0.92	0.88	0.94
$\alpha = 0,95$ $\beta = 0,90$	Normal asintótico	0.88	0.59	0.84	0.88	0.62	0.85
	Hill asintótico	0.36	0.43	0.36	0.31	0.40	0.34
	<i>Data Tilting</i>	0.59	0.73	0.61	0.55	0.60	0.58
	<i>Bootstrap</i> de submuestra	0.95	0.95	0.95	0.60	0.47	0.60
$\alpha = 0,90$ $\beta = 0,90$	Normal asintótico	0.89	0.66	0.84	0.83	0.66	0.82
	Hill asintótico	0.26	0.41	0.27	0.28	0.34	0.28
	<i>Data Tilting</i>	0.64	0.73	0.74	0.65	0.64	0.66
	<i>Bootstrap</i> de submuestra	0.95	0.95	0.96	0.46	0.34	0.47
$\alpha = 0,99$ $\beta = 0,95$	Normal asintótico	0.94	0.57	0.89	0.93	0.61	0.89
	Hill asintótico	0.99	0.97	0.99	0.97	0.95	0.98
	<i>Data Tilting</i>	0.95	0.94	0.96	0.93	0.92	0.94
	<i>Bootstrap</i> de submuestra	0.96	0.96	0.95	0.96	0.93	0.97
$\alpha = 0,99$ $\beta = 0,99$	Normal asintótico	0.98	0.70	0.96	0.99	0.75	0.96
	Hill asintótico	1.00	0.99	1.00	1.00	0.96	0.99
	<i>Data Tilting</i>	0.99	0.98	1.00	0.98	0.96	0.99
	<i>Bootstrap</i> de submuestra	0.99	0.99	0.99	0.99	0.98	0.99

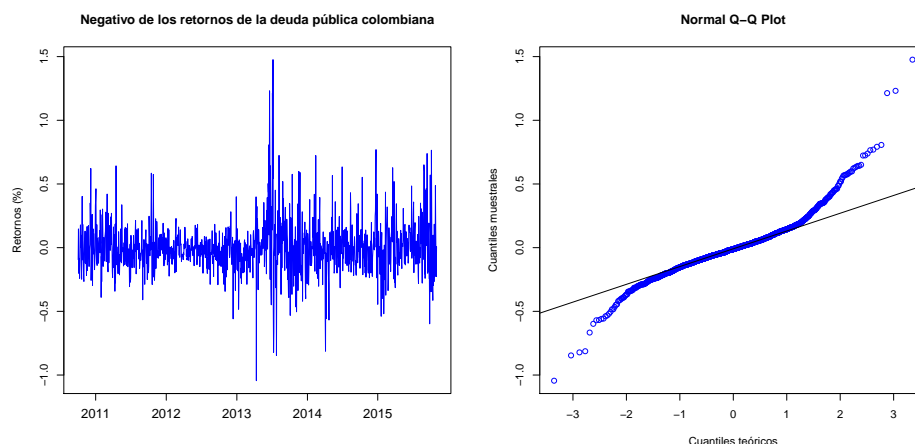
Tasas de cobertura simuladas para los intervalos de confianza al $\beta\%$ del $VaR(\alpha)$ un periodo adelante. El PGD asociado a las columnas 3 hasta la 5 (*iid*) es $r_t \stackrel{iid}{\sim} F$ con media $\mu/(1-\phi_1)$ y varianza $c/(1-b_1-a_1)$, con $\mu = 1$, $\phi_1 = 0,1$, $c = 0,1$, $b_1 = 0,05$ y $a_1 = 0,92$, donde F es normal, t de Student con 3 grados de libertad o t de Student con 10 grados de libertad. Mientras que el PGD asociado a las últimas tres columnas (GARCH) es $r_t = \mu + \phi_1 r_{t-1} + \varepsilon_t$ y $\sigma_t^2 = c + b_1 \varepsilon_{t-1}^2 + a_1 \sigma_{t-1}^2$ con $\mu = 1$, $\phi_1 = 0,1$, $c = 0,1$, $b_1 = 0,05$ and $a_1 = 0,92$. Los residuos estandarizados del modelo GARCH se distribuyen normal, t de Student con 3 grados de libertad o t de Student con 10 grados de libertad.

4. EJERCICIO EMPÍRICO

El mercado de deuda pública en Colombia es el principal vehículo de inversión de las entidades financieras. En promedio, a junio de 2015 el 37% del total de las inversiones se concentraron en títulos de endeudamiento soberano (TES) (Véase Departamento de Estabilidad Financiera [2015]). Por lo anterior es relevante calcular el valor en riesgo de este portafolio y una medida de incertidumbre que evalúe su exactitud.

Esta sección contiene una aplicación empírica de los cuatro métodos para estimar los intervalos de confianza del VaR presentados en la sección 2. En este ejercicio se usaron los retornos diarios negativos de los diferentes vértices de la curva cero cupón en pesos y UVR⁷ ponderados por el saldo reportado en el Depósito Central de Valores (DCV) que tiene el sistema financiero colombiano en cada vértice⁸, desde el 9 de octubre del 2010 hasta el 23 de octubre del 2015, que resulta en un tamaño de muestra de 1250 observaciones. La serie de los retornos y el QQ-plot normal se presentan en la Figura 1. Ambos gráficos exhiben los hechos estilizados clásicos de las series de tiempo financieras, i.e. conglomerados de volatilidad y comportamiento de colas pesadas.

FIGURA 1. Negativo de los retornos de la deuda pública colombiana



El panel izquierdo muestra el negativo de los retornos de la deuda pública (%) para la muestra completa, desde el 9 de Octubre del 2010 hasta el 23 de Octubre del 2015 y el panel derecho presenta el QQ-plot normal de los retornos.

Con el fin de evaluar el desempeño de estos métodos, se computan los intervalos de confianza del 90% para el $VaR(95\%)$ y $VaR(99\%)$ del periodo siguiente. Con este propósito, inicialmente se construye el VaR del 10 de octubre del 2014 estimando un modelo $AR(p)$ -GARCH(1,1) que usa la información disponible hasta el 10 de octubre del 2014. De ahí en adelante, el conjunto de información se aumenta una observación a la vez, por lo cual se estima un nuevo modelo $AR(p)$ -GARCH(1,1). Este procedimiento se implementa recursivamente 250 veces, hasta que el conjunto de información alcance el 22 de octubre del 2015.⁹

Las Figuras 2 y 3 exhiben los retornos (negativos) de la deuda pública colombiana desde el 14 de octubre del 2014 hasta el 23 de octubre del 2015; los 250 pronósticos del VaR (95%) y VaR(99%)

⁷Las curvas cero cupón en pesos y UVR las suministra diariamente el proveedor de precios para valoración INFOVALMER.

⁸Para mayor información sobre la metodología de vértices véase JP Morgan [1996].

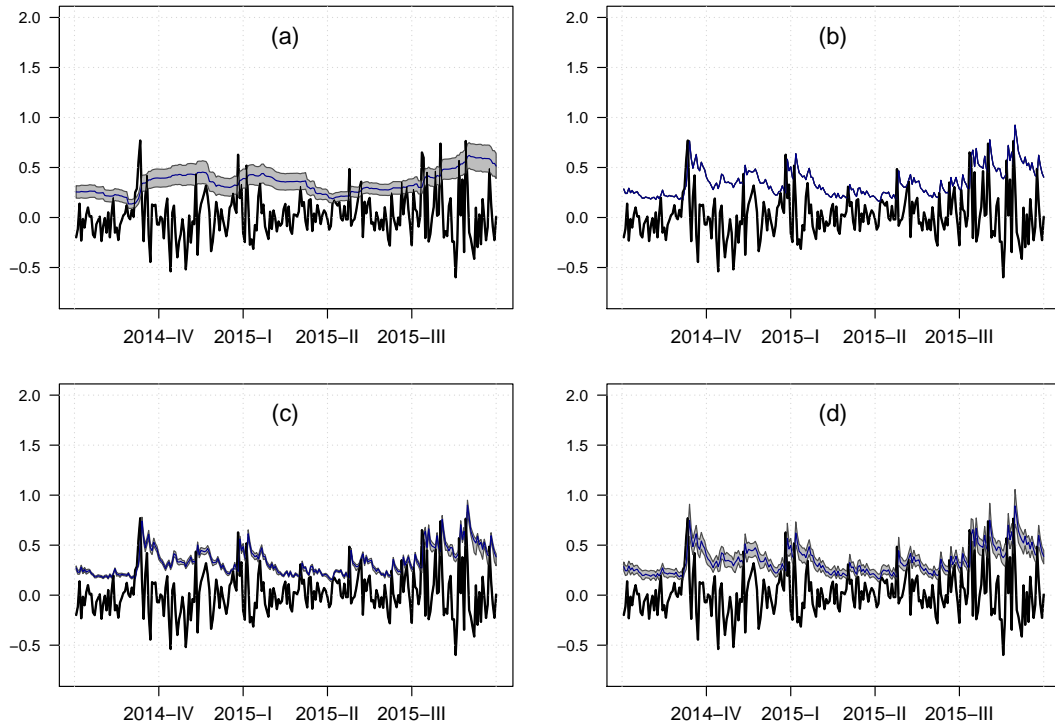
⁹Las pruebas de especificación asociadas al modelo $AR(p)$ -GARCH(1,1) para toda la muestra se presentan en los Cuadros 4, 5 y 6 del Apéndice A. Los resultados de estas pruebas de hipótesis no indican una especificación incorrecta.

un periodo adelante, y los intervalos de confianza al 90% estimados por cada método. Como resultado se aprecia que para escenarios normalmente distribuidos, los intervalos de confianza son más anchos que los intervalos computados usando otras técnicas. Este método también falla al reconocer la dinámica de la volatilidad de la serie de retornos, principalmente porque este método asume que la serie de tiempo es independiente e idénticamente distribuida normal, mientras que los demás métodos sí modelan este hecho.

Para el caso del VaR(95%) en la Figura 2, los intervalos de confianza para, Estimación de Hill asintótica y *Data Tilting*, son extremadamente angostos. Esto implica una estimación más eficiente del VaR; sin embargo, ésta puede ser inconsistente ya que los intervalos muestran una baja tasa de cobertura en el ejercicio de simulación. Por otro lado, los intervalos del *bootstrap* de submuestra no son tan angostos como los intervalos de los otros dos métodos.

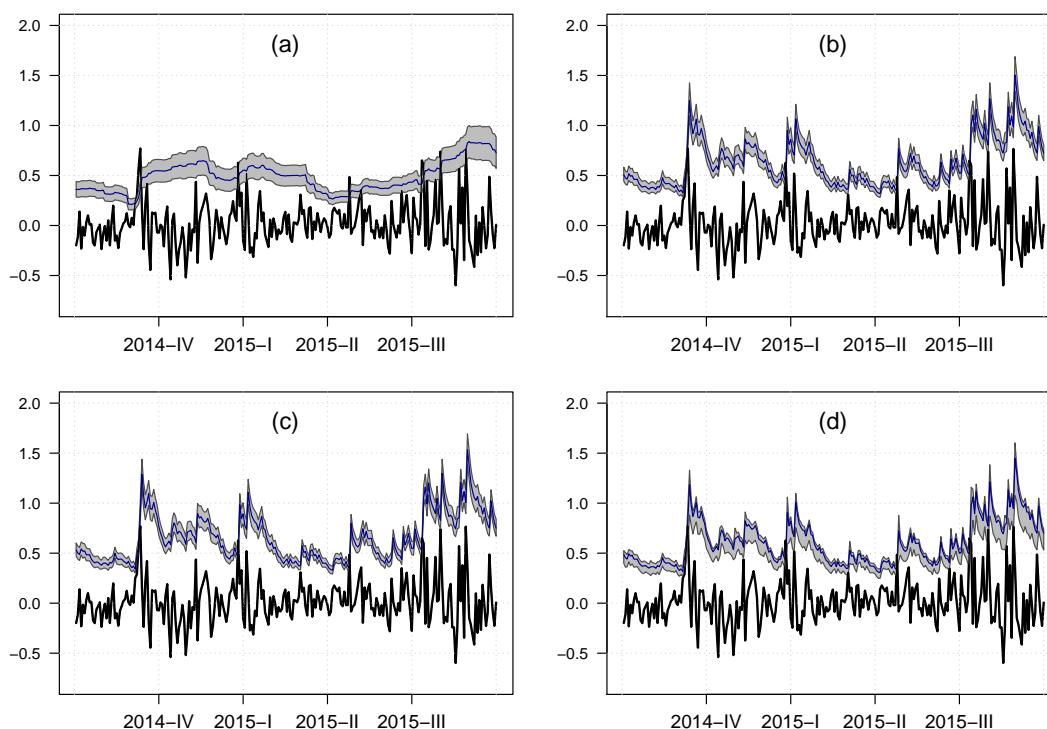
En el caso del VaR(99%) expuesto en la Figura 3, los intervalos de confianza del estimador asintótico de Hill y de *Data Tilting* resultan muy similares. Este resultado es esperado ya que ambos métodos se basan en la estimación por QML y teoría del valor extremo. Para la metodología de *bootstrap* de submuestra, los intervalos obtenidos son más anchos (menos eficientes) que los dos anteriores. Sin embargo, Spierdijk [2014] afirma que el ancho de los intervalos de confianza del *bootstrap* de submuestra se debe a la robustez de este método a la ausencia de normalidad asintótica.

FIGURA 2. Intervalos de confianza al 90% para el $VaR(95\%)$ del negativo de los retornos de la deuda pública colombiana



Intervalos de confianza al 90% para el $VaR(95\%)$ del negativo de los retornos de la deuda pública colombiana desde el 14 de octubre del 2014 hasta el 23 de octubre del 2015. El panel (a) muestra la estimación de normalidad asintótica de Moraux [2011]; el panel (b) la estimación de Hill asintótica de Chan et al. [2007]; el panel (c), la estimación por *Data Tiling* de Chan et al. [2007] y el panel (d), la estimación de *bootstrap* de submuestra de Spierdijk [2014].

FIGURA 3. Intervalos de confianza al 90% para el $VaR(99\%)$ del negativo de los retornos de la deuda pública colombiana



Intervalos de confianza al 90% para el $VaR(99\%)$ del negativo de los retornos de la deuda pública colombiana desde el 14 de octubre del 2014 hasta el 23 de octubre del 2015. El panel (a) muestra la estimación de normalidad asintótica de Moraux [2011]; el panel (b) la estimación de Hill asintótica de Chan et al. [2007]; el panel (c), la estimación por *Data Tilting* de Chan et al. [2007] y el panel (d), la estimación de *bootstrap* de submuestra de Spierdijk [2014].

5. CONCLUSIONES

Este artículo evalúa el desempeño de diferentes metodologías para estimar los intervalos de confianza del Valor en Riesgo usando ejercicios de Monte Carlo. Los métodos evaluados son normalidad asintótica, estimador de Hill asintótico, *Data Tilting* y *bootstrap* de submuestra. El estudio de simulación recae sobre porcentajes de cobertura como medida de desempeño para descifrar la robustez de las metodologías ante diferentes escenarios, como persistencia en media y varianza, niveles de probabilidad del VaR, niveles de confianza de los intervalos y distribuciones de probabilidad.

En general, el método de *bootstrap* de submuestra presenta el mejor desempeño de las cuatro metodologías estudiadas. Como se esperaba, la aproximación de normalidad asintótica tiene las peores tasas de cobertura, ya que asume dinámicas *iid*. También es importante notar que los porcentajes

de cobertura de los cuatro métodos estudiados presentan una ligera caída cuando se enfrentan a distribuciones de colas pesadas. Concluyendo, los resultados de la simulación muestran que los métodos del estimador Hill, *data tilting*, y *bootstrap* de submuestra, son sólo válidos para cuantiles extremos. En particular, hay un buen desempeño del VaR(99%) en términos de tasas de cobertura, y uno desfavorable para el VaR(95%) y el VaR(90%).

El ejercicio empírico confirma los principales resultados encontrados en las simulaciones. El VaR(95%), el VaR(99%) y sus intervalos de confianza obtenidos por el método de normalidad asintótica no capturan la dinámica de la volatilidad de la serie analizada. Adicionalmente, estos intervalos resultan ser considerablemente anchos comparados con los del resto de metodologías estudiadas. Por otro lado, los intervalos de confianza para el VaR(95%) calculados por *data tilting*, estimador de Hill y *bootstrap* de submuestra son extremadamente angostos, lo que podría indicar que estas estimaciones son inconsistentes, como se mostró en los ejercicios de simulación. Por último, los intervalos del VaR(99%) para estas tres metodologías presentan dinámicas similares, lo que sugiere que cualquiera de estos métodos es un estimador adecuado como medida de la incertidumbre del VaR para los retornos de los bonos de deuda pública del sistema financiero colombiano.

REFERENCIAS

- CHAN, N. H., S.-J. DENG, L. PENG, Y Z. XIA (2007): "Interval estimation of value-at-risk based on GARCH models with heavy-tailed innovations," *Journal of Econometrics*, 137(2), 556–576.
- DEPARTAMENTO DE ESTABILIDAD FINANCIERA (2015): "Reporte de estabilidad financiera," Banco de la República, No. 2.
- FRANCQ, C., Y J.-M. ZAKOÏAN (2015): "Risk-parameter estimation in volatility models," *Journal of Econometrics*, 184(1), 158–173.
- GAO, F., Y F. SONG (2008): "Estimation risk in GARCH VaR and ES estimates," *Econometric Theory*, 24(5), 1404–1424.
- HALL, A., Y Q. YAO (2003): "Data tilting for time series," *Journal of the Royal Statistical Society: Series B*, 65(2), 425–442.
- HILL, B. M. (1975): "A simple general approach to inference about the tail of a distribution," *The annals of statistics*, 3(5), 1163–1174.
- JP MORGAN (1996): "Riskmetrics: technical document," Morgan guaranty trust company of new york.
- MORAUX, F. (2011): "How valuable is your VaR? Large sample confidence intervals for normal VaR," *Journal of Risk Management in Financial Institutions*, 4(2), 189–200.
- OWEN, A. (1988): "Empirical likelihood ratio confidence intervals for single function," *Biometrika*, 75(2), 967–975.
- (1990): "Empirical likelihood regions," *Annals of Statistics*, 18(1), 90–120.
- (2001): *Empirical likelihood*. Chapman and Hall, London.
- SPIERDIJK, L. (2014): "Confidence intervals for ARMA-GARCH Value-at-Risk: The case of heavy tails and skewness," *Computation Statistics and Data Analysis*, <http://dx.doi.org/10.1016/j.csda.2014.08.011>.

**APÉNDICE A. PRUEBAS DE DIAGNÓSTICO DEL MODELO AR-GARCH PARA LOS
RETORNOS DE LA DEUDA PÚBLICA COLOMBIANA**

CUADRO 4. Prueba de sesgo de signo

	estadístico-t	p-valor
Sesgo de signo	0.482	0.630
Signo negativo	0.583	0.560
Signo positivo	0.911	0.363

CUADRO 5. Prueba Ljung-Box para los residuos estandarizados

Rezago	estadístico- χ^2	p-valor
20	32.388	0.039
50	57.027	0.230
100	121.710	0.069

CUADRO 6. Prueba Ljung-Box para el cuadrado de los residuos estandarizados

Rezago	estadístico- χ^2	p-valor
20	20.164	0.448
50	52.641	0.372
100	116.810	0.120