

Distribución hiperbólica generalizada: una aplicación en la selección de portafolios y en cuantificación de medidas de riesgo de mercado

Recibido: Octubre 30, 2014 - Aceptado: Abril 5, 2016

Doi: [dx.doi.org/10.12804/rev.econ.rosario.18.02.2015.04](https://doi.org/10.12804/rev.econ.rosario.18.02.2015.04)

José Luis Alayón G.*

Resumen

La distribución hiperbólica generalizada ha sido usada por académicos y profesionales para eliminar los problemas de colas de distribución delgadas en finanzas, y por su utilidad en la modelación de los retornos de los activos y de las medidas de riesgo de mercado. En este trabajo, la distribución hiperbólica generalizada es usada para encontrar el portafolio óptimo y su riesgo de mercado. Igualmente, se desarrolla un método para la Selección de Portafolio Robusto la cual reduce la sensibilidad del portafolio ante variaciones de los parámetros de la distribución. Luego de esto, se muestra un esquema comparativo para determinar cómo la inclusión del nuevo método representa un avance respecto a la teoría de selección de portafolios de Markowitz. Por último, en algunos gráficos se muestra el efecto de los parámetros sobre la forma de la distribución, lo que se usa para generar escenarios de estrés y portafolios óptimos.

Clasificación JEL: C15, C44, C61

Palabras clave: distribución hiperbólica generalizada, selección de portafolio, selección de portafolio robusto, valor en riesgo condicional, Markowitz, , valor en riesgo condicional del peor escenario, asignación de activos, administración de riesgos, multiciclo, expectativa, y momento de estimación condicional.

* Portfolio Manager en Acciones y Valores S.A. Correo electrónico: joseluisalayon@hotmail.com

Para citar este artículo: Alayón González, J. L. (julio-diciembre, 2015). Distribución hiperbólica generalizada: una aplicación en la selección de portafolios y cuantificación de medidas de riesgo de mercado. *Revista de Economía del Rosario*, 18(2), 249-308. Doi: [dx.doi.org/10.12804/rev.econ.rosario.18.02.2015.04](https://doi.org/10.12804/rev.econ.rosario.18.02.2015.04)

Generalized Hyperbolic Distribution: An Application to Portfolio Selection and Measures of Market Risk

Abstract

The generalized hyperbolic distribution has been used for academics and practitioners to eliminate the problems of light tails of distributions in finance, and for its usefulness in the modeling of the assets returns and market risk measurements. In this paper, the generalized hyperbolic distribution is used to find the optimal portfolio and its market risk. Likewise, a methodology to robust portfolio selection is developed to reduce the sensibility in the optimal weights given the uncertainty in the real value of the parameters of the distribution. Thereafter a comparative scheme is showed to determinate the advance with the new methodology with regard to Markowitz Portfolio Selection. Later, in some graphics we can see the effects of the parameters in the shape of the distribution, and we use this to generate hypothetical scenarios to build a stress test analysis and optimal portfolios.

JEL Classifications: C15, C44, C61.

Keywords: generalized hyperbolic distribution, portfolio selection, robust portfolio selection, Conditional Value at Risk (CVaR), Markowitz, worse case Conditional Value at Risk, asset allocation, risk management, multi-cycle, expectation, and conditional estimation method.

Distribuição hiperbólica generalizada: uma aplicação na seleção de portfólios e quantificação de medidas de risco de mercado

Resumo

A distribuição hiperbólica generalizada tem sido usada por acadêmicos e profissionais para eliminar os problemas de filas de distribuição delgadas em finanças, e por utilidade na modelação dos retornos dos ativos e das medidas de risco do mercado. Neste trabalho, a distribuição hiperbólica generalizada é usada para encontrar o portfólio ótimo e seu risco de mercado. Igualmente, desenvolve-se um método para a Seleção de Portfólio Robusto a qual reduz a sensibilidade do portfólio ante variações dos parâmetros da distribuição. Depois disto, se mostra um esquema comparativo para determinar como a inclusão do novo método representa um avanço respeito à teoria de seleção de portfólios de Markowitz. Por último, em alguns gráficos se mostra o efeito dos parâmetros sobre a forma da distribuição, o que se usa para gerar cenários de stress e portfólios ótimos.

Classificações JEL: C15, C44, C61

Palavras-chave: distribuição hiperbólica generalizada, seleção de portfólio, seleção de portfólio robusto, valor em risco condicional, Markowitz, valor em risco condicional do pior cenário, assinatura de ativos, administração de riscos, multiciclo, expectativa, e momento de estimação condicional.

Capítulo 1. Introducción

Los modelos financieros actuales han avanzado significativamente con respecto a aquellos que partían de supuestos gaussianos, al lograrse incluir características importantes de los retornos de los activos, tales como colas pesadas y asimetrías en sus distribuciones. No obstante, a medida que las distribuciones se vuelven más convenientes para este tipo de modelaciones, el costo computacional también se incrementa rápidamente, por lo que la escogencia de un buen modelo debe hacerse siguiendo un análisis de costo-beneficio riguroso, para evitar desgastes innecesarios.

En medio de este contexto se han propuesto modelos bajo distribuciones de colas pesadas o semipesadas que logran una mejor concordancia con el comportamiento de los activos financieros, pasando por la t-Student hasta distribuciones de valores extremos. El problema con la t-Student es que converge demasiado pronto a la normal, y con las distribuciones de valores extremos, que su ajuste no es tan bueno como se desea en la mayoría de los casos, por lo que las distribuciones hiperbólicas generalizadas (GHD) han elevado su importancia en los últimos años, ya que ante especificaciones de algunos de sus parámetros se puede tener una amplia gama de distribuciones particulares. Dentro de esta se encuentran distribuciones que presentan desde colas semipesadas hasta de valores extremos, junto con características de asimetría y curtosis fácilmente ajustables a los objetivos de la investigación que se desee desarrollar, al tiempo que posee propiedades útiles para la modelación de derivados financieros, la cuantificación de riesgos y las simulaciones financieras.

El uso de este tipo de distribuciones fue introducido por Barndorff-Nielsen en 1977, en la ilustración de algunos estudios prácticos, pero fue solo hasta 1995 que Eberlein y Keller lograron implementarla a una serie financiera con un alto grado de precisión, ya que la complejidad que conlleva calibrar los parámetros y la restricción en términos computacionales que existía hasta entonces, no había permitido encontrar un algoritmo óptimo para tal fin. Además, Prause con su trabajo de 1999 logró derivar aspectos importantes de esta distribución sobre la función de máxima verosimilitud, lo que ayudó a que surgieran algoritmos cada vez más precisos para la calibración de esta distribución probabilística. Desde entonces ha sido utilizada ampliamente para la valoración de riesgo de activos financieros [Eberlein (1998), McNeil (2005), Breyman y otros (2003)]; en la selección de portafolios [Hellmitch, Kassberg (2009), en modelos de Lévy [Barndorff-Nielsen(1997) y Prause(1999)], y en múltiples trabajos de modelación empírica.

En el caso particular de selección de portafolios, la literatura también ha dado un salto importante gracias a la nueva generación de medidas de riesgo basadas en los cuantiles de la distribución, tales como el valor en riesgo (VaR) o el valor en riesgo condicional (CVaR), debido a que la optimización se logra bajo conceptos más completos de riesgo de mercado que los alcanzados con el modelo de Markowitz o medidas tradicionales.

Así pues, el objetivo de este trabajo será la calibración de los parámetros de la distribución hiperbólica generalizada con una metodología que asegure una alta precisión, para luego generar un proceso que dé como resultado un portafolio óptimo¹ desde un esquema de retorno objetivo. Todo esto con el fin de hacer comparaciones frente al modelo de Markowitz en términos de desempeño y riesgo, y así determinar el avance cualitativo que genera la inclusión de una distribución de colas pesadas, asimetrías y alta curtosis en el proceso de selección de portafolios. Además, se analizarán los alcances de la distribución hiperbólica generalizada en la modelación de escenarios hipotéticos y en la conformación de portafolios óptimos bajo diferentes requerimientos del investigador, los cuales pueden ir más allá del de mínimo riesgo.

Para tal fin, el trabajo estará dividido de la siguiente forma: en la sección 2 se abordará la caracterización, desarrollo, calibración y ajuste de una distribución hiperbólica generalizada a una serie financiera; posteriormente, se hará el desarrollo de la metodología de portafolios óptimos a partir de la distribución hiperbólica generalizada; en la sección 4, se simularán portafolios óptimos bajo diferentes escenarios; en 5, se presentarán escenarios de estrés (*stress test*) y, por último, se expondrán las conclusiones del trabajo.

Capítulo 2. Distribución hiperbólica generalizada multivariada (mGHD)

La distribución hiperbólica generalizada se encuentra parametrizada de diferentes formas en la literatura.² Una de las especificaciones más famosas es la de Prause (1999), cuyos parámetros están definidos en términos de λ , α , β , μ , Δ y δ , pero en general existen más de tres tipologías. Para este trabajo se utilizará la parametrización $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$, la cual es obtenida por medio de la distribución normal media-varianza mixta (*normal mean-variance mixture*)

¹ El término 'óptimo' va a ser entendido en este contexto como el portafolio de mínimo riesgo asociado.

² Algunas parametrizaciones son expuestas en el anexo 1.

cuando $\omega \sim GI G^3(\lambda, \chi, \psi)$, ya que permite extraer de forma más intuitiva algunas características de esta distribución.

2.1. Normal media-varianza mixta (normal mean-variance mixture)

Suponga una variable aleatoria que se distribuye de la siguiente forma:

$$X^d = \mu + W\gamma + \sqrt{W}AZ \quad (2.1.1)$$

Donde,

$$Z \sim \mathbb{N}_k(0, I_k)$$

$W \in [0, \infty)$ y es independiente de Z

$$A \in \mathbb{R}^{dxk}$$

μ, γ son vectores en \mathbb{R}^d

Así pues, se llega a que la variable condicional sigue una distribución de este tipo.⁴

$$X | W = \omega \sim \mathbb{N}_d(\mu + \omega\gamma, \omega\Sigma) \quad (2.1.2)$$

Donde,

$$\Sigma = AA'.^5$$

Por lo que si la variable latente W ⁶ se distribuye inversa gaussiana generalizada $GIG(\lambda, \chi, \psi)$, entonces X debe seguir una distribución hiperbólica generalizada (GHD, por su sigla en inglés), donde μ hace referencia al parámetro de localización (media), Σ es el parámetro de escala (matriz de varianzas y covarianzas), γ el de asimetría ($\gamma = 0$, entonces la distribución es simétrica) y λ, χ, ψ controlan la forma de la distribución (peso de las colas y empinamiento).

³ GIG hace referencia a la distribución inversa gaussiana generalizada (*generalized inverse gaussian distribution*).

⁴ Ver 2.1.1 Momentos de la variable X^d .

⁵ Es por esta característica que la distribución es conocida como normal media-varianza mixta.

⁶ La variable mixta W puede tener muchas interpretaciones interesantes en economía, en términos de choques que afectan la varianza y la media del modelo normal estándar, por ello son interesantes en muchas áreas de estudio este tipo de distribuciones.

2.1.1. *Momentos de la variable X^d*

Dada la forma funcional de X^d se puede llegar a que los dos primeros momentos centrados están representados de la siguiente forma:

$$\mathbb{E}[X] = \mu + \mathbb{E}[W]\gamma \tag{2.1.1.1}$$

$$var [X] = \mathbb{E}[cov(X | W)] + cov(\mathbb{E}[X | W]) = var(W) \gamma\gamma' + \mathbb{E}[W]\Sigma \tag{2.1.1.2}$$

2.1.2. *Función de densidad condicional de la variable mixta W*

Como la función de $X | W$ es gaussiana con media $\mu + \omega\gamma$ y varianza $\omega\Sigma$, la distribución $W | X$ puede ser hallada fácilmente utilizando el teorema de Bayes.

$$f_{W|X}(\omega) = \frac{f_{X|W}(\omega) f_W(\omega)}{f_X(x)} \tag{2.1.2.1}$$

$$f_{W|X}(\omega) = \frac{f_{X,W}(x, \omega)}{f_X(x)} = \frac{f_{X|W}(\omega) f_{GIG}(\omega)}{\int_0^\infty f_{X|W}(\omega) f_{GIG}(\omega) d\omega} \tag{2.1.2.2}$$

Siguiendo la expresión anterior, con un $W \sim GIG(\lambda, \chi, \psi)$ y con las siguientes fórmulas se puede llegar a una representación para $f_{W|X}(\omega)$:

$$f_{GIG}(\omega) = \left(\frac{\psi}{X}\right)^{\frac{\lambda}{2}} \frac{\omega^{\lambda-1}}{2K_\lambda(\sqrt{\psi X})} \exp\left\{-\frac{1}{2}\left(\frac{X}{\omega} + \psi\omega\right)\right\} \tag{2.1.2.3}$$

$$\begin{aligned} f_{X|W}(\omega) &= \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2} \omega^{d/2}} \exp\left\{-\frac{1}{2}(x - \mu - \omega\gamma)' (\omega\Sigma)^{-1} (x - \mu - \omega\gamma)\right\} \\ &= \frac{e^{(x-\mu)'\Sigma^{-1}\gamma}}{(2\pi)^{d/2} |\Sigma|^{1/2} \omega^{d/2}} \exp\left\{-\frac{Q(x)}{2\omega} - \frac{\gamma'\Sigma^{-1}\gamma}{2/\omega}\right\} \end{aligned} \tag{2.1.2.4}$$

Donde,

$Q(x) = (x - \mu)' \Sigma^{-1} (x - \mu)$ la cual corresponde a la distancia de Mahalanobis.

Así pues, y utilizando la formula (2.1.2.2) se tiene que

$$f_{W|X}(\omega) = \frac{\Omega_1}{\Omega_2} = \frac{f_{X|W}(\omega) f_{GIG}(\omega)}{\int_0^\infty f_{X|W}(\omega) f_{GIG}(\omega) d\omega} \tag{2.1.2.5}$$

Donde,

$$\Omega_1 = \frac{e^{(x-\mu)'\Sigma^{-1}\gamma}}{(2\pi)^{d/2} |\Sigma|^{1/2} K_\lambda(\sqrt{\psi\chi})} \left(\sqrt{\frac{\psi}{\chi}} \right)^\lambda \frac{1}{2} \omega^{\lambda-\frac{d}{2}-1} \exp\left\{-\frac{1}{2}\left(\frac{Q(x)-\chi}{\omega} + (\gamma'\Sigma^{-1}\gamma + \psi)\omega\right)\right\} \tag{2.1.2.6}$$

$$\Omega_2 = \frac{e^{(x-\mu)'\Sigma^{-1}\gamma}}{(2\pi)^{d/2} |\Sigma|^{1/2} K_\lambda(\sqrt{\psi\chi})} \left(\sqrt{\frac{\psi}{\chi}} \right)^\lambda \frac{1}{2} \int_0^\infty \omega^{\lambda-\frac{d}{2}-1} \exp\left\{-\frac{1}{2}\left(\frac{Q(x)-\chi}{\omega} + (\gamma'\Sigma^{-1}\gamma + \psi)\omega\right)\right\} d\omega \tag{2.1.2.7}$$

Luego de simplificar, la expresión de $f_{w|x}(\omega)$ queda representada así:

$$f_{W|X} = \frac{\omega^{\lambda-\frac{d}{2}-1} \exp\left\{-\frac{1}{2}\left(\frac{Q(x)-\chi}{\omega} + (\gamma'\Sigma^{-1}\gamma + \psi)\omega\right)\right\}}{\int_0^\infty \omega^{\lambda-\frac{d}{2}-1} \exp\left\{-\frac{1}{2}\left(\frac{Q(x)-\chi}{\omega} + (\gamma'\Sigma^{-1}\gamma + \psi)\omega\right)\right\} d\omega} \tag{2.1.2.8}$$

Haciendo uso de la función (2.1.3.2) en la expresión Ω_2 , entonces es posible cambiar la integral por la función modificada de Bessel de tercera clase (*modified bessel function of the third kind*), lo que simplifica el término anterior a:

$$f_{W|X}(\omega) = \left(\frac{\gamma'\Sigma^{-1}\gamma + \psi}{Q(x)-\chi} \right)^{\frac{1}{2}\left(\lambda-\frac{d}{2}\right)} \frac{\omega^{\lambda-\frac{d}{2}-1} \exp\left\{-\frac{1}{2}\left(\frac{Q(x)-\chi}{\omega} + (\gamma'\Sigma^{-1}\gamma + \psi)\omega\right)\right\}}{2K_{\lambda-\frac{d}{2}}\left(\sqrt{(Q(x)+\chi)(\gamma'\Sigma^{-1}\gamma + \psi)}\right)} \tag{2.1.2.9}$$

Esta expresión es equivalente a decir que $f_{w|x}(w|x) = \hat{\omega} \sim GIG(\lambda - d/2, Q(x) + \chi, \psi + \gamma'\Sigma^{-1}\gamma)$.

En la ecuación (2.1.2.9) $K_{\lambda-d/2}(\dots)$ hace referencia a la función modificada de Bessel de tercera clase, la cual es expuesta en el siguiente numeral. Este resultado será utilizado más adelante en 2.3.1 cuando se vaya a realizar el ajuste de la distribución a los retornos financieros.

2.1.3. Función modificada de Bessel de tercera clase

Las funciones de Bessel son las soluciones a ecuaciones diferenciales que aparecen en las funciones de Laplace en coordenadas cilíndricas, principalmente. La función modificada de Bessel de tercera clase viene representada de la siguiente forma:

$$K_{\lambda}(x) = \frac{1}{2} \int_0^{\infty} \omega^{\lambda-1} \exp\left\{-\frac{1}{2}x(\omega + \omega^{-1})\right\} d\omega, x > 0 \tag{2.1.3.1}$$

Cabe resaltar que esta integral no tiene solución cerrada y que por su estructura genera algunos problemas algorítmicos cuando se intenta aproximar su solución mediante métodos numéricos, por lo que hay que recurrir a técnicas como las presentadas por Slavinsky y Safouhi (2010), en las que se exhibe un algoritmo recursivo para tal fin.

Sin embargo, se puede hallar una representación cerrada bajo algunas especificaciones de las distribuciones hiperbólicas, así:

Esquema 1. Al sustituir ω por $x\sqrt{\frac{\chi}{\psi}}$ en la función de la distribución hiperbólica generalizada, se tiene que:

$$\begin{aligned} K_{\lambda}(x) &= \frac{1}{2} \int_0^{\infty} \omega^{\lambda-1} \exp\left\{-\frac{1}{2}x\left(\psi\omega + \frac{\chi}{\omega}\right)\right\} d\omega \\ &= 2\left(\frac{\chi}{\psi}\right)^{\frac{\lambda}{2}} K_{\lambda}\left(\sqrt{\chi\psi}\right) \end{aligned} \tag{2.1.3.2}$$

Esta representación resulta especialmente útil en casos donde $\lambda = 1/2$, o, lo que es lo mismo, cuando se tiene una distribución NIG, ya que se puede obtener una representación cerrada para esta función.

$$K_{1/2}(x) = \sqrt{\frac{\pi}{2x}} e^{-x}, x > 0 \tag{2.1.3.3}$$

Este resultado también es válido para $\lambda = -1/2$, gracias a que la función de Bessel es simétrica.⁷

2.2. Función de distribución mGHD

Siguiendo el planteamiento hecho en (2.1.1), y con un Σ no singular, la siguiente función de distribución hiperbólica generalizada conjunta con d dimensiones puede ser formulada mezclando $X | W$ con W .

$$f_x(x) = \int_0^\infty f_{X|W}(x|\omega) f_W(\omega) d\omega \tag{2.2.1}$$

Donde $f_{X|W}(x|\omega)$ hace referencia a la expresión (2.1.2.4) encontrada previamente y $f_W(\omega)$ es la distribución $GIG(\lambda, \chi, \psi)$ expuesta en (2.1.2.3), lo que es igual a hallar la expresión Ω_2 utilizada en (2.1.2.9).

$$f_x(x) = \int_0^\infty \frac{e^{(x-\mu)' \Sigma^{-1} \gamma}}{(2\pi)^{d/2} |\Sigma|^{1/2} \omega^{d/2}} \exp \left\{ -\frac{Q(x)}{2\omega} - \frac{\gamma' \Sigma^{-1} \gamma}{\frac{2}{\omega}} \right\} \left(\frac{\psi}{\chi} \right)^{\frac{\lambda}{2}} \frac{\omega^{\lambda-1}}{2K_\lambda(\sqrt{\psi\chi})} \exp \left\{ -\frac{1}{2} \left(\frac{\chi}{\omega} + \psi\omega \right) \right\} d\omega \tag{2.2.2}$$

Simplificando, se tiene que

$$f_x(x) = \frac{\left(\frac{\psi}{\chi} \right)^{\lambda/2} e^{(x-\mu)' \Sigma^{-1} \gamma}}{(2\pi)^{d/2} |\Sigma|^{1/2} K_\lambda(\sqrt{\psi\chi})} \frac{1}{2} \int_0^\infty \omega^{\lambda-d/2-1} \exp \left\{ -\frac{1}{2} \left(\frac{Q(x)+\chi}{\omega} + \omega(\gamma' \Sigma^{-1} \gamma + \psi) \right) \right\} d\omega \tag{2.2.3}$$

⁷ Ver Abramowitz (1970).

Y al hacer el cambio de variable de $y = \omega \frac{\sqrt{\psi + \gamma' \Sigma^{-1} \gamma}}{\sqrt{(\chi + (x - \mu)' \Sigma^{-1} (x - \mu))}}$ y agrupar algunos términos, entonces:

$$f_X(x) = \frac{\left(\frac{\psi}{\chi}\right)^{\lambda/2} e^{(x-\mu)' \Sigma^{-1} \gamma}}{(2\pi)^{d/2} |\Sigma|^{1/2} K_\lambda(\sqrt{\psi\chi})} * \frac{1}{2} \int_0^\infty \left(\sqrt{\frac{Q(x) + \chi}{\gamma' \Sigma^{-1} \gamma + \psi}}\right)^{\lambda-d/2-1} y^{\lambda-d/2-1} \exp\left\{-\frac{1}{2}\left(\sqrt{(Q(x) + \chi)(\gamma' \Sigma^{-1} \gamma + \psi)}\left(y + \frac{1}{y}\right)\right)\right\} dy \tag{2.2.4}$$

En la fórmula anterior, la expresión

$$\frac{1}{2} \int_0^\infty \left(\sqrt{\frac{Q(x) + \chi}{\gamma' \Sigma^{-1} \gamma + \psi}}\right)^{\lambda-d/2-1} y^{\lambda-d/2-1} \exp\left\{-\frac{1}{2}\left(\sqrt{(Q(x) + \chi)(\gamma' \Sigma^{-1} \gamma + \psi)}\left(y + \frac{1}{y}\right)\right)\right\} dy$$

corresponde exactamente con (2.1.3.2) luego del cambio de variable y expandir Q(x), por lo que se puede hacer uso de (2.1.3.2) para simplificar esta expresión de la misma forma que se usó para (2.1.2.9).

De esta forma se llega a la función de distribución hiperbólica generalizada.

$$f_X(x) = \Delta * e^{(x-\mu)' \Sigma^{-1} \gamma} \frac{K_{\lambda-d/2}\left(\sqrt{(\chi + (x - \mu)' \Sigma^{-1} (x - \mu))(\psi + \gamma' \Sigma^{-1} \gamma)}\right)}{\sqrt{(\chi + (x - \mu)' \Sigma^{-1} (x - \mu))(\psi + \gamma' \Sigma^{-1} \gamma)}^{d/2-\lambda}} \tag{2.2.5}$$

Donde Δ hace las veces de una constante de normalización.

$$\Delta = \frac{(\sqrt{\chi\psi})^{-\lambda} \psi^\lambda (\psi + \gamma' \Sigma^{-1} \gamma)^{d/2-\lambda}}{(2\pi)^{d/2} |\Sigma|^{1/2} K_\lambda(\sqrt{\psi\chi})} \tag{2.2.6}$$

Aquí |Σ| denota el determinante de Σ y K_λ(...) hace alusión a la función modificada de Bessel de tercera clase.

2.2.1. Transformaciones lineales sobre una función de distribución hiperbólica generalizada

La familia de distribuciones hiperbólicas posee esta propiedad particular que las hace muy prácticas y útiles para resolver problemas de maximización de portafolios.

Si se supone que $X \sim GH_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ y $Y = BX + b$, donde B es una matriz de dimensiones $k \times d$, y $b \in \mathbb{R}^{k \times d}$, entonces se tiene que $Y \sim GH_{1d}(\lambda, \chi, \psi, B\mu + b, B\Sigma B', B\gamma)$.

Prueba 1. Transformaciones lineales

La función característica de una distribución normal multivariada es representada como sigue:

$$\phi_X^{(t)} = \mathbb{E}(e^{it'X}) = e^{it'\mu - \frac{t'\Sigma t}{2}} \tag{2.2.1.1}$$

De esta se puede deducir fácilmente la función característica de la distribución hiperbólica, partiendo de la definición de la distribución normal de media-varianza mixta.

$$\begin{aligned} \phi_X^{(t)} &= \mathbb{E}(\mathbb{E}(e^{it'X} | W)) = \mathbb{E}\left(e^{it'(\mu + \omega\gamma) - \frac{\omega t'\Sigma t}{2}} \right) \\ &= e^{it'\mu} \mathbb{H}\left(\frac{t'\Sigma t}{2} - it'\gamma \right) \end{aligned} \tag{2.2.1.2}$$

Donde $\mathbb{H}(\theta) = \mathbb{E}(e^{-\theta\omega})$ es la transformada de Laplace de la función de distribución GIG.

De esta forma se tiene que:

$$\begin{aligned} \phi_X^{(t)} &= \mathbb{E}(e^{it'(BX+b)}) = e^{it'b} \phi_X(B't) \\ &= e^{it'(B\mu+b)} \mathbb{H}\left(\frac{t'B\Sigma B't}{2} - it'B\gamma \right) \end{aligned} \tag{2.2.1.3}$$

Esta propiedad es muy útil en selección de portafolios de inversión cuando se supone $b = 0$, y B como un vector de dimensiones $d \times 1$, donde $\sum_{i=1}^d B_i$, lo cual puede ser visto como los pesos de cada activo dentro del portafolio.

Así pues, la función de distribución conjunta de d-activos puede ser representada dentro de un portafolio como una función univariada facilitando el proceso de optimización.

La versión univariada de (2.2.5) puede ser expresada de la siguiente forma:⁸

$$y \sim GH_1(\lambda, \chi, \psi, \omega', \mu, \omega' \Sigma \omega, \omega' \gamma) \tag{2.2.1.4}$$

$$f_{1d}(y) = \frac{(\sqrt{\chi\psi})^{-\lambda} \psi^\lambda (\psi + \gamma^2/\sigma^2)^{1/2-\lambda}}{\sqrt{2\pi\sigma} K_\lambda(\sqrt{\chi\psi})} * \frac{K_{\lambda-1/2} \left(\sqrt{\left(\frac{\chi + (x-\mu)^2}{\sigma^2} \right) \left(\psi + \frac{\gamma^2}{\sigma^2} \right)} \right)}{\sqrt{\left(\frac{\chi + (x-\mu)^2}{\sigma^2} \right) \left(\psi + \frac{\gamma^2}{\sigma^2} \right)}} \exp \left\{ \frac{\gamma(y-\mu)^2}{\sigma^2} \right\} \tag{2.2.1.5}$$

Donde $B = \omega' = (\omega_1, \dots, \omega_d)$ y $\sum_{i=1}^d \omega_i = 1$, y y representa los retornos del portafolio.

2.2.2. Problemas de calibración

La parametrización $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ tiene la ventaja de que cada uno de sus términos posee un significado más intuitivo con respecto a las demás esquemas, pero presenta una dificultad en su calibración, ya que:

$$fGH_d(y; \lambda, \chi, \psi, \mu, \Sigma, \gamma) = fGH_d(y; \lambda, \chi/\Lambda, \Lambda\psi, \mu, \Lambda\Sigma, \Lambda\gamma), \text{ con } \Lambda > 0^9$$

Este problema de identificación puede generar inestabilidad en el algoritmo de ajuste. No obstante, esto se puede resolver al incluirle algunas restricciones a los parámetros, como atar el valor del determinante de Σ a 1. De esta forma, se elimina el problema de identificación sin pérdida de generalidad, pero los demás parámetros del modelo deben ser modificados.

Otra parametrización usada comúnmente para evitar este problema, sin ser tan restrictiva como el caso anterior, consiste en hacer que $E(W) = 1$, lo que

⁸ Nótese que la calibración de esta función es más sencilla que (2.2.5).

⁹ Ver ecuación (2.2.5).

permite a su vez que el ajuste del modelo converja más rápido, sin perjudicar las características de la distribución.

Entonces, si:

$$\mathbb{E}(W) = \sqrt{\frac{\chi}{\psi}} \frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_{\lambda}(\sqrt{\chi\psi})} = 1 \quad (2.2.2.1)$$

Y se define $\hat{\alpha} = \sqrt{\chi\psi}$, entonces los siguientes parámetros se pueden expresar como:

$$\psi = \hat{\alpha} \frac{K_{\lambda+1}(\hat{\alpha})}{K_{\lambda}(\hat{\alpha})} \quad (2.2.2.2)$$

$$\chi = \frac{\hat{\alpha}^2}{\psi} = \hat{\alpha} \frac{K_{\lambda}(\hat{\alpha})}{K_{\lambda+1}(\hat{\alpha})} \quad (2.2.2.3)$$

Por lo que ahora se puede pasar a una parametrización de la forma $fGH_d(\lambda, \hat{\alpha}, \mu, \Sigma, \gamma)$, que no tiene problemas de identificación, y de la cual se pueden recuperar los parámetros iniciales luego de la calibración de la distribución utilizando (2.2.2.3).

2.2.3. Casos particulares de la distribución hiperbólica generalizada

Al efectuar algunas restricciones a los parámetros de la distribución hiperbólica generalizada se obtienen algunas funciones específicas, las cuales poseen características deseadas en finanzas.

Distribución hiperbólica

Si $\lambda = (d + 1)/2$, entonces se tiene una distribución hiperbólica d-dimensional.

NIG (inversa gaussiana normal)

Si $\lambda = -1/2$, entonces se tiene una distribución inversa gaussiana normal, la cual es muy utilizada gracias a sus colas semipesadas y a que es de las más fáciles de calibrar dentro de la familia HGD, debido a que se tiene una solución cerrada para su función de Bessel de tercera clase, tal como se detalló en el numeral 2.1.3.

Distribución varianza gamma

$\lambda > 0$ y $\chi = 0$, esta distribución hace referencia a uno de los casos extremos de la distribución hiperbólica generalizada, dado que el peso de sus colas desciende muy lentamente. Esta función de densidad también es conocida como la distribución generalizada de Laplace o la función de distribución de Bessel.

Distribución t asimétrica

Con $\lambda = -\nu / 2$, $\chi = \nu$ y $\psi = 0$, donde ν hace referencia a los grados de libertad de la distribución t-Student.

2.3. Calibración de la distribución hiperbólica generalizada

En la literatura existe una gran cantidad de métodos de estimación que podrían ser utilizados para la calibración de esta distribución, sin embargo, el más recomendado por Wolfgang Breyman (2013) es el método denominado MCECM, el cual es una modificación al método EM (*expectation-maximization*).

2.3.1. Método multi-cycle, expectation, conditional estimation (MCECM)

Este método es una extensión del esquema EM (*expectation-maximization*),¹⁰ el cual es muy útil en la estimación de modelos por máxima verosimilitud ante variables latentes o mixtas, como es el caso de las distribuciones hiperbólicas generalizadas.

Para la implementación de este método se requiere un conjunto de observaciones independientes e idénticamente distribuidas:

$$x_1, \dots, x_n \sim iid$$

y un conjunto de parámetros representados como sigue:¹¹

$$\theta Z = (\lambda, \hat{\alpha}, \mu, \Sigma, \gamma) \hat{\alpha} = \sqrt{\chi \psi}$$

¹⁰ Ver Breyman, W. (2013).

¹¹ Aquí se toma la parametrización con $\hat{\alpha}$ para evitar los problemas de identificación descritos en 2.2.2.

Así pues, el problema de maximización quedará definido de la siguiente forma:

$$\max \ln (f_x(x; \theta))$$

Como este problema de optimización no es fácil de solucionar, dada la complejidad de la función y el alto número de parámetros que obligan a maximizar sobre matrices de covarianzas, el problema debe ser re-escrito de la siguiente manera:

$$\max f_{x|W}(x | \omega) f_W(\omega)$$

De esta forma, la función no es maximizada directamente, sino como el producto de dos funciones de distribución. La función objetivo (en logaritmos) resultante es expresada como sigue:

$$\ln L(\theta; x_1, \dots, x_n, \omega_1, \dots, \omega_n) = \sum_{i=1}^n \ln f_{x|W}(x_i | \omega_i; \mu, \Sigma, \gamma) + \sum_{i=1}^n \ln f_W(\omega_i; \lambda, \hat{\alpha}) \quad (2.3.1.1)$$

Así pues, se debe optimizar la función de máxima verosimilitud aumentada (2.3.1.1), donde $\ln f_W(\omega)$ representa la función de densidad $GIG(\alpha, \chi, \psi)$ siguiendo los supuestos de 2.1, y $f_{x|W}$ la función normal multivariada expuesta en (2.1.2.4).

$$\ln f_W(\omega) = \frac{\lambda}{2} \ln \left(\frac{\psi}{\chi} \right) - \ln \left(2K_\lambda(\sqrt{\chi\psi}) \right) + (\lambda - 1) \ln(\omega) - \frac{x}{2\omega} - \frac{\psi}{2} \omega \quad (2.3.1.2)$$

Y al igual que en el método EM donde la maximización es realizada en dos etapas (E y M), acá se deberá separar el proceso en dos, dado que en estas expresiones se tiene una variable latente (ω) que no se puede hallar de forma directa. De este modo, se deberá cambiar esta variable no observada por su valor esperado a medida que se van generando las iteraciones del método para aproximar la función global a su máximo. Para ello se definirán las siguientes variables auxiliares:

$$\eta_i^{[k]} := \mathbb{E}[\omega_i | x_i; \theta^{[k]}], \delta_i^{[k]} := \mathbb{E}[\omega_i^{-1} | x_i; \theta^{[k]}], \xi_i^{[k]} := \mathbb{E}[\log \omega_i | x_i; \theta^{[k]}] \quad (2.3.1.3)$$

Estos nuevos parámetros podrán ser hallados utilizando la función de momentos de la distribución GIG y los resultados encontrados en 2.1.2.

$$\mathbb{E}[X^n] = \left(\frac{\chi}{\psi}\right)^{n/2} \frac{K_{\lambda+n}(\sqrt{\chi\psi})}{K_{\lambda}(\sqrt{\chi\psi})} \tag{2.3.1.4}$$

$$\mathbb{E}[\ln X] = \left. \frac{d\mathbb{E}(x^\alpha)}{d\alpha} \right|_{\alpha=0} \tag{2.3.1.5}$$

Ya determinadas estas variables, se pueden hallar los valores promedio que serán utilizados en la función $\ln f_w(\omega)$.

$$\hat{\eta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \eta_i^{[k]} \tag{2.3.1.6}$$

$$\hat{\delta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} \tag{2.3.1.7}$$

Luego de realizarse la maximización de la primera parte de la función de máxima verosimilitud (2.3.1.2), se encuentran las expresiones óptimas de (μ, Σ, γ) , las cuales son halladas de la función normal multivariada descrita en (2.1.2.4), y cuyos estimadores de máxima verosimilitud son:

$$\gamma^{[k+1]} = \frac{1}{n} \frac{\sum_{i=1}^n \delta_i^{[k]} (\bar{x} - x_i)}{\hat{\eta}^{[k]} \hat{\delta}^{[k]} - 1} \tag{2.3.1.8}$$

$$\mu^{[k+1]} = \frac{1}{n} \frac{\sum_{i=1}^n \delta_i^{[k]} x_i - \gamma^{[k+1]}}{\hat{\delta}^{[k]}} \tag{2.3.1.9}$$

$$\Sigma^{[k+1]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} (x_i - \mu^{[k+1]})(x_i - \mu^{[k+1]})' - \eta^{[k]} \gamma^{[k+1]} \gamma^{[k+1]'} \tag{2.3.1.10}$$

Por último, se deberá maximizar la segunda parte de la función de máxima verosimilitud con respecto a $\lambda, \bar{\alpha}$ o, lo que es lo mismo, a (λ, χ, ψ) , completando el cálculo de $\theta^{[k,2]}$, para posteriormente repetir el algoritmo de forma iterativa hasta que los valores de los parámetros converjan a un valor tolerable.

2.4. Aplicación a series financieras

En esta sección se hará el ajuste de la distribución hiperbólica generalizada a series financieras y se compararán sus resultados con respecto a la distribución normal estándar. Los retornos financieros que se utilizarán para este propósito serán los correspondientes a las series de los ETF sectoriales del mercado financiero estadounidense de SPDR.¹²

Así pues, los ETF serán: XLE (*energy select sector SPDR*), XLI (*industrial select sector SPDR*), XLP (*consumer staples select sector SPDR*), XLU (*utilities select sector SPDR*), XLF (*financial select sector SPDR*), XLY (*consumer discretionary select sector SPDR*) y XLK (*technology select sector SPDR*), con una periodicidad semanal desde 2010.

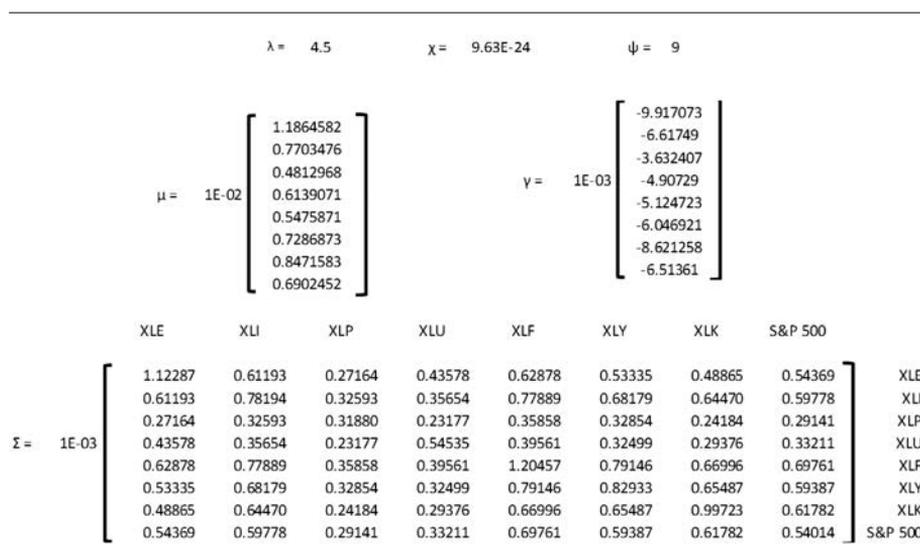


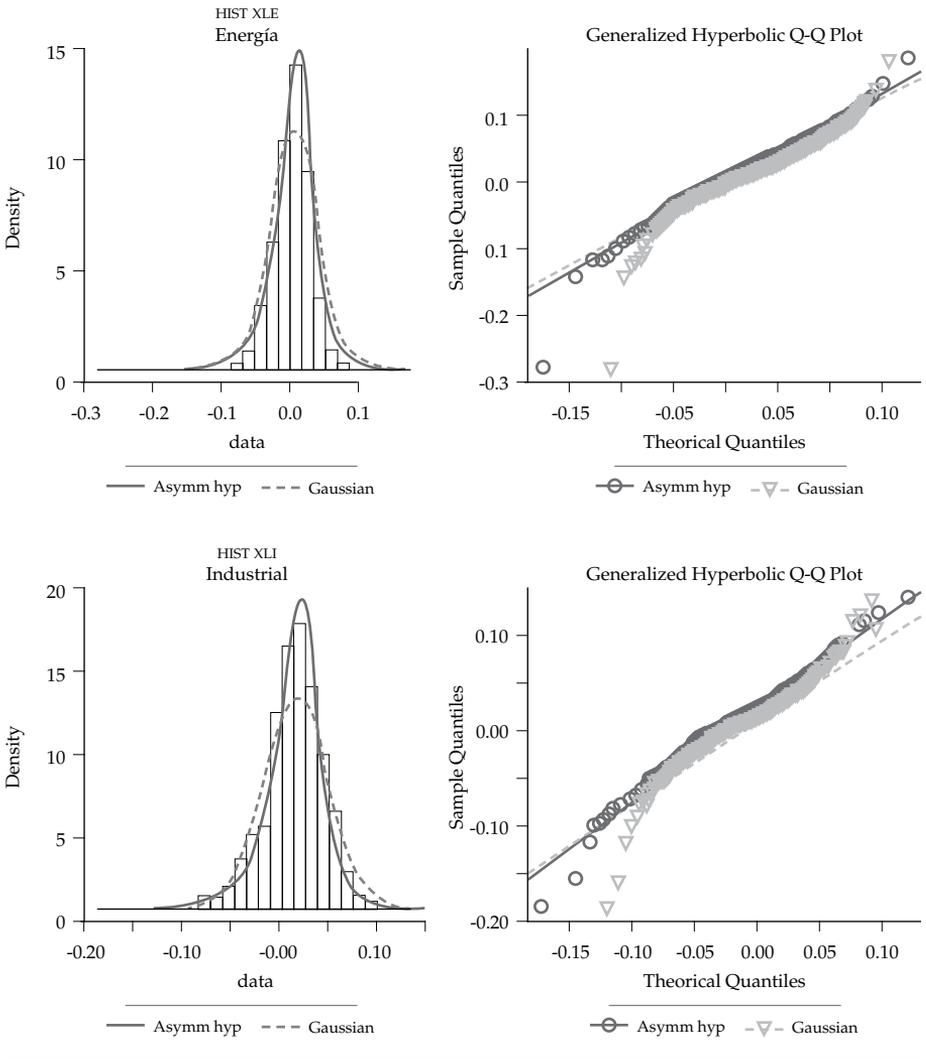
Figura 1.1. Parámetros de la distribución hiperbólica generalizada

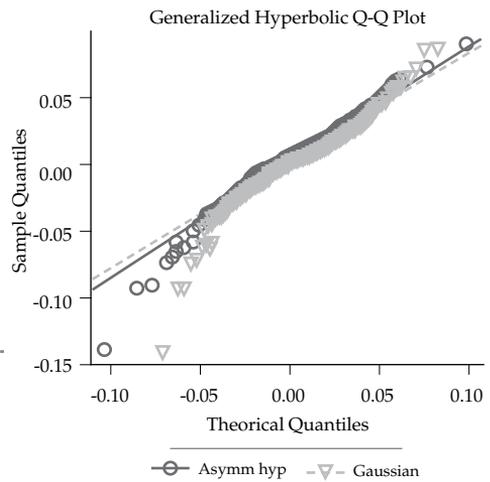
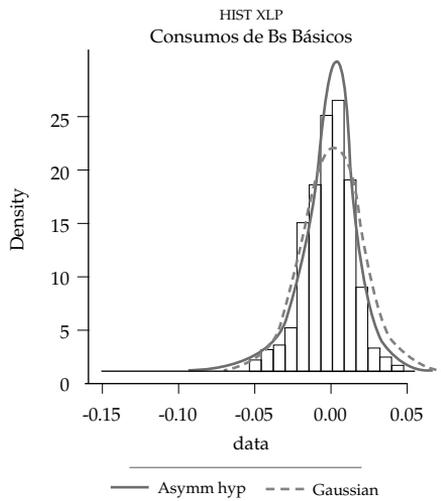
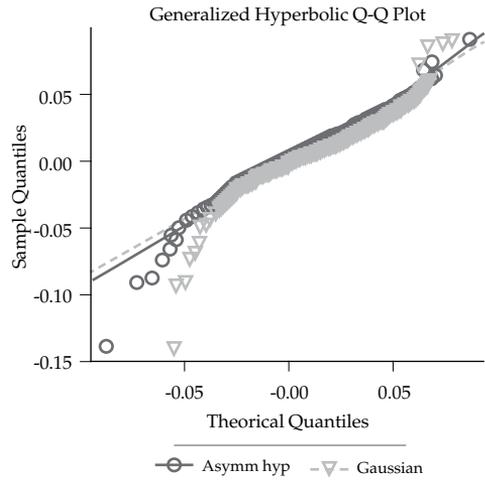
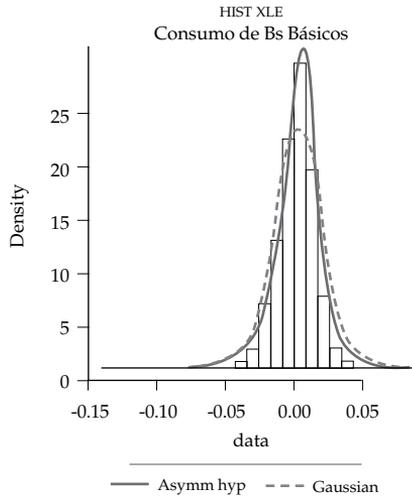
Fuente: cálculos del autor.

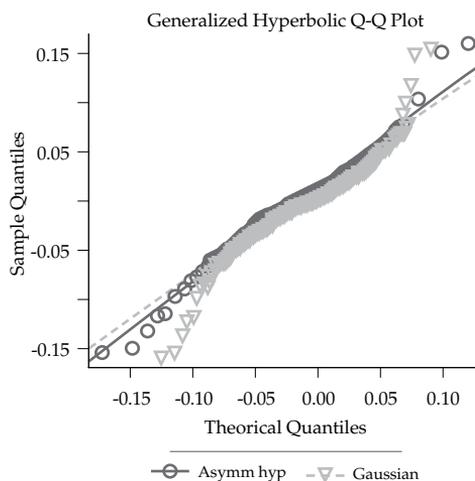
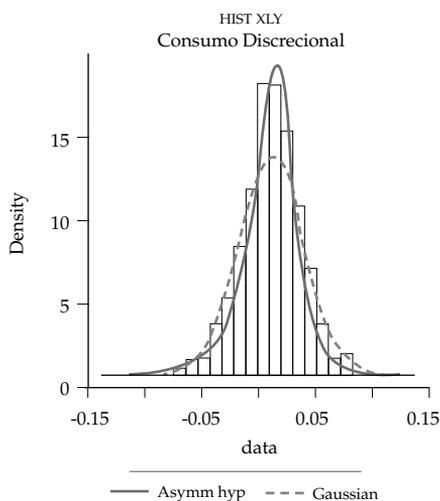
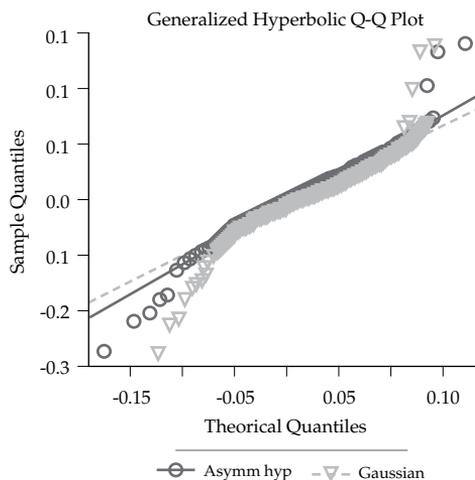
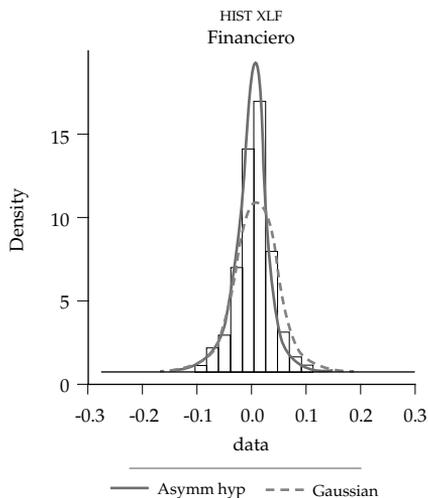
Los coeficientes encontrados en este ejercicio y los resultados gráficos del ajuste de la distribución hiperbólica generalizada a los datos financieros versus el ajuste de una distribución normal son presentados en las figuras 2.1 y 2.2.

¹² "SPDR is a registered trademark of Standard & Poor's Financial Services LLC ("S&P") and has been licensed for use by State Street Corporation. Standard & Poor's, S&P, S&P 500 and S&P Midcap 400 are registered trademarks of Standard & Poor's Financial Services LLC". (Ver www.spdrs.com).

En este punto es importante recalcar algunos resultados. El primero es que la distribución hiperbólica presenta una mayor curtosis y colas más pesadas que la distribución normal, lo que genera una mejor aproximación a los cuantiles de la distribución empírica de los retornos; del mismo modo, la distribución hiperbólica muestra una ligera asimetría positiva (hacia el lado derecho de la distribución), hecho que no puede recoger la distribución gaussiana, por lo que se logra un avance cualitativo relevante frente a los modelos que asumen normalidad.







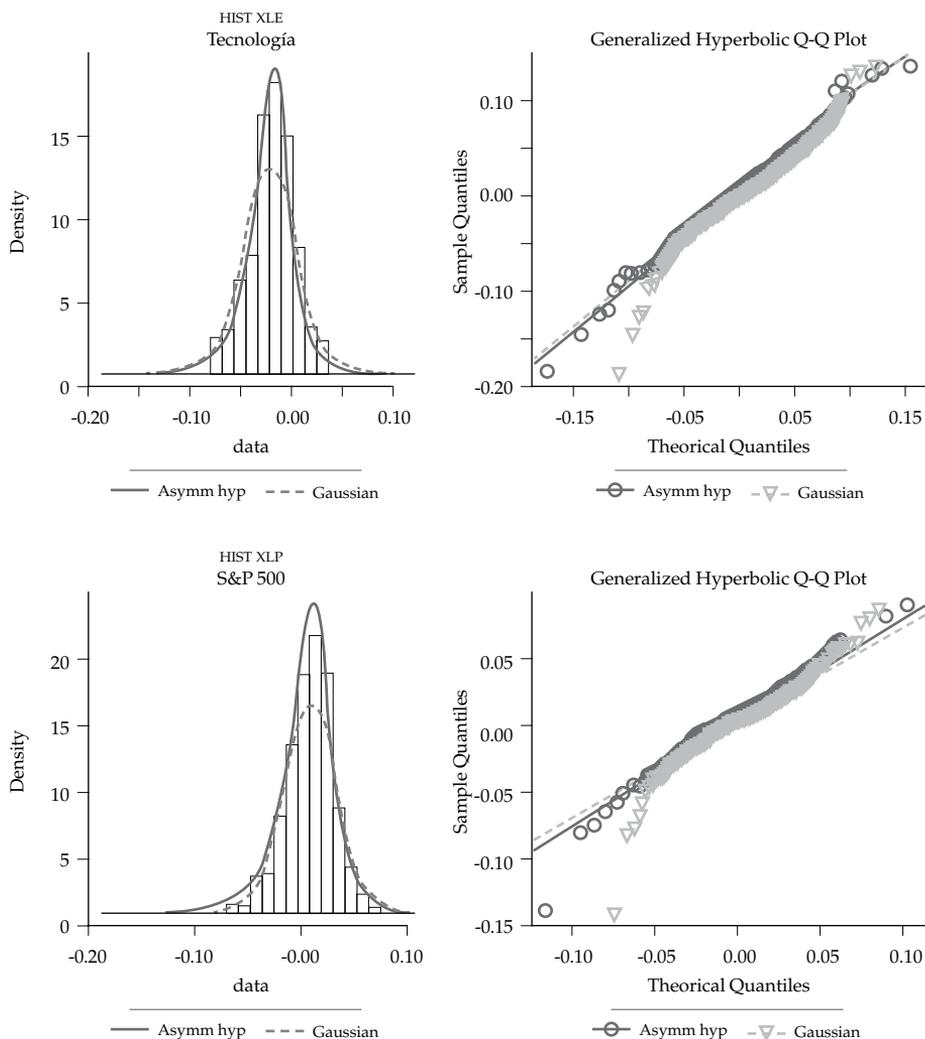


Figura 1.2. Gráficas de ajuste de las distribuciones

Fuente: cálculos del autor.

No obstante, los mayores beneficios están centrados alrededor de la media de la distribución, donde la diferenciación con la gaussiana es más significativa. En cuanto a los datos más extremos, la diferenciación entre las colas de ambas distribución es marginal, pero lo suficientemente significativa para obtener una mejor valoración del riesgo.

Capítulo 3. Selección de portafolio

La teoría de portafolios ha evolucionado considerablemente desde el trabajo de Markowitz de 1952, cuando se publicó su texto "Portfolio Selection", ya que, aunque persiste su idea de obtener el mejor nivel riesgo-ganancia de una canasta de bienes, el concepto de riesgo ha ido modificándose progresivamente, al pasar de una matriz de varianzas y covarianzas a medidas más completas que dependen de las colas de la distribución (cuantiles extremos). De igual forma, las distribuciones que se utilizan, cada vez logran capturar más hechos estilizados del mercado, como colas pesadas, alta curtosis y asimetrías, en una carrera por mejorar los estándares bajo los cuales los administradores de portafolio deberían tomar decisiones ante mercados cada vez más volátiles y eventos más extremos.

Es así que bajo los axiomas presentados en 1999 por Artzner, Delbaen y Heath [ADEH:99] sobre medidas coherentes de riesgo, el CVaR (VaR condicional) tomó vital importancia, dado que cumplía con las siguientes características:

- Es invariante ante traslaciones:

Si L es una pérdida aleatoria y l es una constante, entonces la función de pérdida mantiene que

$$\phi(L + l) = \phi(L) + l$$

- Subaditividad:

Sean L_1 y L_2 dos variables aleatorias (pérdidas), se tiene que

$$\phi(L_1 + L_2) \leq \phi(L_1) + \phi(L_2)$$

- Homogeneidad positiva:

Si λ es una constante mayor que cero, entonces

$$\phi(\lambda L) = \lambda \phi(L)$$

- Monotonicidad:

Si $\phi(L_1) \leq \phi(L_2)$, entonces debe ser cierto que $L_1 \leq L_2$.

Otras medidas de riesgo, como el VaR, que son muy utilizadas en la actualidad por su buen desempeño y facilidad, presentan fallas con respecto al axioma de subaditividad bajo distribuciones no elípticas, por lo que es el CVaR la medida más idónea en cuanto a medición de riesgo de mercado se refiere.

El CVaR está definido de la siguiente manera:

$$CVaR_{\beta}(L) = E[L | L \geq VaR_{\beta}(L)] \quad (3.0.1)$$

Como se puede ver, el CVaR también contiene más información que el VaR, ya que el valor en riesgo mide un nivel máximo de pérdidas dentro de un rango de confianza dado, mientras que el CVaR cuantifica el valor promedio de pérdidas cuando la variable aleatoria (pérdidas) sobrepasa dicho nivel de VaR. No obstante, esto conlleva un costo computacional más alto, como se verá a continuación.

3.1. Selección de portafolio usando la medida de riesgo CVaR

Si se define $R(x)$ como la función (lineal) de retornos del portafolio, donde x hace referencia al vector de pesos de este, y $\phi(x)$ a la medida de riesgo asociada a tal canasta de bienes, se tiene entonces el siguiente problema de optimización:

$$\begin{aligned} & \underset{x \in \mathbb{R}}{\text{minimize}} && \phi(x) \\ & \text{subject to} && R(x) \geq \rho \end{aligned} \quad (3.1.1)$$

Donde ρ es definido como el retorno objetivo.

Este problema es convexo, siempre y cuando $\phi(x)$ esté definido como el VaR condicional del portafolio bajo un vector x dado, y la suma de tal vector sea igual a 1 (no apalancamiento). Además, se supondrá en un principio que no se puede tener una posición corta en tales activos, lo cual es una característica básica en muchos fondos de inversión mutuos, fondos bursátiles y en la normatividad de la mayoría de los mercados. Así pues, se tiene la siguiente representación:

$$\begin{aligned} & \underset{x \in \mathbb{R}_+}{\text{minimize}} && CV_a R(-xY) \\ & \text{subject to} && x \in \mathbb{R} \{x : \Gamma'x \geq \rho, 1'x = 1\} \end{aligned} \quad (3.1.2)$$

Donde $Y \sim GH_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ y Γ representa el valor esperado de los retornos de cada activo dentro del portafolio.

Esta representación es muy útil, gracias a que al tener un problema de minimización convexa se puede hacer uso de un gran conjunto de herramientas numéricas para encontrar el valor objetivo de la función. No obstante, evaluar el CVaR de una función de distribución tan compleja como la de la hiperbólica generalizada conlleva un problema de gran envergadura. Por lo tanto, se deberá recurrir al uso de una función auxiliar para su aproximación. Para ello, se seguirá la función propuesta por Rockafellar y Uryasev en 1999 [RU:99], en su texto "Optimization of Conditional Value-at-Risk", en el cual se demuestra que optimizar (minimizar) el valor de esta medida de riesgo, sin importar su distribución, es equivalente a optimizar la siguiente expresión:

$$F_\beta(x, \alpha) = \alpha + (1 - \beta)^{-1} \int_{y \in \mathbb{R}^n} [-x'y - \alpha]^+ p(y) dy \tag{3.1.3}$$

Donde β hace alusión al valor del cuantil; x , al vector de pesos; y , al vector de retornos de los activos; $p(y)$, a la función de probabilidad; y α , hace referencia al valor del VaR.

De esta forma, la función objetivo queda representada como sigue:

$$F_\beta(x, \alpha) = \alpha + (1 - \beta)^{-1} \int_{-\infty}^{-\alpha} (-x'y - \alpha) fGH_d(y; \lambda, \chi, \psi, \mu, \Sigma, \gamma) dy \tag{3.1.4}$$

De igual forma, esta función también puede ser representada mediante simulaciones de la variable aleatoria en vez de utilizar funciones complejas de distribución:

$$F_\beta(x, \alpha) = \alpha + \frac{1}{q(1 - \beta)} \sum_{k=1}^n [-x'y_k - \alpha]^+ \tag{3.1.5}$$

Pero en este caso en específico se debe tener un buen mecanismo generador de datos aleatorios que sigan una función hiperbólica generalizada, lo cual es difícil por la complejidad que conlleva hallar la inversa de esta función, al tiempo que estos algoritmos tienden a tomar mucho tiempo.

Del mismo modo, para hacer el algoritmo de optimización más eficiente, se hará uso de la propiedad de transformaciones lineales sobre la distribución hiperbólica generalizadas, expuesta en la sección 2.2.2. Así pues, la función anterior puede ser fácilmente simplificada:

$$F_{\beta}(x, \alpha) = \alpha + (1 - \beta)^{-1} \int_{-\infty}^{-\alpha} (-z - \alpha) fGH_1(z; \lambda, \chi, \psi, x' \mu, x' \Sigma x, x' \gamma) dz \quad (3.1.6)$$

$$F_{\beta}(x, \alpha) = \alpha + \frac{1}{q(1 - \beta)} \sum_{k=1}^n [f(x, \theta) - \alpha]^+ \quad (3.1.7)$$

Y el proceso de optimización va a quedar establecido de una forma bastante práctica en términos computacionales:

$$\begin{aligned} & \underset{x, \alpha}{\text{minimize}} && F_{\beta}(x, \alpha) \\ & \text{subject to} && x \in \mathbb{R}_+ \{x : \Gamma'x \geq \rho, 1'x = 1\} \\ & && \alpha \in \mathbb{R} \end{aligned} \quad (3.1.8)$$

3.2. Implementación de selección de portafolios bajo una distribución hiperbólica generalizada

El problema de optimización expuesto en la sección anterior puede ubicarse dentro de la categoría de funciones convexas, lo cual es una gran ventaja dada la cantidad de *softwares* estadísticos que permiten encontrar los mínimos de dichas funciones. Esta característica fue demostrada por Rockafellar y Uryasev [RU:99] en su trabajo, por lo que el problema logra una gran simplicidad, unida al uso de la función alterna (3.1.6) que mejora el algoritmo y mantiene todas las propiedades deseadas de la función original.

Dadas estas características, el proceso de optimización fue implementado en R (*software*) utilizando la misma base de datos de la sección 2.4. Cada uno de los portafolios fue hallados dejando lapsos de 20 semanas entre sí, lo que serían los tiempos entre recomposiciones de la canasta, y se impuso un retorno objetivo igual o superior¹³ al marcado por el S & P 500 en ese período.

Así pues, luego de haberse implementado los algoritmos descritos anteriormente, se llegó a que la composición de los portafolios óptimos a través del

¹³ Cuando se restringe el modelo a un retorno igual al del índice, en mercado bajistas puede llevar a que la solución presente un nivel de CVaR mayor que si se estableciera la restricción con un mayor o igual, dado que el algoritmo puede encontrar soluciones menos negativas que el índice. De esta forma, la mejor opción es maximizar la función sujeto a un valor igual o superior, para que el algoritmo pueda acceder a un conjunto más amplio de portafolios, y así pueda alcanzar mejores canastas.

tiempo privilegiase a algunos activos como el XLP, mientras que otros apenas aparecieron por breves períodos, como el XLF.

En este punto vale la pena resaltar que el nivel de riesgo del portafolio es inferior al del índice y en algunos períodos es inferior hasta al total de los activos que lo componen, gracias a la diversificación, mientras logra replicar por lo menos el comportamiento del S &P 500, lo que muestra su pertinencia y utilidad.

Por otro lado, en cuanto a la utilización del algoritmo por simulaciones de la variable aleatoria, este resultó ser más lento en su implementación que la función (3.1.6), ya que su convergencia toma un alto número de simulaciones, por lo que no resultó ser óptima su implementación en este contexto. Así pues, de ahora en adelante los portafolios óptimos serán hallados utilizando la expresión (3.1.6) en vez de la (3.1.7).

3.3. Portafolio robusto bajo la distribución hiperbólica generalizada

Como se vio en la sección anterior, el problema numérico para encontrar las ponderaciones óptimas de un portafolio de mínimo CVaR atado a una rentabilidad específica mejora significativamente haciendo uso de la función de Rockafellar y Uryasev. No obstante, el proceso para hallar los parámetros de la distribución hiperbólica, unido a la complejidad de la función $f_{GH_1}(z; \lambda, \chi, \psi, x' \mu, x' \Sigma x, x' \gamma)$, usada dentro del proceso de maximización, pueden generar problemas de robustez, o, lo que es lo mismo, a que pequeños cambios en los parámetros puedan llevar a grandes variaciones en el resultado final del problema. Por esta razón, se introducirá a continuación el establecimiento de un algoritmo que permita el desarrollo de un modelo robusto de selección de portafolio.

3.3.1. El método del CVaR del peor escenario ('worst case CVaR -WCVaR)

Al desarrollar un modelo robusto de optimización de portafolio se debe crear un rango en el cual se espera esté el valor del parámetro desconocido, y no un valor puntual que puede no ser el óptimo dado los problemas numéricos que pueden surgir en el algoritmo utilizado por el investigador. En este sentido, si el portafolio es óptimo bajo el peor de los casos, entonces debe ser óptimo también para los demás escenarios.

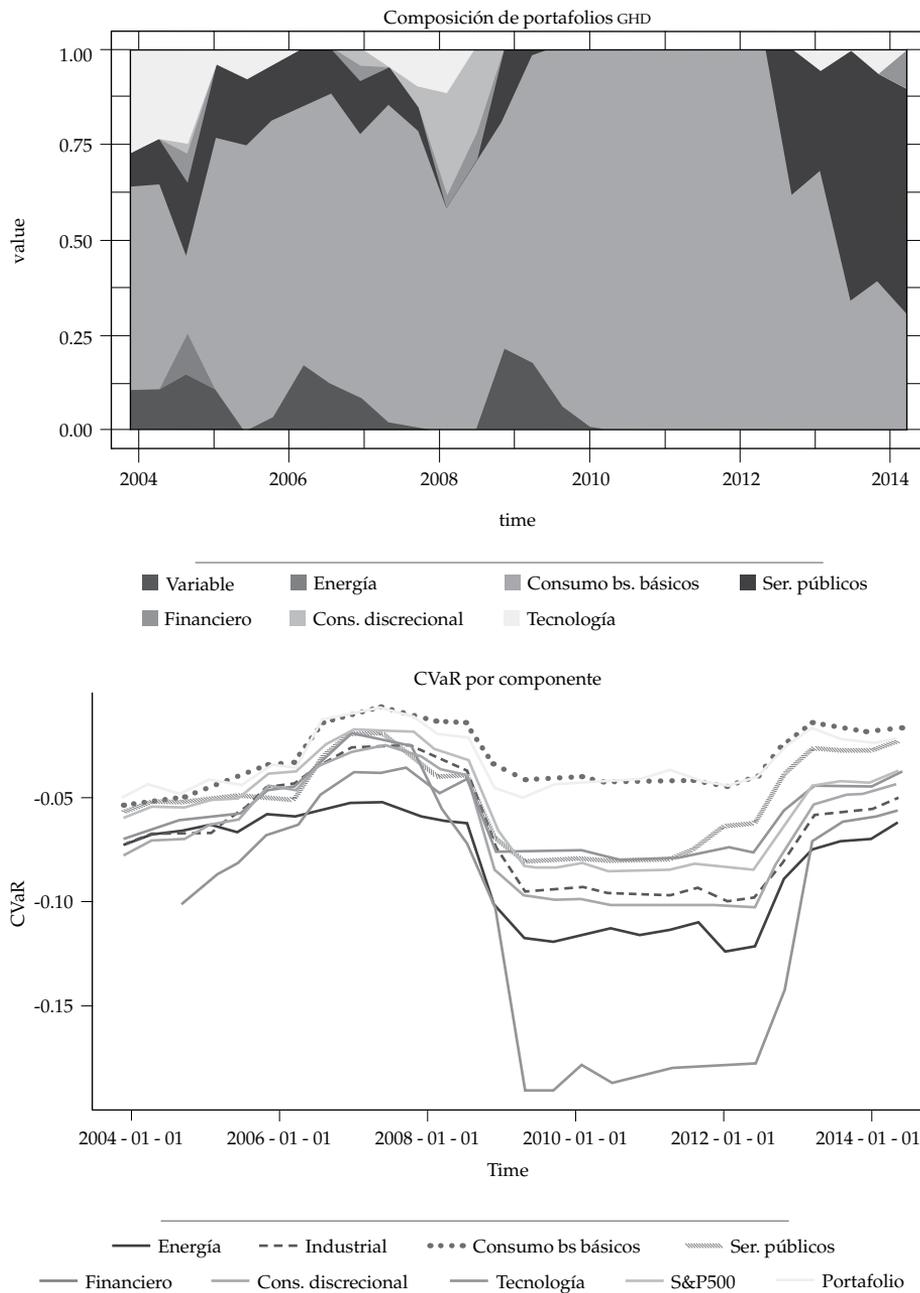


Figura 2.1. Portafolios óptimos usando la distribución hiperbólica generalizada

Fuente: cálculos del autor.

De igual modo, y siguiendo a Hellmitch y Kassberg (2009), el nuevo proceso de optimización seguirá los siguientes supuestos:

Permita que el portafolio siga una distribución hiperbólica multivariada ϑ , y que Y^p sea el vector de retornos del portafolio y x los ponderadores de los activos dentro de la canasta. Por lo tanto, el WCVaR del portafolio bajo un nivel de significancia β quedará definido así:

$$WCVaR_{\beta}^{\vartheta}(x) \triangleq \sup CVaR(-x'Y^p) \tag{3.3.1.1}$$

Y al igual que el CVaR, el WCVaR posee todos los axiomas de las medidas de riesgo coherentes expuestas al inicio de esta sección, al tiempo que mantiene la característica de convexidad, lo que ayuda a mantener la simplicidad del problema, por lo que solo se requiere cambiar la medida de riesgo dentro del problema de optimización.

Para encontrar tal medida se deberán introducir algunas propiedades adicionales. La primera consiste en hallar intervalos de confianza para cada parámetro de forma independiente con respecto a los demás, hecho que permite simplificar el algoritmo sin violar los supuestos del modelo, según los trabajos de Halldórsson y Tütüncü (2003) [HT:03], Tütüncü y Koenig (2004) [TK:04] y Kim y Boyd (2007) [KB:07].

Así pues,

$$\vartheta \triangleq \{GH_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma) : (\mu, \gamma, \Sigma) \in \mathfrak{S}\} \tag{3.3.1.2}$$

Donde \mathfrak{S} hace referencia al conjunto de intervalos de confianza de cada parámetro.

$$I_{\mu} \triangleq \{\mu \in \mathbb{R}^d : \mu_l \leq \mu \leq \mu_h\} \tag{3.3.1.3}$$

$$I_{\gamma} \triangleq \{\gamma \in \mathbb{R}^d : \gamma_l \leq \gamma \leq \gamma_h\} \tag{3.3.1.4}$$

$$I_{\Sigma} \triangleq \{\Sigma \in \mathbb{R}^d : \Sigma_l \leq \Sigma \leq \Sigma_h\} \tag{3.3.1.5}$$

Se debe verificar que cada Σ sea definido positivo, y de ser así y dado que cada uno de estos conjuntos son cerrados, entonces se tiene que:

$$WCVaR_{\beta}^{\vartheta}(x) \triangleq \sup CVaR(-x' Y^p) \triangleq \max CVaR(-x' Y^p) \tag{3.3.1.6}$$

De manera que, combinado con el problema de optimización original $\min_{x,\alpha}(x, \alpha)$, el nuevo algoritmo puede establecerse como un proceso donde se mini-

mice con respecto a x y α , mientras se maximiza en relación con los valores de μ , γ y Σ , que son los que permiten la transformación entre la medida CVaR y el WCVaR. No obstante, estos valores de μ , γ y Σ deben estar acotados por un conjunto posible de valores que podrían tomar estos parámetros.

De esta forma, el problema puede ser reescrito de la siguiente forma:

$$\min_x \max_{\mu, \gamma, \Sigma} \min_{\alpha} F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu, \Sigma, \gamma) \quad (3.3.1.7)$$

El valor $\max_{\mu, \gamma, \Sigma}(F_{\beta})$, por su parte, puede ser alcanzado de una forma muy simple, dada la característica de convexidad. Ya que $F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu, \Sigma, \gamma)$ es monótona creciente bajo Σ , y monótona decreciente en μ y γ , entonces se tiene la siguiente igualdad,

$$\max_{\mu, \gamma, \Sigma} F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu, \Sigma, \gamma) = F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu_{low}, \Sigma_{high}, \gamma_{low}) \quad (3.3.1.8)$$

Prueba 2. WCVaR como medida de riesgo en la optimización del portafolio

Dado que F_{β} es una función convexa bajo x, α ; y monótona decreciente en μ, γ y monótona creciente en Σ , se tiene que

$$\begin{aligned} \max_{\mu, \gamma, \Sigma} \min_{\alpha} F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu, \Sigma, \gamma) &\leq \\ \min_{\alpha} \max_{\mu, \gamma, \Sigma} F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu, \Sigma, \gamma) &= \min_{\alpha} F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu_{low}, \Sigma_{high}, \gamma_{low}) \end{aligned} \quad (3.3.1.9)$$

Además,

$$\max_{\mu, \gamma, \Sigma} \min_{\alpha} F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu, \Sigma, \gamma) \geq \min_{\alpha} F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu_{low}, \Sigma_{high}, \gamma_{low}) \quad (3.3.1.10)$$

Por lo que la única solución a esta desigualdad es que

$$\max_{\mu, \gamma, \Sigma} \min_{\alpha} F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu, \Sigma, \gamma) = \min_{\alpha} F_{\beta}(x, \alpha; \lambda, \chi, \psi, \mu_{low}, \Sigma_{high}, \gamma_{low}) \quad (3.3.1.11)$$

Corroborada la paridad anterior, ahora se puede incluir en el proceso de optimización.

$$\begin{aligned} \min_x WCV_a R_\beta(x) &= \min_x \min_\alpha F_\beta(x, \alpha; \lambda, \chi, \psi, \mu_{low}, \Sigma_{high}, \gamma_{low}) \\ &= \min_{x, \alpha} F_\beta(x, \alpha; \lambda, \chi, \psi, \mu_{low}, \Sigma_{high}, \gamma_{low}) \end{aligned} \quad (3.3.1.12)$$

La expresión (3.3.1.12) sigue conservando la practicidad del problema original, pero ahora para un portafolio robusto, siempre que se mantenga que la función F_β es convexa bajo x, α . No obstante, abre una nueva problemática con respecto a cómo hallar los valores de $\mu_{low}, \gamma_{low}, \Sigma_{high}$, a lo que Hellmich y Kassberg (2009) [HK:09] se remontan a estudios de Kim y Boyd (2007) [KB:07] para su solución, en los que se sugiere la creación de intervalos de confianza con una desviación estándar del 10% para cada parámetro. No obstante, en este trabajo se irá un poco más allá y se correrá un proceso histórico desde enero de 2000 hasta marzo de 2014 en muestras de 200 semanas para determinar el valor de estos parámetros y así encontrar estos valores, tal como se podrá apreciar en la siguiente sección.

3.4. Implementación de la metodología de portafolios robustos bajo una distribución hiperbólica generalizada

Los pesos óptimos bajo la metodología de portafolio robusto desarrollada en 3.3.1 dan como resultado una abultada ponderación en XLY y XLE. Por su parte, XLK y XLU solo están presentes por un corto período, y XLP, a diferencia de los demás portafolios, presenta una ponderación marginal a través del tiempo.

En la figura (2.2) se ve un detrimento en el nivel de riesgo del portafolio con respecto al hallado inicialmente. Esto se debe a que bajo las restricciones impuestas para hallar este portafolio, el algoritmo no logra recoger de forma eficiente la actualidad del mercado y esto lleva a que no sea totalmente eficiente a la hora de minimizar el riesgo bajo el panorama actual, ya que es óptimo en un escenario hipotético donde el mercado se encuentra en un nivel de estrés máximo, lo cual no es la constante.

No obstante, este portafolio fue construido para disminuir la sensibilidad de los pesos óptimos ante variaciones de los parámetros, por lo que los cambios en los ponderadores corresponden a modificaciones únicamente de λ, χ y ψ . Por lo tanto, la composición del portafolio durante la crisis permaneció muy estable, y solo en los períodos de transición es que se aprecian cambios significativos, como resultado de su ajuste a un nuevo régimen de mercado, lo cual es coherente con el proceso de maximización que se llevó a cabo. De esta forma, se cumple con el cometido inicial de construir un portafolio robusto, pero se pierde en términos de riesgo con respecto al portafolio inicial.

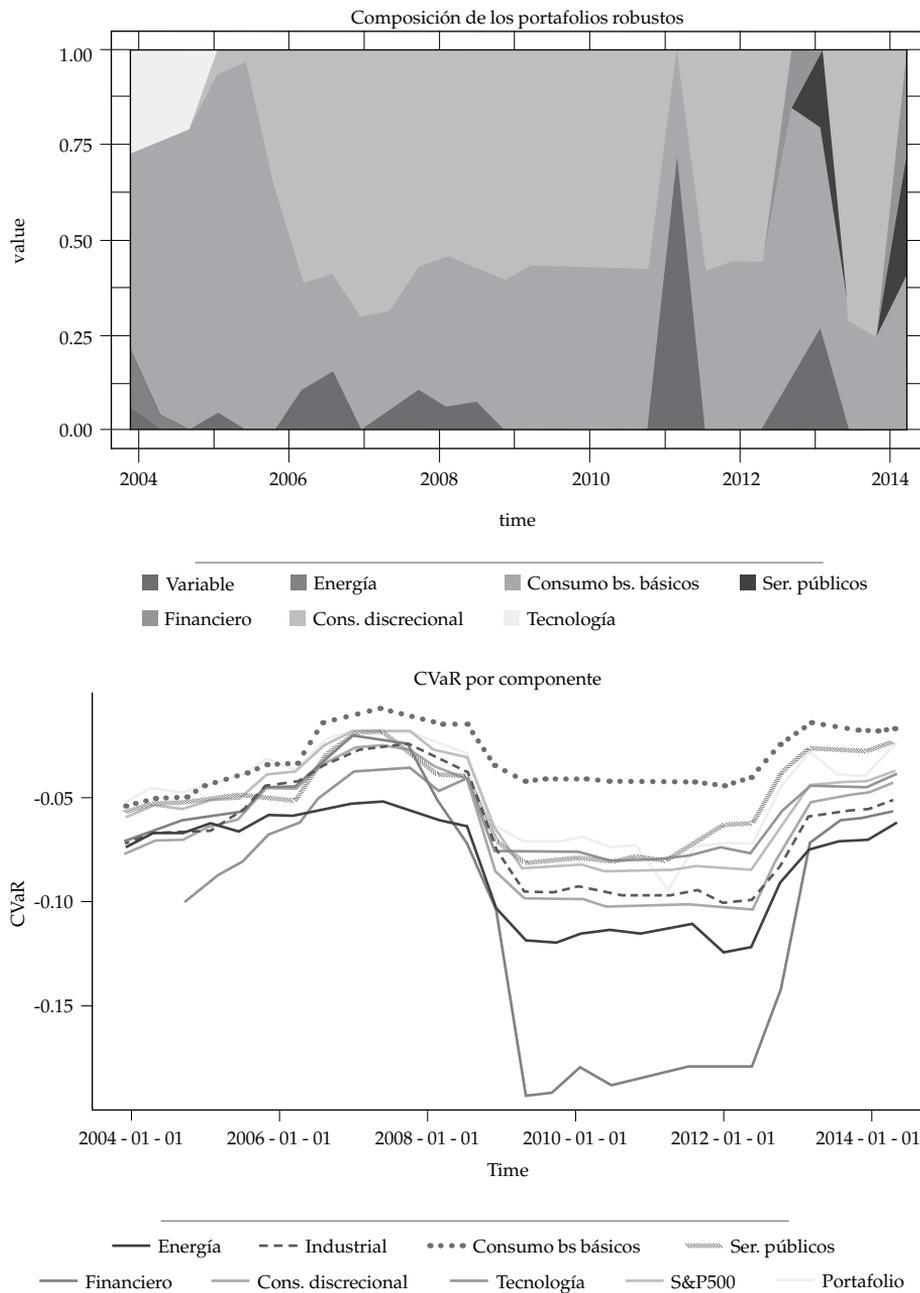


Figura 2.2. Portafolios robustos usando la distribución hiperbólica generalizada

Fuente: cálculos del autor.

3.5. Comparación de los diferentes portafolios óptimos

Luego de haberse implementado los algoritmos para hallar portafolios óptimos usando distribuciones de probabilidad hiperbólicas generalizadas, se generaron los mismos portafolios bajo la teoría de Markowitz para tener un comparativo entre los tres, y poder obtener así una mayor claridad sobre los avances que ha tenido la teoría de optimización de portafolios con la inclusión de distribuciones más avanzadas y medidas de riesgo más robustas.

El primer resultado visible es que el portafolio de Markowitz, al igual que el portafolio inicial, le da una mayor importancia a XLP (consumo) y a XLU (*utilities*), activos poco volátiles, con los que el portafolio logra mantener niveles de volatilidad bajos a través del tiempo. También es importante resaltar cómo el portafolio con la distribución hiperbólica y el de media-varianza consiguen con XLP el retorno objetivo y el mínimo riesgos simultáneamente, alcanzando en algunas ocasiones hasta el 100% de participación.

Así pues, el único portafolio que presenta una composición significativamente diferente es el robusto, dado que su proceso de optimización parte de un principio muy diferente al de los otros dos, al hallarse bajo una nueva función que permite que el portafolio sea más estable en su composición.

Del mismo modo, la similitud entre el portafolio de Markowitz y el hallado bajo la distribución hiperbólica generalizada hace que, al estimar los CVaR empíricos de los datos en el interior de la muestra, estos no sean muy diferentes entre sí, pero sí más eficientes con respecto al portafolio robusto, el cual presenta una gran brecha en su nivel de riesgo empírico.

Lo anterior es un resultado muy significativo, ya que muestra que, al *optimizar* un portafolio bajo el modelo de Markowitz de 1952 y el modelo sobre la distribución hiperbólica generalizada, el resultado no es muy diferente en términos de riesgo, como se esperaba, o, por lo menos, en el interior de la muestra.

Este fenómeno se explica porque el proceso de optimización sigue estando centrado en la matriz de varianzas y covarianzas, a pesar de que la distribución hiperbólica generalizada permite un mejor ajuste de los datos que la normal, tal como se puede apreciar en las figuras (2.2), y de que con ello se llega a una mejor valoración del riesgo. Por tal motivo, los resultados no difieren en gran medida de los alcanzados bajo el portafolio de media-varianza, pues este estimador (Σ) no es muy diferente entre estas dos funciones. Esto muestra que, pese a incluir otros momentos en el proceso de optimización, los efectos de estos sobre los datos más extremos no son tan significativos como en los cuantiles más cercanos a la media, que es donde se logra un mayor avance con el ajuste de la distribución hiperbólica.

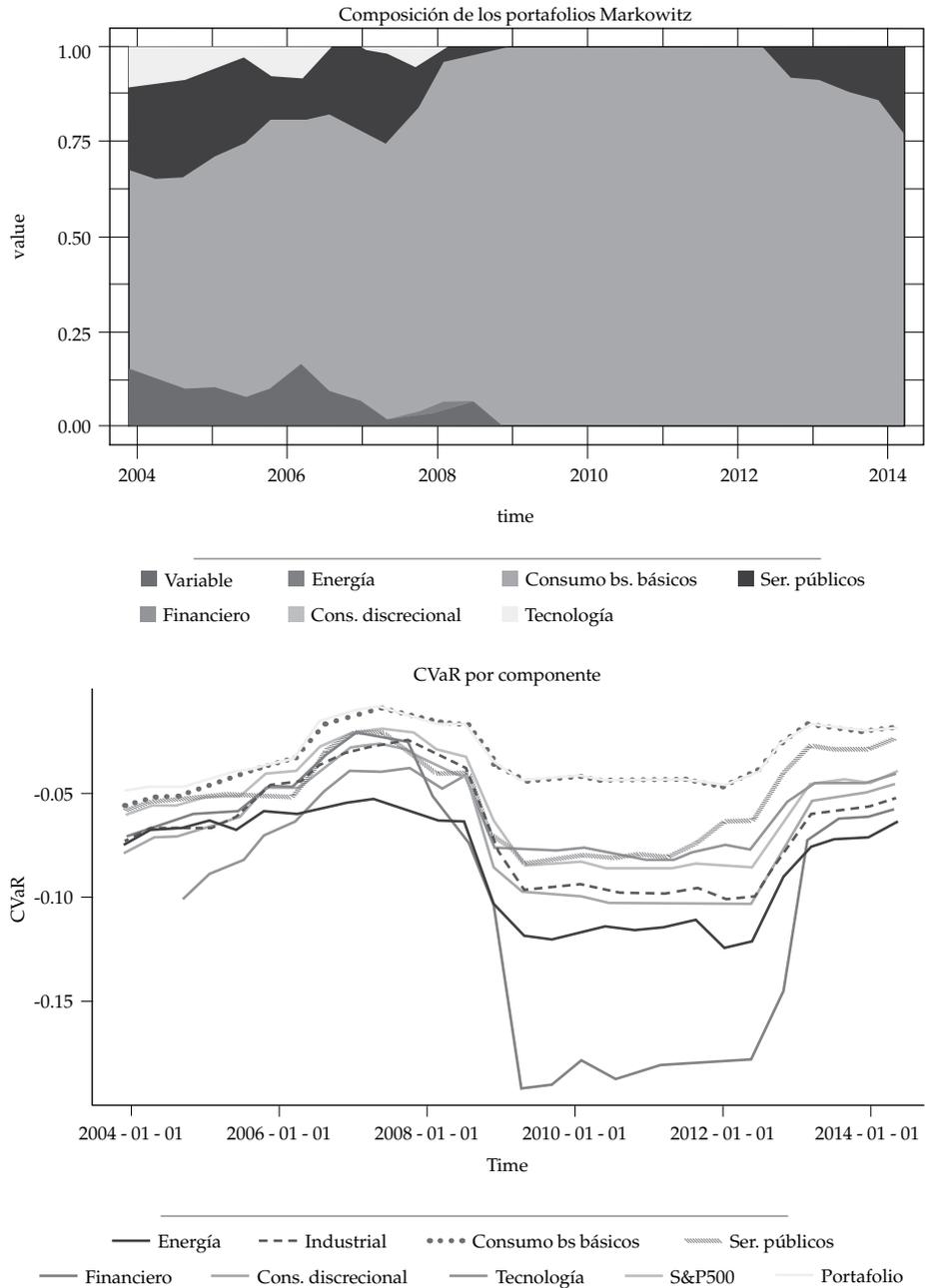


Figura 2.3. Portafolios óptimo bajo m. Markowitz

Fuente: cálculos del autor.

De esta forma, las dos medidas de riesgo tienden a ser muy parecidas, a pesar de que se esté minimizando el riesgo de eventos extremos sobre dos funciones de distribución diferentes y sobre dos medidas distintas.

El anterior es un resultado bastante importante en este trabajo, y para determinar que este fuera un resultado consistente, se cambiaron los retornos promedio exigidos para cada portafolio con el ánimo de eliminar la posibilidad de que los retornos exigidos fueran demasiado altos, y esto llevara a que los dos portafolios convergieran en términos de ponderaciones para alcanzar tal nivel de rentabilidad, pero los resultados no fueron muy diferentes a los que aquí se presentan.

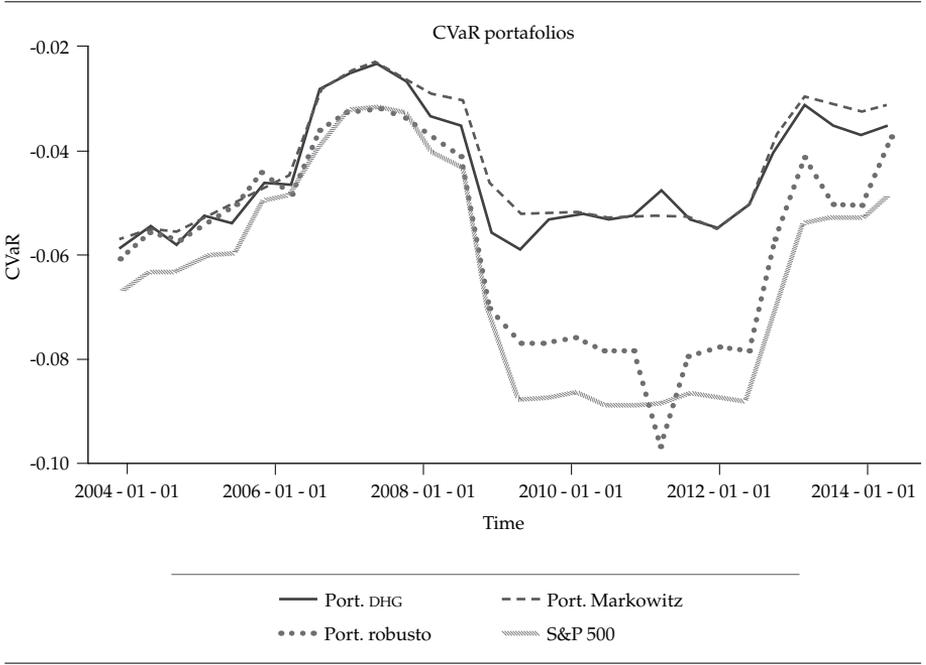


Figura 2.4. Comparación del CVaR histórico de las dif. metodologías

Fuente: cálculos del autor.

Entonces, dentro de la muestra los mejores portafolios son el de media-varianza y el que surge de la distribución hiperbólica, gracias a que reaccionan muy rápido ante cambios en el mercado; posteriormente, está el portafolio robusto, el cual tiene más restricciones y, por ende, tarda más en acomodarse frente a eventos puntuales, pero ofrece una mayor protección fuera de la muestra debido a que tiende a evitar los activos históricamente con mayor CVaR (ya que está restringido siempre al peor escenario). Otro resultado relevante

es que el portafolio GHD y el de Markowitz logran replicar el desempeño del S &P 500 a un nivel de riesgo significativamente inferior, mostrando la pertinencia de estos modelos, y el portafolio robusto consigue tal objetivo en algunos períodos.

Capítulo 4. Simulación de escenarios y portafolios

Los resultados anteriores muestran que no hay un salto cualitativo importante entre la metodología que utiliza la distribución hiperbólica generalizada y el modelo de media-varianza de Markowitz, por lo menos en el interior de la muestra. No obstante, el mejor ajuste de la distribución hiperbólica a los datos del mercado lleva a avances significativos en la valoración del riesgo del portafolio, pues, a pesar de que ambas canastas exponen niveles de riesgo semejantes, solo bajo la distribución de colas pesadas se logra una mejor cuantificación de las posibles pérdidas de los activos.

Del mismo modo, otro de los avances obtenidos al utilizar distribuciones hiperbólicas es que permiten la inclusión de momentos de orden superior, y, por lo tanto, se puede alcanzar un mejor modelaje de los activos ante diversos escenarios. Esto no es posible con la distribución normal, ya que esta solo puede ser estresada a partir del parámetro de volatilidad, sin poder aumentar el peso de las colas, modificar la asimetría de los retornos o su curtosis. De esta forma, una distribución como la hiperbólica generalizada es mucho más manejable ante investigaciones puntuales, gracias a sus múltiples parámetros y a su ajuste a las diferentes series financieras.

Por consiguiente, un investigador podría analizar cómo casos específicos del mercado o de un activo podrían afectar el desempeño del portafolio, a través del pronóstico de cuánto sería el valor de cada uno de los parámetros que componen su distribución. Estos valores pueden ser fácilmente extrapolados de experiencias o datos históricos, en una especie de prueba de estrés. De la misma forma, este algoritmo permite al investigador hallar el mejor portafolio posible ante un escenario hipotético, lo que le daría una hoja de ruta desde el portafolio actual.

Así pues, aunque los avances encontrados con la distribución hiperbólica en la optimización de portafolios no sean significativos, estas pueden llegar a ser muy útiles para los administradores de portafolios e investigadores en cuanto a simulación de escenarios y hallazgo de portafolios óptimos bajo cambios de régimen en el mercado.

4.1. Parámetros históricos

Los parámetros de los diferentes portafolios óptimos encontrados a través de la historia son de gran interés para el investigador a la hora de analizar los cambios de régimen en el mercado y las variaciones en las composiciones de cada uno de los portafolios, por lo que su exposición a través de series de tiempo es de gran valor a la hora de evaluar escenarios futuros del mercado. Por lo tanto, se esbozará el comportamiento de cada uno de estos parámetros dentro de las series de retornos de los portafolios óptimos encontrados en 3.2.

El valor del parámetro μ del portafolio, bajo la distribución hiperbólica generalizada, está asociado al promedio de los retornos de la canasta de bienes. Este ha presentado un comportamiento irregular a través de su historia alrededor de 0,006. Cabe resaltar que nunca ha tocado terreno negativo, pese a presentarse en 2008 y 2010 una de las peores crisis financieras de la historia, y que su punto más bajo se encuentra en 2013.

El parámetro ψ , por su parte, ha presentado algunos picos muy pronunciados en los últimos diez años, donde uno de ellos corresponde al inicio de la crisis de 2008, después del cual cayó rápidamente a niveles cercanos a cero. Esto se traduce en que previo a la recesión las colas de la distribución se estaban haciendo cada vez más livianas, y cuando empezó la inestabilidad en septiembre de ese año, se vio revertida esta tendencia. El período siguiente, comprendido entre 2008 y 2012, se caracterizó por presentar distribuciones de colas pesadas, lo que es coherente con lo vivido en el mercado de valores a nivel mundial.

Por su parte, el valor de χ ha sido bastante volátil a través de la historia, y al contario que los parámetros anteriores, este alcanzó su nivel máximo en un período cercano a 2008 (a menor valor de este parámetro, mayor es la concentración de datos alrededor de la media y menor es su dispersión). La función de χ viene dada por el tamaño de las colas y la curtosis de la función.

λ , al igual que μ , ψ y χ , alcanzó uno de sus puntos más altos en el período previo a la crisis de 2008 y actualmente se encuentra en niveles superiores al promedio histórico. En suma, esto refleja que estos parámetros, previo a la crisis, le imprimieron a la distribución características de colas livianas a la izquierda de la distribución (pérdidas), una alta curtosis y un desplazamiento hacia la derecha, mostrando el buen momento que vivían las bolsas mundiales entonces. Pero, posteriormente, se puede apreciar que los retornos tendieron a una mayor dispersión, a reducir el promedio de los valores de los retornos (menos positivos) y a elevar el tamaño de las colas de la distribución.

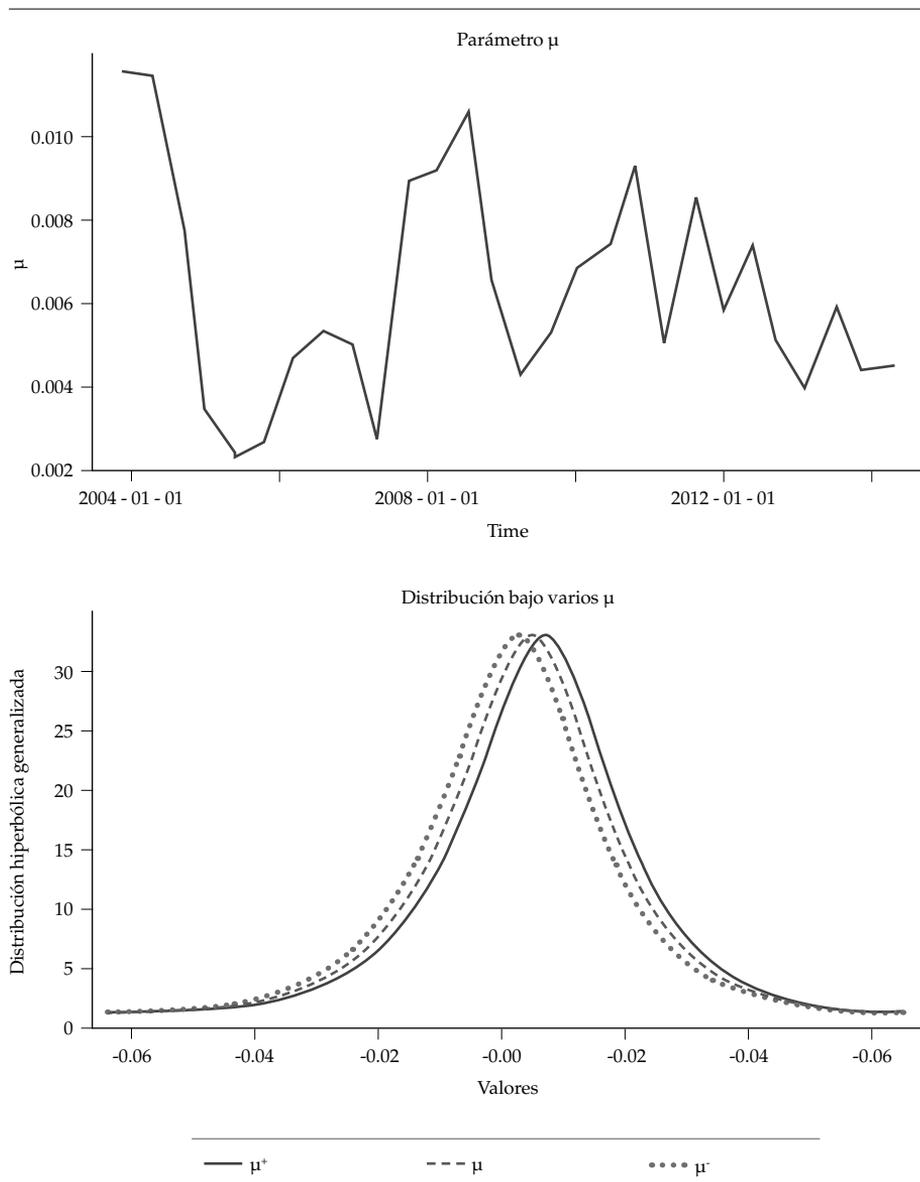


Figura 3.1. Valores históricos del parámetro μ

Fuente: cálculos del autor.

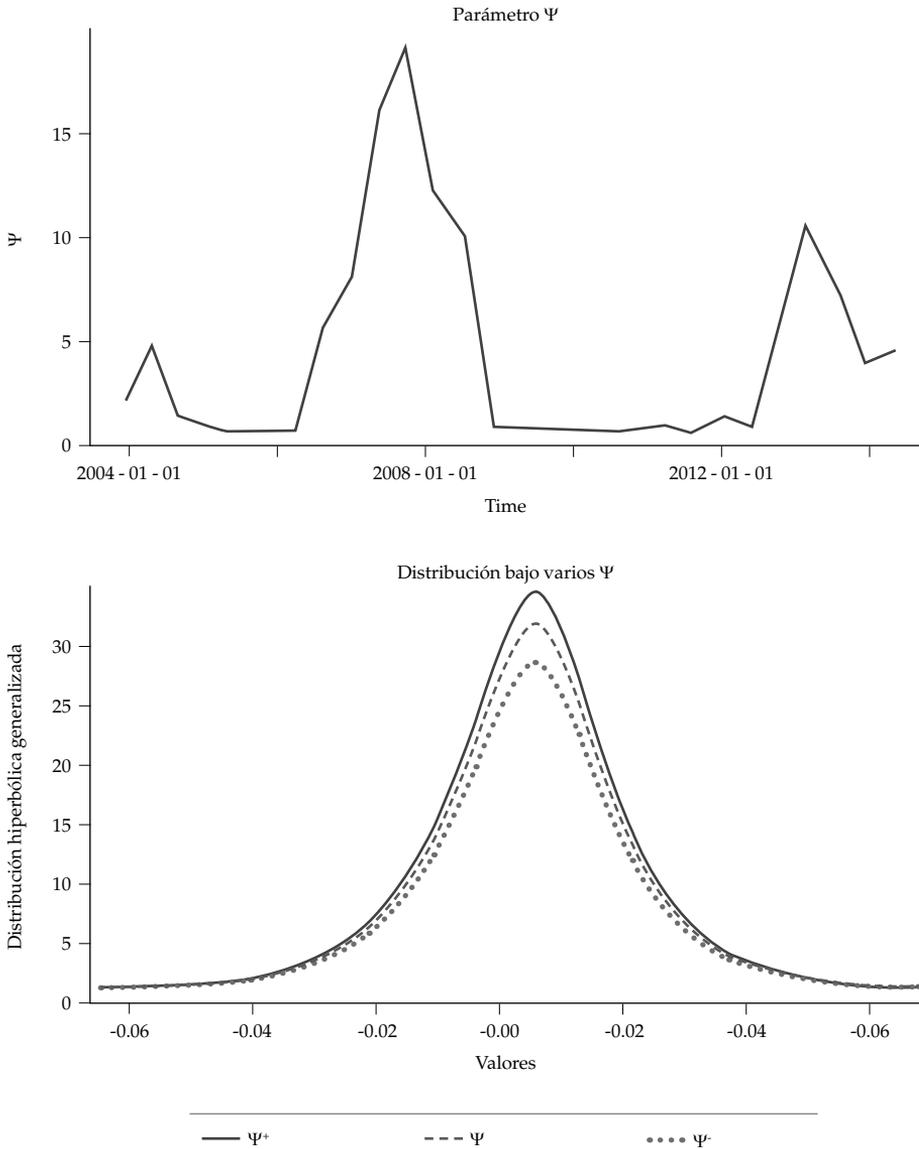


Figura 3.2. Valores históricos del parámetro ψ

Fuente: cálculos del autor.

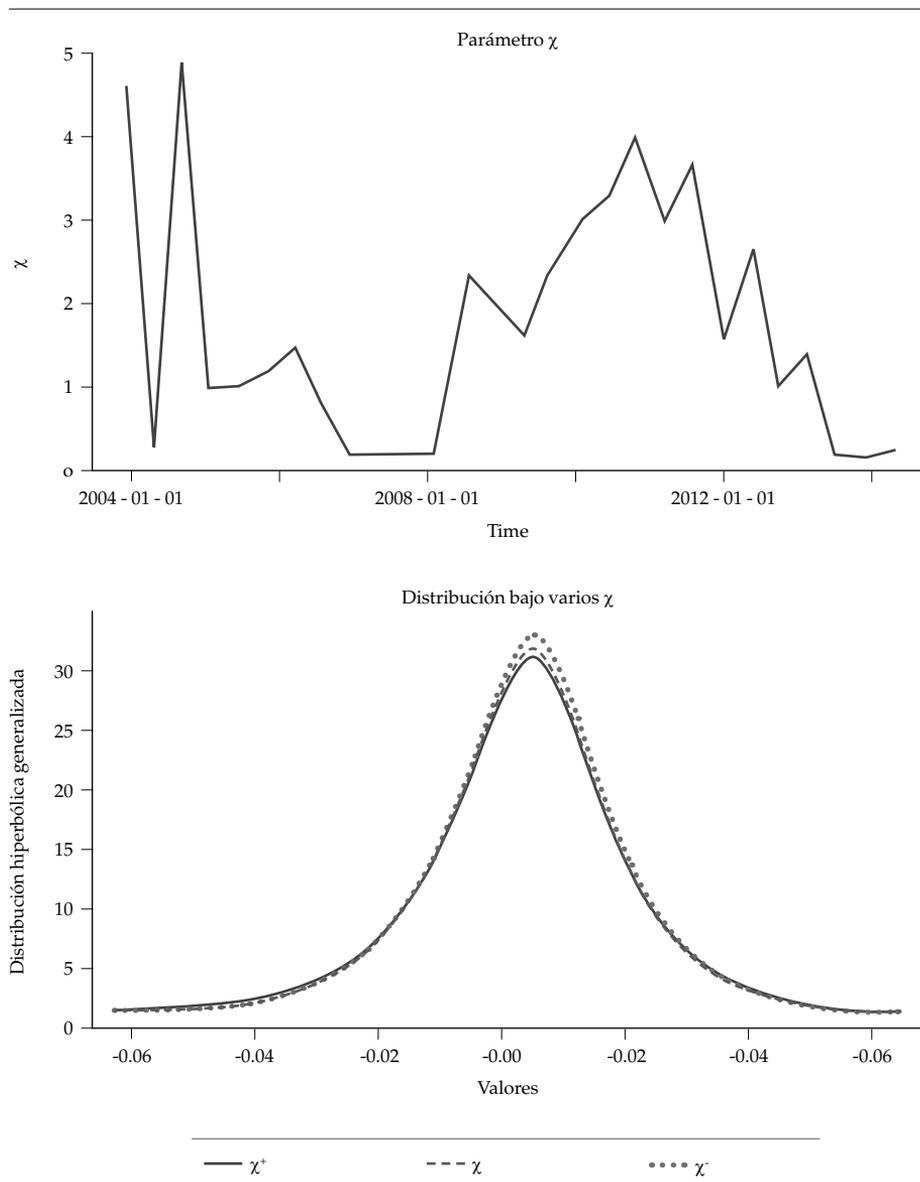


Figura 3.3. Valores históricos del parámetro χ

Fuente: cálculos del autor.

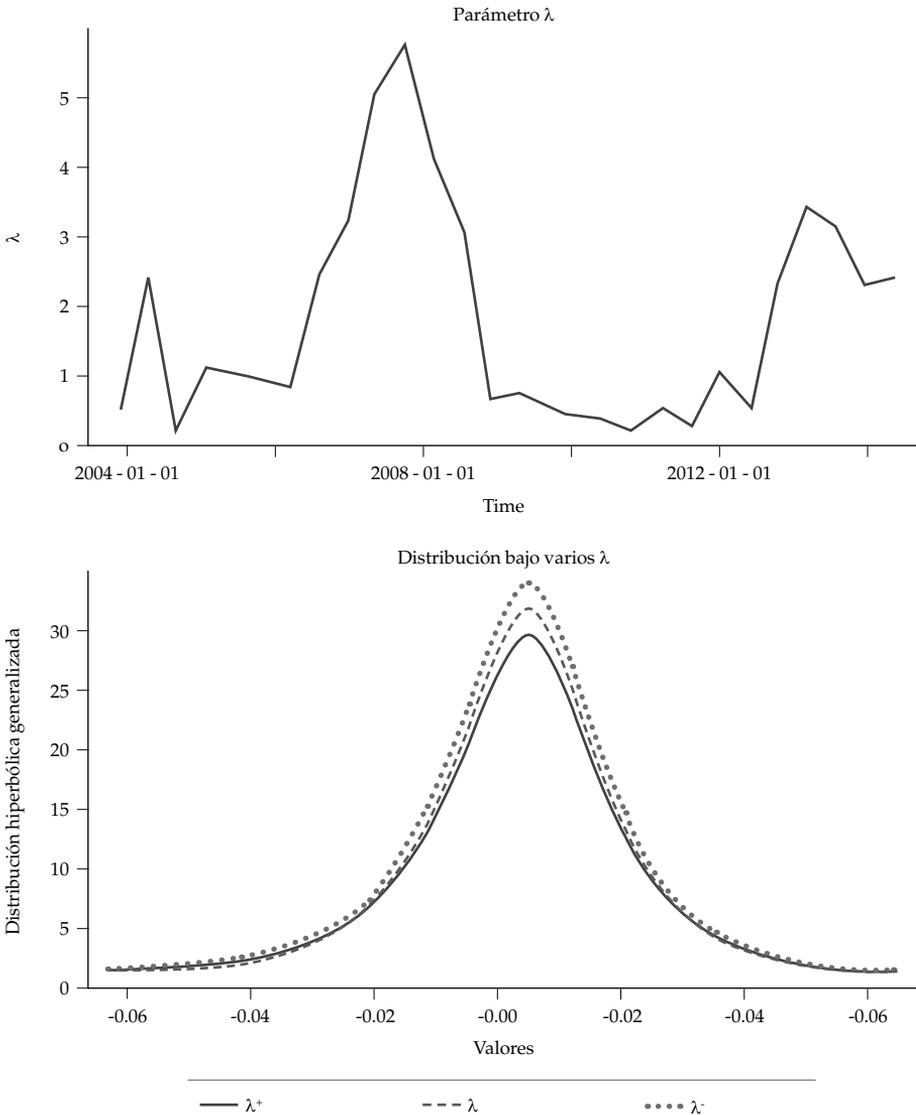


Figura 3.4. Valores históricos del parámetro λ

Fuente: cálculos del autor.

Por su parte, el término de volatilidad mostró un repunte evidente a partir de la crisis inmobiliaria y no paró hasta estabilizarse por encima de 0,02. Posterior al 2012 la volatilidad ha venido disminuyendo, hasta ubicarse en niveles históricamente bajos, comparables con los vistos hace siete años, pero superiores a los observados en el período previo a la crisis.

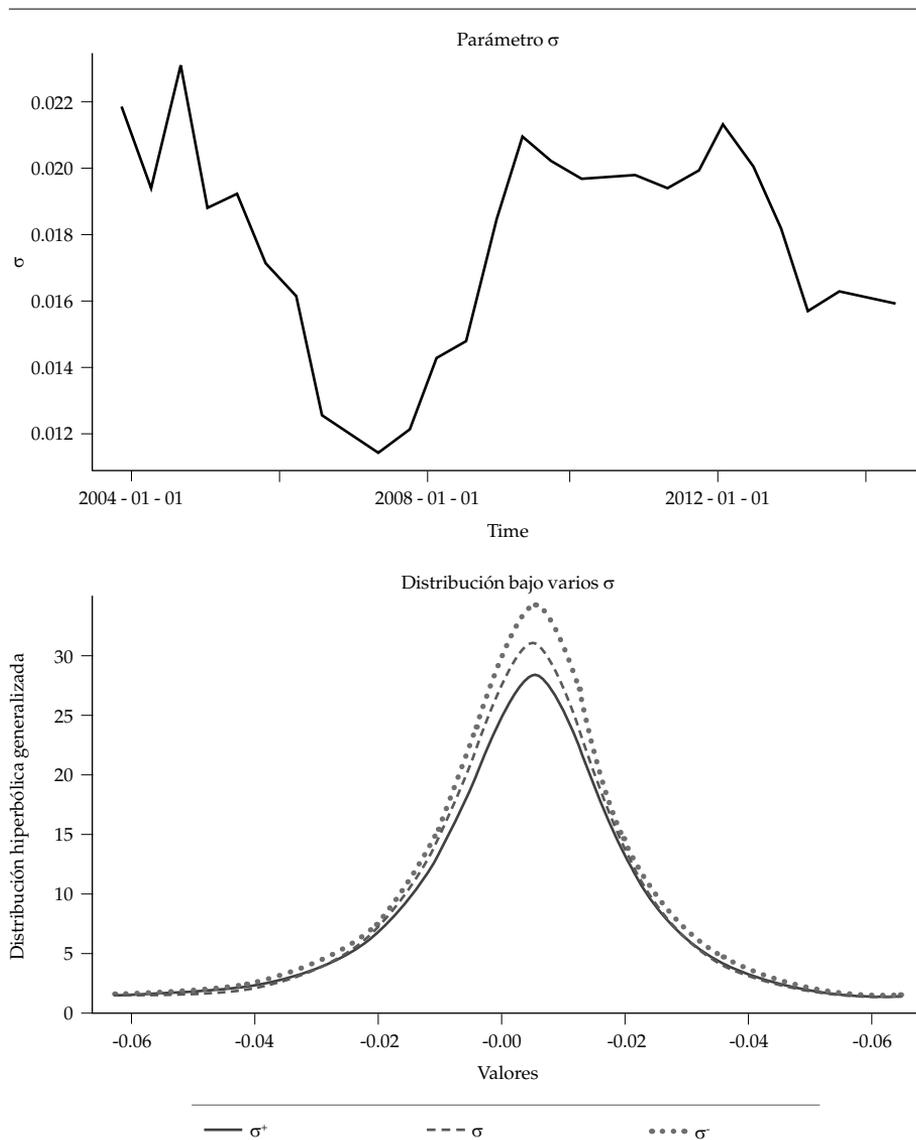


Figura 3.5. Valores históricos del parámetro Σ

Fuente: cálculos del autor.

Por último, el parámetro de asimetría ha estado en terreno negativo en los últimos diez años, hecho estilizado que no se podía recoger con la distribución gaussiana. Actualmente, el valor de γ es alto en comparación con su historia reciente, lo que refleja el buen momento que viven muchos mercados a nivel mundial, entre ellos el americano.

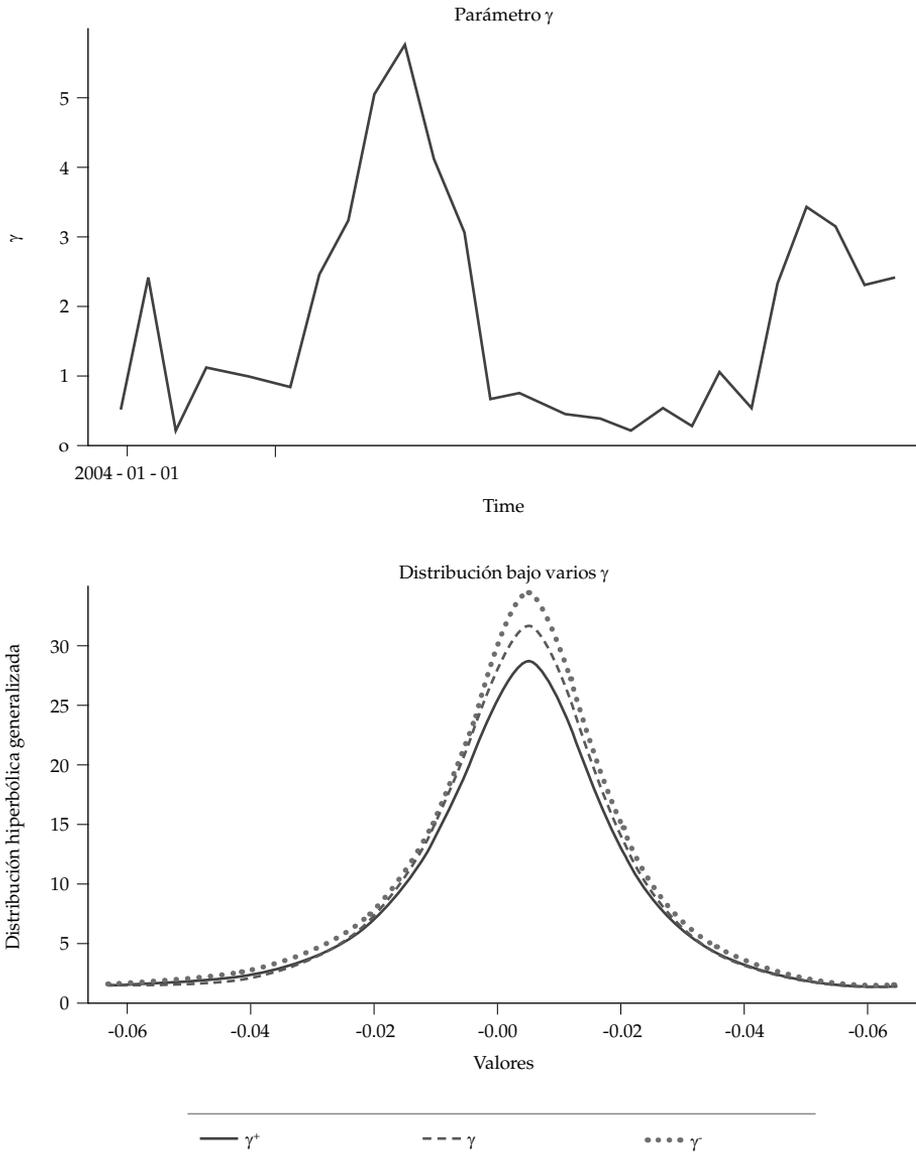


Figura 3.6. Valores históricos del parámetro γ

Fuente: cálculos del autor.

Así pues, un portafolio óptimo sería aquel que mantuviera un nivel alto en su parámetro μ , ψ , γ , χ y λ y bajo en Σ .

4.2. Portafolios óptimos

Luego de lo encontrado en el numeral anterior, el interés del investigador se torna en la optimización de portafolios bajo criterios de maximización de μ , ψ , γ , χ y λ y minimización de Σ , sujeto a un nivel de rentabilidad objetivo.

En cuanto a los valores de los parámetros de χ , ψ y λ , estos vienen dados por el ajuste de la distribución global de los datos, es decir, estos son parámetros comunes a toda la distribución compuesta por los n -activos. De esta forma, un mejoramiento en el desempeño de un portafolio por medio de estas medidas solo es posible si se eligen mejores activos dentro de la canasta sujeta de inversión.

De este modo, en este punto el proceso de optimización debe ser entendido en dos etapas. En la primera se busca maximizar χ , ψ y λ , en una etapa denominada como proceso de selección de activos; y hay otra llamada proceso de optimización, donde se busca encontrar los pesos óptimos de los activos que configuren el portafolio de menor riesgo dadas unas especificaciones en sus retornos. A partir de este momento, el documento se centrará en la segunda etapa.

4.2.1. Portafolio por minimización de varianza

Uno de los principales intereses de un administrador es mantener un portafolio de bajo riesgo, dadas algunas especificaciones sobre la rentabilidad. El primer acercamiento que se hizo en este contexto fue la minimización del CVaR de un portafolio que seguía una distribución hiperbólica en sus retornos; y, más adelante, se realizó un procedimiento en el cual se minimizaba la matriz de varianzas y covarianzas de un problema de optimización cuadrático (modelo de Markowitz); ambos con una exigencia mínima de rentabilidad. En este caso se efectuó una optimización bajo la minimización de la matriz Σ que arroja la distribución hiperbólica generalizada, es decir, se replicó el algoritmo de Markowitz pero usando la matriz de varianzas y covarianzas dada por el ajuste de los retornos a una distribución con momentos de orden superior.

El resultado encontrado muestra que el portafolio obtenido bajo el modelo de Markowitz y el expuesto en este numeral son muy parecidos, pero no exactos. Esto se debe a que la matriz de varianzas y covarianzas obtenida bajo el ajuste de la distribución hiperbólica generalizada es ligeramente inferior a la utilizada en el modelo de Markowitz, dado que la primera logra depurar efectos como asimetría y curtosis, los cuales son recogidos como ruido en la estimación de esta matriz bajo la distribución gaussiana. No obstante, las variaciones de los dos portafolios no son significativas.

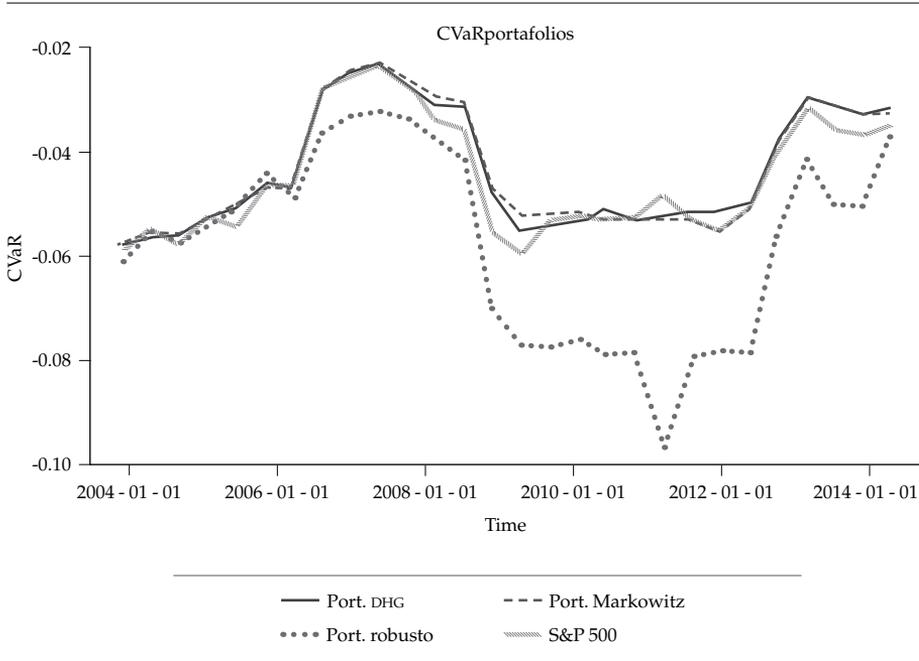


Figura 3.7. CVaR de portafolios óptimos

Fuente: cálculos del autor.

Por lo tanto, esta metodología es eficiente con respecto a la hallada en el numeral 3.2, debido a que logra conformar un portafolio con niveles de riesgo reducidos y a un costo computacional muy bajo, pero no se consigue un avance significativo con respecto al modelo de Markowitz.

4.2.2. Portafolio óptimo con asimetrías

El portafolio anterior es óptimo en minimización de riesgo y alcanza una alta semejanza con respecto a la metodología determinada en 1952 por Markowitz, pero cae en el problema de dejar fuera de su optimización características básicas como asimetría y curtosis. De este modo, para introducir momentos de orden superior en el algoritmo, se va a incorporar primero una función de utilidad, así:

$$U_t = \alpha_t - \frac{1}{2} \kappa \sigma_t^2 \tag{4.2.2.1}$$

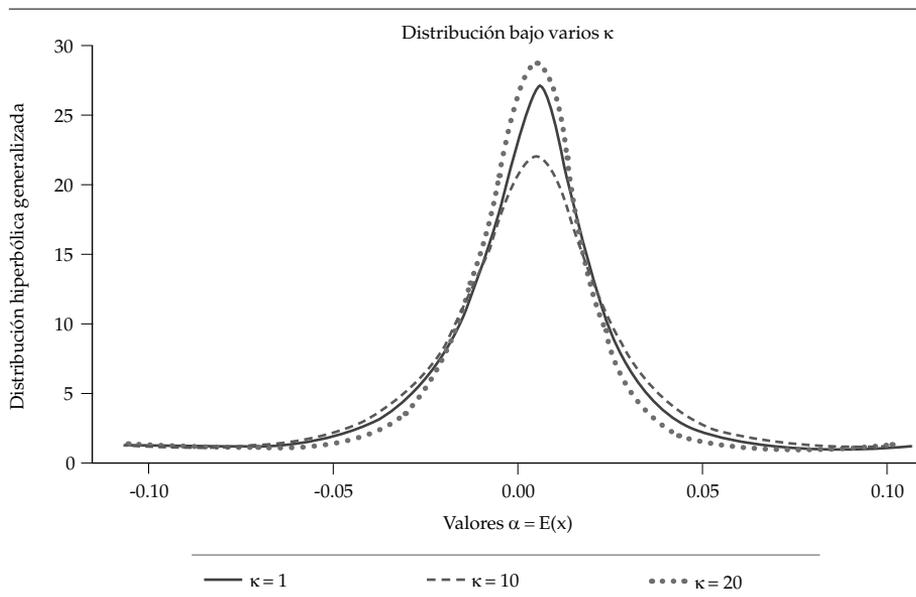


Figura 3.8. Distribuciones de retornos que se obtienen al variar la aversión al riesgo del agente

Fuente: cálculos del autor.

Donde κ hace alusión a la aversión al riesgo del inversionista, la cual no había sido tomada en cuenta en el desarrollo de los algoritmos de los portafolios óptimos anteriores, ya que solo se minimizaba el riesgo; σ_i^2 es el factor de riesgo del portafolio y α_i es la deriva del modelo.

α_i ha sido reemplazado comúnmente por $E[r_i]$, o el valor esperado de los activos, convirtiendo a (4.2.2.1) en un problema de optimización dual.

Así pues, el problema de selección de portafolio queda establecido como:

$$\min_{\alpha, \sigma^2} U_i \tag{4.2.2.2}$$

O, lo que es lo mismo,

$$\min_w -w' \Gamma_i + \kappa w' \Sigma_i w \tag{4.2.2.3}$$

Este problema puede ser resuelto fácilmente siguiendo los supuestos de Goldfarb e Idnani [GI:83] en su texto de 1983, en el cual se minimiza el riesgo medido por la matriz de varianzas y covarianzas, mientras se maximiza el retorno esperado del portafolio.

Luego de realizarse algunas simulaciones, se llegó a que entre más averso es el inversionista al riesgo, la distribución de los retornos resultante del proceso de maximización presenta una mayor curtosis, ya que en la función objetivo el riesgo pesa significativamente más que el retorno del portafolio. En caso contrario, el peso de las colas se eleva y la concentración de los datos alrededor de la media se hace menor (mayor dispersión), ya que el algoritmo permite al portafolio la existencia de retornos extremos en procura de un nivel superior de rentabilidad, so pena de un aumento en los niveles de riesgo. Este resultado es coherente con la intuición financiera, pero aún se mantiene en línea con la teoría de Markowitz, donde solo son incluidos los dos primeros momentos de la distribución de los retornos en la optimización de portafolios.

De manera que se hará uso del parámetro γ de la distribución hiperbólica, para introducir el concepto de asimetría a este problema (momento de orden superior).

De este modo, y haciendo uso de las propiedades de transformaciones lineales vistas en 2.2.2, el problema puede ser reescrito de la siguiente forma:

$$\min_w -w'\gamma_t + \kappa w'\Sigma_t w \tag{4.2.2.4}$$

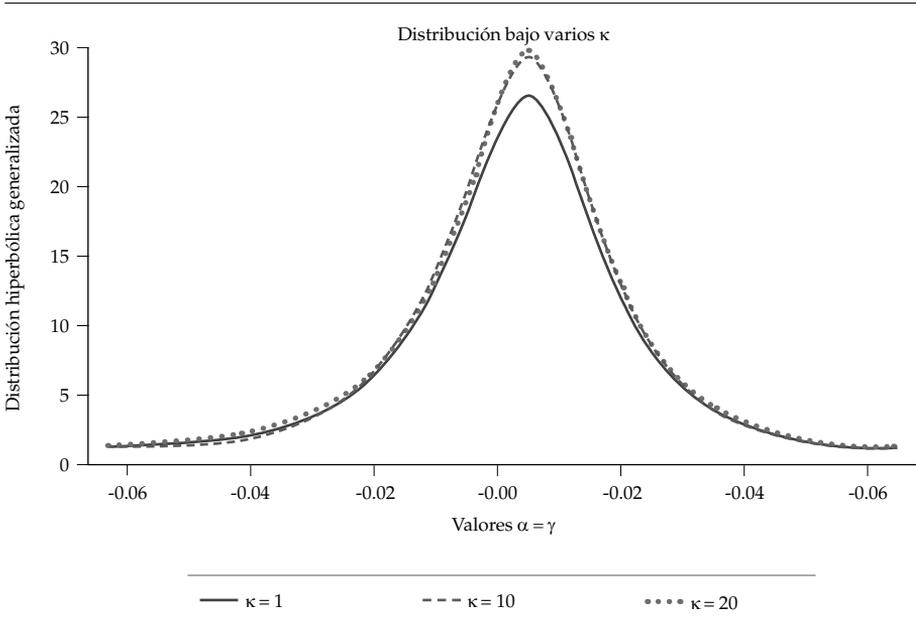


Figura 3.9. Distribuciones de retornos que se obtienen al variar la aversión al riesgo del agente

Fuente: cálculos del autor.

Donde γ es el coeficiente de asimetría de la distribución.

Al desarrollar este proceso de optimización para varios niveles de κ , se tiene que a mayor ponderación de riesgo dentro del algoritmo (mayor aversión al riesgo) con respecto al parámetro α , la distribución de los retornos se torna más simétrica y de mayor curtosis, mientras que cuando γ pesa más (relativamente) dentro de la función de utilidad (4.2.2.4), las colas de la distribución se hacen más pesadas, sobre todo a la derecha, ya que se desea que una mayor concentración de datos de corte positivo estén dentro de los resultados del portafolio. Así pues, entre mayor sea el peso de γ dentro del algoritmo, más agresivo va a ser el portafolio.

Este resultado es coherente con la teoría financiera, pues a mayor retorno esperado o asimetría positiva, mayor riesgo asociado.

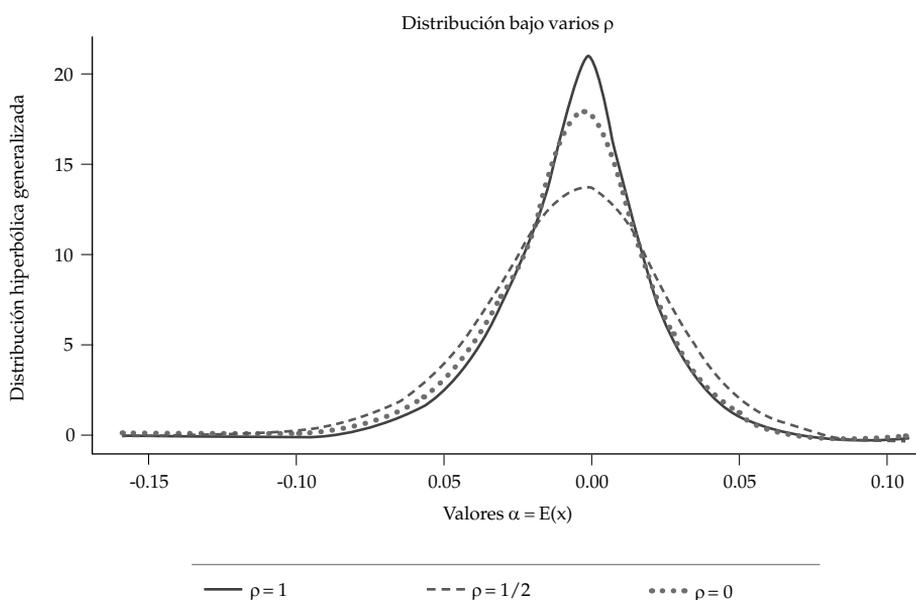


Figura 3.10. Distribuciones de retornos que se obtienen al variar el parámetro ρ

Fuente: cálculos del autor.

No obstante, para no tener que elegir necesariamente entre retorno esperado y asimetría en el parámetro α de la expresión (4.2.2.1), se puede construir un índice ponderado donde se recopilen los dos. De esta forma, se pueden incorporar en el proceso de optimización todos los momentos que proporciona la distribución hiperbólica generalizada, cada uno valorados a la necesidad del investigador o al gusto del administrador de portafolios.

$$I_t = \rho E[r_t] + (1 - \rho)\gamma_t \quad (4.2.2.5)$$

$$\underset{w}{\text{mín}} - w'I_t + \kappa w'\Sigma_t w \quad (4.2.2.6)$$

De esta forma, se puede construir una amplia gama de portafolios óptimos en términos de riesgo, dadas algunas demandas del administrador sobre los momentos de orden superior de la distribución de los retornos, lo que genera un avance importante con respecto a la teoría clásica, ya que esta solo tiene en cuenta las medidas de riesgo y retornos esperados a la hora de seleccionar un portafolio.

Así pues, la distribución hiperbólica generalizada conlleva un avance cualitativo importante al permitir la expansión de la selección de portafolios a otros momentos de orden superior, los cuales no eran tenidos en cuenta en la teoría clásica pero que son de vital relevancia a la hora de gestionar inversiones de capital.

Capítulo 5. Stress test

En este numeral se va a dividir en dos el concepto de escenario de estrés. En el primero, se estresará la función global de retornos del portafolio para simular eventos que llevan a un choque generalizado en el mercado, como el vivido en 2008; y, en un segundo plano, se generarán impactos a ciertos activos o estimadores, con el fin de precisar el efecto que puede tener un activo sobre la composición global del portafolio cuando este ha sufrido un efecto adverso importante, sin que esto afecte a la totalidad del mercado –*ceteris paribus*–.

5.1. Escenarios de estrés global

Para generar escenarios de estrés, dadas la relación entre los parámetros y la forma de la distribución de los retornos, basta con elevar los valores de σ , y reducir las magnitudes de μ , χ , ψ , λ y γ , lo que es similar al proceso realizado anteriormente para determinar el portafolio robusto en 3.3.1.

Así pues, se va a suponer un ejercicio simple en el cual todos los parámetros van a variar un 10% en procura de estresar la distribución de los retornos para ver cuánto se modifica el CVaR del portafolio.

Luego de realizado el experimento, los resultados arrojan que el CVaR original es del 1,15%, y bajo el escenario de estrés este aumenta hasta 1,37%, lo que implica un aumento de 22 puntos básicos.

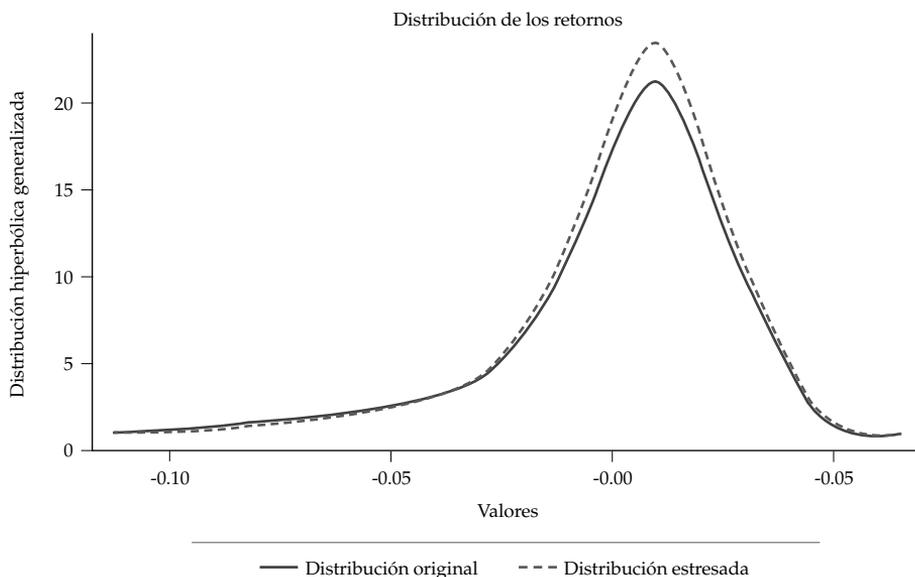


Figura 4.1. Distribución de los retornos antes y después del choque a los parámetros

Fuente: cálculos del autor.

Esta es una forma muy práctica de valorar el riesgo de mercado de un portafolio ante un evento de crisis. No obstante, el análisis puede ser llevado un poco más lejos, al permitirse generar un nuevo portafolio óptimo bajo este escenario, y así dotar al administrador de una hoja de ruta, en dado caso que se llegue a presentar tal situación.

5.2. Escenarios de estrés sobre activos puntuales

Un ejercicio adicional de estrés se podría realizar modificando los parámetros de cada activo dentro de la canasta de títulos sujetos de inversión, para luego determinar el nuevo portafolio óptimo, lo que abre un gran pliego de opciones para que el analista pueda simular y determinar portafolios ante diversos escenarios, o la opción de tener diversas hojas de ruta para los administradores de portafolio ante eventos hipotéticos sobre cada activo.

De esta forma, si se eleva el valor de la varianza del activo con mayor participación histórica dentro de los portafolios, su efecto en la composición podría llevar a cambios muy importantes, y, por ende, le sería de gran ayuda al administrador saber cuál sería la nueva composición óptima a la que debería llevar la canasta de bienes, de presentarse tal situación. Estos choques sobre activos particulares podrían deberse a diversas situación de mercado, tales

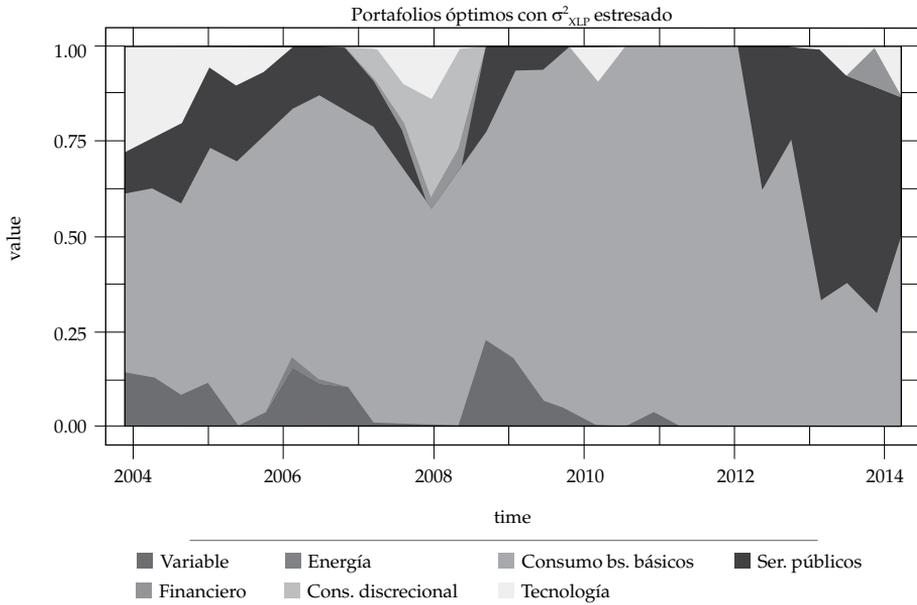


Figura 4.2. Composición de los portafolios óptimos bajo un escenario estresado en XLP
Fuente: cálculos del autor.

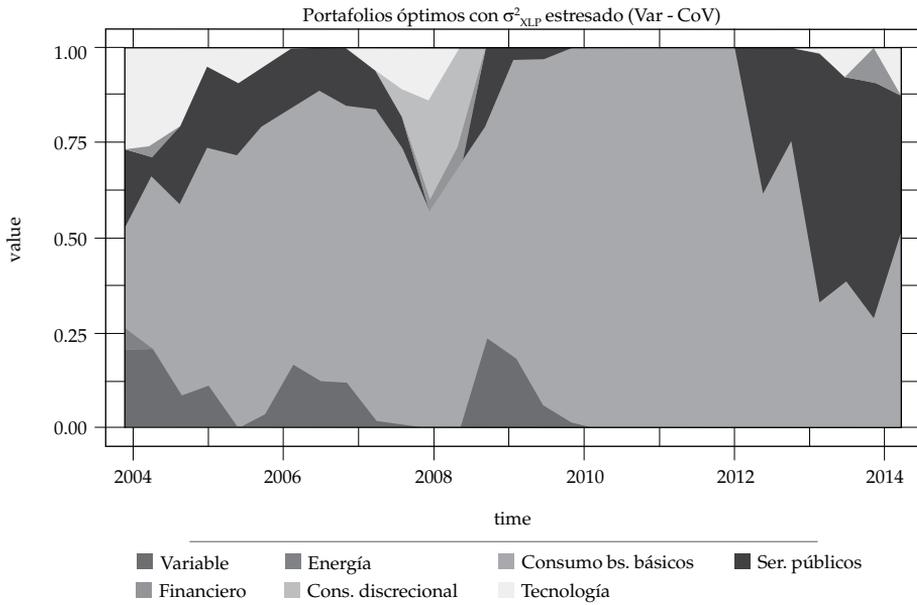


Figura 4.3. Composición de los portafolios óptimos bajo un escenario estresado en XLP (10%)
Fuente: cálculos del autor.

como cambios en la estructura organizacional, fallas en la ejecución del algún proyecto o alguna problemática legal o financiera derivada de su actividad económica, sin que esto afecte la totalidad del mercado.

Así pues, si se toma la varianza de XLP (Sector de consumo de bienes básicos) y se eleva en un 10% a través de la historia, la nueva composición óptima sería la expresada en la figura 4.2.

En esta se puede apreciar cómo hay solo una pequeña variación en la composición de este ETF en los diferentes portafolios óptimos, lo que muestra la gran importancia de este activo en la canasta seleccionada, ya que su relación de riesgo-rentabilidad sigue siendo atractiva a pesar de elevarse su cuantificación de riesgo en un 10%.

Dado este resultado, se elevaron también las covarianzas en un 10% y se volvieron a calcular los portafolios óptimos. Los resultados arrojaron que las variaciones no son significativas a lo largo del tiempo, salvo en un par de ocasiones; y en cinco casos su variación fue nula para los tres escenarios. También se encontró que las modificaciones más grandes se dan cuando se eleva solo la varianza del activo, sin modificar su estructura de covarianzas.

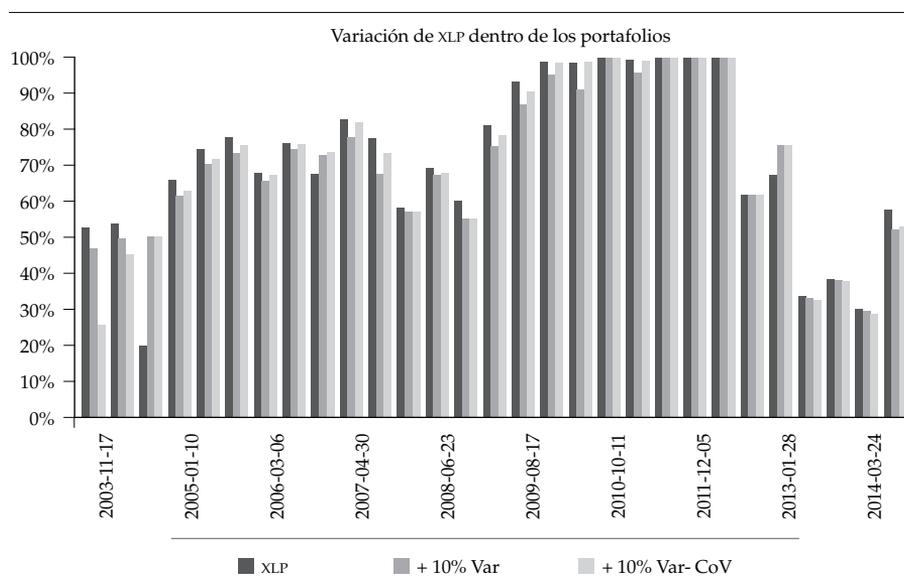


Figura 4.4. Variación de la participación de XLP dentro de los portafolios óptimos

Fuente: cálculos del autor.

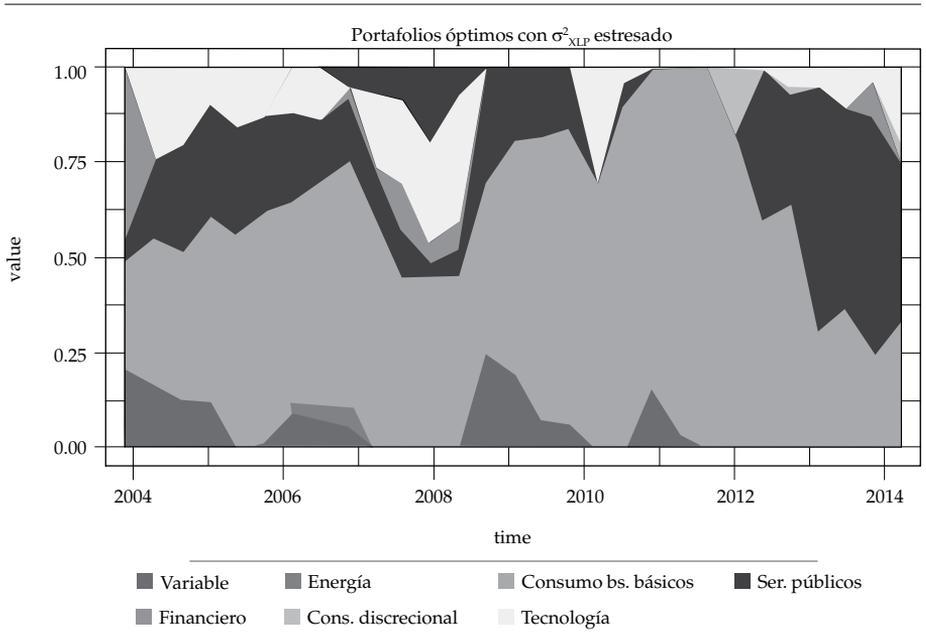
Si se genera un escenario de más estrés, como elevar la varianza en un 50%, se puede apreciar un cambio importante en la composición del portafolio, aunque se mantiene XLP como el ETF con mayor participación.

Y al aumentar la varianza y covarianza de este activo en un 50 %, la composición dista bastante de los dos escenarios anteriores. Entonces, el nivel de estrés generado en este escenario hace que la hoja de ruta del administrador de portafolio sea más difícil de seguir, dado que, al darse estos choques, la recomposición sería más significativa.

Estos resultados son totalizados en la figura 5.6, donde se aprecia una disminución importante de XLP en todos los portafolios, con una mayor sensibilidad durante 2003 y 2008. También se puede ver cómo el aumento de las covarianzas puede mermar el efecto, tal como se presentó entre 2010 y 2012, o incrementar su impacto como entre 2003 y 2005, lo que es un caso interesante de análisis.

De otro lado, si en vez de elevar la matriz de varianzas y covarianzas, se simulan unos choques a la asimetría de XLP, los efectos son considerablemente diferentes. Así pues, si se hace que el parámetro γ sea un 10 % más negativo que en sus estimaciones reales, las variaciones sobre la canasta total son más fuertes que cuando se eleva la volatilidad.

Del mismo modo, si el choque sobre γ es un 50 %, el efecto total es considerable, al punto que ya XLP deja de ser el activo más importante dentro de estos portafolios, para cederle el paso a XLU (servicios públicos) y XLK (Tecnología), y desplazarse a ser un activo secundario que desaparece durante largos períodos de tiempo.



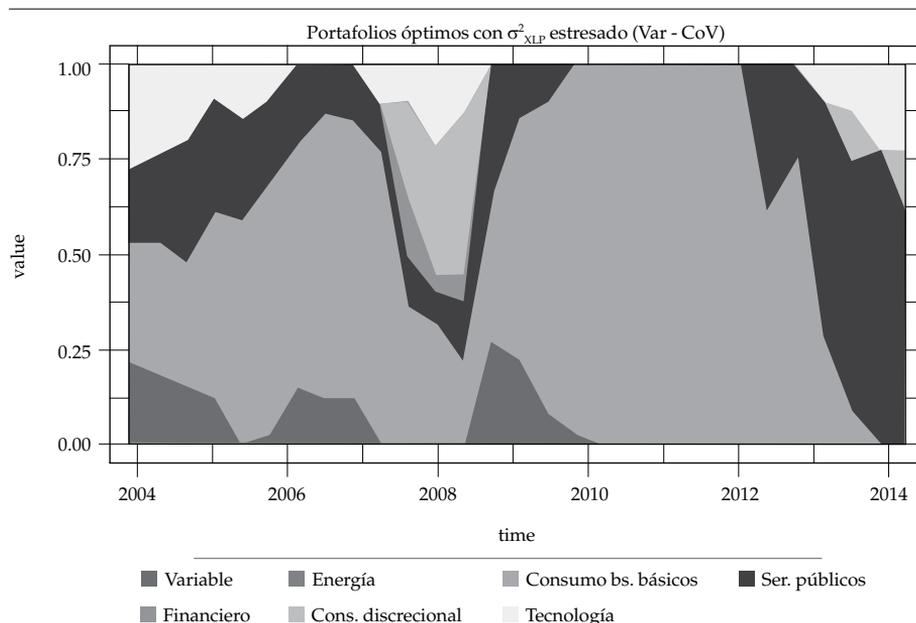


Figura 4.5. Composición de los portafolios óptimos bajo un escenano estresado en XLP (50%)

Fuente: cálculos del autor.

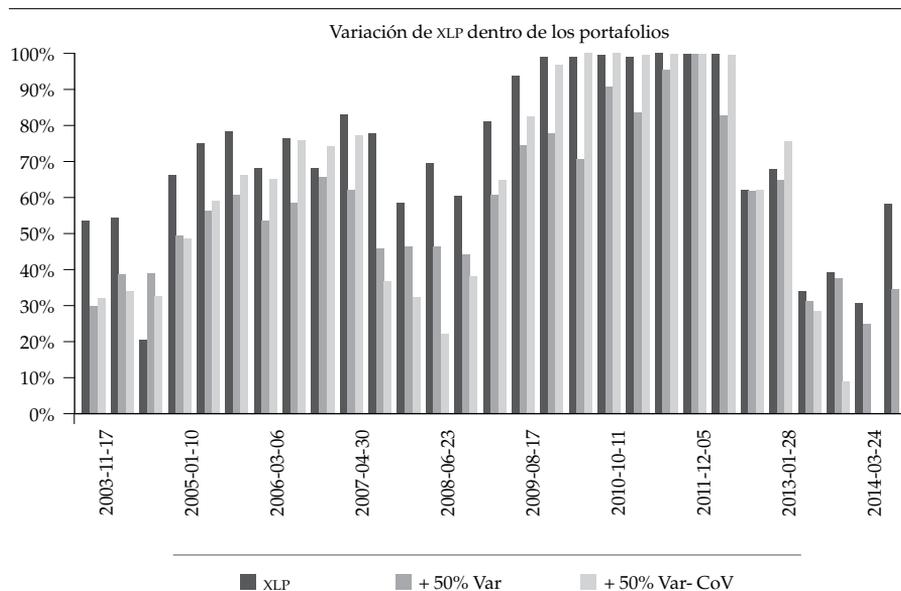


Figura 4.6. Variación de la participación de XLP dentro de los portafolios óptimos

Fuente: cálculos del autor.

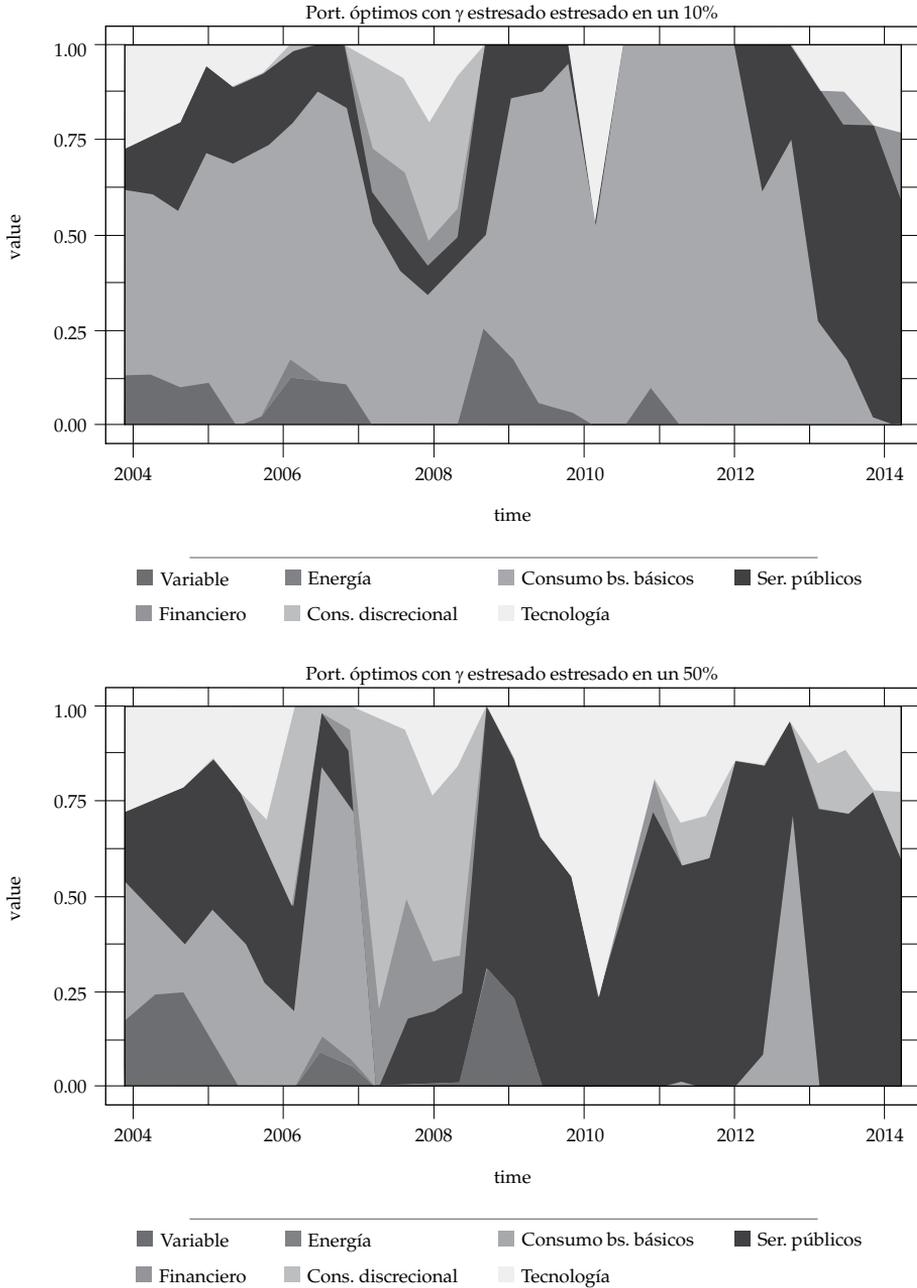


Figura 4.7. Composición de los portafolios óptimos bajo un escenano estresado en XLP (50%)

Fuente: cálculos del autor.

Así pues, si un administrador de portafolio conoce esta información, posee una hoja de ruta que lo lleva a tomar mejores decisiones en caso de que algún escenario de estos se realice, o, por lo menos, lo lleva a conocer a cuáles riesgos se podría enfrentar. Sabría que ante un aumento de la varianza del 10% en XLP en este contexto, las modificaciones que debería hacer en su portafolio serían limitadas, en comparación a cuando se da un aumento de la asimetría del mismo activo.

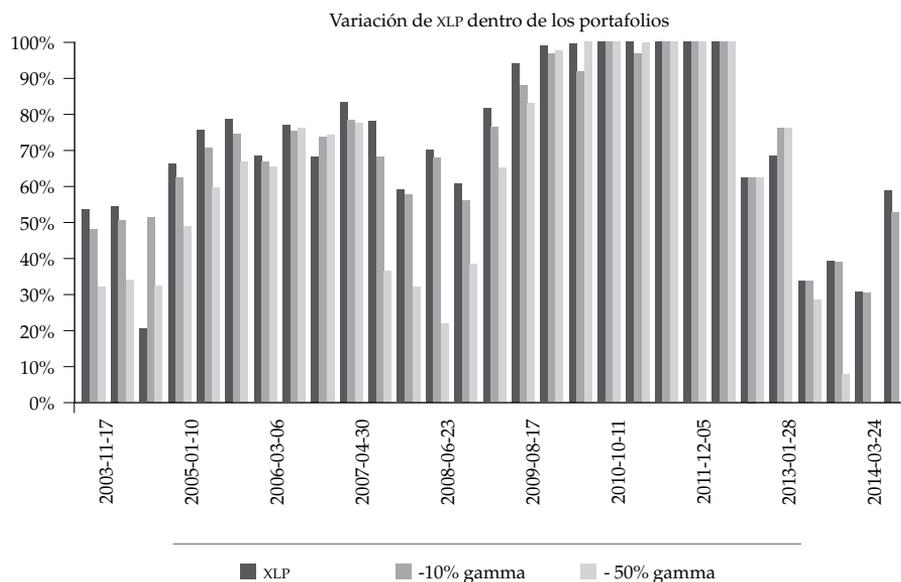


Figura 4.8. Variación de la participación de XLP dentro de los portafolios óptimos

Fuente: cálculos del autor.

Capítulo 6. Conclusiones

Con la utilización de la distribución hiperbólica generalizada en la optimización de portafolios y el uso de medidas de riesgo robustas, no se logra un gran avance en el interior de la muestra con respecto al portafolio de media-varianza, ya que las ponderaciones óptimas tienden a ser muy homogéneas entre estos dos portafolios, pese a que provienen de procesos totalmente diferentes. Es por esto que muchos administradores de portafolio siguen centrados en la minimización de la varianza, dado que, si son eficientes bajo esta medida, también lo son bajo las demás.

Sin embargo, un avance importante que se da con la distribución hiperbólica generalizada es que al utilizar esta función de distribución se tiene una mejor medición de los verdaderos niveles de riesgo a los que se expone un determinado activo o portafolio, puesto que con sus múltiples parámetros se logra un mejor ajuste de lo que sería la verdadera forma funcional de la distribución de pérdidas, hecho en el que falla la distribución normal. Otra de las ventajas que se obtiene al utilizar esta distribución es que se pueden generar diversos escenarios, con los cuales se consigue simular numerosos choques sobre los niveles de riesgo a los que se expone el administrador y sus efectos sobre los portafolios.

Así pues, en la parte de simulación el avance es significativo, ya que se pueden generar escenarios de estrés muy particulares, como cambios en la asimetría, en la curtosis o en el tamaño de las colas de la distribución, mientras se mantienen las otras características invariantes, o escenarios de estrés generalizados, como el vivido en 2008, donde todos los parámetros de la distribución sufrieron perturbaciones negativas. Esto permite no solo generar simulaciones para épocas de estrés (*stress test*), sino también cuantificar ciertos impactos ante cambios estructurales de los activos, tales como adquisiciones, emisiones o nuevos regímenes de mercado.

Además, es posible generar procesos de optimización para hallar portafolios con ciertas características particulares en términos de asimetría, curtosis, valor esperado y riesgo, lo cual es una extensión a la optimización de portafolio tradicional, dado que se logra incluir momentos de orden superior a un costo computacional muy bajo, con resultados de alto interés para los analistas.

Así mismo se desarrolló e implementó una metodología de portafolios óptimos siguiendo el CVaR del peor escenario, bajo la cual se consigue disminuir la sensibilidad de las ponderaciones ante posibles efectos de error de estimación de los parámetros de la distribución. El resultado en la práctica fue satisfactorio, dadas las restricciones que se le imponen dentro del proceso de optimización, ya que, aunque los niveles de riesgo aumentaron, la dispersión de las ponderaciones a través del tiempo fueron bajas, y se elevaron solo en períodos de transición en el mercado.

Por último, el tiempo empleado para correr cada uno de los portafolios históricos sobre una base de 200 semanas empleando las tres metodologías (bajo distribución hiperbólica generalizada, el portafolio robusto y el de Markowitz) muestra que el algoritmo más eficiente es el de media-varianza con un tiempo promedio muy cercano a cero; luego, le siguen el portafolio robusto y el portafolio hallado bajo la distribución hiperbólica generalizada, debido a que cuentan con estructuras muy similares. No obstante, el tiempo de los portafolios no gaussianos no es elevado en términos absolutos, al tardar

en promedio entre 3 y 7 segundos. En cuanto a los portafolios hallados en el numeral 4.2, sus tiempos fueron equivalentes a los de Markowitz, dado que siguen el mismo proceso de optimización bajo funciones convexas cuadráticas.

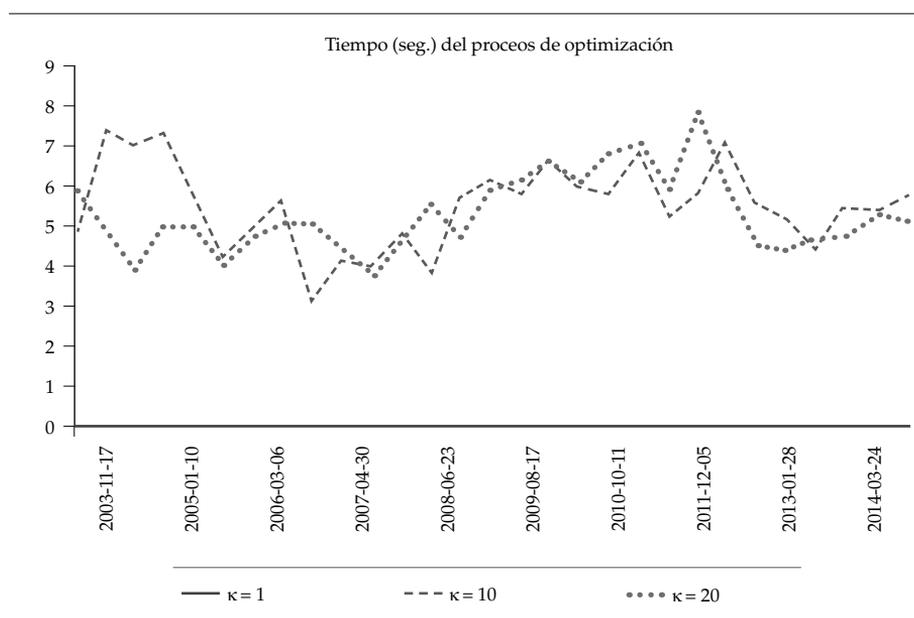


Figura 5.1. Tiempo empleado (en segundos) para obtener cada uno de los portafolios
Fuente: cálculos del autor.

Dados estos resultados, una estrategia óptima para un administrador de portafolios podría ser: determinar las ponderaciones óptimas bajo el modelo de media-varianza propuesto por Markowitz, y, posteriormente, hacer el ajuste de los retornos de tal portafolio a la distribución hiperbólica generalizada para determinar así sus niveles de riesgo y los escenarios de estrés de interés del investigador, obteniendo de esta forma lo mejor de cada metodología. O si, por el contrario, desea canastas con características más allá de los portafolios de mínimo riesgo, puede ajustar el proceso siguiendo las metodologías expuestas en 4.2, donde puede ajustar una canasta de bienes a momentos de orden superior.

Referencias

K. Aas (1995), *NIG and Skew Student's t: Two Special Cases of the Generalized Hyperbolic Distribution*, Norwegian Computing Center.

- M. Abramowitz (1970), *Handbook of mathematical functions*, NY: Dover Publications.
- P. Artzner, F. Delbaen, J. Eber, D. Heath (1999), *Coherent measures of risk*, Math., pp. 203-228.
- T. Bali (2007), *Value at risk and the cross-section of hedge*, Journal of Banking and Finance, pp. 1135-1166.
- T. Bali (2004), *Alternative approaches to estimating VaR for hedge fund*, Risk Books.
- Barndorff-Nielsen (1977), *Exponentially decreasing distributions for the logarithm of the particle size*, Proceedings of the Royal Society.
- Barndorff-Nielsen (1997), *Normal inverse Gaussian distributions and stochastic*, Scandinavian Journal of Statistics, pp. 1-13.
- Barndorff-Nielsen (1997), *Processes of normal inverse Gaussian type*, Springer, pp. 41-68.
- N. Bingham, R. Kiesel (1991), *Modeling asset returns with hyperbolic distributions*.
- W. Breymann, A. Dias, P. Embrechts (2003), *Dependence Structures for Multivariate High Frequency Data in Finance* Quantitative Finance, pp. 1-14.
- W. Breymann (2008), *Measuring Risk of Short Return Series with an Application to Fund of Hedge Funds Data*.
- W. Breymann (2013), *GHYP*, Recuperado de http://cran.r-project.org/web/packages/ghyp/vignettes/Generalized_Hyperbolic_Distribution.pdf
- R. Browne, P. McNicholas (2013), *A mixture of generalized hyperbolic distributions*, Cornell University Library.
- E. Eberlein, U. Keller, K. Prause (1998), *New insights into smile, mispricing and value at risk*, Journal of Business, pp. 371-405.
- E. Giorgi (2014), *On the computation of multivariate scenario sets for the skew-t and generalized hyperbolic families*.
- D. Goldfarb, A. Idnani (1983), *A numerically stable dual method for solving strictly convex quadratic programs*, Mathematical Programming.
- B. Halldórsson, R. Tütüncü (2003), *An interior-point method for a class of saddle-point*, Journal of Optimization Theory and Applications, pp. 559-590.
- M. Hellmich, S. Kassberger (2009), *Efficient and Robust Portfolio Optimization in the Multivariate Generalized Hyperbolic Distribution*.
- W. Hu (2005), *Calibration of multivariate generalized hyperbolic distribution using EM algorithm, with applications in risk management, portfolio optimization and portfolio credit risk*, Electronic Theses, Treatises and Dissertations.
- S. Kassberger (2006), *A fully parametric approach to return modeling Financial Markets and Portfolio Management*, pp. 472-491.
- S. Kim, S. Boyd (2007), *Robust Efficient Frontier Analysis with a Separable Uncertainty Model*, Stanford University.
- R. Korn (2005), *Optimal portfolios: new variations of an old theme*, Springer.

- C. Liu (1994), *The ECME algorithm a simple extension of EM and ECM with faster*, *Biometrika*, pp. 633-648.
- A. McNeil (2005), *Quantitative risk management: concepts, techniques and tools*, Princeton U.
- M. Paoletta (2007), *Intermediate probability: a computational approach*.
- K. Prause (1999), *The generalized hyperbolic model: estimation financial derivatives*.
- T. Rockafellar, S. Uryasev (1999), *Optimization of conditional value at risk*.
- D. Scott, D. Würtz (2009), *Moments of the Generalized Hyperbolic Distribution*, Zürich.
- R. Slevinsky, D. Würtz (2010), *A recursive algorithm for the G transformation and accurate computation of incomplete Bessel function*, *Applied Numeric Mathematics*.
- B. Taylor (2011), *Nonparametric goodness-of-fit tests for discrete null distributions*, *The R Journal*, pp. 34-39.
- R. Tütüncü, M. Koenig (2004), *Robust asset allocation*, *Annals of Operations Research*, pp. 157-187.

Apéndice A

Diferentes parametrizaciones de la distribución hiperbólica generalizada¹⁴

		$(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ -Parametrization					
		λ	χ	ψ	μ	Σ	γ
ghyp		$\lambda \in \mathbb{R}$	$\chi > 0$	$\psi > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^{\Sigma}$	$\gamma \in \mathbb{R}^d$
hyp		$\lambda = \frac{d+1}{2}$	$\chi > 0$	$\psi > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^{\Sigma}$	$\gamma \in \mathbb{R}^d$
NIG		$\lambda = -\frac{1}{2}$	$\chi > 0$	$\psi > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^{\Sigma}$	$\gamma \in \mathbb{R}^d$
t		$\lambda < 0$	$\chi > 0$	$\psi = 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^{\Sigma}$	$\gamma \in \mathbb{R}^d$
VG		$\lambda > 0$	$\chi = 0$	$\psi > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^{\Sigma}$	$\gamma \in \mathbb{R}^d$
		$(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ -Parametrization					
		λ	$\bar{\alpha}$	μ	Σ	γ	
ghyp		$\lambda \in \mathbb{R}$	$\bar{\alpha} > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^{\Sigma}$	$\gamma \in \mathbb{R}^d$	
hyp		$\lambda = \frac{d+1}{2}$	$\bar{\alpha} > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^{\Sigma}$	$\gamma \in \mathbb{R}^d$	
NIG		$\lambda = \frac{1}{2}$	$\bar{\alpha} > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^{\Sigma}$	$\gamma \in \mathbb{R}^d$	
t		$\lambda = -\frac{d}{2} < -1$	$\bar{\alpha} = 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^{\Sigma}$	$\gamma \in \mathbb{R}^d$	
VG		$\lambda > 0$	$\bar{\alpha} = 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^{\Sigma}$	$\gamma \in \mathbb{R}^d$	
		$(\lambda, \alpha, \mu, \Delta, \delta, \beta)$ -Parametrization					
		λ	α	δ	μ	Δ	β
ghyp		$\lambda \in \mathbb{R}$	$\alpha > 0$	$\delta > 0$	$\mu \in \mathbb{R}^d$	$\Delta \in \mathbb{R}^{\Delta}$	$\beta \in \{x \in \mathbb{R}^d : \alpha^2 - x' \Delta x > 0\}$
hyp		$\lambda = \frac{d+1}{2}$	$\alpha > 0$	$\delta > 0$	$\mu \in \mathbb{R}^d$	$\Delta \in \mathbb{R}^{\Delta}$	$\beta \in \{x \in \mathbb{R}^d : \alpha^2 - x' \Delta x > 0\}$
NIG		$\lambda = -\frac{1}{2}$	$\alpha > 0$	$\delta > 0$	$\mu \in \mathbb{R}^d$	$\Delta \in \mathbb{R}^{\Delta}$	$\beta \in \{x \in \mathbb{R}^d : \alpha^2 - x' \Delta x > 0\}$
t		$\lambda < 0$	$\alpha = \sqrt{\beta' \Delta \beta}$	$\delta > 0$	$\mu \in \mathbb{R}^d$	$\Delta \in \mathbb{R}^{\Delta}$	$\beta \in \mathbb{R}^d$
VG		$\lambda > 0$	$\alpha > 0$	$\delta = 0$	$\mu \in \mathbb{R}^d$	$\Delta \in \mathbb{R}^{\Delta}$	$\beta \in \{x \in \mathbb{R}^d : \alpha^2 - x' \Delta x > 0\}$

¹⁴ Esta tabla fue tomada de Breyman, W. (2013).