

NEGOCIACION POR PARES: UN ENFOQUE MEDIANTE CÓPULAS

Luis Fernández

Trabajo de investigación 017/009

Master en Banca y Finanzas Cuantitativas

Tutores: Dr. Oscar Carchano Alcina

Universidad Complutense de Madrid

Universidad del País Vasco

Universidad de Valencia

Universidad de Castilla-La Mancha

www.finanzasquantitativas.com

**Negociación por pares:
un enfoque mediante cópulas**

Autor: Luis Fernández

Contacto: luisfernandezpayeras@hotmail.com

Supervisor: Dr. Oscar Carchano

RESUMEN

Desde que en los años ochenta se popularizó por primera vez la negociación por pares gracias a un grupo de *quants* de Morgan Stanley, esta estrategia de arbitraje estadístico ha ido despertando el interés de multitud de inversores, tanto de fondos de inversión como de particulares, principalmente atraídos entre otras cosas por su teórica neutralidad frente al mercado y su alto ratio rentabilidad/riesgo.

En la actualidad, los trabajos relacionados con esta estrategia muestran como han disminuido tanto el número de señales generadas como los rendimientos obtenidos de aplicar dicha estrategia. Las causas están relacionadas con lo extendida que se encuentra entre los fondos de inversión y lo rápido que estos se aprovechan de dichas ineficiencias, dejando un margen muy pequeño de entrada al inversor minorista. También es cierto que este proceso de reducción de oportunidades ha venido acompañado de nuevos enfoques que permiten seguir aprovechándose de estas ineficiencias del mercado generadas a partir de pares de activos altamente correlacionados.

En este trabajo se va a realizar un estudio empírico de la estrategia de negociación por pares bajo el enfoque de cópulas. La motivación de modelizar la dependencia entre dos activos a través de cópulas es que permite recoger posibles dependencias asimétricas y no lineales, y de este modo es posible ser más precisos al llevar a cabo esta estrategia y solventar algunas deficiencias de los enfoques previos.

ÍNDICE

Resumen

1	Introducción	3
2	Literatura previa	5
2.1	Método de la distancia	5
2.2	Método de cointegración	7
2.3	Método de series temporales	9
2.4	Método de control estocástico	10
2.5	Otros enfoques	11
2.5.1	<i>Machine learning</i>	12
2.5.2	Componentes principales	12
2.5.3	Cópulas	13
3	Metodología	17
3.1	Introducción y selección de pares	17
3.2	Modelización de la dependencia mediante cópulas	18
3.3	Tipologías de cópulas	19
3.4	Probabilidades condicionales y estrategia de negociación	22
4	Resultados empíricos	26
4.1	Datos utilizados y estadísticos descriptivos	26
4.2	Resultados con datos diarios	29
4.3	Resultados con datos diarios en 3 períodos	37
4.4	Resultados ventana móvil con datos de alta frecuencia vs datos diarios	43
4.5	Optimización de las señales de entrada	48
5	Conclusión	52
	Anexo	54
	Bibliografía	76

1.INTRODUCCIÓN

La idea que radica detrás de la negociación por pares es encontrar dos activos que presenten una elevada dependencia en su evolución a lo largo del tiempo, detectar posibles divergencias que se produzcan entre ambos y aprovechar esos *gaps* tomando posición simultáneamente larga en el activo relativamente sobrevalorado y corta en el activo relativamente infravalorado con el objetivo de cerrarlas cuando se produzca la reversión al “equilibrio” (por este motivo se considera que se trata de una estrategia con “reversión a la media”).

Las dos características principales presentes en este tipo de estrategia son que es considerada como una estrategia de arbitraje estadístico y es neutra frente al mercado, véase Nath (2003) [22] y Lin et al. (2006) [20].

Si la estrategia se lleva a cabo con dos acciones del mismo país y del mismo sector y dado que se toman posiciones opuestas en ambas acciones, se elimina de esta forma el riesgo sistemático procedente del país y del sector al que pertenezcan, asumiendo de esta forma un riesgo idiosincrático. Esto se conoce como una cartera u operación cero-beta respecto al mercado ya que su evolución no depende de la evolución del mercado al que pertenecen.

Desde principios de los años ochenta, la industria y la literatura han investigado y perfeccionado esta estrategia. Se han sofisticado las técnicas econométricas y estadísticas para detectar estos pares de activos, la modelización de sus dependencias, así como las diferentes estrategias y metodologías con las que detectar oportunidades de trading.

En los últimos años, muchos trabajos han puesto en evidencia como esta estrategia cada vez es menos rentable y se generan menos oportunidades bajo enfoques como el de distancia y el de cointegración. Sin embargo, nuevos enfoques como el de cópulas o “*machine learning*” podrían permitir seguir aprovechándose de los desequilibrios surgidos ya que modelizan las dependencias entre ambos activos de una forma más precisa respecto a los enfoques anteriores.

La introducción de cópulas en el estudio de la estrategia de negociación por pares aporta la posibilidad de modelizar la dependencia entre dos activos de manera más precisa en comparación con las metodologías empleadas hasta el momento.

Además, es común pensar que en períodos de alta volatilidad se van a generar más oportunidades de tomar posiciones y se van a obtener rendimientos más elevados ya que las divergencias deberían ser más pronunciadas. En este trabajo se va a estudiar y analizar la estrategia en diferentes períodos para comprobar si esta afirmación está justificada o no.

El objetivo principal de este trabajo es responder a la pregunta de si el empleo de cópulas en una estrategia de negociación por pares permite obtener beneficios una vez descontados los costes de transacción. Así mismo, se va a responder a la pregunta de si estos beneficios,

en el caso de ser positivos, son consistentes y robustos si se divide la muestra en períodos diferentes y se aplica la estrategia a cada período. Una vez respondidas estas preguntas para datos diarios, se llevará a cabo un estudio con datos de alta frecuencia (datos de cada 30 minutos) y se procederá a comparar a los resultados de aplicar la estrategia para ambas frecuencias de datos en un mismo período de tiempo.

Una vez realizado este primer análisis, se estudiará el comportamiento de las oportunidades de trading en función de la volatilidad y se optimizará la estrategia clasificando dichas oportunidades en potencialmente beneficiosas o no.

Después de la introducción, donde se describe brevemente el trabajo realizado, la motivación del mismo y las preguntas que se pretenden responder, el trabajo se va a estructurar de la siguiente manera: en el segundo apartado se realiza un análisis de la literatura previa segmentado en función de las diferentes metodologías utilizadas hasta el momento para este tipo de estrategia. Después en el tercer apartado se lleva a cabo una descripción detallada de la metodología utilizada para el trabajo seguida del cuarto apartado donde se presentan en un primer punto los datos utilizados y sus estadísticos descriptivos para posteriormente analizar los resultados del análisis empírico realizado. Por último, se procederá a realizar la conclusión final en el quinto apartado.

2.LITERATURA PREVIA

2.1 MÉTODO DE LA DISTANCIA

Este método ha sido el más trabajado en la literatura, así como el que más se ha llevado a la práctica y, por tanto, cada vez es más difícil obtener rendimientos elevados con él.

Se toma como referencia el trabajo de Gatev et al. (1999, 2006) [11] y [12], cuyos autores fueron los primeros autores en desarrollar el método de la distancia y lo aplicaron al mercado americano.

Para llevarlo a cabo, lo primero que hacen es construir un índice de rendimiento total acumulado P_{it} para cada acción,

$$P_{it} = \prod_{s=1}^t (1 + r_{is})$$

Donde r_{is} representa el rendimiento del día s , teniendo en cuenta los dividendos.

Posteriormente, se normaliza este índice respecto al primer día de los 12 meses que componen el período de formación. A continuación, se calcula el SSD (la suma de la distancia Euclídea al cuadrado o “*Sum of squared deviation*”) para cada posible combinación de dos activos, así como su desviación típica, donde i y j representan activos diferentes y t representa el tiempo:

$$SSD = \sum_{t=1}^T (P_{it} - P_{jt})^2$$

Para un total de $n(n - 1)/2$ posibles combinaciones.

A partir de los resultados, se considera que los 20 pares con menor SSD son los pares adecuados para realizar el análisis ya que históricamente son los que presentan mayor relación entre sí. Definido el período de formación, se define el período de *trading*, donde 6 meses son considerados.

Durante este período de *trading*, cuando el *spread* diverge más de 2 desviaciones típicas, se crea una señal de entrada. La desviación del *spread* tiene la siguiente expresión,

$$\sqrt{V(P_{it} - P_{jt})} = \sqrt{\frac{1}{T} \sum_{t=1}^T (P_{it} - P_{jt})^2 - \left(\frac{1}{T} \sum_{t=1}^T (P_{it} - P_{jt})\right)^2}$$

donde V es la varianza muestral y T es el número total de días del período de formación.

Al tratarse de una estrategia de reversión a la media, se espera que este *spread* converja a la relación de “equilibrio” y así cerrar la posición abierta. Por lo que se abre una posición

larga en el activo relativamente infravalorado y una posición corta en el activo relativamente sobrevalorado. Hay que tener en cuenta que al no realizar un test de cointegración para los pares seleccionados es posible que dichos pares tengan relación espúrea, lo cual podría provocar que dicho *spread* no convergiera y generar de este modo un riesgo adicional.

En esta dirección, Do y Faff (2010) [5], utilizando los mismos datos y metodología que Gatev et al. (2006), demostraron que el 32% de los pares identificados con el SSD únicamente no convergen. Además, Huck (2015) [15] demuestra que los pares seleccionados en función de si están cointegrados o no, presentan mayor reversión a la media que los pares seleccionados en función de su distancia. Finalmente, Do y Faff (2010, 2012) [6] demostraron como esta metodología es cada vez menos beneficiosa, y una vez añadidos los costes por transacción, tiene resultados altamente negativos. Engelberg et al. (2009) [9] también demuestran como la rentabilidad de este método es cada vez menor con el paso del tiempo.

Otros autores han utilizado el trabajo de Gatev et al. (1999, 2006) como punto de partida y a partir de ahí han introducido variaciones como es el caso de Chen et al. (2012) [4], donde dichos autores desarrollan el método de la distancia, pero en este caso se decantan por utilizar el índice de correlación lineal de Pearson sobre los rendimientos para identificar los pares.

Los resultados de aplicar este método en el caso de Gatev et al. (2006) son positivos, hasta un 11% anual, si bien es cierto que una vez ajustados estos resultados por riesgo, y ver que las rentabilidades de diferentes pares presentan alta correlación, concluyen que se podría estar recompensando otro riesgo que no se está teniendo en cuenta.

Más adelante, Papadakis y Wysocki (2007) [23], siguiendo el método de Gatev et al. (2006) encontraron que las señales de entrada están altamente relacionadas con las publicaciones de resultados financieros y con las publicaciones de las previsiones de los analistas. Las operaciones inmediatamente realizadas después de ambos eventos son considerablemente menos beneficiosas que las operaciones en un entorno sin este tipo de acontecimiento.

Un resultado significativamente importante se encuentra en el trabajo que realizaron Chen et al. (2012) donde se confirma que los rendimientos obtenidos gracias a esta estrategia se deben a un desfase temporal en cuanto a la absorción de información de cada activo del par se refiere. Ya que se percataron de que en entornos donde la información no fluye con facilidad los rendimientos son mayores. Más tarde, Jacobs y Weber (2013) [16] usando una variante del algoritmo de Gatev et al. (2006) demuestran que efectivamente, una de las causas que generan estos beneficios, son los desfases en la difusión de la información.

Recientemente, Huck (2015) estudia el impacto de la volatilidad a la hora de llevar a cabo la negociación por pares. Después de analizar la estrategia en el S&P 500 y Nikkei 225, llega a la conclusión de que los rendimientos no se pueden ver incrementados una vez se introduce un índice de volatilidad como el VIX en el análisis.

Es interesante destacar que Nath (2003) realiza un primer estudio de esta estrategia en un entorno de datos de alta frecuencia para el mercado secundario de deuda de Estados Unidos, en este caso se toman 10 cotizaciones dentro de un mismo día para cada valor y se consideran 40 días como período de formación, para examinar la estrategia fijan de igual modo 40 días de período de negociación. Este autor concluye que su estrategia ofrece buenos resultados en cuanto a ratios de *Sharpe* y Ganancia-Pérdida se refiere.

Aunque en otro trabajo con datos intradiarios, Bowen et al. (2010) [3], al estudiar el comportamiento de la estrategia de Gatev et al. (2006) en el FTSE 100 para datos de cada 60 minutos, concluyen que después de costes de transacción y asumiendo costes de horquilla, los beneficios desaparecen.

2.2 MÉTODO DE COINTEGRACIÓN

Este método se caracteriza por identificar pares que presentan una alta relación a partir de un test de cointegración como el de Engle-Granger o Johansen.

Dos variables se considera que están cointegradas cuando comparten una tendencia estocástica en común. De este modo, si se tienen dos variables como pueden ser dos series financieras, no estacionarias y cuyo máximo orden de integración es $I(q)$, se dice que ambas variables están cointegradas si se puede crear una combinación lineal de las mismas de manera que se obtenga un orden de integración $I(p)$ con $p < q$.

Para este método, se tiene como referencia el trabajo realizado por Vidyamurthy (2004) [27]. En dicho trabajo, donde se trabaja de una forma teórica y no empírica, el autor procede de la siguiente manera: primero preselecciona los pares potencialmente cointegrados a partir de información fundamental y estadística, segundo estudia las posibilidades que permite esta metodología y tercero, implementa una estrategia de negociación basada en métodos no paramétricos. Do et al. (2006) [7] y Puspaningrum (2012) [24] han realizado trabajos similares a Vidyamurthy.

Para la preselección de los pares, se lleva a cabo una descomposición de los precios logarítmicos, p_{it} , en una parte no estacionaria, n_{it} y en una parte estacionaria, ϵ_{it} . De igual modo se procede a separar el rendimiento r_{it} en una parte común, r_{it}^c y una parte específica, r_{it}^s .

Una vez hecho esto, se construye una cartera con una posición larga en una acción del activo i y corta en γ acciones de j ,

$$m_{ijt} = p_{it} - \gamma p_{jt} = n_{it} - \gamma n_{jt} + \epsilon_{it} - \gamma \epsilon_{jt}$$

$$\Delta m_{ijt} = r_{ijt} = r_{it}^c - \gamma r_{jt}^c + r_{it}^s - \gamma r_{jt}^s$$

Donde m_{ijt} representa la evolución del precio de la cartera y Δm_{ijt} su rentabilidad.

Entonces, se busca que estas rentabilidades comunes se cancelen de manera que el spread resultante sea estacionario. En el trabajo llevado a cabo por Vidyamurthy, se utiliza el modelo APT de Ross (1976) [25] para identificar acciones con similares r_{it}^c , con este fin se apoya de igual modo en el índice de correlación lineal de Pearson, creando un ranking de menor a mayor de los pares de acciones y su índice de correlación de estos factores de rentabilidades comunes, donde los pares con mayor correlación, son los que se consideran que tienen una mayor probabilidad de estar cointegrados.

Chen et al. (2012) demuestra de manera empírica como una estrategia donde los pares de activos son seleccionados en función de la correlación de sus factores comunes genera mejores resultados que si se seleccionan en función de la correlación residual.

Una vez se tienen identificados los pares, se lleva a cabo mediante OLS la regresión,

$$p_{it} = \mu + \gamma p_{jt} + \epsilon_{ijt} \quad \epsilon_{ijt} \sim N(0, \sigma_c^2)$$

Donde p_{it} y p_{jt} representan los precios logarítmicos, μ es la prima pagada por mantener la acción i respecto la acción j y γ es el coeficiente de cointegración.

Una vez resuelta la regresión, se tiene que

$$p_{it} - \hat{\gamma} p_{jt} = \hat{\mu} + \hat{\epsilon}_{ijt}$$

A partir de aquí, lo habitual sería contrastar si $\hat{\epsilon}_{ijt}$ es estacionario con el test de Dickey-Fuller (ADF), es decir si es $I(0)$. Si resulta que es estacionario entonces se puede decir que se ha encontrado una combinación lineal de dos series $I(1)$ no estacionarias que están cointegradas.

Cabe resaltar la importancia de que variable seleccionamos como explicativa y que variable como explicada, ya que los residuos resultantes serán diferentes.

Vidyamurthy sugiere estudiar si existe reversión a la media en el *spread* o no, donde

$$[\widehat{Spread}_t] = E[p_{it} - \hat{\gamma} p_{jt}] = E[\hat{\mu} + \hat{\epsilon}_{ijt}]$$

Y dada la propiedad $E[\hat{\epsilon}_{ijt}] = 0$ se tiene que $\hat{\mu}$ será el valor de “equilibrio” a largo plazo tomado como referencia.

Por último, se generan oportunidades de entrada de una manera no paramétrica similar a Gatev et al. (1999), en función de cuando $\hat{\mu}$ diverge n desviaciones típicas. También se lleva a cabo una modelización paramétrica con un ARMA de los residuos $\hat{\epsilon}_{ijt}$ y después de aplicar la fórmula de Rice (1945) se calcula el valor óptimo de divergencia que maximiza los beneficios.

Lin et al. (2006) llevaron a cabo un enfoque que incluye una condición de beneficios mínimos, por la cual mediante una regresión similar a la llevada a cabo por Vidyamurthy pero sacando de la ecuación $\hat{\mu}$,

$$p_{it} + \hat{\gamma}p_{jt} = \hat{\epsilon}_{ijt}$$

Se tiene que el beneficio total por cada *trade* es,

$$TP_{ijt_c} = \frac{n}{c} [(\epsilon_{ijt_c} - P_{it_c}) - (\epsilon_{ijt_0} - P_{it_0})] + \frac{n}{|\gamma|} [P_{it_0} - P_{it_c}] = \frac{n(\epsilon_{ijt_0} - \epsilon_{ijt_c})}{|\gamma|} > K$$

Donde t_0 representa cuando se abre posición y t_c cuando se cierra.

2.3 MÉTODO DE SERIES TEMPORALES

La finalidad de este método es modelizar la reversión a la media del *spread* presente en los pares de acciones altamente correlacionados con el fin de maximizar beneficios.

El trabajo de referencia en esta metodología es el que realizaron Elliott et al. (2005) [8]. Aportaciones posteriores son las hechas por Do et al. (2006) y Herlemont (2008). En él se describe el *spread* como la diferencia de precios a partir de una cadena de Markov Gaussiana con reversión a la media, observado en un proceso que sigue una distribución normal.

Se parte de una ecuación de estado, donde la variable de estado x_k sigue un proceso de reversión a la media tal que,

$$x_{k+1} - x_k = (a - bx_k)\tau + \sigma\sqrt{\tau}\epsilon_{k+1} \quad \epsilon_k \sim N(0,1) \text{ iid}$$

Con media $\mu = a/b$, $\sigma, a, b > 0$. Esta ecuación de estado se expresa en tiempo continuo mediante el proceso de Ornstein-Uhlenbeck de la siguiente forma,

$$dx_t = b(\mu - x_t)dt + \sigma dW_t$$

Donde μ representa la media, dW_t es un movimiento Browniano estándar y b indica la velocidad de reversión a la media.

Una vez definida la ecuación de estado, se describe la ecuación de medida, de este modo tenemos el *spread* que resulta ser,

$$y_k = x_k + Dw_k \quad D > 0, w_t \sim N(0,1) \text{ iid}$$

A partir de aquí, se producen señales de entrada cuando,

$$y_k \geq \mu + c\left(\frac{\sigma}{\sqrt{2b}}\right) \text{ ó } y_k \leq \mu - c\left(\frac{\sigma}{\sqrt{2b}}\right)$$

Donde c representa un parámetro fijo. La posición se cierra una vez llegado a T , el tiempo necesario para que el proceso revierta a la media μ , obtenido en el proceso de Ornstein-Uhlenbeck mencionado anteriormente, cuya expresión es,

$$T = \inf\left\{t \geq 0, x_t = \mu \left| x_0 = \mu + c \left(\frac{\sigma}{\sqrt{2b}} \right) \right.\right\} = \hat{t}b$$

$$\hat{t} = 0.5 \ln\{1 + 0.5[\sqrt{(c^2 - 3)^2 + 4C^2} + C^2 - 3]\}$$

Durante dicho proceso, se utiliza el filtro de *Kalman* para estimar los parámetros.

Al tratarse de una modelización directa del *spread*, este método tiene la ventaja de que una vez estimados los parámetros, se pueden realizar estudios sobre expectativas y rentabilidades esperadas al llevar a cabo una estrategia de negociación por pares. En Do et al. (2006) se lleva a cabo un análisis de ventajas y desventajas que pueden surgir de este modelo, como por ejemplo el trabajar con precios directamente en lugar de precios logarítmicos. En el trabajo de Do et al. (2006) se propone trabajar con rentabilidades en lugar de con precios, lo cual conlleva a una ecuación de estado y de medida diferentes.

Recientemente en Bertram (2010) [2] se lleva a cabo una estrategia de arbitraje estadístico donde se parte del modelo en tiempo continuo derivado de un proceso Ornstein-Uhlenbeck para el *spread* x_t ,

$$dx_t = -kx_t dt + \sigma dW_t$$

Una vez definida la ecuación de estado, se tiene que el ciclo de trading se puede dividir en T_1 y T_2 donde el primer período es el tiempo entre la apertura y el cierre de una posición y el segundo período es el tiempo entre esta salida y la siguiente entrada. El objetivo que se tiene en este modelo es obtener los umbrales de entrada y de salida óptimos (los que maximicen el beneficio).

Es importante tener en cuenta que una desventaja de este modelo es que, al trabajar con series financieras, muy posiblemente los datos no sigan una distribución normal por lo que esto podría distorsionar los resultados.

2.4 MÉTODO DE CONTROL ESTOCÁSTICO

Este método se centra en encontrar la inversión óptima en cada uno de los activos que componen un par cuando hay otros activos en los que invertir disponibles.

El trabajo de referencia es el realizado por Jurek y Yang (2007) [17], y su objetivo es diseñar la estrategia de inversión óptima para pares de activos que tengan *spreads* con reversión a la media a través de un proceso Ornstein-Uhlenbeck. Los autores se plantean dos escenarios en cuanto a las preferencias de los individuos, por un lado, asumen que son adversos al riesgo de forma permanente y por otro lado que siguen una función de utilidad del tipo Epstein-Zin. A partir de la teoría de valoración de activos y teniendo en cuenta las restricciones en cuanto a capital disponible se refiere, aplican la teoría del control estocástico para obtener la ecuación Hamilton-Jacobi-Bellmann (HJB) a partir de la cual se hallan las ecuaciones de valoración.

En cuanto a los resultados, concluyen que se obtienen mejores resultados que en Gatev et al. (2006) en términos de rentabilidades absolutas y ratio de *Sharpe* si se analizan pares con *spreads* que tengan una rápida reversión a la media.

En la misma línea que Jurek y Yang (2007), Liu y Timmermann (2013) [21] realizan un análisis sobre los valores óptimos a tener en cartera a partir de un análisis de cointegración para la dinámica que siguen los precios, dicha dinámica sigue un paseo aleatorio geométrico tal que,

$$\frac{dP_{mt}}{P_{mt}} = (r + \mu_m)dt + \sigma_m dB_t$$

donde P_{mt} es el índice de mercado, μ_m es la prima constante del riesgo de mercado, σ_m la volatilidad, B_t es un movimiento browniano y r es la tasa libre de riesgo.

Una vez definida la dinámica que sigue el índice, se consideran dos activos con las siguientes dinámicas,

$$\begin{aligned} \frac{dP_{1t}}{P_{1t}} &= (r + \beta\mu_m)dt + \beta\sigma_m dB_t + \sigma dZ_t + bd - \lambda_1 x_t dt \\ \frac{dP_{2t}}{P_{2t}} &= (r + \beta\mu_m)dt + \beta\sigma_m dB_t + \sigma dZ_t + bdZ_{1t}Z_{2t} - \lambda_2 x_t dt \end{aligned}$$

$$x_t = \ln(P_{1t}) - \ln(P_{2t})$$

donde x_t es el término de error y es estacionario, Z_t , Z_{1t} y Z_{2t} son movimientos Brownianos independientes entre sí, λ_1 , λ_2 , β , b y σ son constantes y P_{1t} , P_{2t} representan los precios y sus logaritmos están cointegrados.

Una vez se tienen las dinámicas, el objetivo de los autores es encontrar los pesos óptimos a asignar a cada uno de los activos disponibles (2 activos y el índice).

Las conclusiones de este trabajo son relevantes ya que sugieren que a pesar de que el *spread* de precios haya convergido, es posible que sea óptimo mantener posiciones largas o cortas en ambos activos y también puede ser óptimo de igual modo mantener uno de los activos únicamente. Este resultado es contradictorio con el clásico enfoque de negociación por pares donde se forma una estrategia delta neutral.

2.5 OTROS ENFOQUES

Dentro de este grupo se encuentran todo el resto de enfoques mediante los cuales se llevan a cabo estrategias de negociación por pares. Se podrían englobar en 3 subgrupos, de reciente investigación, en un primer grupo se tiene todo lo relacionado con el *machine learning* o aprendizaje automático que pueda tener una aplicación dentro de este tipo de estrategias. El segundo grupo sería el de componentes principales. Y ya por último el

tercer enfoque haría referencia a un enfoque por cópulas, metodología analizada en este trabajo.

2.5.1 MACHINE LEARNING

El trabajo más relevante en cuanto a metodologías de aprendizaje automático se refiere es el realizado por Huck (2009) [13] y Huck (2010) [14]. En este estudio se abandona el concepto de “modelo de equilibrio” que estaba presente en los enfoques previos donde las metodologías estaban basadas en la reversión a la media en un entorno en el que se asumía un “equilibrio” entre ambos activos de los pares formados.

En el trabajo de Huck se parte de un número total de activos n a partir de los cuales se forman los pares, para ello utiliza la técnica de redes neuronales de Elman para generar predicciones de rentabilidad a una semana para cada activo,

$$\hat{x}_{|X_{i,t}, X_{j,t}}^{i,t+1}$$

Estas predicciones están condicionadas a las rentabilidades pasadas tanto de i como de j . Realizando esto de forma recursiva se llega a que para cada activo tenemos un total de $n-1$ predicciones para cada período.

A continuación, Huck aplica el método llamado Electre III, el objetivo de esta técnica es ordenar una serie de alternativas a partir de una serie de criterios. En un trabajo como el llevado a cabo por Huck, el número de activos representa las alternativas, así como los criterios diferentes a tener en cuenta.

Entonces se modeliza el *spread* de forma que,

$$x_{ij} = \hat{x}_{|X_{i,t}, X_{j,t}}^{i,t+1} - \hat{x}_{|X_{i,t}, X_{j,t}}^{j,t+1}$$

Donde x_{ij} es el *spread* anticipado. Una vez modelizados los *spreads*, mediante el método Electre III y a partir de unos umbrales fijados previamente, se lleva a cabo el proceso de detección de activos infra y sobrevalorados a partir de los cuales se toman posiciones durante una semana que es el período de *trading*, una vez finalizada esta semana se repite el procedimiento.

2.5.2 COMPONENTES PRINCIPALES

Avellaneda y Lee (2010) [1] son los autores referencia en este caso. En dicho trabajo se llevan a cabo 2 formas diferentes de descomponer las rentabilidades de los activos en el componente idiosincrático y el sistemático.

La primera alternativa que sugieren es realizar una regresión tal que,

$$R_i = \beta_i F + \epsilon_i$$

donde R_i es la rentabilidad del activo i , $\beta_i F$ representa el componente sistemático (F es el rendimiento del ETF correspondiente al sector pertinente) y ϵ_i representa el componente idiosincrático.

La segunda alternativa es un modelo multifactorial donde hay m factores tal que,

$$R_i = \sum_{j=1}^m \beta_{ij} F_j + \epsilon_i$$

A partir del modelo multifactorial, se asume que las rentabilidades de las acciones cumplen la siguiente ecuación diferencial,

$$\frac{dP_{it}}{P_{it}} = \mu_i dt + \sum_{j=1}^m \beta_{ij} \frac{dI_{jt}}{I_{jt}} + dX_{it}$$

donde X_{it} sigue un proceso Ornstein-Uhlenbeck y μ_i es la deriva del precio de la acción.

Entonces se aplica una estrategia similar a Elliott et al. (2005) y se concluye que a pesar de obtener buenos resultados en términos de ratio de Sharpe, dichos resultados carecen de robustez. También sugieren que sería conveniente introducir el estudio de cointegración para completar el análisis.

2.5.3 CÓPULAS

Los principales trabajos bajo este enfoque son Ferreira (2008) [10], Liew y Wu (2013) [19], Stander et al. (2013) [26], Wu et al. (2014) [28], Krauss y Stübinger (2015) [18].

La motivación de utilizar cópulas para modelizar las dependencias entre dos activos surge debido a que los activos financieros tienen dependencias no lineales y posiblemente asimétricas. La asimetría es debido a que en ocasiones la correlación entre dos activos no es la misma cuando estos tienen rentabilidades positivas que cuando tienen rentabilidades negativas, precisamente en este último contexto, suelen presentar mayor correlación.

Las cópulas son una herramienta de mucha utilidad y su uso se ha extendido de forma considerable en la gestión de riesgos como se puede ver por ejemplo en McNeil, Frey y Embrechts (2005) y su aplicación en el trabajo sobre riesgo de crédito, de mercado y operacional o en Andersen, Bollerslev, Christoffersen y Diebold (2007) en el que se lleva a cabo una aplicación en la modelización del riesgo de mercado.

La forma de proceder empieza con la selección de pares con alta relación entre sí. En Stander et al. (2013) se lleva a cabo una descripción de las diferentes metodologías para identificar pares con una dependencia potencialmente elevada. Por un lado, se tiene el análisis fundamental, mediante el cual se buscan dependencias a partir de los libros y resultados contables históricos de la empresa, lo cual conlleva un elevado coste de tiempo.

Sin embargo, la forma más habitual de proceder es realizar un análisis técnico/estadístico con el fin de encontrar activos que presenten por ejemplo una elevada correlación histórica o un bajo valor en el SSD, como se hace en Stander et al. (2013) y en Wu et al. (2014).

Una vez seleccionados los pares, la muestra se separa en dos partes, a la primera parte se le conoce como “*período de formación*”, en donde se estiman las distribuciones marginales, así como los parámetros de la cópula, mientras que la segunda parte de la muestra es el “*período de trading*”, en el que se pone en práctica la cópula estimada con el fin de obtener resultados.

En el período de formación se calculan en primera instancia las rentabilidades logarítmicas y a continuación se fijan las distribuciones marginales que encajan mejor para cada activo, con las cuales se llevará a cabo la transformación integral de probabilidad y se obtendrá de esta forma el vector de uniformes para ambos activos, u y v , que serán el input de la cópula. Una vez se tienen las distribuciones marginales, se obtienen los parámetros de la cópula que recoge de manera más precisa la relación entre ambos activos. En Wu et al. (2014) se analizan 5 tipos de cópulas (Gumbel, Frank, Clayton, Normal y Student-T) y se ordenan en función de su valor de verosimilitud de mayor a menor. Dichos autores justifican el no incluir más cópulas razonando que con estas 5 se recogen todas las posibles estructuras de dependencia.

En Liew, Wu (2013) se realiza un ranking para estas 5 cópulas a partir de más criterios de clasificación como SIC, Akaike Information Criterion (AIC) y Hannan-Quinn Information Criterion (HQIC). Una vez analizadas las diferentes cópulas concluyen que la cópula Gumbel es la que recoge de mejor manera las dependencias entre ambos activos.

Una vez se tienen los parámetros de la cópula, se derivan las probabilidades condicionales tal que,

$$P(U \leq u|V = v) = \frac{\partial C(u, v)}{\partial v}; \quad P(V \leq v|U = u) = \frac{\partial C(u, v)}{\partial u}$$

Y se definen los umbrales a partir de los cuales se tomarán posiciones. Según ambas probabilidades condicionales, se considera que ambos activos se encuentran en “equilibrio” cuando las dos probabilidades son del 0.5. En Stander et al. (2013) se decantan por un nivel de confianza tal que $P(U \leq u|V = v) = 0.99$ y $P(V \leq v|U = u) = 0.01$ a partir del cual tomarían una posición corta en el activo que se identifica con el vector u y larga en el activo identificado con el vector v . En los próximos capítulos se justificará de manera más profunda y técnica este desarrollo y su interpretación.

En Wu et al. (2014) se calculan las señales de entrada del siguiente modo. Primero se definen “FlagX” y “FlagY” en donde se van acumulando las probabilidades condicionales

$$P(U \leq u|V = v) - 0.5 \text{ y } P(V \leq v|U = u) - 0.5$$

respectivamente en cada período. Entonces se fijan los umbrales D y S. Cuando se alcanza “D” o “-D” se abren posiciones y se cierran al alcanzar “S” o “-S”, que se identifica con un “Stop-Loss” o bien cuando “FlagX” o “FlagY” vuelven a 0. Los autores fijan estos valores en $D=0.6$ y $S=2$.

Wu et al. (2014) realizan el análisis en 17 períodos de formación/*trading* diferentes en los cuales se consideran para cada uno un total de 89 activos posibles con los que realizar combinaciones. Cada período de formación se compone de 12 meses, 252 días de cotización, mientras que el período de trading está compuesto por 6 meses, 126 días de cotización.

A partir de este punto, en Stander et al. (2013) se discute el número de acciones a comprar y vender en cada señal de entrada. Los autores sugieren la posibilidad de utilizar el ratio de *Sharpe* como indicador, pero advierten de una inestabilidad temporal como principal inconveniente. Finalmente, ellos se decantan por utilizar,

$$drift_x = \frac{1}{n} \sum_{j=2}^n (P_{X,j} - P_{X,j-1})$$

$$drift_y = \frac{1}{n} \sum_{j=2}^n (P_{Y,j} - P_{Y,j-1})$$

Y para cada unidad comprada del activo X, $\frac{|drift_x|}{|drift_y|}$ unidades del activo Y son vendidas.

En cuanto a los costes de transacción, en Stander et al. (2013) se detallan los costes por operar en Suráfrica en el JSE (“Johannesburg Stock Exchange”), donde según se cita en Sanlam (2012) son la cuota mensual de mantenimiento de la cuenta, cuota fija sobre el nominal de la operación, STT (“Securities Transfer Tax”) de 0,25%, un impuesto de 0,0002% sobre todos los *trades*, un VAT (“Value added tax”) de 14%.

En Wu et al. (2014) se mencionan dos posibles costes de transacción adicionales, por una parte, se tiene el coste relacionado con tener abiertas posiciones en corto y por último y más relevante para los autores, el coste por horquilla (bid/ask) al realizar operaciones.

En Krauss y Stübinger (2015) se consideran 10 bps por transacción, ya que argumentan que a pesar de que Do y Faff (2012) ya habían considerado que los costes de transacción habían disminuido considerablemente recientemente, pasando de 16 bps a 7-9 bps por transacción, prefieren ser conservadores.

Los principales resultados y conclusiones presentes en la literatura previa al aplicar una estrategia de negociación por pares en un entorno de modelización con cópulas indican que la estrategia es capaz de producir beneficios. Wu et al. (2014) concluyen que, dado el descenso de oportunidades de realizar negociación por pares bajo el método de la distancia, introducir cópulas a la estrategia mejora los resultados considerablemente tanto en el número de oportunidades como en sus rendimientos. En este sentido, Stander et al.

(2013) sugieren que dado las dependencias no lineales que se producen entre los activos de un par, el introducir cópulas es el camino a seguir. En función del nivel de confianza elegido y en especial en función de si hay entornos de elevada volatilidad, afirman que hay un número considerable de oportunidades para llevar a cabo la estrategia, sin embargo, a pesar de que las operaciones generan rendimientos positivos, los costes de transacción compensan esas ganancias. Pero en Krauss y Stübinger (2015) se obtiene un 7.98% de rentabilidad anual en los 5 mejores pares según el criterio de reversión a la media y un 7.22% anual en los 5 mejores pares elegidos según el *momentum*. Resaltan que estos resultados son robustos una vez aplicados costes de transacción, así como si se aplica la regla de trading de cerrar la posición un día después de cuando se produce la señal de salida.

3.METODOLOGÍA

3.1 INTRODUCCIÓN Y SELECCIÓN DE PARES

Como ya se ha comentado anteriormente, la motivación de introducir cópulas en la estrategia de negociación por pares encuentra su razón de ser en que habitualmente las distribuciones marginales de las rentabilidades de los activos financieros son asimétricas y la dependencia entre ellos no es lineal. Modelizando la dependencia entre dos activos mediante cópulas, es posible construir una distribución de probabilidad conjunta que describa la dependencia entre dos series de rentabilidades de manera más precisa, solventando este tipo de inconvenientes.

Es importante resaltar que, en un contexto de este tipo, se tiene que interpretar con cautela el coeficiente de correlación lineal ya que su cálculo es justificado cuando las variables aleatorias siguen una distribución elíptica¹.

El procedimiento a seguir será el siguiente: la muestra total de datos será dividida en dos partes, en la primera parte, llamada período de formación, se seleccionarán los pares y se estimarán los parámetros de la cópula. A continuación, en la segunda parte de la muestra, el conocido como período de negociación o período de *trading*, se pondrá en práctica la estrategia en todos los pares seleccionados a partir de los parámetros estimados en el período previo.

Lo primero a tener en cuenta antes de modelizar la dependencia entre dos activos mediante cópulas es el procedimiento a realizar para seleccionar dicho par. Stander et al. (2013) describe los dos posibles caminos a seguir como se ha comentado en el punto anterior y lleva a cabo la selección en función del coeficiente de correlación lineal, al igual que en Wu y Liew (2013). En este trabajo se tienen en cuenta el índice de correlación lineal² y el SSD.

Una vez se han calculado ambos indicadores para todos los posibles pares de activos, se eligen aquellos que tengan un SSD más pequeño y un índice de correlación lineal mayor. Los pares seleccionados son los que potencialmente presentan una elevada relación, y se procede a modelizar esta dependencia mediante cópulas.

¹ Una distribución elíptica se caracteriza por tener líneas de equiprobabilidad elípticas, como sucede en la Normal o la t-Student multivariantes.

² Al utilizar una muestra considerablemente grande en el cálculo del índice de correlación lineal entre ambos activos, se considera que dicho índice identifica activos que se haya movido de forma conjunta a largo de la muestra analizada.

3.2. MODELIZACIÓN DE LA DEPENDENCIA MEDIANTE CÓPULAS

A partir de las series de precios, se calculan las rentabilidades logarítmicas. A continuación, es importante destacar que, para la modelización mediante cópulas, es preciso tener una serie de uniformes (0,1) y no una serie de rentabilidades. Por este motivo, es preciso ajustar la muestra de datos a la distribución que mejor se adapte a la muestra de cada serie para realizar posteriormente la transformación integral de probabilidad y obtener de esta forma los vectores de uniformes para cada activo del par que posteriormente serán el input de la cópula.

Es relevante recordar que dado que la distribución $U(0,1)$ tiene una función de distribución lineal se tiene que,

$$P(F(X) < u) = P(X < F^{-1}(u)) = F(F^{-1}(u)) = u$$

Por lo que se demuestra que para cualquier variable aleatoria X , como puede ser una serie de rentabilidades en este caso, su función de distribución define otra variable aleatoria, $F(X)$, con distribución $U(0,1)$. De la misma forma se tendría,

$$P(F^{-1}(U) < x) = P(U < F(x)) = F(x)$$

Entonces se definen R_1 y R_2 como las rentabilidades logarítmicas, que en este caso son las variables aleatorias con funciones de distribución marginales continuas, a partir de las cuales se tienen las uniformes $u = F_1(r_1)$ y $v = F_2(r_2)$.

A la hora de determinar las funciones de distribución marginales es posible optar por un enfoque no paramétrico como se sugiere en Stander et al. (2013) y en Krauss y Stübinger (2015), donde la transformación integral de probabilidad se realiza asumiendo una distribución empírica para la muestra, o bien es posible realizar un enfoque paramétrico, como se plantea en Ferreira (2008) o en Wu y Liew (2013).

En este trabajo se realiza un enfoque paramétrico para determinar las distribuciones marginales, mediante el cual se estudian un total de 17 posibles distribuciones diferentes y se elige la distribución óptima que mejor se adapta a cada muestra en función del valor del BIC (Bayesian information criterion), otros posibles criterios de selección pueden ser NLogl (Negative of the log likelihood), AIC (Akaike information criterion) o bien AICc (AIC con corrección para muestras finitas).

Una vez se tienen las distribuciones marginales, comienza el proceso de modelizar la dependencia a través de la cópula.

Se define cópula como la función de distribución bivalente definida sobre el cuadrado $[0,1] \times [0,1]$ con distribuciones marginales que son Uniformes (0,1), con propiedades tal que:

1. $C: [0,1] \times [0,1] \rightarrow [0,1]$
2. $C(u, 1) = u; C(1, v) = v$

$$3. C(u, 0) = C(0, v) = 0$$

4. Para todo u_1, u_2, v_1, v_2 con $u_1 \leq u_2, v_1 \leq v_2$ se tiene:

$$C(u_2, v_2) - C(u_1, v_2) \geq C(u_2, v_1) - C(u_1, v_1)$$

De trascendental importancia fue el “Teorema de Sklar” (1959), el cual dice que dada una función de distribución bivalente $F(r_1, r_2)$ con distribuciones marginales F_1 y F_2 , existe una cópula $C: [0,1] \times [0,1] \rightarrow [0,1]$ tal que $\forall r_1, r_2$ en R^2 se tiene:

$$\forall (r_1, r_2) \in R^2 \rightarrow F(r_1, r_2) = C(F_1(r_1), F_2(r_2))$$

Se dice que C es la cópula de la distribución F y si las distribuciones marginales son continuas, es única.

De esta forma, es posible generar una distribución conjunta a partir de las marginales mediante la cópula.

3.3 TIPOLOGÍAS DE CÓPULAS

El siguiente paso es determinar el abanico de posibles cópulas a tener en consideración para determinar cuál es la óptima en cada caso.

En la literatura hay diferentes opiniones al respecto, Stander et al. (2013) considera un total de 22 tipos de cópulas y mediante una prueba de Kolmogorov-Smirnov se selecciona la cópula óptima. En Wu y Liew (2013) estudian 5 diferentes tipos de cópulas y las clasifican según los valores en AIC, BIC y HQIC. Mientras que en Krauss y Stübinger (2015) se lleva a cabo la selección ordenando las cópulas en función de su valor en el test “Cramér-von Mises” así como en función de los valores AIC y BIC.

Siguiendo Wu y Liew (2013), en este trabajo se van a considerar un total de 5 cópulas diferentes. Dos de ellas no Arquimedianas, la Gaussiana y la t-Student y 3 cópulas Arquimedianas, Gumbel, Clayton y Frank. Si bien es cierto que para completar el análisis es posible incluir más tipologías de cópulas, el elevado coste computacional y el hecho de que con estas 5 cópulas se recoge prácticamente toda la estructura de dependencia que pueda existir entre 2 activos, no se considera necesario ampliar el número de cópulas considerado por el momento.

Cópulas no Arquimedianas

La cópula Gaussiana se caracteriza por ser simétrica, $C(u, v) = C(v, u)$ y presenta una dependencia en las colas nula o muy baja, a no ser que la correlación entre ambas series esté próxima a 1. Por estos dos motivos y dada la asimetría comúnmente observada en las series financieras, esta cópula es posible que no figure como una de las candidatas a modelizar la dependencia entre 2 activos financieros.

Sin embargo, es incluida en el análisis debido a que, en ocasiones, en este tipo de estrategia se pueden mantener operaciones abiertas durante espacios temporales relativamente largos. Hay que recordar que las series financieras siguen una agregación gaussiana, por lo que con datos de menor frecuencia se distribuyen como una normal.

Ejemplo de una cópula Gaussiana para diferentes grados de correlación entre ambos activos³:

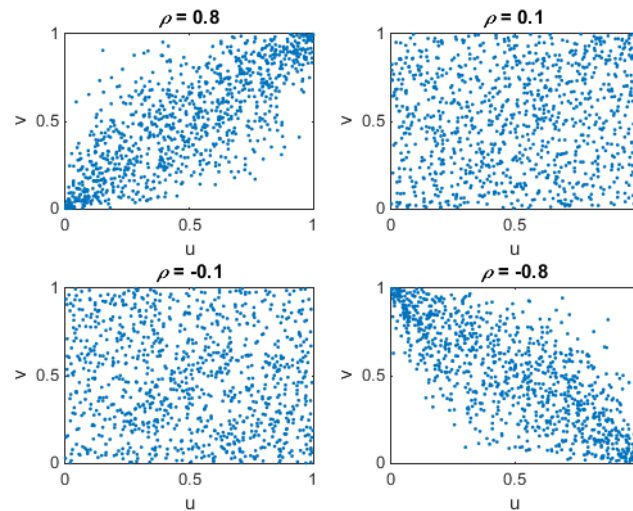


Figura 1. Estructura de dependencia de una cópula Gaussiana para diferentes niveles de correlación.

La otra cópula no Arquimediana es la t-Student, modelizando mediante esta cópula, se acentúa la dependencia en las colas por lo que es más adecuada cuando se trabaja con series financieras.

Ejemplo de una cópula t-Student para diferentes grados de correlación entre ambos activos:

³ Cada ejemplo se ha realizado a partir de 1000 números Uniformes (0,1) generados de forma aleatoria para u y v respectivamente asumiendo un determinado grado de correlación.

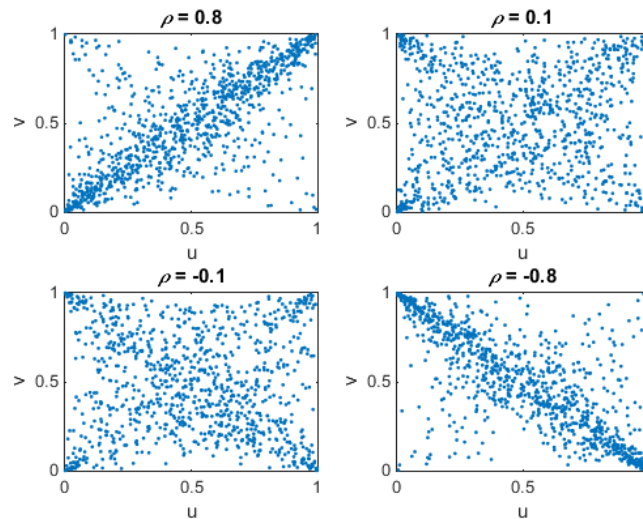


Figura 2. Estructura de dependencia de una cópula t -Student para diferentes niveles de correlación.

Ambos tipos de cópulas son conocidas como “cópulas implícitas”, ya que sus densidades se obtienen mediante el “método de inversión”.

Cópulas Arquimedianas

Dentro del grupo de las Arquimedianas se pueden encontrar alrededor de 22 tipos diferentes de cópulas. En este trabajo se utilizan 3 de ellas, Clayton, Gumbel y Frank.

El elemento diferenciador más relevante respecto las no Arquimedianas es el uso de la función generatriz, $\varphi(u)$, para la obtención de sus densidades de cópula.

La principal ventaja de las cópulas Arquimedianas es que permiten recoger distintos grados de dependencia, pudiendo llegar a modelizar una dependencia asimétrica en las colas. Esto permite mucha flexibilidad y junto con las no Arquimedianas descritas anteriormente abarcan las diferentes estructuras de dependencia que pueden llegar a tener las series financieras.

Ejemplo de cópulas Clayton, Frank y Gumbel para una correlación de 0.8:

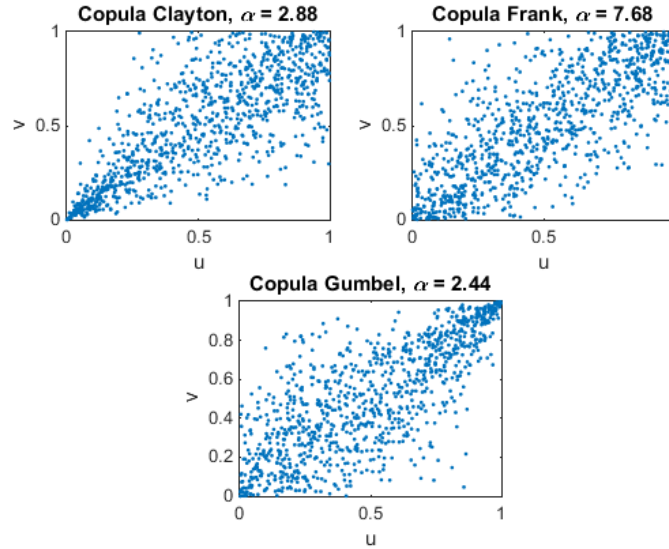


Figura 3. Estructura de dependencia de 3 cópulas Arquimedianas.

Una vez conocidas las diferentes posibles cópulas, es preciso determinar cuál de ellas es la que mejor se ajusta al par de activos en cuestión. Para ello se ordenan en función de los valores máximos de la función de verosimilitud, obteniendo de este modo la cópula óptima. Otro criterio posible sería elegir la cópula que presente menores valores en el criterio de información de Akaike (AIC) o en el criterio Bayesiano de información (BIC), los cuales también se tienen en cuenta y tienen la siguiente expresión,

$$AIC = 2k - 2\ln L$$

$$BIC = T^{-1}(k\ln T - 2\ln L)$$

3.4 PROBABILIDADES CONDICIONALES Y ESTRATEGIA DE NEGOCIACIÓN

En este punto se detalla el procedimiento a seguir a la hora de aplicar la estrategia de negociación en la parte de la muestra correspondiente al período de trading.

Como punto de partida, es necesario introducir el concepto de cópulas condicionales. Una vez se tiene la cópula estimada, derivando parcialmente⁴ se tiene,

$$P(U < u|V = v) = \frac{\partial C(u, v)}{\partial v}$$

$$P(V < v|U = u) = \frac{\partial C(u, v)}{\partial u}$$

A la hora de interpretar estas probabilidades condicionales se tiene que, como se introduce en Ferreira (2008), cuando $P(U < u|V = v) > 0.5$ se considera que el activo u está sobrevalorado respecto v y se espera que el activo, en este caso la acción, pierda valor

⁴ En el [ANEXO 1](#) se pueden consultar las derivadas parciales para las cinco cópulas estudiadas.

hasta que la probabilidad condicional recupere un valor cercano a 0.5. De forma análoga, se tiene que si $P(U < u|V = v) < 0.5$ se considera que el activo está infravalorado, ya que la probabilidad condicionada a un valor fijo de v de que se den valores de u menores al analizado en ese momento en concreto es baja, por lo que se espera que el valor de ese activo aumente.

Del mismo modo se interpreta la probabilidad condicionada a un valor fijo de u , $P(V < v|U = u)$.

Una vez se tienen estas probabilidades condicionales es posible establecer límites para las mismas a partir de los cuales, en una estrategia de negociación por pares, se generaría la señal de posicionarse largo en un activo y corto en otro de manera que si ambos presentan un movimiento dependiente entre sí a lo largo del tiempo se espera que ambas probabilidades vuelvan a un nivel cercano al 0.5.

De este modo, si se siguiese una estrategia según Ferreira (2008), se abriría una posición cuando simultáneamente se diera una de las siguientes alternativas,

1. $P(U < u|V = v) > 0.5$ y $P(V < v|U = u) < 0.5$
2. $P(V < v|U = u) > 0.5$ y $P(U < u|V = v) < 0.5$

Donde en el primer caso por ejemplo se tomaría una posición corta en el activo u y una posición larga en el activo v , para posteriormente cerrar dicha posición cuando ambas probabilidades volvieran a un valor próximo a 0.5.

En este punto, fijar dichos límites que delimitan las señales de entrada y de salida es subjetivo, como se demuestra en la literatura, donde en Liew y Wu (2013) se considera una cota superior de 0.95 y una cota inferior de 0.05 necesarias para abrir posiciones mientras que se establece que el momento de cerrar posiciones es cuando cruzan ambas la frontera del 0.5.

Krauss y Stübinger (2015) proponen incluir un Stop-Loss en la estrategia mientras que en Stander et al. (2013) señalan que una posible alternativa para cerrar posiciones es cuando dicha operación genera rentabilidad o bien si transcurre cierto período de tiempo desde cuando se produjo la señal de entrada.

Para este trabajo, dado que se trabaja con más de un par de activos y se llevan a cabo diferentes tipos de análisis, se fijan unos límites diferentes en función del tipo de análisis en el que se aplica la estrategia. Es importante señalar que una forma de interpretar estos límites es su función como indicador de confianza como se señala en Ferreira (2008), por lo que al fijar dichos límites es relevante tener en consideración varios factores como la correlación que presentan ambos activos o la cópula óptima para cada par. En el apartado de resultados empíricos se detalla con más profundidad los límites para cada par de activos.

Una vez se tienen definidas las señales de entrada y de salida, es relevante concretar el número de acciones a comprar/vender en cada operación. En la literatura previa, no se ha llegado a una conclusión común acerca del número óptimo de acciones a comprar/vender en cada operación. Para este trabajo se han fijado 10000 euros como cantidad por operación, es decir, cuando se genere una oportunidad de entrada, se tomarán posiciones de 10000 euros en largo para un activo del par y 10000 euros en corto para el otro activo.

Otros autores como por ejemplo Stander et al. (2013) sugieren otras posibles alternativas como utilizar un ratio una vez se ha calculado el ratio de Sharpe de cada activo o utilizar como indicador otro ratio diferente relacionado con las derivadas de cada acción como se ha señalado en el apartado correspondiente a cópulas en el punto de literatura previa.

Para finalizar el análisis, en los casos en los que existe un número suficientemente grande de oportunidades de entrada, se lleva a cabo un modelo de clasificación *logit* para las mismas. La finalidad de modelizar las señales de entrada es detectar aquellas que con mayor probabilidad generaran rentabilidades positivas a partir de la volatilidad presente en los 20 días previos al día de la señal, de esta manera, se identifica con 1 a las señales que generan rentabilidades positivas y con 0 las que generan rentabilidades negativas.

Se pretende hallar la probabilidad de que las rentabilidades de las oportunidades de trading sean positivas, es decir, que la variable dependiente Y que en este caso es dicotómica sea 1,

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

donde X_i representa la desviación típica, β_1 es la constante y u_i es el término de error. Y_i toma el valor 1 cuando la rentabilidad es positiva y 0 cuando resulta ser negativa.

Esto se conoce como modelo lineal de probabilidad. En dicho modelo es posible interpretar la esperanza condicional de Y dado X como la probabilidad condicional.

Es importante destacar que en dicho modelo a pesar de que el término de error no siga una distribución normal, es posible estimar mediante MCO ya que las estimaciones siguen siendo insesgadas.

Sin embargo, la restricción más fuerte de este modelo dado que la variable dependiente Y es una variable que se interpreta como una probabilidad, debe estar acotada tal que,

$$0 \leq E(Y_i|X_i) \leq 1$$

Dado que MCO no garantiza esta restricción es preciso realizar una transformación al modelo, de este modo surge el modelo *logit*.

Entonces se tiene,

$$E(y) = \frac{e^{x'\beta}}{1 + e^{x'\beta}} = \frac{1}{1 + e^{-x'\beta}} = \pi_i$$

donde x es el vector de variables explicativas, en este caso la desviación típica, y β es el vector de parámetros.

Una vez se aplican logaritmos se obtiene que,

$$\text{Ln}\left(\frac{\pi_i}{1 - \pi_i}\right) = x' \beta$$

La estimación de los parámetros se realiza mediante máxima verosimilitud (Green, 2001).

Una vez se tienen las probabilidades de que la rentabilidad de las señales sea positiva condicionales a la volatilidad, se descartan aquellas que no presentan una probabilidad superior a 0.5.

4. RESULTADOS EMPÍRICOS

4.1 DATOS UTILIZADOS Y ESTADÍSTICOS DESCRIPTIVOS

Los datos utilizados para este trabajo se han extraído de la base de datos de *Thomson Reuters*. Los datos disponibles son:

-Datos de cotización diaria (de cierre y de apertura) comprendidos entre 30/12/1999 y 26/05/2017 para los pares Santander/BBVA, REE/Iberdrola, ACS/Abertis, FCC/Sacyr, ExxonMobil/Chevron y Coca-Cola/Pepsi. Los pares pertenecientes al mercado español tienen un total de 4417 datos mientras que los pertenecientes al mercado americano constan de un total de 4380 cotizaciones diarias.

Para el par ETF Ibex/Futuro Ibex se dispone de los datos comprendidos entre las fechas 20/07/2006 y 26/05/2017 (2770 cotizaciones diarias). Mientras que en el par Abengoa A/Abengoa B se tienen un total de 1171 datos comprendidos entre el 25/10/2012 y el 26/05/2017.

-En cuanto a los datos disponibles de alta frecuencia (precios de *trade*, *bid* y *ask*), cuya cotización es con frecuencia de 30 minutos, se dispone de un total de 4626 datos que corresponden a las cotizaciones presentes en el período comprendido entre las fechas 27/05/2016 y 26/05/2017 para los pares Santander/BBVA, REE/Iberdrola, ACS/Abertis, FCC/Sacyr. Para los dos pares correspondientes al mercado americano se tienen un total de 3765 datos comprendidos entre el 31/05/2016 y el 26/05/2017.

Es importante destacar el hecho de que cada día de cotización en los pares correspondientes al mercado español posee 18 cotizaciones intradía mientras que los pares correspondientes al mercado americano tienen 15 cotizaciones por día. De ahí se deriva la diferencia en el número de datos disponibles para ambos mercados.

A continuación, se presentan los estadísticos descriptivos más relevantes para las rentabilidades logarítmicas presentes en ambas muestras completas (datos diarios de cierre y datos de *trade* de alta frecuencia).

	Santander	BBVA	REE	Iberdrola	ACS	Abertis	FCC	Sacyr
Media	0.00%	-0.01%	0.06%	0.02%	0.04%	0.03%	-0.01%	-0.02%
Mediana	0.03%	0.00%	0.04%	0.04%	0.08%	0.04%	0.00%	0.00%
Desviación típica	2.25%	2.17%	1.57%	1.66%	1.85%	1.50%	2.18%	2.78%
Máximo	20.88%	19.91%	34.29%	17.23%	16.29%	11.48%	24.80%	19.35%
Mínimo	-22.16%	-17.64%	-9.12%	-13.43%	-11.41%	-10.91%	-9.73%	-13.98%
Asimetría	-0.077	0.180	2.540	0.310	0.053	-0.030	0.599	0.068
Kurtosis	10.184	8.504	59.133	15.383	7.836	7.151	9.862	6.646
Autocorrelación Lag 1	0.003	<u>0.061</u>	0.015	0.005	<u>0.032</u>	<u>-0.031</u>	<u>0.035</u>	<u>0.041</u>
Autocorrelación Lag 2	<u>-0.041</u>	<u>-0.042</u>	-0.018	<u>-0.065</u>	-0.030	-0.016	-0.008	0.014
Bandas autocorrelación	[0,0301;-0,0301]	[0,0301;-0,0301]	[0,0301;-0,0301]	[0,0301;-0,0301]	[0,0301;-0,0301]	[0,0301;-0,0301]	[0,0302;-0,0302]	[0,0302;-0,0302]
	Abengoa A	Abengoa B	ExxonMobil	Chevron	Coca-Cola	Pepsi	ETF Ibex	Futuro Ibex
Media	-0.36%	-0.44%	0.02%	0.02%	0.01%	0.03%	0.00%	0.00%
Mediana	-0.24%	-0.14%	0.03%	0.06%	0.02%	0.02%	0.05%	0.07%
Desviación típica	7.79%	7.78%	1.55%	1.63%	1.33%	1.25%	1.60%	1.60%
Máximo	50.63%	45.37%	15.86%	18.94%	13.00%	13.86%	13.15%	13.79%
Mínimo	-159.85%	-152.24%	-15.03%	-13.34%	-10.60%	-12.71%	-12.89%	-13.04%
Asimetría	-8.036	-7.541	0.033	0.069	0.046	0.065	-0.172	-0.222
Kurtosis	167.166	141.497	13.351	13.713	11.818	14.726	9.080	9.840
Autocorrelación Lag 1	<u>0.288</u>	<u>0.280</u>	<u>-0.107</u>	<u>-0.071</u>	-0.017	<u>-0.056</u>	0.011	0.019
Autocorrelación Lag 2	0.054	-0.021	<u>-0.092</u>	<u>-0.056</u>	<u>-0.056</u>	<u>-0.040</u>	<u>-0.060</u>	<u>-0.056</u>
Bandas autocorrelación	[0,0585;-0,0585]	[0,0585;-0,0585]	[0,0302;-0,0302]	[0,0302;-0,0302]	[0,0302;-0,0302]	[0,0302;-0,0302]	[0,0380;-0,0380]	[0,0380;-0,0380]

Tablas 1 y 2. Estadísticos descriptivos de las rentabilidades diarias. Los elementos subrayados hacen referencia a aquellos retardos que sobrepasan las bandas de autocorrelación, lo cual indica que es posible que dicha autocorrelación se encuentre presente en la correspondiente serie financiera.

	Santander	BBVA	REE	Iberdrola	ACS	Abertis	FCC	Sacyr
Media	0.01%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Mediana	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Desviación típica	0.66%	0.62%	0.27%	0.44%	0.56%	0.32%	0.46%	0.72%
Máximo	6.41%	6.43%	2.53%	9.23%	13.70%	6.40%	4.18%	8.23%
Mínimo	-28.70%	-25.02%	-5.16%	-22.15%	-25.65%	-10.44%	-3.02%	-22.95%
Asimetría	-17.673	-13.880	-1.740	-25.246	-17.830	-6.181	0.262	-6.207
Kurtosis	797.307	575.409	45.952	1420.300	1053.400	290.689	12.563	237.201
Autocorrelación Lag 1	<u>-0.109</u>	<u>-0.112</u>	<u>-0.081</u>	<u>-0.252</u>	<u>-0.230</u>	<u>-0.112</u>	<u>-0.231</u>	<u>-0.052</u>
Autocorrelación Lag 2	<u>-0.045</u>	<u>-0.088</u>	-0.027	0.015	<u>-0.072</u>	<u>-0.040</u>	-0.007	<u>-0.108</u>
Bandas autocorrelación	[0,0294;-0,0294]	[0,0294;-0,0294]	[0,0294;-0,0294]	[0,0294;-0,0294]	[0,0294;-0,0294]	[0,0294;-0,0294]	[0,0294;-0,0294]	[0,0294;-0,0294]
	Abengoa A	Abengoa B	ExxonMobil	Chevron	Coca-Cola	Pepsi		
Media	0.05%	-0.06%	0.00%	0.00%	0.00%	0.00%		
Mediana	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%		
Desviación típica	3.96%	2.52%	0.25%	0.28%	0.20%	0.20%		
Máximo	99.60%	28.77%	1.96%	2.70%	1.29%	2.43%		
Mínimo	-189.22%	-62.29%	-2.91%	-2.89%	-2.65%	-3.32%		
Asimetría	-20.634	-6.887	-0.365	0.127	-1.974	-1.377		
Kurtosis	1232.500	171.222	16.347	16.323	28.645	42.039		
Autocorrelación Lag 1	<u>-0.205</u>	<u>-0.067</u>	-0.007	0.029	-0.020	<u>-0.047</u>		
Autocorrelación Lag 2	<u>-0.126</u>	<u>0.073</u>	-0.005	-0.019	-0.003	-0.010		
Bandas autocorrelación	[0,0294;-0,0294]	[0,0294;-0,0294]	[0,0326;-0,0326]	[0,0326;-0,0326]	[0,0326;-0,0326]	[0,0326;-0,0326]		

Tablas 3 y 4. Estadísticos descriptivos de las rentabilidades correspondientes a los datos de alta frecuencia. Los elementos subrayados hacen referencia a aquellos retardos que sobrepasan las bandas de autocorrelación, lo cual indica que es posible que dicha autocorrelación se encuentre presente en la correspondiente serie financiera.

A raíz de estos estadísticos descriptivos, donde se observan indicios que indican la posible existencia de autocorrelación, así como asimetría y exceso de *kurtosis* en prácticamente todas las series financieras, el introducir cópulas tiene una justificación añadida, ya que en el caso de encontrar pares de activos que presenten estas características y además tengan una dependencia en su evolución, modelizar cada par mediante cópulas permite realizar un análisis más preciso de dicha dependencia.

Una vez se tienen los estadísticos descriptivos, es interesante si se comparan ambas frecuencias ver como se tiene de forma sistemática una mayor *kurtosis* y asimetría en los datos de alta frecuencia.

Si se analiza la estructura de autocorrelación, se observa que, en los datos de alta frecuencia, el 79% de los pares presenta indicios de autocorrelación si se analiza el primer retardo mientras que en los datos diarios dicho porcentaje se sitúa en el 63%.

Destacar los valores muy elevados de la rentabilidad máxima y mínima en las series de Abengoa A y Abengoa B. Dichas rentabilidades se deben a una dilución en el valor de ambas acciones consecuencia de una ampliación de capital.

Es muy interesante el resultado que se obtiene al comparar la volatilidad y los valores máximos y mínimos registrados en ambas frecuencias para los diferentes pares. Se puede observar como en los datos de alta frecuencia se obtienen unos valores más bajos para estos estadísticos, lo cual podría perjudicar a la hora de poner en práctica la estrategia para dicha frecuencia de datos.

4.2 RESULTADOS CON DATOS DIARIOS

En este primer apartado se lleva a cabo un análisis de la estrategia en 8 pares de acciones, donde se incluyen pares pertenecientes a diferentes sectores: banca, construcción, energía, petróleo, bienes de consumo y por último se analiza la estrategia en el par compuesto por un ETF Ibex y el Futuro Ibex.

De estos pares, 6 pares pertenecen al mercado español y los otros 2 pares están formados por empresas internacionales con un amplio recorrido histórico como ExxonMobil/Chevron y Coca-Cola/Pepsi.

La muestra de datos con frecuencia diaria analizada está comprendida entre las fechas del 30/12/1999 y el 26/05/2017 salvo en los pares ETF Ibex/Futuro Ibex y Abengoa A/Abengoa B donde la muestra de datos disponible no es tan grande. El número de cotizaciones diarias es de aproximadamente 4400 datos por acción, este número varía en función del par ya que una vez seleccionados los pares y filtrados los datos es posible que haya ligeras diferencias entre el mercado español y americano, como ya se ha comentado previamente.

El objetivo principal es comprobar si es posible obtener beneficios netos al aplicar una estrategia de negociación por pares en el mercado español, no obstante, es interesante y enriquece el análisis incorporar pares adicionales pertenecientes a otros sectores y mercados diferentes a los estudiados en los pares previos. De esta forma se completa el análisis añadiendo dos sectores de relevancia como pueden ser el petróleo y el de bienes de consumo.

Una vez se tienen los datos limpios, se analiza la evolución histórica de los precios indexados para posteriormente calcular los rendimientos logarítmicos. Se procede a indexar todas las series de manera que todas partan de un mismo valor, 100, entonces se aplican las rentabilidades sucesivamente sobre los precios indexados y de esta forma se puede comparar la evolución de todos los pares a escala de una forma más precisa.

Se han seleccionado los pares teniendo en cuenta principalmente el coeficiente de correlación de Pearson y el indicador SSD presentes en la muestra disponible. En cuanto al mercado español, se han calculado ambos indicadores para cada combinación posible entre un total de 53 acciones comprendidas entre el mercado continuo y el Ibex-35. De esta forma es posible realizar un primer filtro y a partir de ahí es posible seleccionar los pares. A parte de las acciones correspondientes al mercado español, se han añadido los pares ExxonMobil/Chevron y Coca-Cola/Pepsi, de esta manera se incluyen pares pertenecientes a diferentes mercados al analizar la estrategia. Para este trabajo los pares seleccionados son los siguientes:

Valores	SAN-BBVA	REE-IBE	ACS-ABE	FCC-SACYR
Coeficiente de correlación de Pearson	0.88	0.74	0.70	0.45
SSD	33.53	9.62	13.34	2.22

Valores	ABEA-ABEB	EXXON-CHEV	COLA-PEPSI	ETF-FUT(IBEX)
Coeficiente de correlación de Pearson	0.87	0.83	0.54	0.96
SSD	0.04	253.30	17303.00	0.48

Tabla 5. Correlación y SSD de cada par para toda la muestra disponible

A pesar de que los indicadores mencionados son importantes y es preferible que el coeficiente de correlación y el SSD sean lo mayor y lo menor posible respectivamente, conviene subrayar que el primero de ellos es un índice de correlación lineal y el segundo es un indicador de distancia relativa histórica, por lo que podría ser que en ocasiones estos indicadores no fueran capaces de detectar de forma muy eficiente los pares óptimos. Por ello, es deseable completar el análisis generando los gráficos de la evolución histórica de los precios estandarizados para cada uno de los pares y de esta forma evidenciar su dependencia.

A modo de ejemplo, es posible encontrar pares que no presenten una correlación muy elevada en un momento determinado pero su evolución histórica a lo largo de toda la muestra siga una misma tendencia y exista un “equilibrio” entre ambos activos. Del

mismo modo el SSD entre las dos series a comparar es preferible que sea lo menor posible, a pesar de ello, se puede dar que dos activos sigan una evolución similar, pero a cierta distancia una serie de la otra, por lo que el SSD resultante sería elevado y no por ello habría que desestimar el par.

En los siguientes gráficos de la evolución de los precios indexados para Santander/BBVA y Coca-Cola/Pepsi donde se puede ver su movimiento conjunto a lo largo de toda la muestra (véase [ANEXO 2.1](#)),

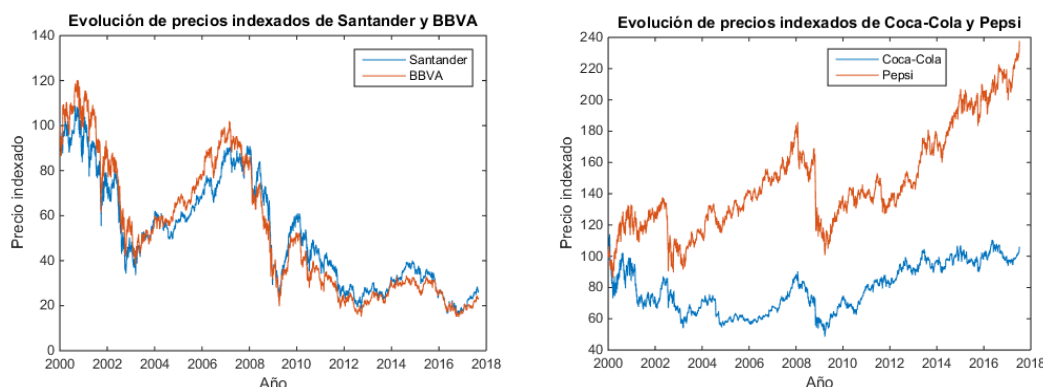


Figura 4. Evolución de los precios de Santander/BBVA y Coca-Cola/Pepsi. Se han indexado de forma que todas las series parten de un mismo valor, 100, elegido arbitrariamente.

En estos gráficos se observan los ejemplos anteriormente comentados, el par Coca-Cola/Pepsi se mantiene a cierta distancia relativa, pero sigue una evolución similar, mientras que Santander y BBVA se cruzan constantemente siguiendo una evolución prácticamente calcada.

Por todo lo comentado, hay evidencias suficientes para concluir que los pares seleccionados son buenos candidatos sobre los cuales poner en práctica la estrategia de negociación por pares deseada⁵.

Una vez se tienen los pares seleccionados, se diferencia la muestra disponible de datos en dos. La primera parte de la muestra se conoce como el período de formación y la segunda como período de negociación o período de trading.

Para este primer punto, se considera que el período de formación está compuesto por las rentabilidades calculadas con los precios de cierre correspondientes a los cuatro primeros años de la muestra (años 2000 a 2003). Mientras que para el período de negociación se tiene en cuenta una muestra considerablemente amplia comprendida entre 2004 y 2017. Teniendo de este modo una muestra amplia con la que trabajar, se pretende tener un mayor margen temporal a la hora de abrir y cerrar operaciones, poder prescindir de medidas como stop-loss/stop-profit y de este modo cerrar operaciones únicamente en función de cuando se produce la reversión al “equilibrio”. Por otro lado, una desventaja

⁵ Por motivos de espacio no se han incluido las tablas de correlaciones y SSD. En caso de que sea necesario, es posible facilitar dichas tablas con solicitud previa del lector.

que puede surgir de esta forma de proceder es que en determinados momentos la cópula no esté bien ajustada.

PERÍODO DE FORMACIÓN

El primer punto a desarrollar durante el período de formación es estudiar la distribución que siguen los datos una vez transformados en rentabilidades logarítmicas. Esto es necesario y de relevante importancia ya que posteriormente se procederá a la transformación integral de probabilidad de estas rentabilidades en función de la distribución que se ajuste de mejor manera a los datos con el fin de obtener una serie de uniformes comprendidas entre 0 y 1 necesarias para la cópula. Debido a esto, el transformar los datos con una distribución de probabilidad u otra es un elemento que puede generar unos resultados u otros.

En un primer momento, se ajustan para cada muestra hasta un total de 17 distribuciones continuas, entre las que se encuentran: Beta, Birnbaum-Saunders, Exponencial, Valor extremo, Gamma, Valor extremo generalizado, Pareto generalizado, Normal inversa, Logística, Log-logistic, Lognormal, Nakagami, Normal, Rayleigh, Rician, t location-scale, Weibull.

Se ordenan según el criterio BIC (“Bayesian Information Criterion”) y se estiman los parámetros para cada serie.

Como se puede observar, las distribuciones seleccionadas son la logística y la t location-scale (véase [ANEXO 2.2](#)), ambas distribuciones son similares a la distribución normal, pero con colas más gruesas.

Una vez conocidas las distribuciones de probabilidad, se realiza la transformación integral de probabilidad gracias al teorema de inversión. El cual dice lo siguiente:

Teorema de inversión: Sea X una variable aleatoria con función de distribución de probabilidad acumulada F , continua e invertible, y sea F^{-1} su función inversa. Entonces, la variable aleatoria $U = F(X)$ tiene distribución uniforme en $(0,1)$. Como consecuencia, si U es una variable aleatoria uniforme en $(0,1)$ entonces la variable aleatoria $X = F^{-1}(U)$ satisface la distribución F .

Realizar la transformación integral de probabilidad es necesario dado que el input de la cópula deben ser 2 vectores de variables uniformes en $(0,1)$. Para este trabajo se ha optado por la transformación mediante una distribución de probabilidad ya conocida, como puede ser en este caso la distribución logística o la t location-scale. Es posible realizar esta transformación aplicando la distribución empírica, pero el problema principal de llevar a cabo este procedimiento es que si únicamente se consideran los datos observados es difícil poder modelizar sucesos que no se han producido en esos datos, por lo que el riesgo asumido puede llegar a ser mayor que si se asume cierta distribución fija.

Una vez se tienen los números uniformes para el período de formación, se procede a ajustar la cópula que modelice de forma más precisa la dependencia entre cada par de activos. Para ello, se consideran un total de 5 posibles cópulas. Se estiman los parámetros de cada una y se ordenan las cópulas según el valor máximo de verosimilitud obtenido en la construcción de la cópula.

Es importante subrayar que, ante posibles cambios del ciclo económico, puede suceder que en ocasiones la cópula que mejor describe las dependencias durante el período de formación no sea la más adecuada para el período de *trading* correspondiente, por lo que en ocasiones puede ser recomendable considerar otros criterios que no sean estrictamente estadísticos.

PERÍODO DE NEGOCIACIÓN

Una vez se tienen las distribuciones y la cópula óptimas para cada par de acciones en el período de formación se lleva a cabo la estrategia en el período de negociación. Para ello se dispone de una muestra amplia (2004-2017) en la que se procede de la siguiente manera:

- 1) A partir de los parámetros estimados de las distribuciones óptimas para cada serie de rentabilidades se realiza la transformación integral de probabilidad para el dato nuevo día a día del período de negociación. Esto mismo se lleva a cabo en cada activo de los 8 pares.
- 2) Con las cópulas obtenidas para cada par en el período de formación, se calculan las probabilidades condicionales (véase [ANEXO 1](#)).
- 3) Se fijan los límites que delimitan las señales de entrada y de salida a partir de las probabilidades condicionales,

$$P(U \leq u|V = v) = \frac{\partial C}{\partial v} \quad \text{y} \quad P(V \leq v|U = u) = \frac{\partial C}{\partial u}$$

- 4) Se computan las órdenes de entrada y de salida. En cada señal de entrada, se toman simultáneamente una posición larga en el activo sobrevalorado y corta en el infravalorado. Cada posición larga o corta se lleva a cabo con un capital límite de 10000 euros en acciones.
- 5) Se calculan los beneficios antes de comisiones y después de comisiones.

COSTES DE TRANSACCIÓN

Se tienen en cuenta los siguientes costes de transacción:

-Comisión de 0.15% sobre el nominal de cada operación de compra y de venta.

-Coste de horquilla: No existe al cruzar las órdenes en las subastas de apertura o de cierre. Lo que sí que se tiene en cuenta, para hacer lo más real posible la operativa, es que cuando se genera una oportunidad de entrada, se lleva a cabo utilizando el precio de apertura del día siguiente, ya que el dato de cierre no es conocido hasta que ésta no

concluye, y por lo tanto es irreal que un sistema de una señal de entrada con la subasta de cierre y que el inversor pueda entrar a ese precio. De igual manera, cuando se produce una señal de salida, se efectúa con el precio de apertura del día siguiente.

RESULTADOS

Teniendo en cuenta los costes de transacción, los resultados son los siguientes:

	SAN-BBVA	REE-IBE	ACS-ABE	FCC-SACYR	ABEA-ABEB*	EXXON-CHEV	COLA-PEPSI	ETF-FUT (IBEX)**
Resultado Bruto (euros)	7004.59	10733.19	5474.91	21085.71	7066.99	6738.72	4028.75	1136.72
Resultado Neto (euros)	2696.47	7848.46	1581.59	15317.97	5093.47	-2121.09	2428.14	-7209.60
Rentab. Media bruta por operación	0.51%	1.12%	0.46%	1.02%	1.10%	0.23%	0.78%	0.04%
Rentab. Media neta por operación	0.21%	0.82%	0.16%	0.72%	0.80%	-0.07%	0.48%	-0.26%
Desviación típica	4.75%	5.77%	4.99%	11.86%	6.01%	2.00%	2.23%	0.82%
Cópula	Clayton	Clayton	t-Student	Clayton	t-Student	Frank	t-Student	Gumbel
Número de operaciones compra-venta	69	48	30	86	32	147	13	132
Señal de entrada	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1
Señal de salida	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45
Correlación período de formación	0.83	0.17	0.30	0.19	0.82	0.96	0.74	0.96
Correlación período de trading	0.91	0.56	0.63	0.49	0.94	0.86	0.65	0.97

*Período de formación: 10/2012-09/2016 (1000 datos). Período de trading: 10/2016-05/2017 (170 datos)

**Período de formación: 07/2006-06/2010 (1000 datos). Período de trading: 07/2010-05/2017 (1769 datos)

**Capital por operación: 12000 euros

Tabla 6. Resultados de aplicar la estrategia en los 8 pares. Período de formación: 2000-2003. Período de negociación: 2004-2017.

Es relevante destacar que en el desarrollo y puesta en práctica de esta estrategia no se tiene en cuenta ningún límite de capital, esto significa que es posible que haya más de una operación abierta simultáneamente. El objetivo de este trabajo es analizar si la estrategia es rentable una vez deducidos los costes de transacción, por lo que introducir medidas de trading como “stop-loss”, “stop-profit” o restringir el número de operaciones abiertas podría llevar a mayores beneficios y es importante tenerlas en cuenta de cara a una implementación real de la estrategia.

Siguiendo esta misma línea, no se ha introducido un límite de días para las operaciones abiertas, sino que se procede a cerrarlas una vez han revertido al “equilibrio” al alcanzar los límites fijados que señalan que existe una opción de salida.

Como se observa en la tabla de resultados, se han aplicado los siguientes niveles de confianza para detectar señales de entrada y de salida,

$$\text{Señal 1:} \quad P(U \leq u|V = v) > 0.9 \quad y \quad P(V \leq v|U = u) < 0.1$$

$$\text{Señal 2:} \quad P(U \leq u|V = v) < 0.1 \quad y \quad P(V \leq v|U = u) > 0.9$$

En la señal 1, se toman 2 posiciones, una corta en el activo u y larga en el activo v , cada una de ellas de un número de acciones hasta llegar a un límite de 10000 euros sin sobrepasarlo. De forma análoga, en la señal 2 se tomarían las 2 posiciones contrarias, cerrando la operación.

Se procede al cierre de estas posiciones una vez se alcanzan las siguientes probabilidades simultáneamente,

$$P(U \leq u|V = v) > 0.45$$

$$P(U \leq u|V = v) < 0.55$$

$$P(V \leq v|U = u) > 0.45$$

$$P(V \leq v|U = u) < 0.55$$

La elección de estos límites es totalmente subjetiva. Una forma de interpretarlos es como nivel de confianza que se pretende obtener a la hora de abrir y cerrar posiciones en relación al nivel en que se encuentran en ese momento ambas acciones respectivamente. A medida que se selecciona unos límites menores, se producen mayores señales de entrada y de salida, pero con un peor nivel de confianza y por lo tanto se asume un mayor riesgo a que dicho desequilibrio no converja.

Es muy importante advertir del riesgo de la sobreoptimización en este tipo de estrategias. Sobrecargar una estrategia con stops o limitaciones de capital, así como elegir las probabilidades condicionales que delimitan las señales de entrada o de salida en función de los resultados puede llevar a unos resultados mucho más elevados, pero con un elevado riesgo de que no sean consistentes en el tiempo.

Se observan beneficios brutos positivos para todos los pares, una vez deducidas las comisiones se ven reducidos considerablemente sin embargo siguen siendo elevados en casi todos los pares. Los pares Exxon/Chevron y ETF Ibex/Futuro Ibex son los únicos pares en los que se observan resultados netos negativos. Sorprende como los dos pares que partían con una mayor correlación, son los que presentan peores resultados. Una posible explicación podría ser que una correlación demasiado alta, no permite encontrar divergencias lo suficientemente grandes que permitan compensar los costes de la operativa.

En cambio, para los dos pares que partían con las correlaciones más bajas, como son los pares FCC/Sacyr y REE/Iberdrola, presentan altos beneficios incluso después de costes. Si se observa la estructura de dependencia en el período de formación en estos pares se puede ver como no presentan una estructura de dependencia tan clara como otro par que pueda tener una mayor claridad en dicha estructura como puede ser el par compuesto por Santander/BBVA (véase [ANEXO 2.3](#)).

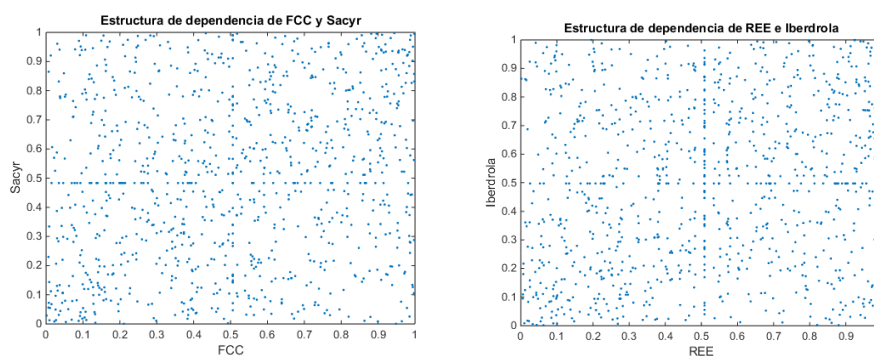


Figura 5. Estructuras de dependencias entre las rentabilidades de FCC/Sacyr y REE/Iberdrola una vez realizada la transformación integral de probabilidad para obtener números uniformes.

También se puede dar que pares en los que esta estrategia es muy popular y está muy extendida, cada vez existen menos ineficiencias que permitan aprovecharse de obtener beneficios aplicando una estrategia de este tipo.

Una vez analizados los resultados de aplicar la estrategia para un período de larga duración, se procede a estudiar la robustez de la estrategia dividiendo la muestra en 3 partes, teniendo de esta forma 3 períodos de formación y 3 períodos de trading diferentes.

4.3 RESULTADOS CON DATOS DIARIOS EN 3 PERÍODOS

Para este punto se consideran los 7 pares para los cuales se dispone de una muestra de datos lo suficientemente grande como para poder llevar a cabo el análisis.

Se mantiene el mismo criterio respecto al punto anterior en cuanto al límite que denota que existen señales de entrada y de salida. No obstante, para este análisis se considera un período de formación compuesto por 2 años, esto se debe a que el período de trading

se ve reducido a su vez a 4 años para cada tramo de este análisis, por lo que se considera que 2 años son suficientes para modelizar el comportamiento de cada par. En la literatura previa no se justifica de manera precisa el período a tener en cuenta para el período de formación/trading, lo relevante en este sentido es modelizar de manera adecuada el período de formación. Ante el riesgo de que exista un cambio estructural que pueda modificar la estructura de dependencia entre el par de activos, el número de años a tener en cuenta no tiene porque ser cuanto más mejor, sino que debe ser capaz de modelizar adecuadamente dicho cambio estructural.

En el ANEXO [3.1](#), [3.2](#), [3.3](#) se pueden consultar todas las distribuciones, así como las estructuras de dependencia de todos los pares para cada uno de los 3 períodos de formación en los que se divide la muestra.

Los resultados de aplicar la estrategia en el primer, segundo y tercer período respectivamente se recogen en la Tabla 7, 8 y 9.

Primer período

Período de formación: 2 Años, 2000-2001 (500 datos)

Período de negociación: 2002-2005 (1000 datos)

Segundo período

Período de formación: 2 Años, 2006-2007 (500 datos)

Período de negociación: 2008-2011 (1000 datos)

Tercer período

Período de formación: 2 Años, 2012-2013 (500 datos)

Período de negociación: 2014-2017 (1000 datos)

	SAN-BBVA	REE-IBE	ACS-ABE	FCC-SACYR	EXXON-CHEV	COLA-PEPSI
Resultado Bruto (euros)	1571.50	678.10	2597.34	4741.16	3151.93	-264.82
Resultado Neto (euros)	298.51	360.15	2213.61	4152.73	1597.52	-447.67
Rentab. Media bruta por operación	0.35%	0.68%	2.17%	2.37%	0.61%	-0.44%
Rentab. Media neta por operación	0.05%	0.38%	1.87%	2.07%	0.31%	-0.74%
Desviación típica	3.39%	1.73%	8.05%	2.51%	2.96%	1.00%
Cópula	t-Student	t-Student	Gumbel	Clayton	Frank	t-Student
Número de operaciones compra-venta	22	5	6	10	26	3
Señal de entrada	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1
Señal de salida	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45
Correlación período de formación	0.75	0.14	0.32	0.20	0.70	0.34
Correlación período de trading	0.88	0.32	0.32	0.25	0.81	0.51

Tabla 7. Resultados de aplicar la estrategia en el primer período. Período de formación: 2000-2001. Período de trading: 2002-2005.

	SAN-BBVA	REE-IBE	ACS-ABE	FCC-SACYR	EXXON-CHEV	COLA-PEPSI	ETF-FUT (IBEX)
Resultado Bruto (euros)	20043.13	14213.34	3116.00	15575.16	9070.48	7667.48	799.21
Resultado Neto (euros)	10365.14	10890.04	-5439.37	8371.81	-3095.57	1187.29	-5166.31
Rentab. Media bruta por operación	0.78%	1.27%	0.11%	0.74%	0.23%	0.37%	0.08%
Rentab. Media neta por operación	0.48%	0.97%	-0.19%	0.44%	-0.07%	0.07%	-0.22%
Desviación típica	6.58%	7.58%	5.13%	10.77%	2.56%	5.42%	1.41%
Cópula	Clayton	Clayton	Normal	t-Student	t-Student	Frank	Gumbel
Número de operaciones compra-venta	128	56	137	105	194	103	88
Señal de entrada	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1
Señal de salida	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45
Correlación período de formación	0.79	0.30	0.65	0.57	0.87	0.54	0.94
Correlación período de trading	0.94	0.61	0.69	0.54	0.91	0.70	0.96

Tabla 8. Resultados de aplicar la estrategia en el segundo período. Período de formación: 2006-2007. Período de trading: 2008-2011.

	SAN-BBVA	REE-IBE	ACS-ABE	FCC-SACYR	EXXON-CHEV	COLA-PEPSI	ETF-FUT (IBEX)
Resultado Bruto (euros)	12556.11	298.05	3496.27	2214.05	7988.13	6981.05	1940.23
Resultado Neto (euros)	2144.10	55.41	2234.68	-231.13	2222.54	2519.06	-1575.63
Rentab. Media bruta por operación	0.36%	0.37%	0.83%	0.28%	0.45%	0.47%	0.18%
Rentab. Media neta por operación	0.06%	0.07%	0.53%	-0.02%	0.15%	0.17%	-0.12%
Desviación típica	2.60%	2.05%	4.04%	14.07%	2.33%	1.20%	0.93%
Cópula	Gumbel	Clayton	Clayton	Clayton	t-Student	Frank	Gumbel
Número de operaciones compra-venta	174	4	21	39	89	74	52
Señal de entrada	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1	0.9-0.1
Señal de salida	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45
Correlación período de formación	0.95	0.55	0.58	0.50	0.76	0.57	0.95
Correlación período de trading	0.89	0.71	0.64	0.40	0.79	0.69	0.98

Tabla 9. Resultados de aplicar la estrategia en el tercer período. Período de formación: 2012-2013. Período de trading: 2014-2017.

En el primer período se observan resultados positivos para todos los pares menos para Coca-Cola/Pepsi, donde únicamente se generan 3 operaciones de compra-venta. Un resultado muy interesante son los beneficios obtenidos en pares como REE/Iberdrola, FCC/Sacyr o ACS/Abertis, donde a pesar de la baja correlación que presentan y a la vista de los resultados, se puede concluir que es posible que haya períodos en los que no se tenga una elevada correlación entre los 2 valores que forman el par, pero a pesar de ello se obtengan resultados positivos al aplicar la estrategia. Esto puede ser debido a que, dos activos que presentan una elevada correlación en una muestra amplia de datos, es posible que no presenten una elevada correlación en determinados momentos del tiempo, sin embargo, y fruto de esta menor correlación, se producen mayores divergencias y más pronunciadas entre ambos activos de forma que al pertenecer al pertenecer al mismo sector y país, y al tener una elevada correlación histórica, se espera una reversión a la media.

El segundo período coincide con el auge de la crisis económica. En dicho período se observa como los índices de correlación aumentan, lo cual es habitual en grandes shocks como puede ser este tipo de crisis. En este segundo tramo analizado, se pone de manifiesto que, al incrementarse la volatilidad en ambos mercados, el número de operaciones de compra venta aumenta de manera considerable, generando estas mismas unos mayores beneficios respecto el primer período en pares como Santander/BBVA o REE/Iberdrola.

No obstante, en otros pares como ACS/Abertis, ExxonMobil/Chevron y ETF Ibex/Futuro Ibex, los beneficios brutos positivos que se obtienen inicialmente se convierten en pérdidas una vez deducidos los costes de transacción.

Destacar el incremento en la volatilidad de las rentabilidades obtenidas en las operaciones respecto al primer período, esto podría ser causado por un aumento en la inestabilidad de los mercados provocado por la crisis.

Durante el tercer período, correspondiente al período después de la crisis, se observa como el número de operaciones de compra-venta así como los resultados vuelven a niveles pre-crisis (primer período). En este tercer período las correlaciones de los pares son significativamente superiores al primer período y se puede observar como los resultados netos de aplicar la estrategia en el par ETF Ibex/Futuro Ibex siguen siendo negativos. Esto es debido posiblemente a que un par que presente una elevada correlación, como puede ser la correlación que existe entre un ETF que replica un índice y el futuro sobre este mismo índice, si ambos activos no tienen divergencias lo suficientemente grandes como para compensar los costes de transacción puede ser que no sea un par óptimo para aplicar este tipo de estrategia.

En cuanto a las cópulas utilizadas, se puede observar que tanto las cópulas Arquimedianas como las no Arquimedianas son utilizadas para modelizar las dependencias, si bien es cierto que las primeras son seleccionadas en mayor medida.

De este modo se pone de manifiesto la robustez de la estrategia en datos diarios a lo largo del tiempo y se puede concluir que, al dividir la muestra en períodos más cortos, es posible modelizar mejor las dependencias entre los pares y de este modo obtener mayores beneficios. Este resultado obtenido sirve como motivación para llevar a cabo un análisis como el realizado en el siguiente punto.

4.4 RESULTADOS VENTANA MÓVIL CON DATOS DE ALTA FRECUENCIA VS DATOS DIARIOS

En este punto se introducen dos nuevos elementos en el análisis. A diferencia de utilizar un único período de formación estático como se ha utilizado en los puntos anteriores para generar la cópula, ahora se procede a construir la cópula cada día del período trading con los 500 datos previos a ese mismo día, de manera que cada día se actualizan los parámetros de dicha cópula así como las distribuciones marginales, pudiéndose obtener de este modo cópulas diferentes para un mismo par a lo largo del período analizado.

El segundo elemento a introducir es el trabajar con datos de alta frecuencia y comparar los resultados con los obtenidos con datos diarios para el mismo período de tiempo. De este modo se pone en práctica el análisis de la estrategia con datos de cotización de cada 30 minutos, esto es interesante porque de este modo se introduce la posibilidad de que el número de oportunidades de trading aumente y de este modo optimizar la estrategia. En el caso de datos de alta frecuencia se considera un período de formación compuesto por 360 datos, 1 mes.

Para el análisis de los datos de alta frecuencia, se ha incorporado el coste provocado por la horquilla al realizar operaciones de compra-venta. De esta forma cada vez que se generan oportunidades de entrada y de salida, se llevan a cabo con los precios de bid-ask que haya en ese mismo momento. Aparte de este coste, permanece el 0,15% sobre el nominal por operación que se tenía en cuenta en los puntos anteriores.

El análisis se resume de la siguiente manera:

-Pares analizados: Santander/BBVA, REE/Iberdrola, ACS/Abertis, FCC/Sacyr, Abengoa A/Abengoa B, ExxonMobil/Chevron y Coca-Cola/Pepsi.

-Para datos de alta frecuencia:

- Período de formación a partir del cual se realiza la ventana móvil: 27/05/2016 a 23/06/2016 (1 mes, 360 datos).
- Período de trading: 24/06/2016 a 26/05/2017 (1 año, 4266 datos).
- Cantidad a invertir en cada compra y venta: 10000 euros.
- Costes de transacción: horquilla (bid/ask) y 0.15% sobre el nominal de cada operación de compra y de venta.

-Para datos diarios:

- Período de formación a partir del cual se realiza la ventana móvil: 9/07/2014 a 23/06/2016 (2 años, 500 datos).
- Período de trading: 24/06/2016 a 26/05/2017 (1 año, 237 datos).
- Cantidad a invertir en cada compra y venta: 10000 euros.
- Costes de transacción: 0.15% sobre el nominal de cada operación de compra y de venta y el coste implícito de efectuar las operaciones con el precio de apertura del día siguiente al que se produce la señal.

En ambos casos se consideran las siguientes probabilidades condicionales para las oportunidades de entrada,

$$\text{Señal 1: } P(U \leq u|V = v) > 0.95 \quad y \quad P(V \leq v|U = u) < 0.05$$

$$\text{Señal 2: } P(U \leq u|V = v) < 0.05 \quad y \quad P(V \leq v|U = u) > 0.95$$

Y de salida,

$$P(U \leq u|V = v) > 0.45$$

$$P(U \leq u|V = v) < 0.55$$

$$P(V \leq v|U = u) > 0.45$$

$$P(V \leq v|U = u) < 0.55$$

Se ha decidido aumentar el nivel de confianza para la señal de entrada de 0.9-0.1 a 0.95-0.05 debido a la mayor inestabilidad y dificultad de modelizar las dependencias de los datos de alta frecuencia a raíz de los resultados obtenidos en el punto 4.1.

Los resultados para datos diarios y de alta frecuencia respectivamente son los siguientes,

	SAN-BBVA	REE-IBE	ACS-ABE	FCC-SACYR	ABEA-ABEB	EXXON-CHEV	COLA-PEPSI
Resultado Bruto (euros)	2841.01	475.17	1272.81	1239.26	-1577.19	798.78	-303.76
Resultado Neto (euros)	1864.47	23.92	817.22	986.22	-1978.04	-272.26	-842.60
Rentab. Media bruta por operación	1.02%	0.46%	1.60%	1.55%	-1.58%	0.22%	-0.24%
Rentab. Media neta por operación	0.72%	0.16%	1.30%	1.25%	-1.88%	-0.08%	-0.54%
Desviación típica	4.89%	2.84%	1.77%	4.18%	4.01%	1.11%	1.00%
Cópula	t-Student	t-Student	t-Student	t-Student (57%) Frank (43%)	t-Student	Frank	t-Student
Número de operaciones compra-venta	14	5	4	4	5	18	8
Señal de entrada	0.95-0.05	0.95-0.05	0.95-0.05	0.95-0.05	0.95-0.05	0.95-0.05	0.95-0.05
Señal de salida	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45
Correlación (período de formación y trading)	0.89	0.75	0.65	0.38	0.88	0.81	0.71

Tabla 10. Resultados de aplicar la estrategia con datos diarios durante un año con período de formación en forma de ventana móvil. El % presente en los diferentes tipos de cópulas hace referencia al número de días de trading para los cuales ese tipo de cópula es la óptima.

	SAN-BBVA	REE-IBE	ACS-ABE	FCC-SACYR	ABEA-ABEB	EXXON-CHEV	COLA-PEPSI
Resultado Bruto (euros)	1940.21	1485.02	1380.10	390.39	-11906.20	1102.72	-818.29
Resultado Neto (euros)	-9445.13	-3990.34	-2354.04	-5641.59	-13410.89	-5626.67	-8880.99
Rentab. Media bruta por operación	0.05%	0.08%	0.11%	0.02%	-2.38%	0.23%	-0.03%
Rentab. Media neta por operación	-0.25%	-0.22%	-0.19%	-0.28%	-2.68%	-0.07%	-0.33%
Desviación típica	0.96%	0.73%	1.50%	2.45%	2.42%	0.98%	0.76%
Cópula	t-Student	t-Student (99%) Frank (1%)	t-Student (97%) Frank (2%) Gumbel(1%)	Frank (36%) Clayton (40%) Normal (24%)	t-Student (95%) Frank (2.5%) Clayton (1.3%) Normal (1.2%)	t-Student	Clayton
Número de operaciones compra-venta	189	91	62	101	25	104	135
Señal de entrada	0.95-0.05	0.95-0.05	0.95-0.05	0.95-0.05	0.95-0.05	0.95-0.05	0.95-0.05
Señal de salida	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45	0.55-0.45
Correlación (período de formación y trading)	0.91	0.49	0.58	0.05	0.47	0.72	0.60

Tabla 11. Resultados de aplicar la estrategia con datos de alta frecuencia durante un año con período de formación en forma de ventana móvil. El % presente en los diferentes tipos de cópulas hace referencia al número de días de trading para los cuales ese tipo de cópula es la óptima.

En ambas tablas se observa que el par Abengoa A/Abengoa B genera pérdidas considerables al aplicar la estrategia. Esto se debe a que durante el 2017 Abengoa ha sido sometida a una ampliación de capital con una consecuente dilución del 95% en el valor de sus acciones, provocando que ambos activos del par pasen a cotizar con un valor muy bajo. Dada una cotización tan baja, los costes por horquilla cobran mucha importancia y pueden llegar a generar pérdidas muy elevadas.

Este tipo de eventos evidencia que en una estrategia de trading algorítmico como puede ser la de negociación por pares, es conveniente realizar un seguimiento individualizado de todos los pares que se quiera tener en cartera.

En el resto de los pares para datos diarios se observan beneficios netos en todos los pares de pertenecientes al mercado español, mientras que los pares pertenecientes al mercado americano ambos generan resultados netos negativos.

En cuanto a los datos de alta frecuencia, se puede observar que una vez aplicado el coste por horquilla y deducidos los costes por transacción, se obtienen unos resultados negativos en todos los casos. Esto es debido a que la baja rentabilidad media por operación no compensa los elevados costes que supone llevar a cabo la estrategia.

Una vez realizada la comparación, es interesante considerar la posibilidad de combinar ambas cópulas (con datos diarios y con datos intradiarios) con el fin de ver si es posible que combinando ambas se puedan optimizar los resultados de la cópula con datos diarios.

La idea al realizar este análisis es que cuando se genere una oportunidad de entrada en la cópula con datos diarios, comprobar si en la cópula con datos de alta frecuencia también se genera una oportunidad de entrada y tratar de optimizar la compra o venta ajustando con datos de alta frecuencia.

Comentar que los resultados⁶ obtenidos una vez aplicado este análisis no mejoran los resultados obtenidos inicialmente por la cópula con datos diarios. Los posibles motivos son los siguientes:

- Al generarse una oportunidad de entrada, el mercado no reacciona de manera instantánea por lo que es posible que en datos diarios se necesiten varios períodos para obtener resultados.
- Las estructuras de dependencia no son exactamente las mismas para las dos frecuencias de datos por lo que es posible que se produzcan divergencias en cuanto a la modelización de la dependencia entre ambas series para cada una de las frecuencias.

⁶ Debido a la extensión del trabajo se ha decidido no incluir estos resultados en el trabajo. Se pueden facilitar al lector con solicitud previa.

- Es importante resaltar que, al introducir la cópula con datos de alta frecuencia, esto incorpora el coste de horquilla que en ocasiones puede llevar a añadir costes considerables.

4.5 OPTIMIZACIÓN DE LAS SEÑALES DE ENTRADA

Al llevar a cabo una estrategia de este tipo, es relevante plantearse la posibilidad de, dado el número elevado de *trades* realizados como por ejemplo en el período de tiempo comprendido entre 2008 y 2011, ser capaz de caracterizar de algún modo aquellas oportunidades de entrada que generan mayor rentabilidad al revertir al “equilibrio”. Como se ha podido comprobar en el punto 4.2, la volatilidad presente en el momento de realizar la operación es un elemento importante y parece estar relacionado tanto con el número de oportunidades de entrada como con la rentabilidad obtenida de dichas operaciones.

En este punto se estudia la relación que existe entre las rentabilidades obtenidas en cada señal de entrada y las desviaciones típicas de los 2 activos del par en cuestión presentes en el momento que se genera dicha señal. La desviación típica se calcula con las 20 rentabilidades previas a cada señal de entrada.

Una razón por la que la desviación típica puede ser determinante a la hora de poder clasificar estas señales de entrada en potencialmente rentables o no es que parece razonable pensar que los desequilibrios generados a partir de un aumento en la volatilidad (se identifica volatilidad con desviación típica) estarán menos relacionados con cambios en la estructura de dependencia entre ambos activos y será más posible y más rápida la reversión al “equilibrio”.

Para el siguiente análisis se han tenido en cuenta lo siguientes pares y períodos en función de que hubiera un número suficiente de señales de entrada como para que resulte robusto realizar el análisis y encontrar algún tipo de relación:

-Santander/BBVA en los períodos 2008-2011 y 2014-2017.

-ExxonMobil/Chevron en los períodos 2008-2011 y 2014-2017.

-Coca-Cola/Pepsi en los períodos 2008-2011 y 2014-2017.

En los siguientes dos ejemplos se puede observar gráficamente (véase ANEXO [5.1](#), [5.2](#), [5.3](#)) la relación existente entre la rentabilidad de cada señal de entrada y la volatilidad presente en el momento de la señal de entrada:

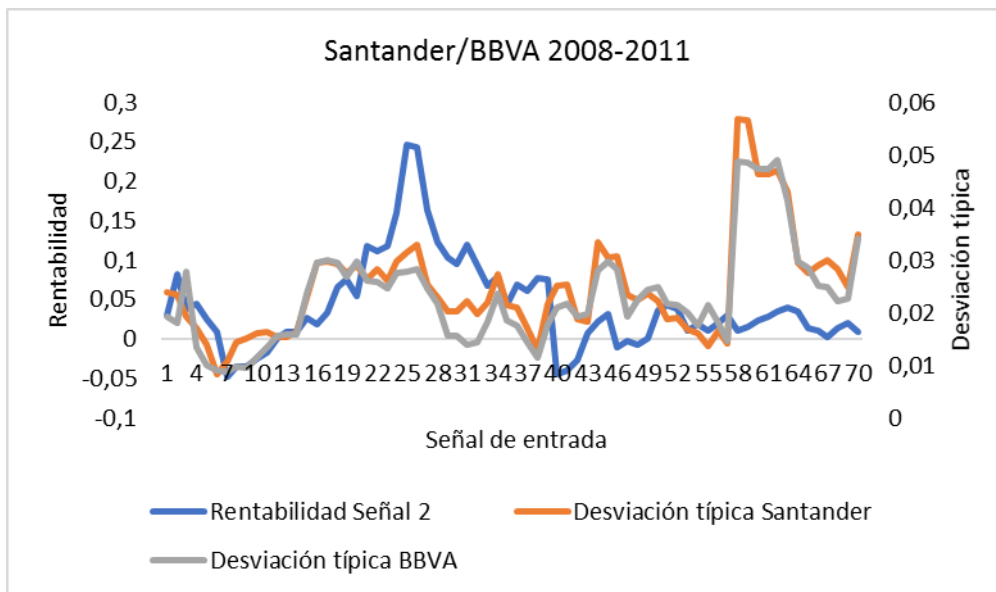


Figura 6. Gráfica de las rentabilidades obtenidas en cada oportunidad de trading correspondientes a la señal 2 de Santander/BBVA durante el período de trading 2008-2011 y las volatilidades correspondientes a las 20 rentabilidades previas de cada uno de los activos del par en el momento de generarse la correspondiente señal de entrada.

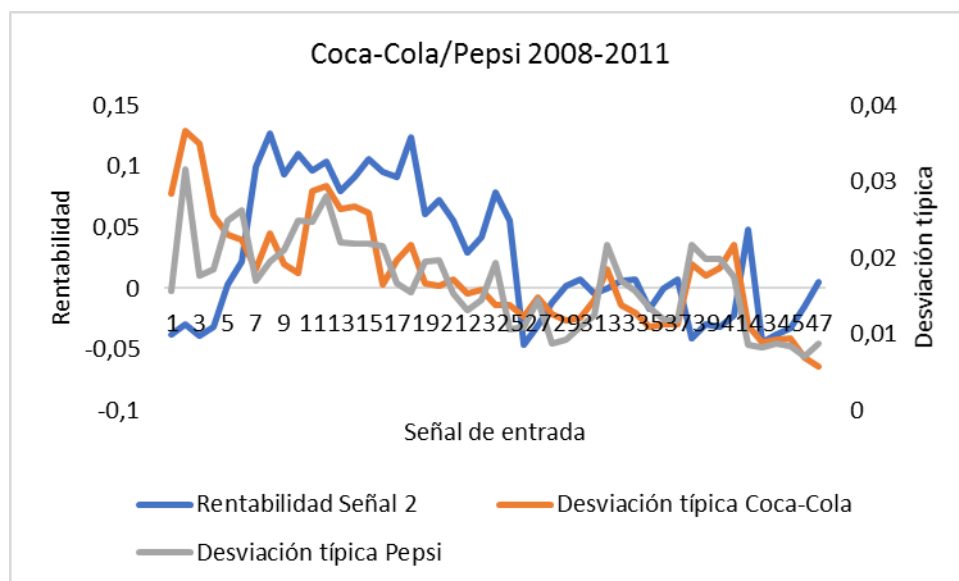


Figura 7. Gráfica de las rentabilidades obtenidas en cada oportunidad de trading correspondientes a la señal 2 de Coca-Cola/Pepsi durante el período de trading 2008-2011 y las volatilidades correspondientes a las 20 rentabilidades previas de cada uno de los activos del par en el momento de generarse la correspondiente señal de entrada.

Ambos gráficos hacen referencia a las rentabilidades obtenidas de la señal 2, donde como se ha citado anteriormente,

Señal 1: $P(U \leq u|V = v) > 0.9$ y $P(V \leq v|U = u) < 0.1$

Señal 2: $P(U \leq u|V = v) < 0.1$ y $P(V \leq v|U = u) > 0.9$

En este caso “ u ” se identifica con el primer activo del par, Santander y Coca-Cola en este caso, y “ v ” con el segundo, BBVA y Pepsi en este ejemplo.

El eje horizontal del gráfico representa el número de señales de entrada, mientras que los ejes verticales, se identifican con la rentabilidad obtenida en dichas señales y la volatilidad presente en el momento que se generaron dichas señales.

Como se observa en ambos gráficos, la tendencia es muy similar. Es relevante destacar que la volatilidad para ambos activos del par es la misma, lo cual es señal de la dependencia presente entre ambas acciones del par.

Una vez realizado este análisis, la intención es poder modelizar estas señales en función de la volatilidad de manera que sea posible optimizar la estrategia y poder detectar señales de entrada en las que no sea conveniente llevar a cabo ninguna operación y por tanto descartar.

Para ello, se aplica un modelo *logit* mediante el cual se obtiene una probabilidad entre 0 y 1 donde 1 se identifica con que la operación llevada a cabo a partir de la señal de entrada tiene una rentabilidad positiva y 0 con una rentabilidad negativa.

Para llevar a cabo este modelo, es necesario disponer de una muestra lo suficientemente grande. A partir de dicha muestra, se utilizan las primeras rentabilidades para generar el modelo y se trata de predecir la rentabilidad de las últimas 40 señales a partir de la volatilidad presente en el momento en el que aparecen.

De este modo se tiene que,

-La volatilidad de uno de los activos (dado su elevada correlación) es la variable predictora a partir de la cual se clasifica la variable a predecir que es la rentabilidad.

-La variable a predecir es la rentabilidad, donde 1 se identifica con una operación que genera una rentabilidad positiva y 0 con una rentabilidad negativa.

-La muestra a analizar es la compuesta por las rentabilidades y volatilidades correspondientes a cada señal de entrada. Dicha muestra se divide en dos partes, con la primera se genera el modelo *logit* y se estiman todos los parámetros y coeficientes. Una vez se tiene el modelo, se evalúa su capacidad predictiva en la segunda parte de la muestra, descartando las señales de entrada que predicen que a partir de la volatilidad presente en ese momento existe una menor probabilidad de que generen rentabilidades positivas.

Los resultados son los siguientes,

2008-2011	SAN-BBVA	EXXON-CHEV	COLA-PEPSI
Rentabilidad media por señal	0.78%	0.23%	0.37%
Rentabilidad media por señal (logit)	1.12%	0.37%	0.08%
Señales totales	128	194	103
Señales descartadas	20	22	15

Tabla 12. Resultados de aplicar el modelo de clasificación logit para predecir las 40 últimas señales de entrada en el período de trading 2008-2011 en los pares Santander/BBVA, ExxonMobil/Chevron y Coca-Cola/Pepsi.

2014-2017	SAN-BBVA	EXXON-CHEV	COLA-PEPSI
Rentabilidad media por señal	0.36%	0.45%	0.47%
Rentabilidad media por señal (logit)	0.39%	0.60%	0.36%
Señales totales	174	89	74
Señales descartadas	14	12	8

Tabla 13. Resultados de aplicar el modelo de clasificación logit para predecir las 40 últimas señales de entrada en el período de trading 2014-2017 en los pares Santander/BBVA, ExxonMobil/Chevron y Coca-Cola/Pepsi.

La rentabilidad media por operación mejora considerablemente en los pares Santander/BBVA y ExxonMobil/Chevron, mientras que para el par formado por Coca-Cola/Pepsi se descartan señales de entrada con rendimientos positivos, lo cual no es recomendable.

5.CONCLUSIONES

Este trabajo tenía como objetivo responder a la pregunta principal de si es posible obtener beneficios netos una vez deducidos los costes por transacción al aplicar una estrategia de negociación por pares bajo el entorno de cópulas. Se ha tratado de responder esta pregunta utilizando datos diarios y datos de alta frecuencia mediante cópulas estáticas y cópulas dinámicas. De igual modo se ha estudiado la robustez de la estrategia en datos diarios dividiendo la muestra en 3 períodos y aplicando la estrategia a cada uno de ellos.

Como consecuencia de responder a la pregunta principal ha surgido la pregunta secundaria de si es posible optimizar la estrategia discriminando aquellas señales de entrada que potencialmente tienen menores probabilidades de generar rendimientos positivos, para ello se ha aplicado un modelo *logit*.

Como conclusión final, es importante resaltar el hecho de que si bien es cierto que los resultados para datos diarios dejan ver que sí es posible llevar a cabo esta estrategia de arbitraje estadístico y obtener beneficios netos, una vez estudiado el funcionamiento de la estrategia en datos de alta frecuencia, no se puede concluir de la misma manera para estos, ya que una vez deducidos los costes por transacción y el coste de horquilla, este último en ocasiones supone un coste muy elevado, se generan beneficios netos negativos.

En cuanto a los resultados de aplicar el modelo *logit* para optimizar la estrategia, no es posible garantizar que dicho modelo suponga una maximización de beneficios con total seguridad, pero sí conviene tenerlo en cuenta ya que en 2 pares de los 3 en los que se ha analizado este modelo se han obtenido mejores resultados que sin tenerlo en cuenta. Este tipo de modelos de clasificación o de técnicas de *machine learning* como pueden ser árboles de clasificación son un recurso que puede ayudar a optimizar la estrategia de negociación por pares y parece sensato tenerlos presentes para futuras investigaciones.

Otro camino a seguir en futuras investigaciones es el de tratar las rentabilidades antes de realizar la transformación integral de probabilidad, modelizándolas por ejemplo de la siguiente manera,

$$\begin{aligned}r_{i,t} &= c_0 + c_1 r_{i,t-1} + e_{i,t} \\ e_{i,t} &= \sigma_{i,t} \varepsilon_{i,t} \quad \varepsilon_{i,t} \sim t - Student \\ \sigma_{i,t}^2 &= w + \alpha e_{i,t-1}^2 + \beta \sigma_{i,t-1}^2\end{aligned}$$

De este modo se pretende eliminar el problema de autocorrelación presente en las series financieras, generando series iid, $e_{i,t}$, a partir de las cuales es posible llevar a cabo la citada transformación integral de probabilidad y usar estas nuevas series uniformes como input de la cópula. Es importante destacar el hecho de que, al llevar a cabo este modelo para su uso en una estrategia como la de negociación por pares, hay que tener en cuenta que, en el momento de interpretar la probabilidad condicional derivada de la cópula para hallar las probabilidades condicionales, estas harían relación a $e_{i,t}$ por lo que habría que

realizar los cálculos oportunos para hallar la probabilidad de $r_{i,t}$ que es a partir de la cual se generarían las señales de entrada y de salida.

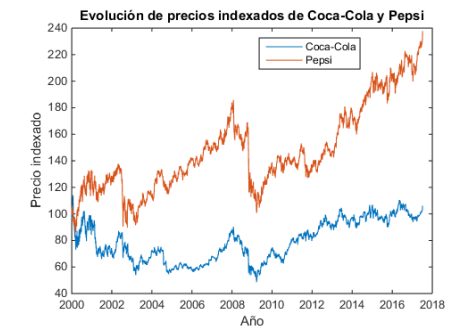
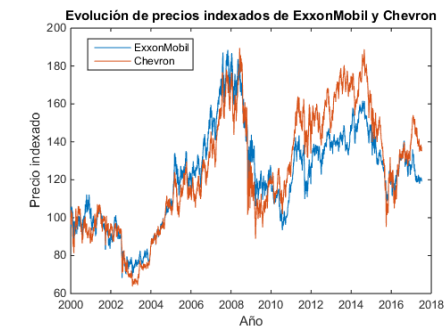
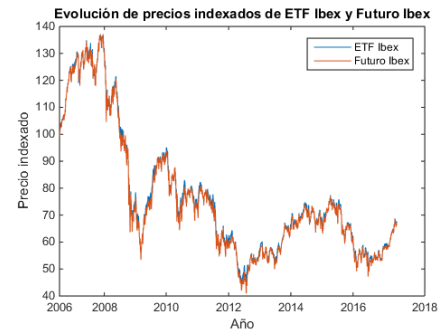
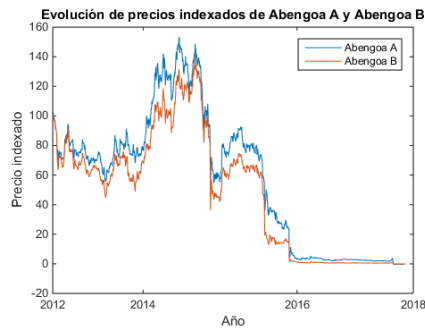
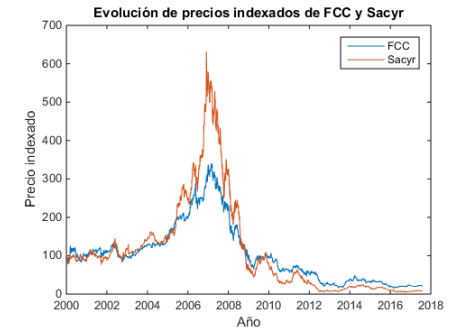
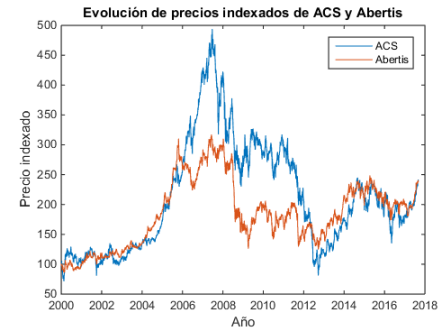
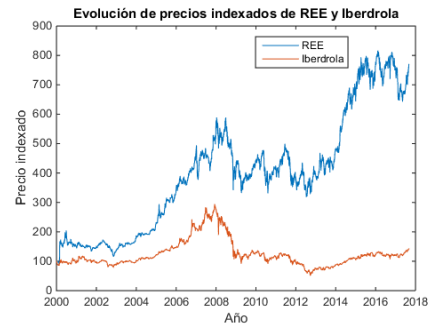
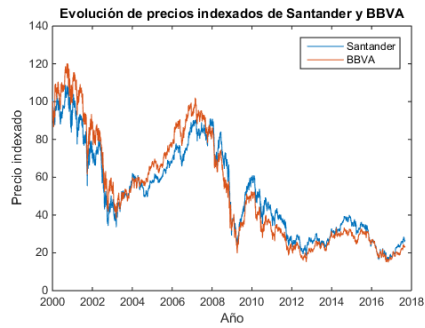
ANEXO 1

EXPRESIONES DE LAS PROBABILIDADES CONDICIONALES DE LAS CÓPULAS UTILIZADAS

CÓPULA	PARÁMETROS	$P(U \leq u V = v) = \frac{\partial C}{\partial v}$	$P(V \leq v U = u) = \frac{\partial C}{\partial u}$
1.NORMAL	$\theta \in (-1,1)$	$\phi\left(\frac{\phi^{-1}(u) - \theta \phi^{-1}(v)}{\sqrt{1-\theta^2}}\right)$	$\phi\left(\frac{\phi^{-1}(v) - \theta \phi^{-1}(u)}{\sqrt{1-\theta^2}}\right)$
2.T-STUDENT	$\theta_1 \in (0, \infty), \theta_2 \in (-1,1)$	$t_{\theta_1+1}\left(\sqrt{\frac{\theta_1+1}{\theta_1+t_{\theta_1}^{-1}(v)^2}} * \frac{t_{\theta_1}^{-1}(u) - \theta_2 t_{\theta_1}^{-1}(v)}{\sqrt{1-\theta_2^2}}\right)$	$t_{\theta_1+1}\left(\sqrt{\frac{\theta_1+1}{\theta_1+t_{\theta_1}^{-1}(u)^2}} * \frac{t_{\theta_1}^{-1}(v) - \theta_2 t_{\theta_1}^{-1}(u)}{\sqrt{1-\theta_2^2}}\right)$
3.CLAYTON	$\theta \in [-1, \infty) \setminus \{0\}$	$v^{-(\theta+1)}(u^{-\theta} + v^{-\theta} - 1)^{\frac{1}{v}-1}$	$u^{-(\theta+1)}(u^{-\theta} + v^{-\theta} - 1)^{\frac{1}{v}-1}$
4.GUMBEL	$\theta \in [1, \infty)$	$C_\theta(u, v) * [(-\ln u)^\theta + (-\ln v)^\theta]^{\left(\frac{1-\theta}{\theta}\right)} * (-\ln v)^{\theta-1} * \frac{1}{v}$	$C_\theta(u, v) * [(-\ln u)^\theta + (-\ln v)^\theta]^{\left(\frac{1-\theta}{\theta}\right)} * (-\ln u)^{\theta-1} * \frac{1}{u}$
5.FRANK	$\theta \in (-\infty, \infty) \setminus \{0\}$	$\frac{e^{-\theta v}(e^{-\theta u}-1)}{(e^{-\theta}-1)+(e^{-\theta u}-1)(e^{-\theta v}-1)}$	$\frac{e^{-\theta u}(e^{-\theta v}-1)}{(e^{-\theta}-1)+(e^{-\theta u}-1)(e^{-\theta v}-1)}$

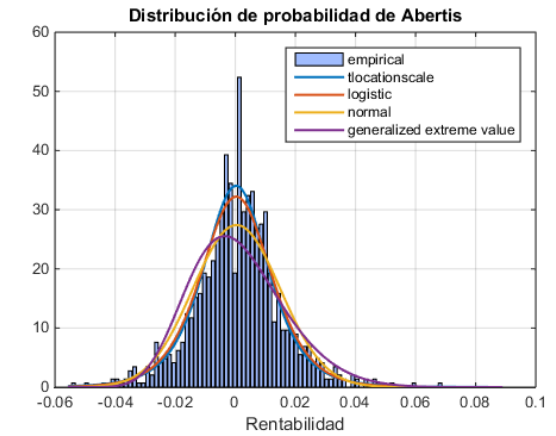
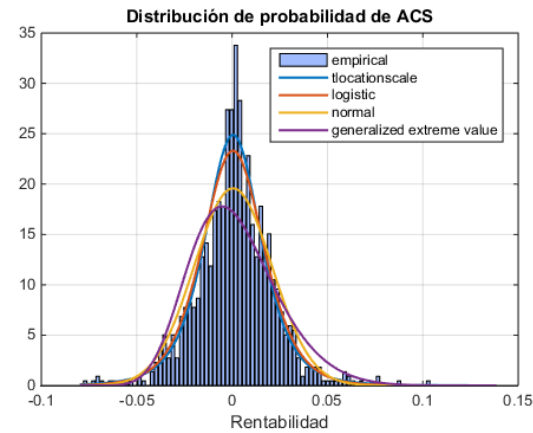
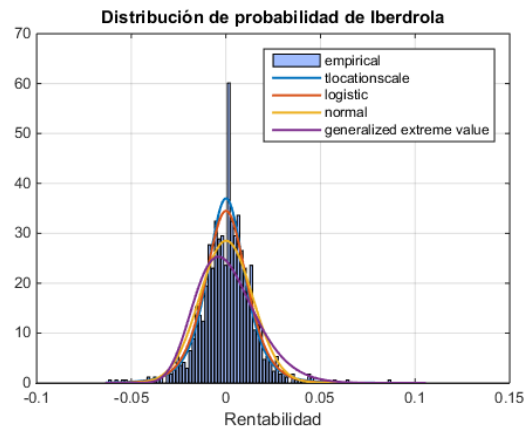
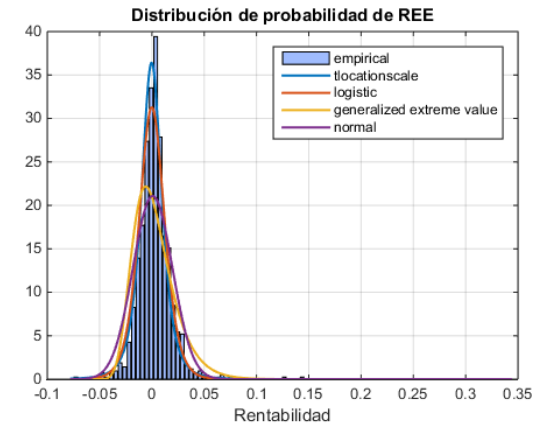
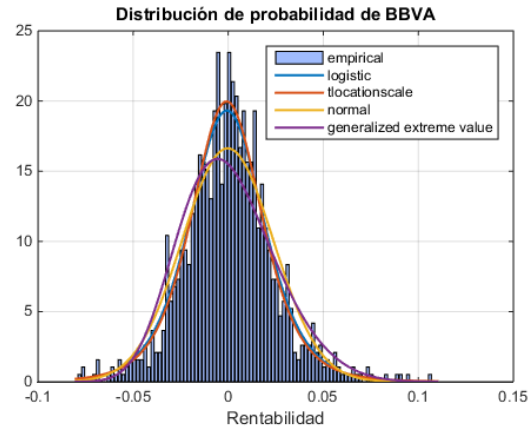
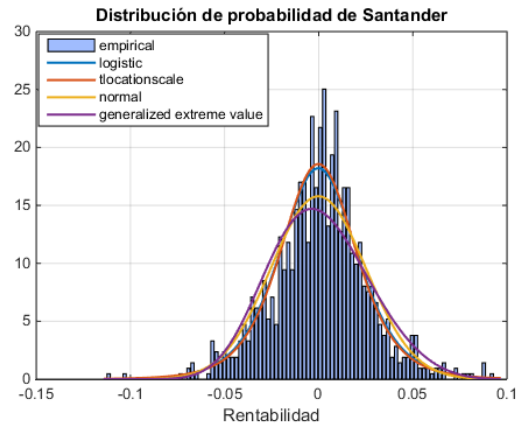
ANEXO 2.1

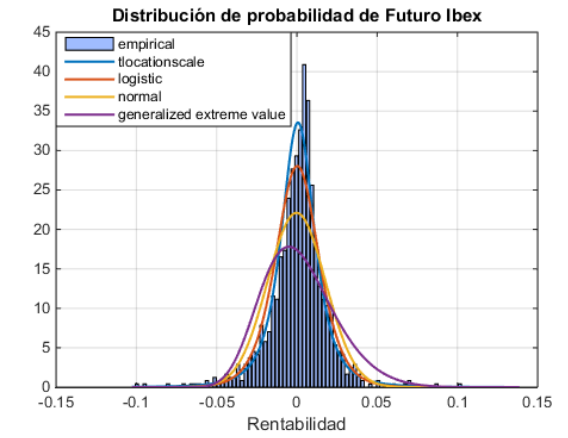
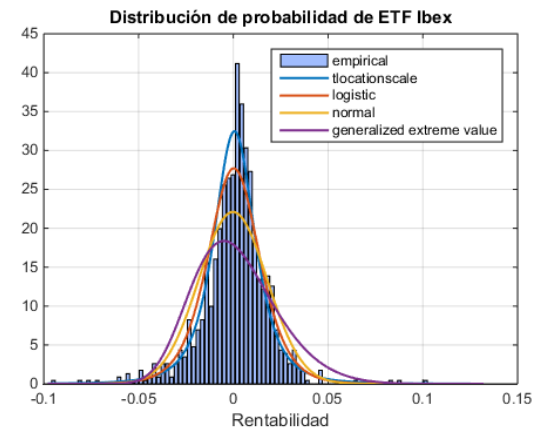
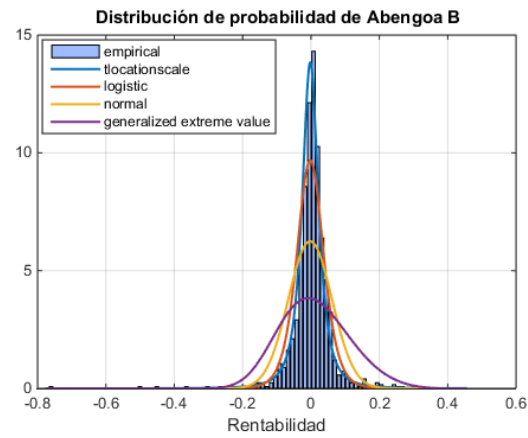
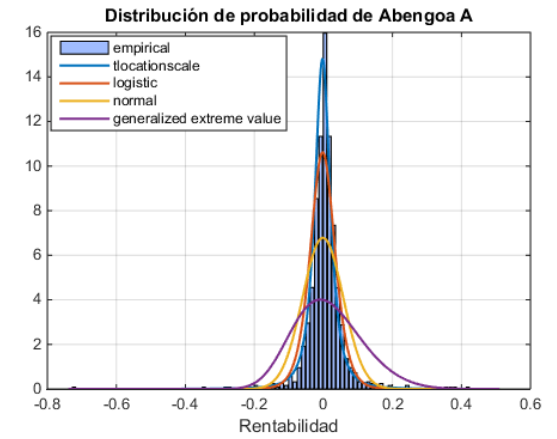
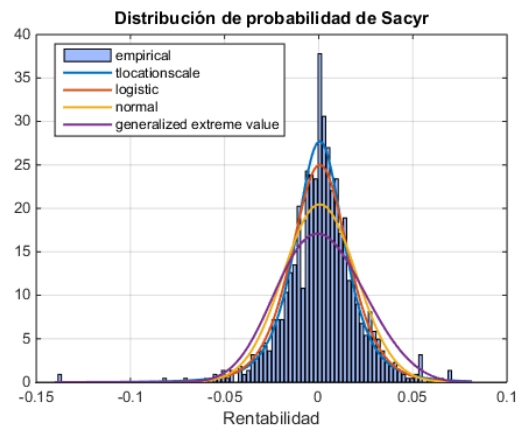
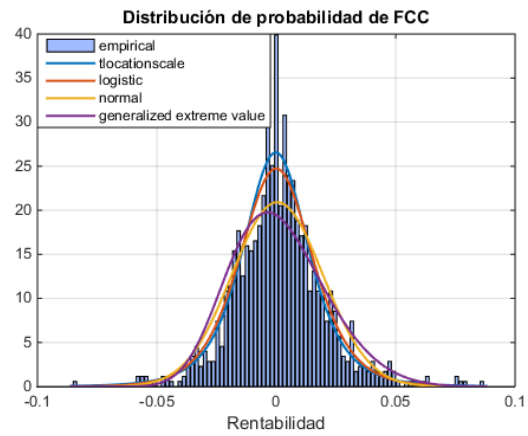
SERIES HISTÓRICAS DE PRECIOS ESTANDARIZADOS DE LOS 8 PARES (INCLUYE PERÍODO DE FORMACIÓN Y DE NEGOCIACIÓN)

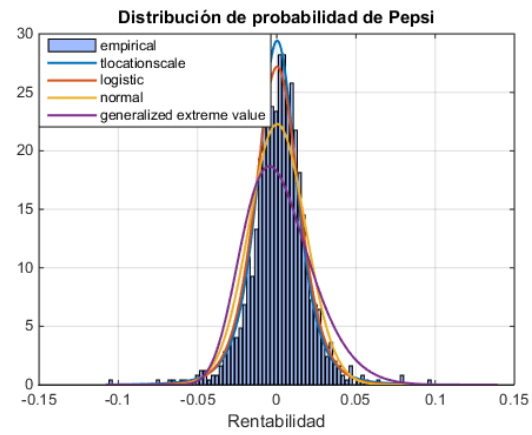
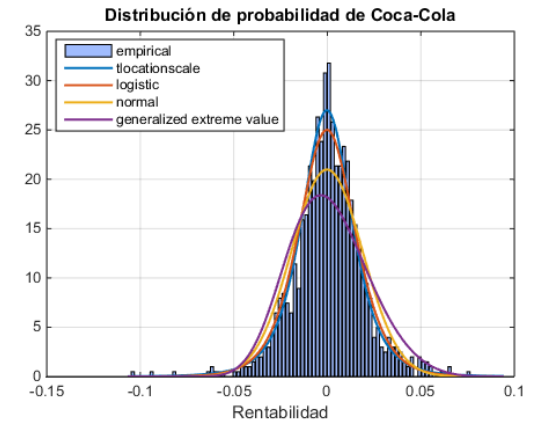
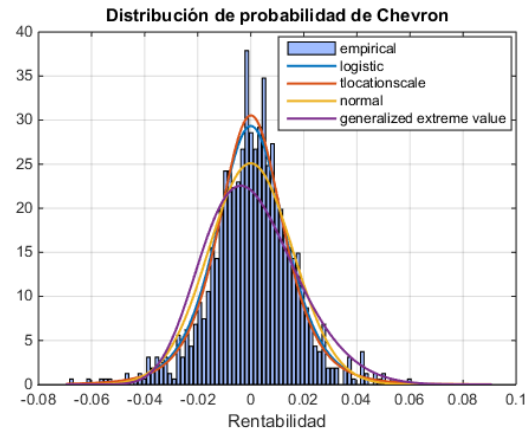
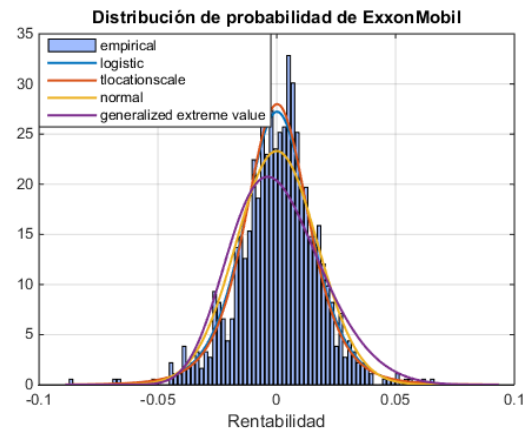


ANEXO 2.2

DISTRIBUCIONES DE PROBABILIDAD PARA CADA ACTIVO EN EL PERÍODO DE FORMACIÓN CON FRECUENCIA DE DATOS DIARIOS (2000-2003)

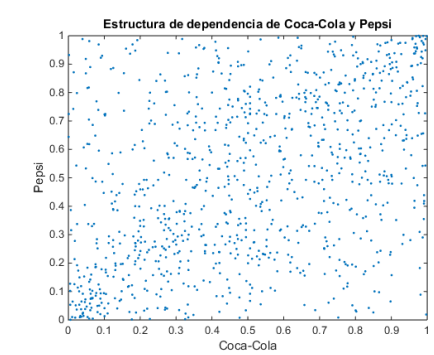
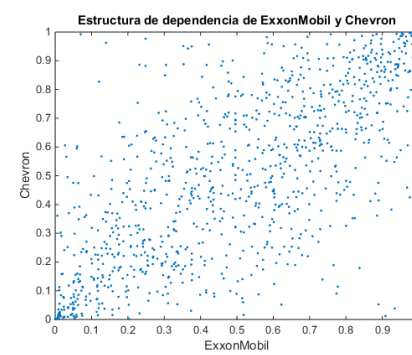
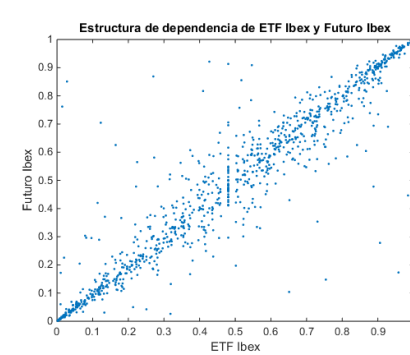
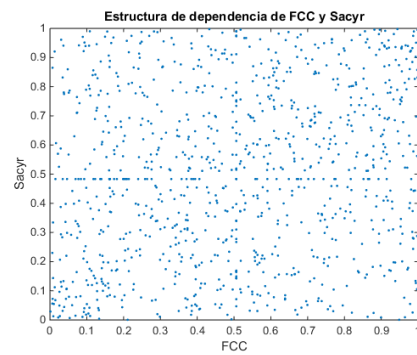
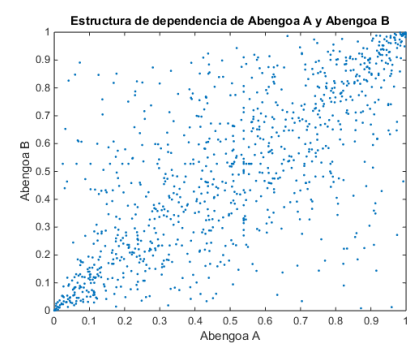
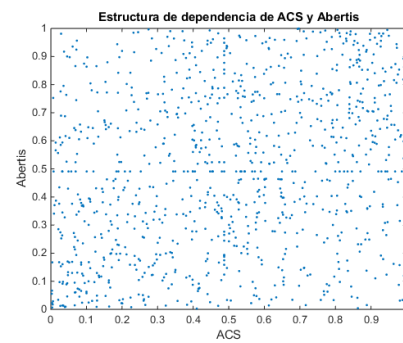
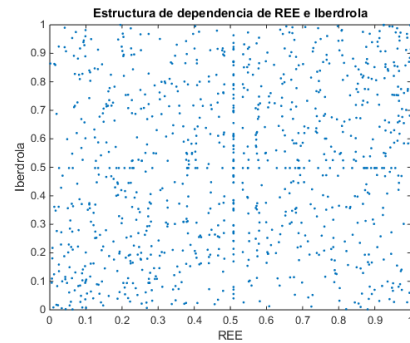
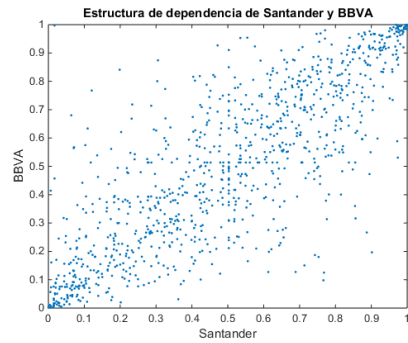






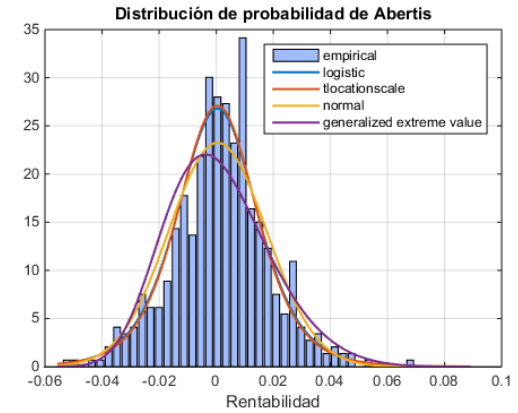
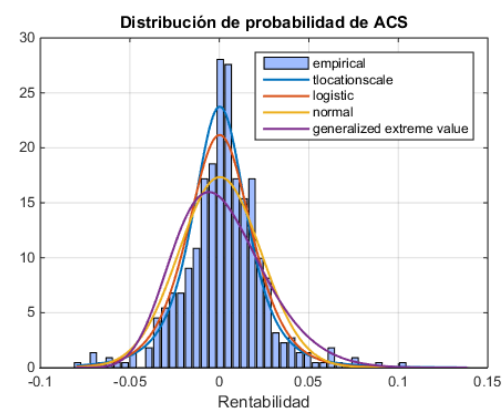
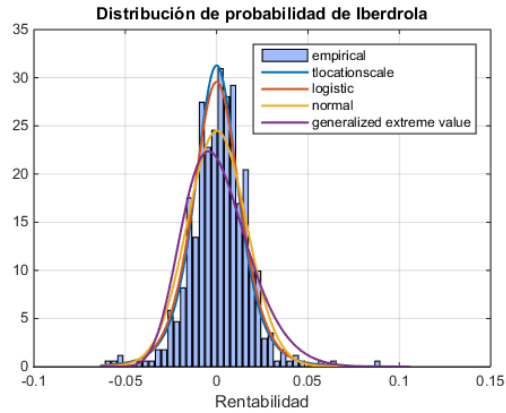
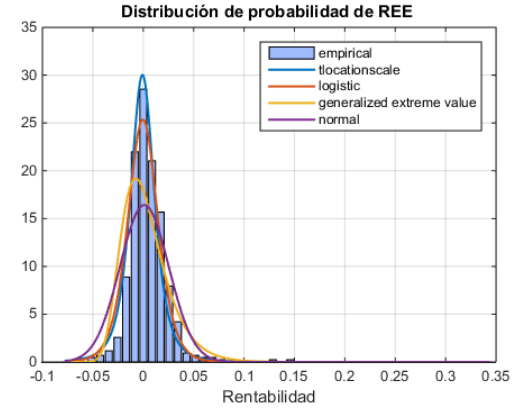
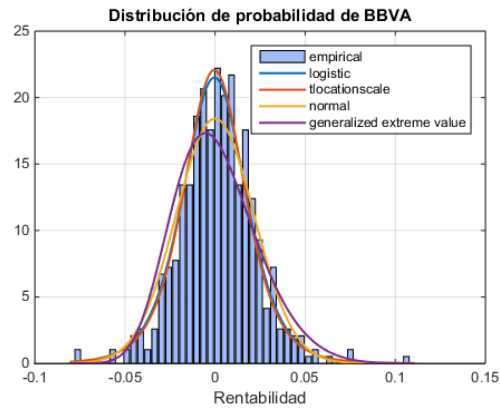
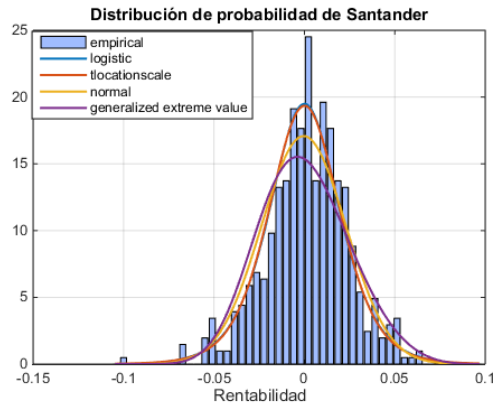
ANEXO 2.3

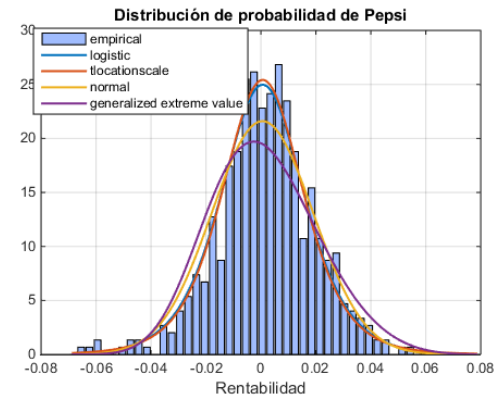
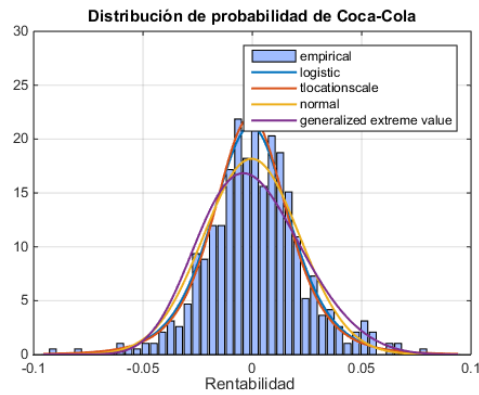
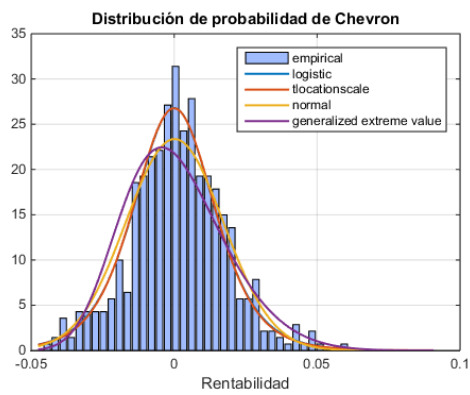
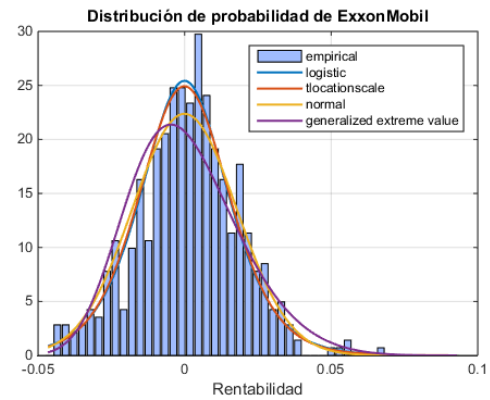
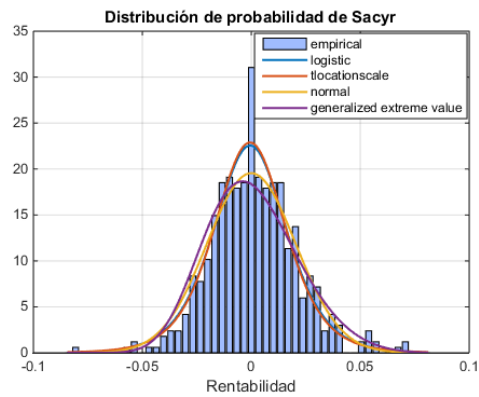
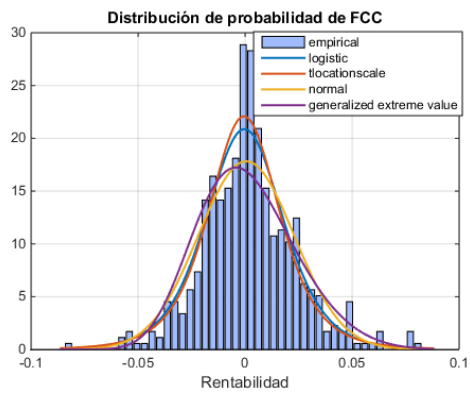
ESTRUCTURAS DE DEPENDENCIAS ENTRE CADA PAR DE ACTIVOS PARA EL PERÍODO DE FORMACIÓN EN DATOS DIARIOS (2000-2003)



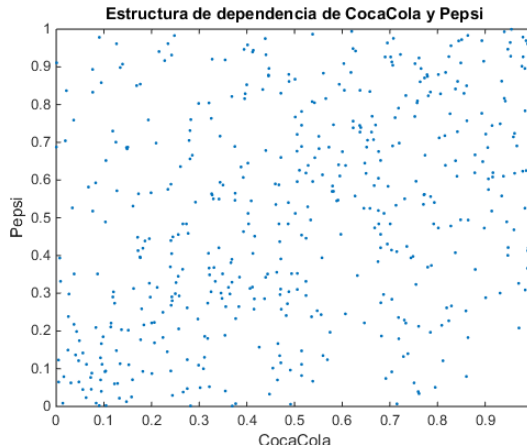
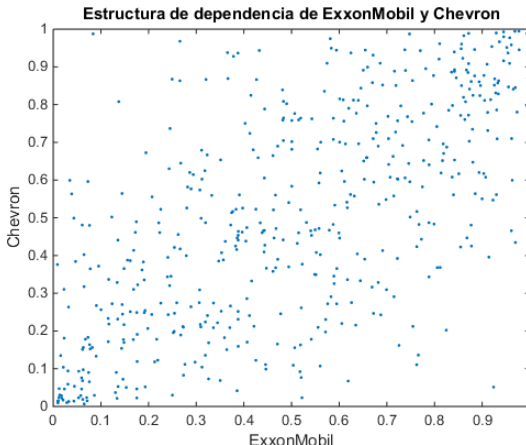
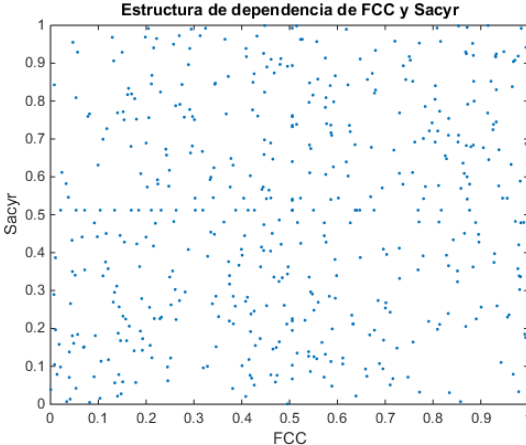
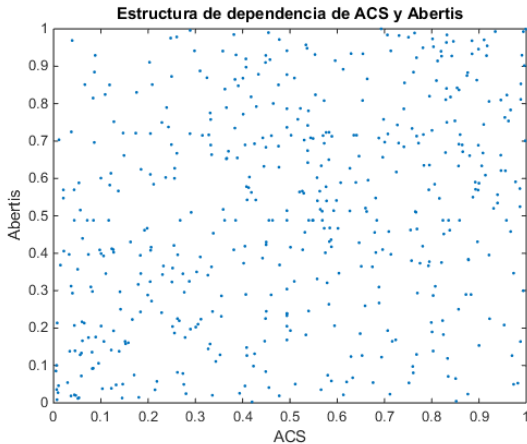
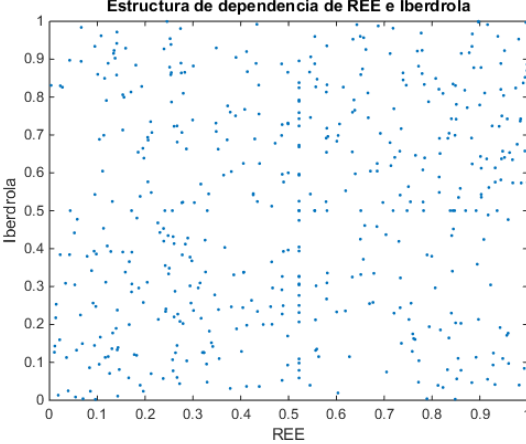
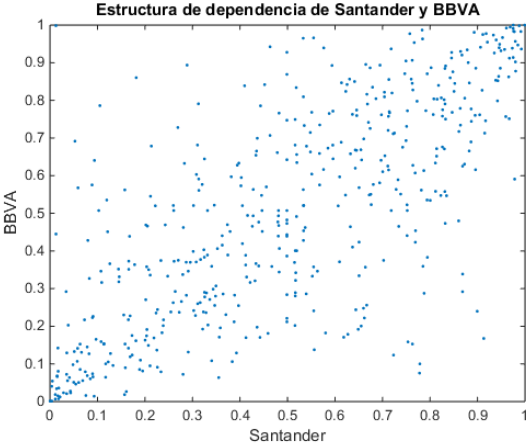
ANEXO 3.1

DISTRIBUCIONES DE PROBABILIDAD PARA CADA ACTIVO EN EL PRIMER PERÍODO DE FORMACIÓN CON FRECUENCIA DE DATOS DIARIOS (2000-2001)



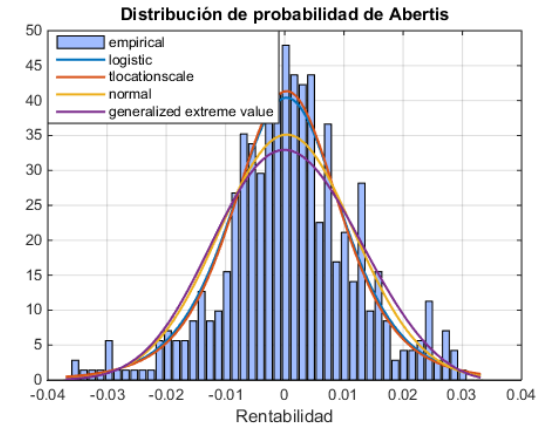
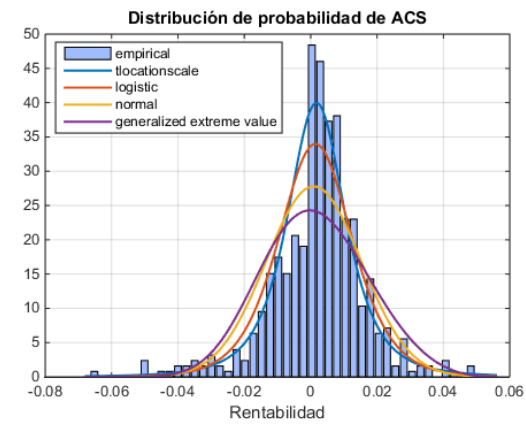
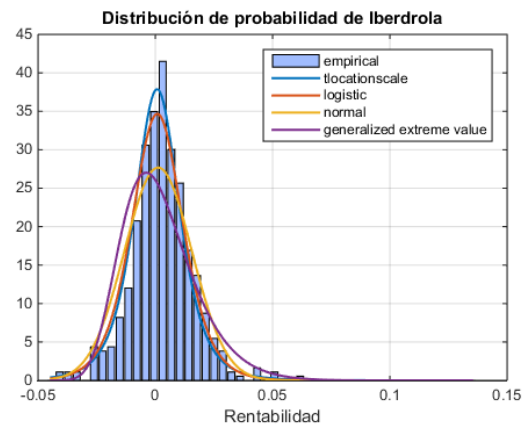
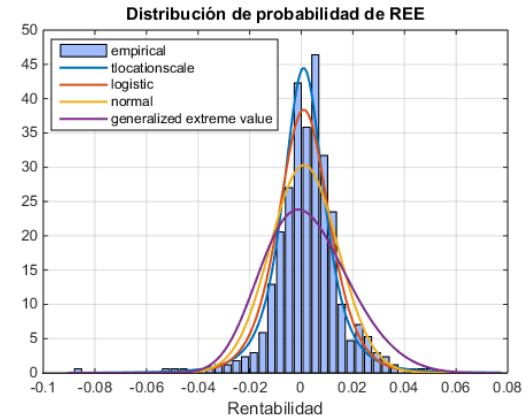
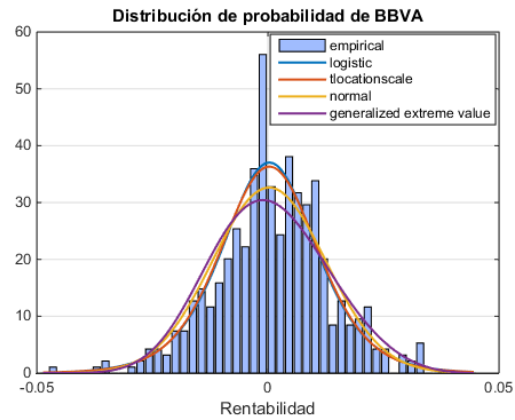
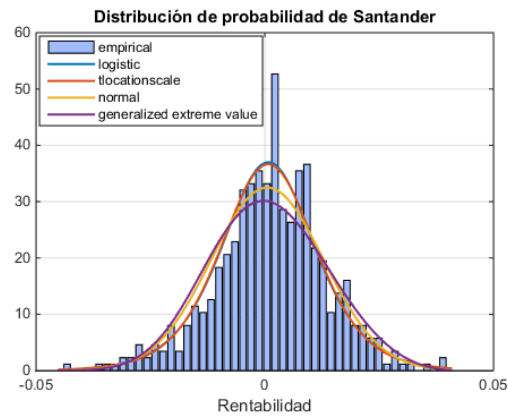


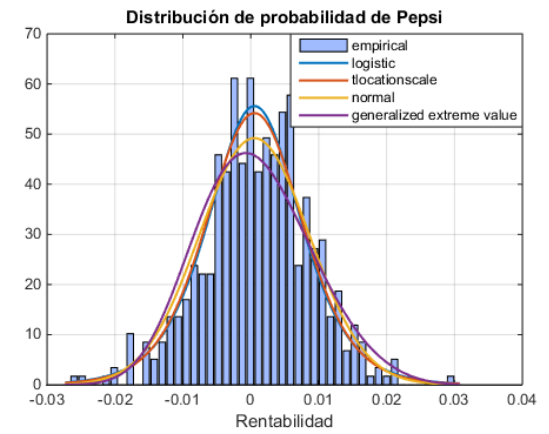
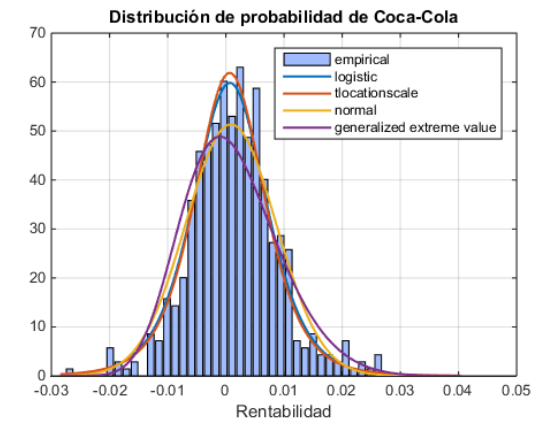
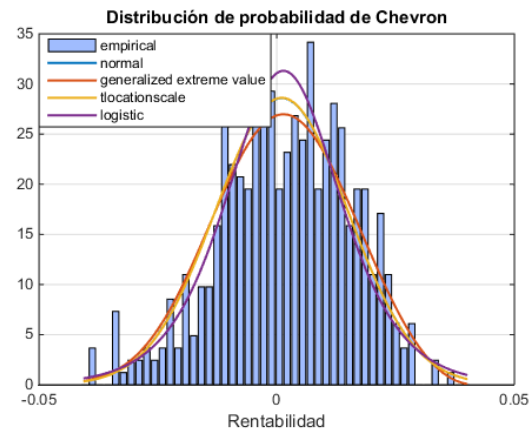
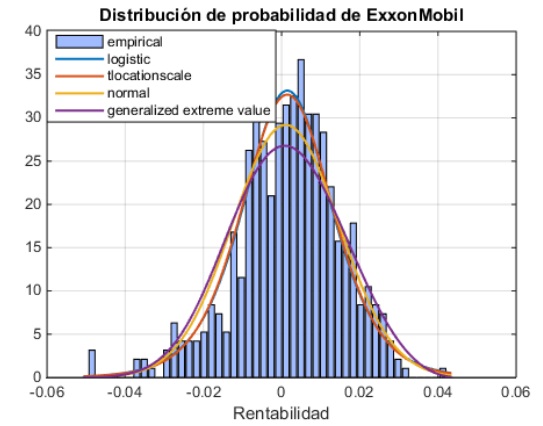
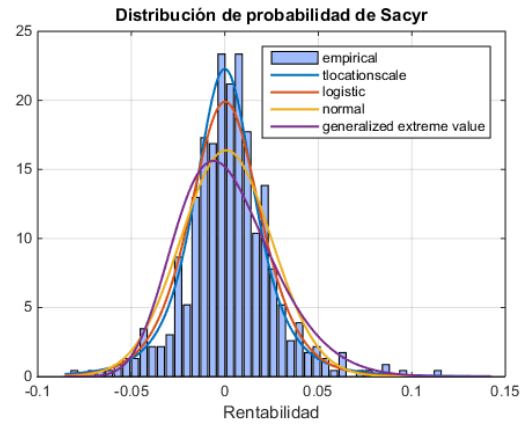
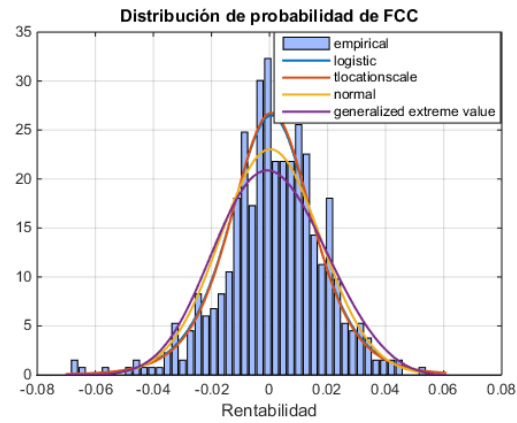
ESTRUCTURAS DE DEPENDENCIAS ENTRE CADA PAR DE ACTIVOS PARA EL PRIMER PERÍODO DE FORMACIÓN CON DATOS DIARIOS (2000-2001)



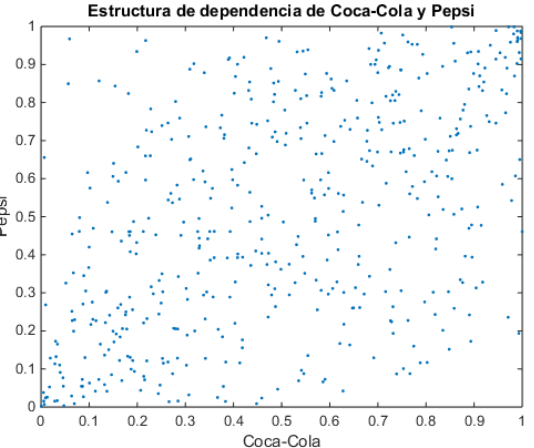
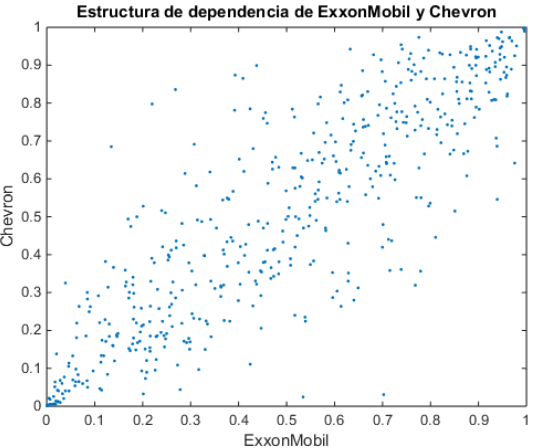
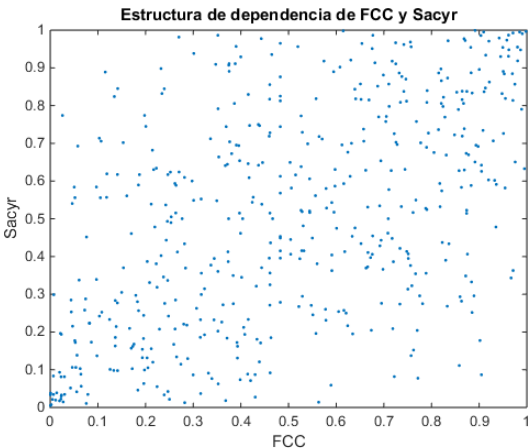
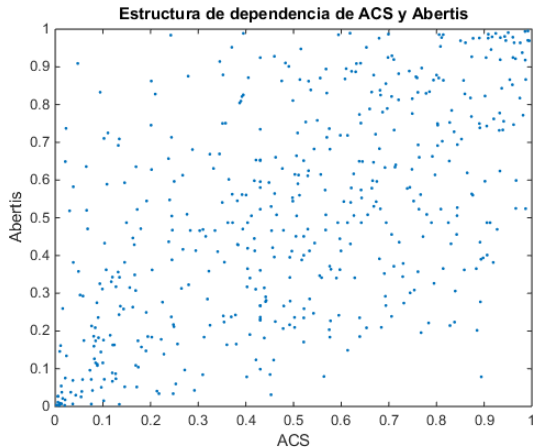
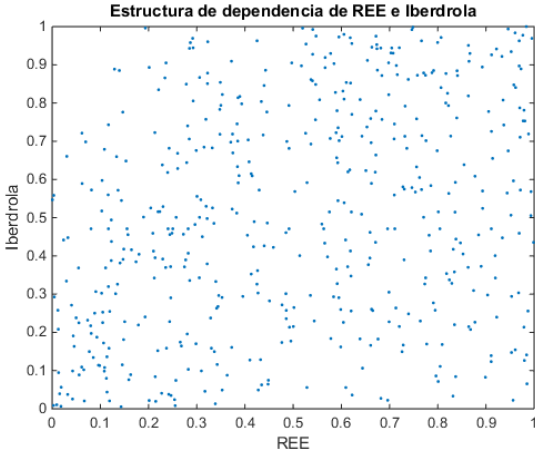
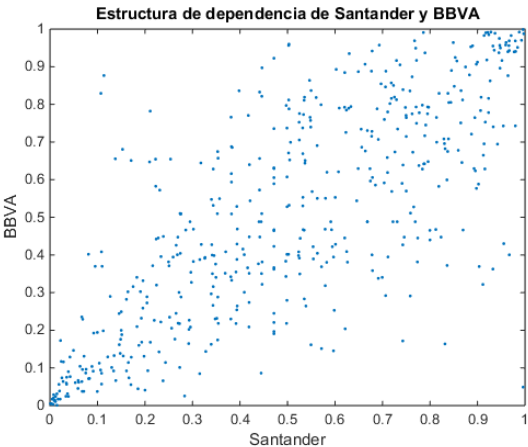
ANEXO 3.2

DISTRIBUCIONES DE PROBABILIDAD PARA CADA ACTIVO EN EL SEGUNDO PERÍODO DE FORMACIÓN CON FRECUENCIA DE DATOS DIARIOS (2006-2007)



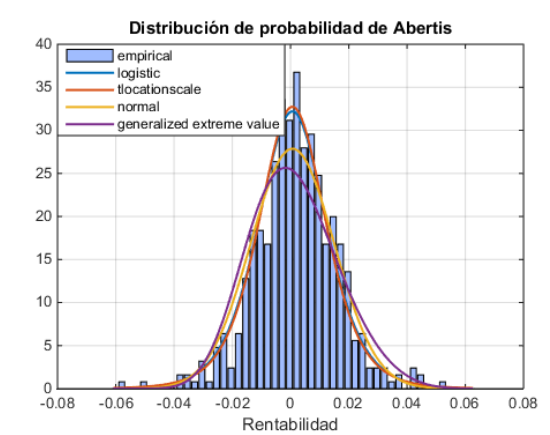
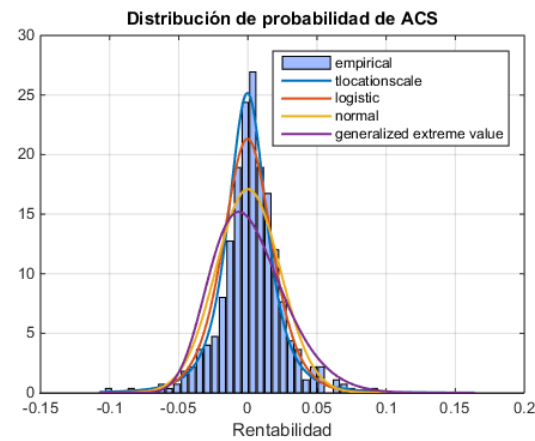
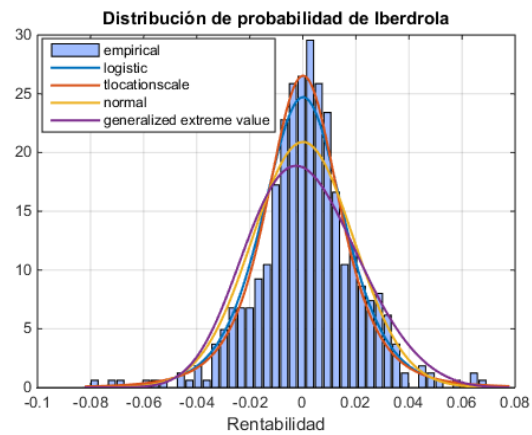
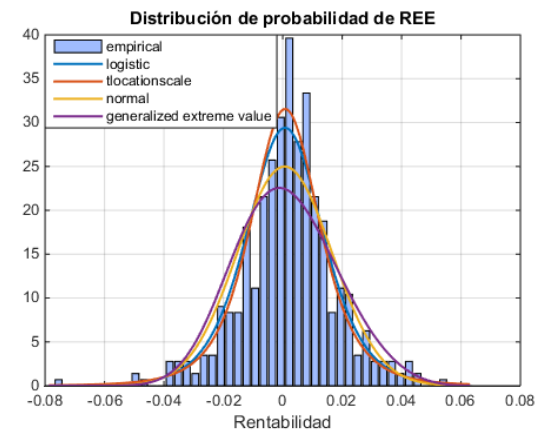
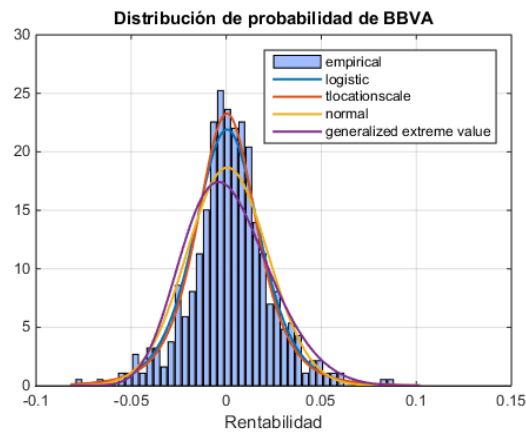
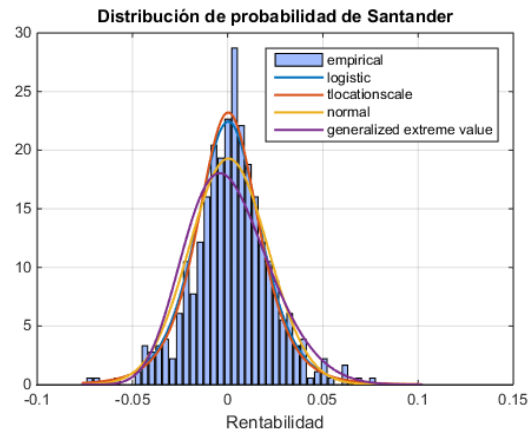


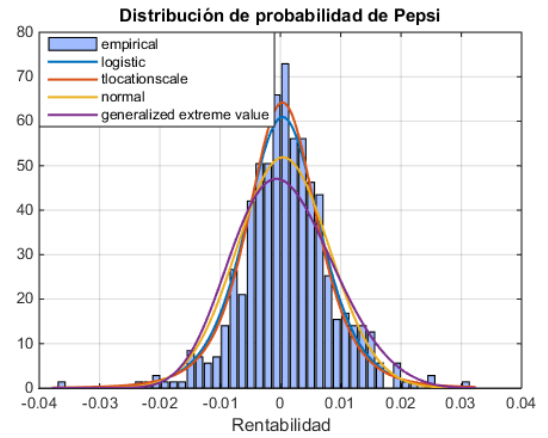
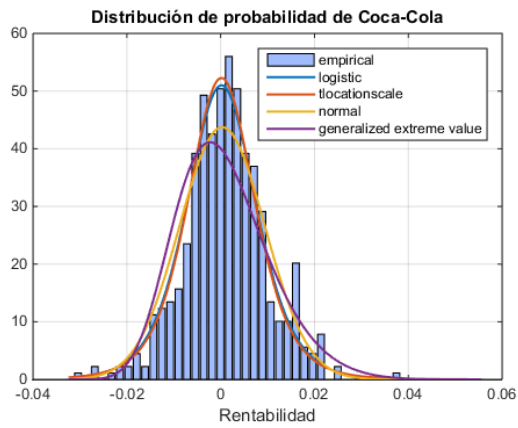
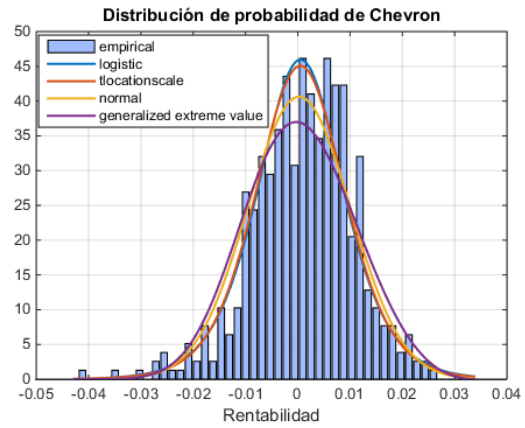
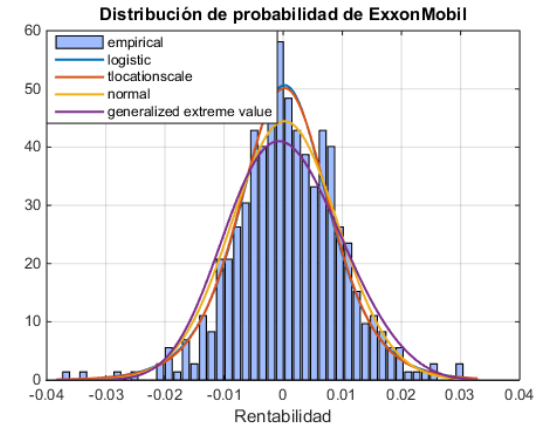
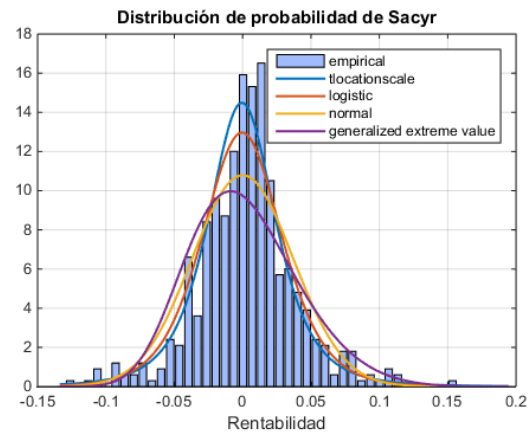
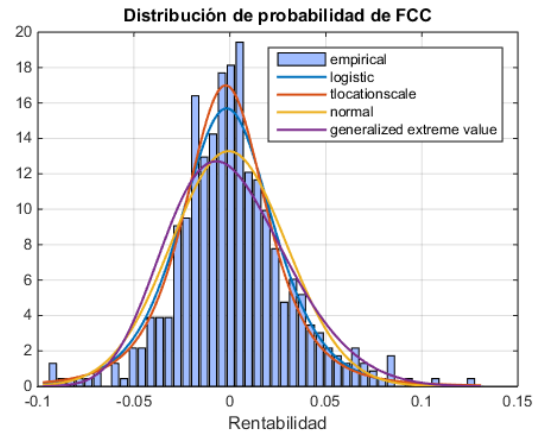
ESTRUCTURAS DE DEPENDENCIAS ENTRE CADA PAR DE ACTIVOS PARA EL SEGUNDO PERÍODO DE FORMACIÓN CON DATOS DIARIOS (2006-2007)



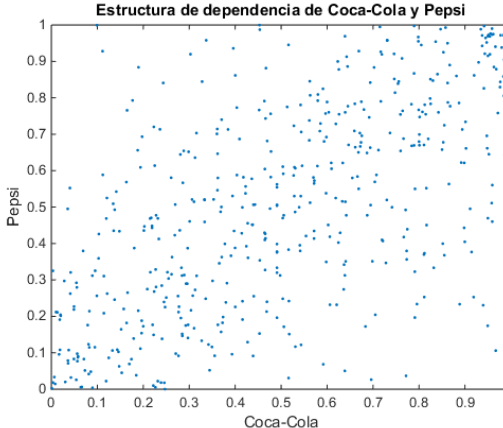
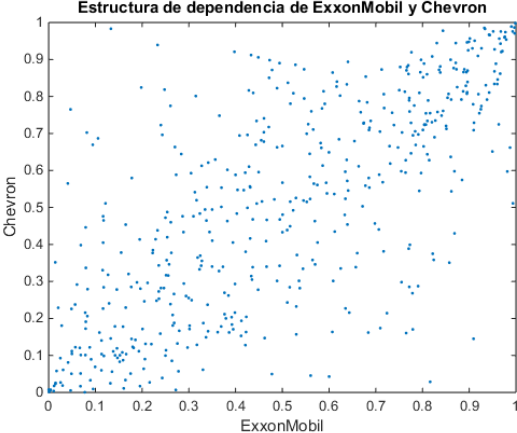
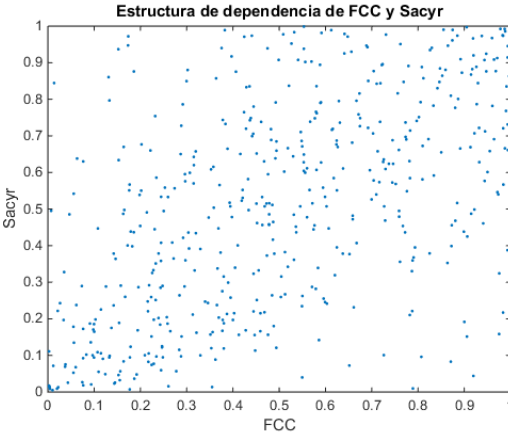
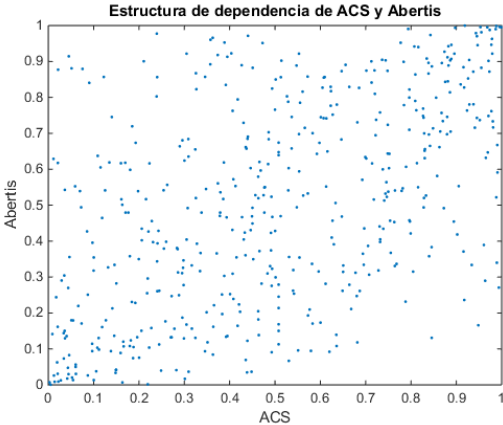
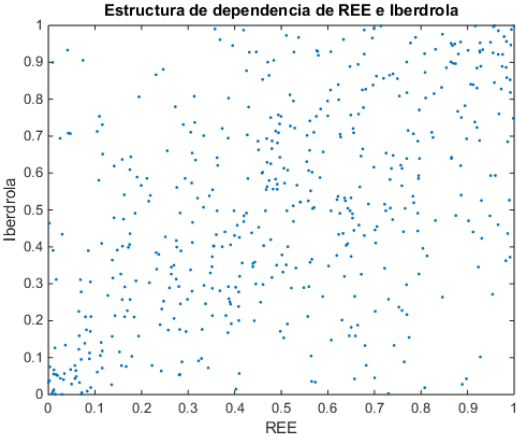
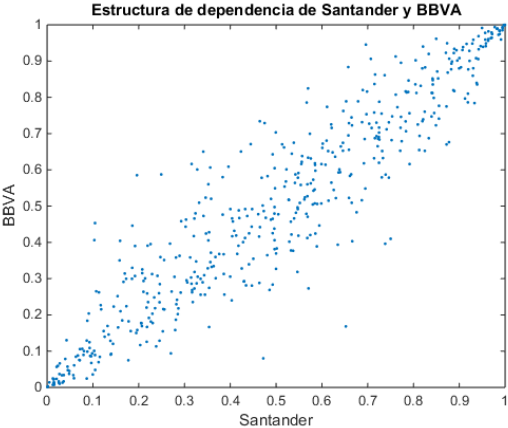
ANEXO 3.3

DISTRIBUCIONES DE PROBABILIDAD PARA CADA ACTIVO EN EL TERCER PERÍODO DE FORMACIÓN CON FRECUENCIA DE DATOS DIARIOS (2012-2013)



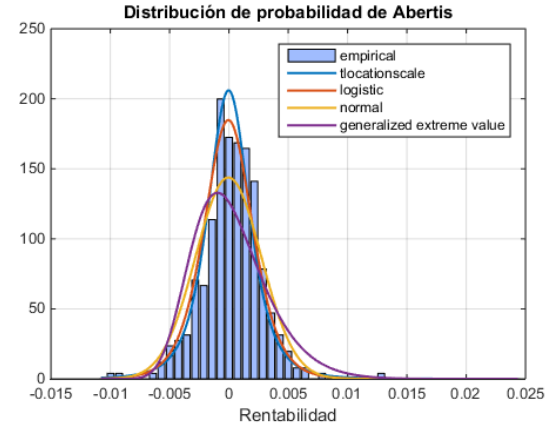
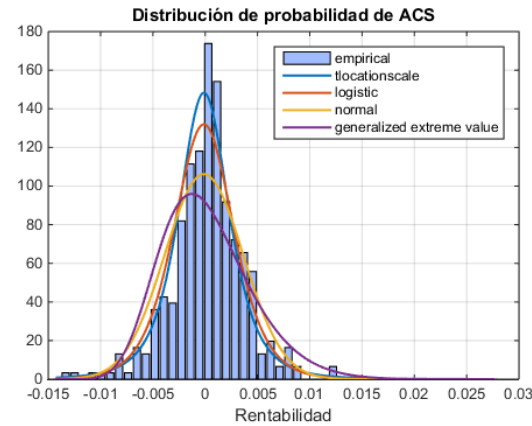
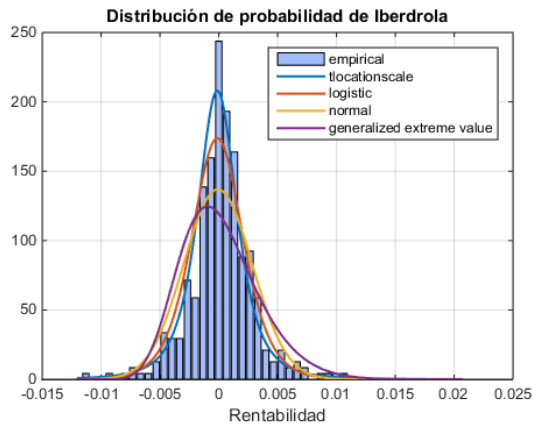
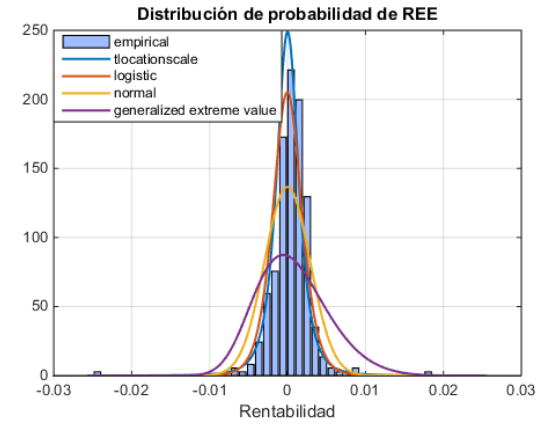
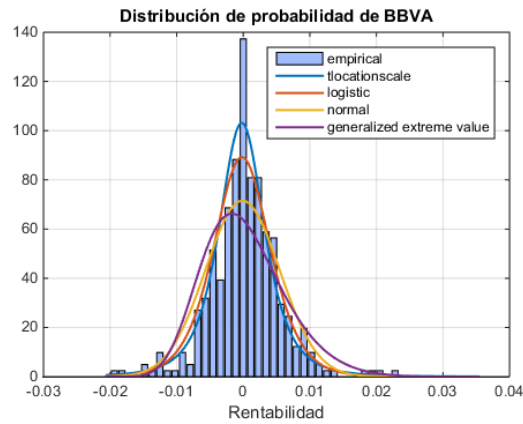
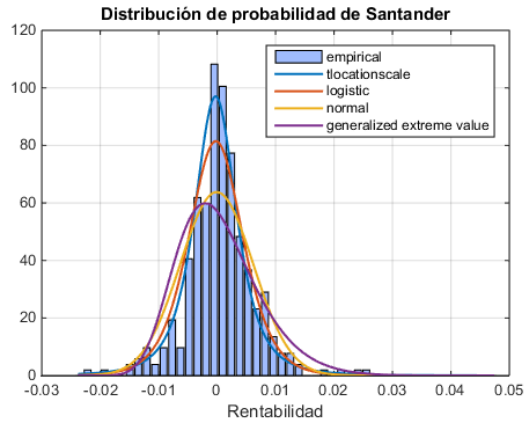


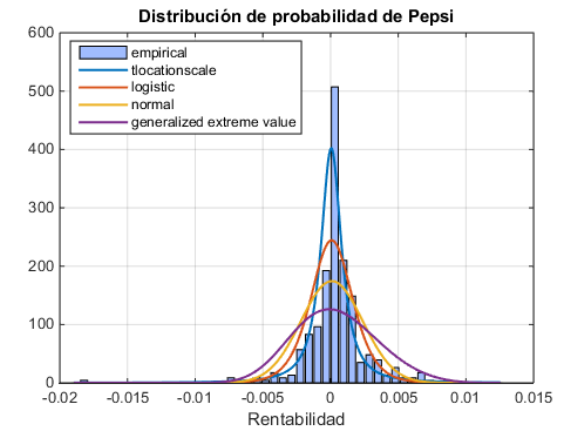
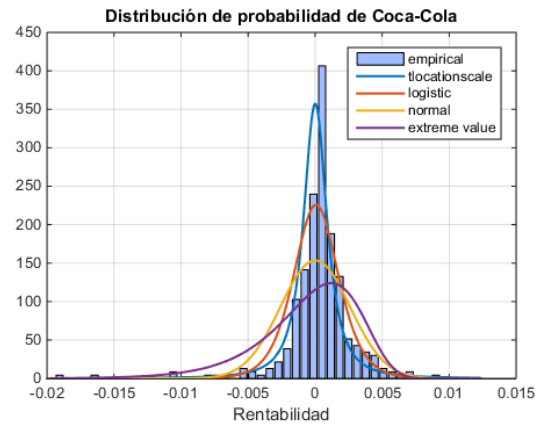
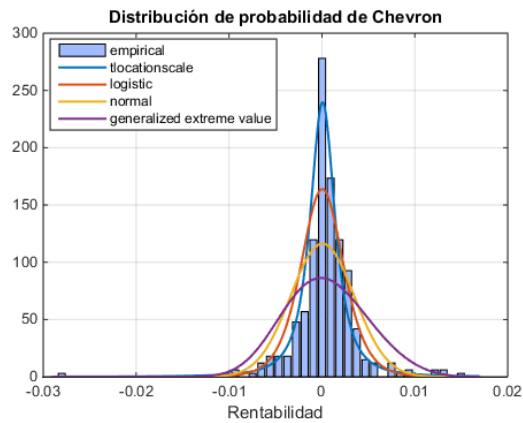
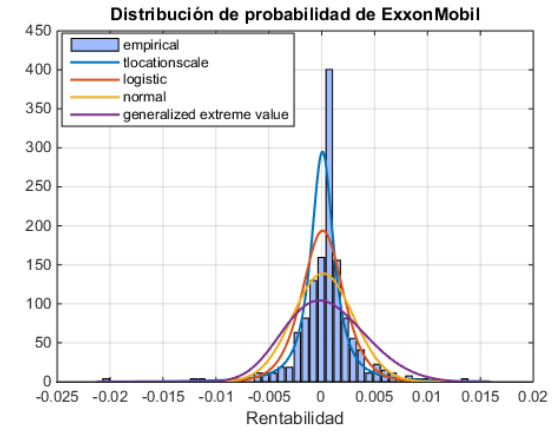
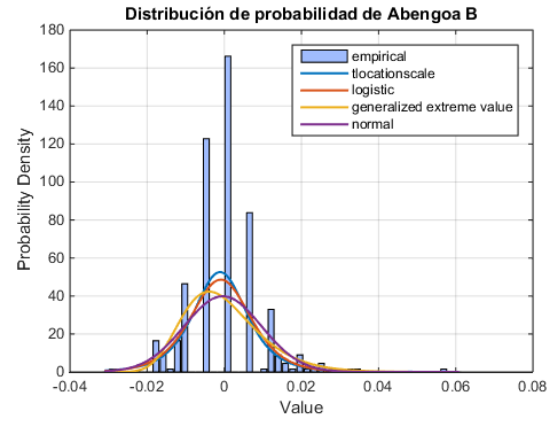
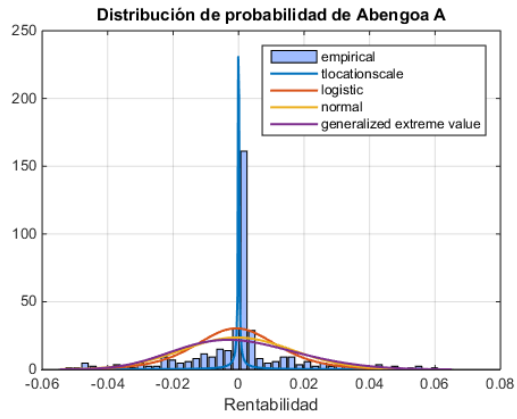
ESTRUCTURAS DE DEPENDENCIAS ENTRE CADA PAR DE ACTIVOS PARA EL TERCER PERÍODO DE FORMACIÓN CON DATOS DIARIOS (2012-2013)

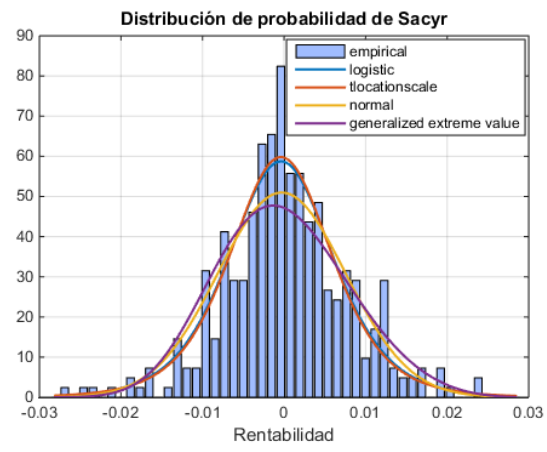
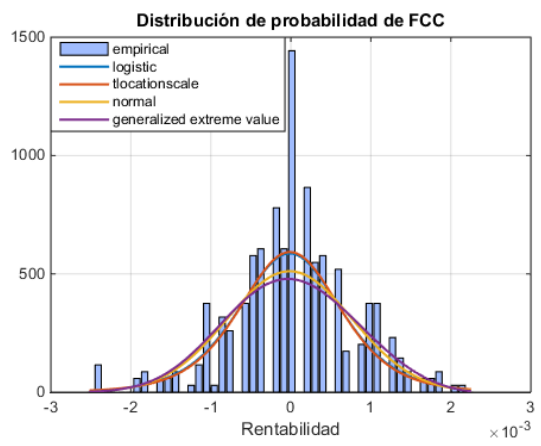


ANEXO 4.1

DISTRIBUCIONES DE PROBABILIDAD PARA CADA ACTIVO CON DATOS DE ALTA FRECUENCIA (30 MINUTOS)
CORRESPONDIENTE AL PRIMER PERÍODO DE FORMACIÓN (359 DATOS) DE LA VENTANA MÓVIL

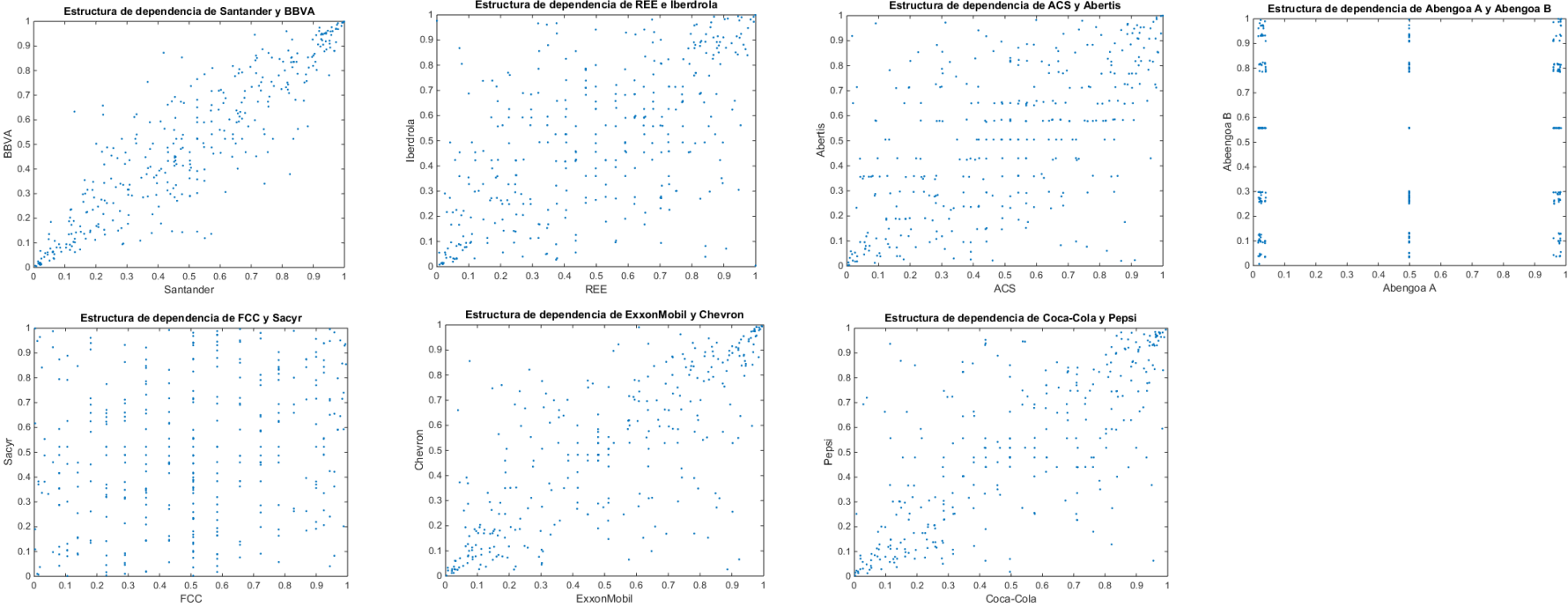






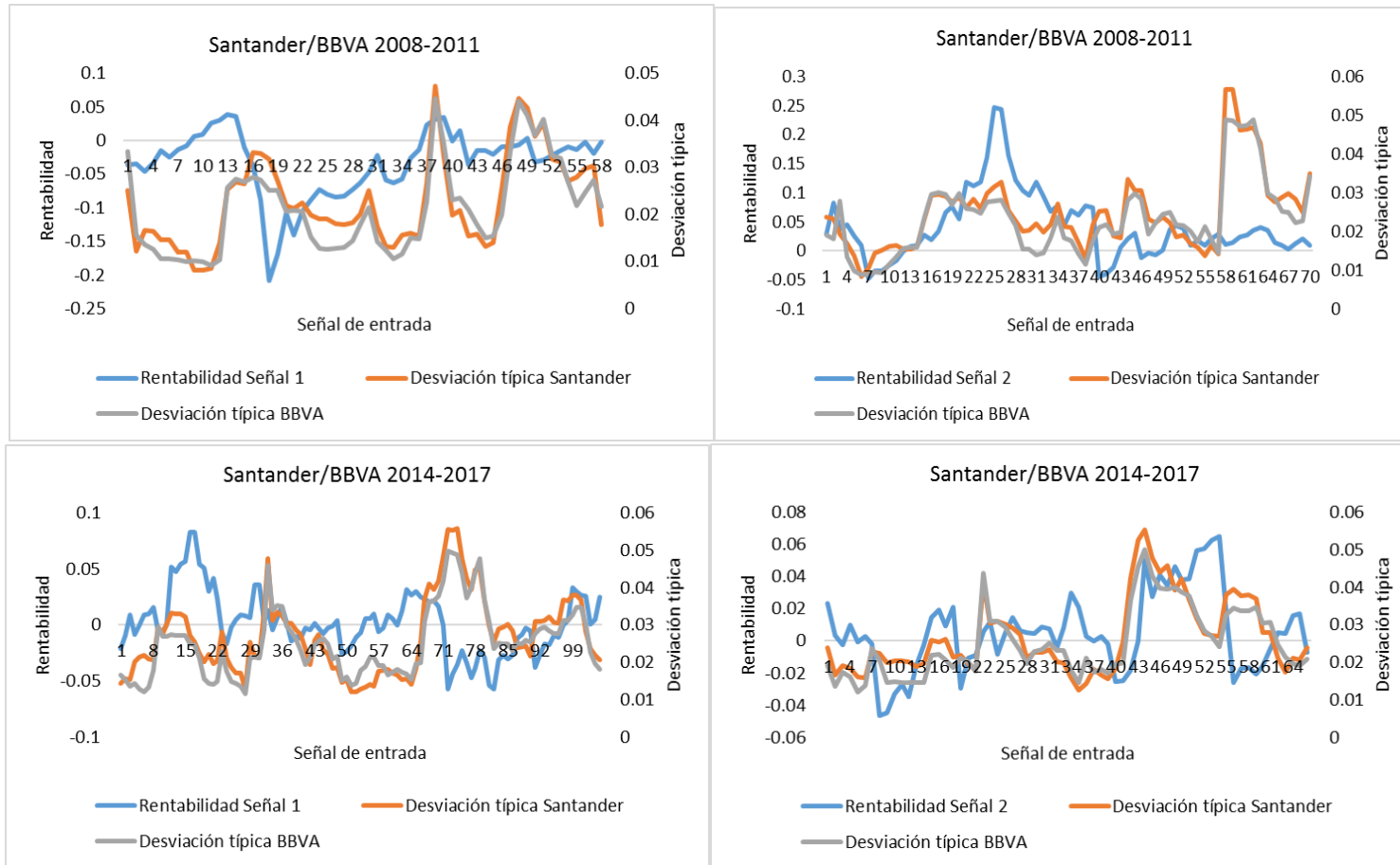
ANEXO 4.2

ESTRUCTURAS DE DEPENDENCIAS ENTRE CADA PAR DE ACTIVOS PARA EL PRIMER PERÍODO DE FORMACIÓN (1 MES, 359 OBSERVACIONES) DE LA VENTANA MÓVIL CON DATOS DE ALTA FRECUENCIA (30 MINUTOS)



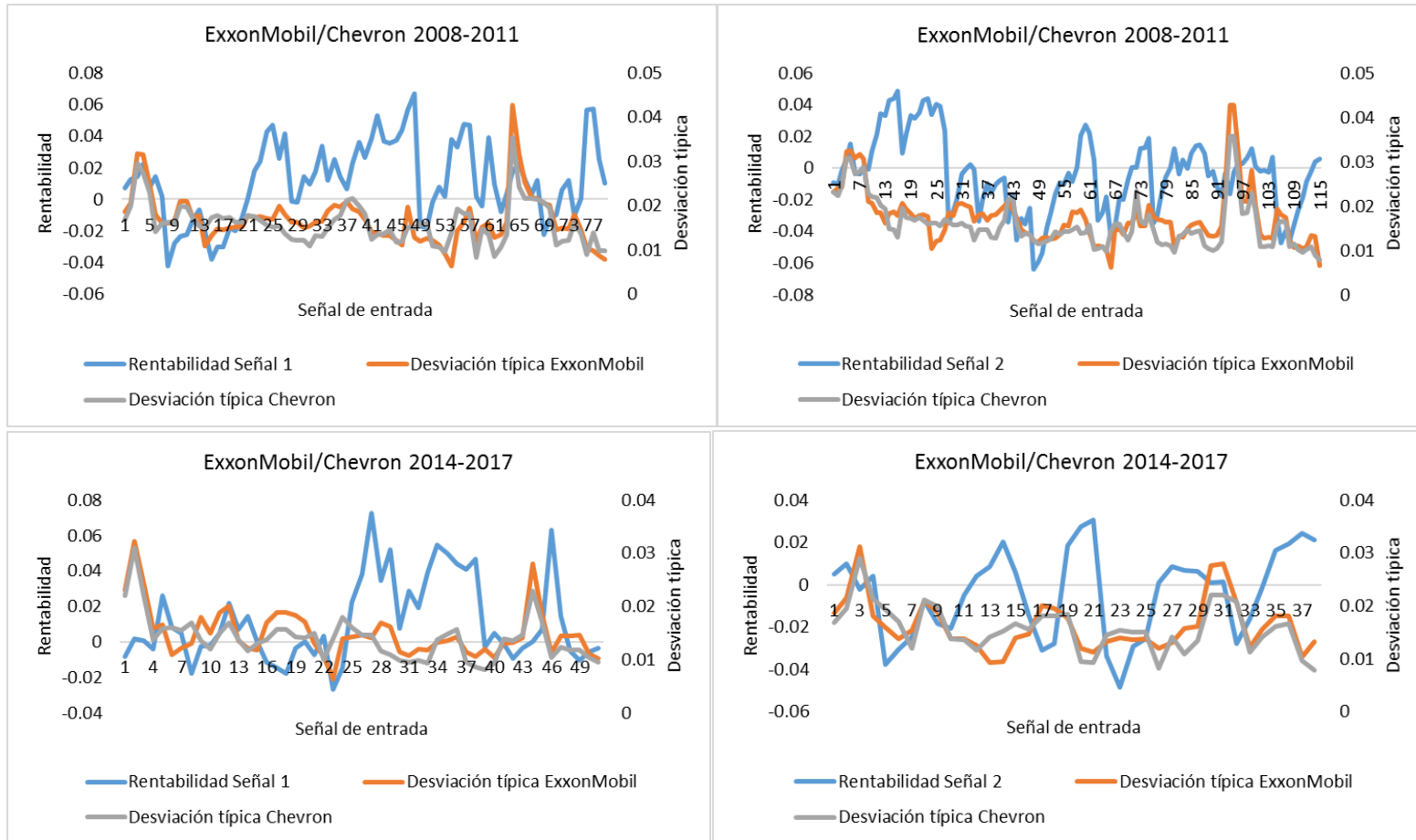
ANEXO 5.1

COMPARACIÓN ENTRE RENTABILIDAD Y DESVIACIÓN TÍPICA DE SANTANDER/BBVA



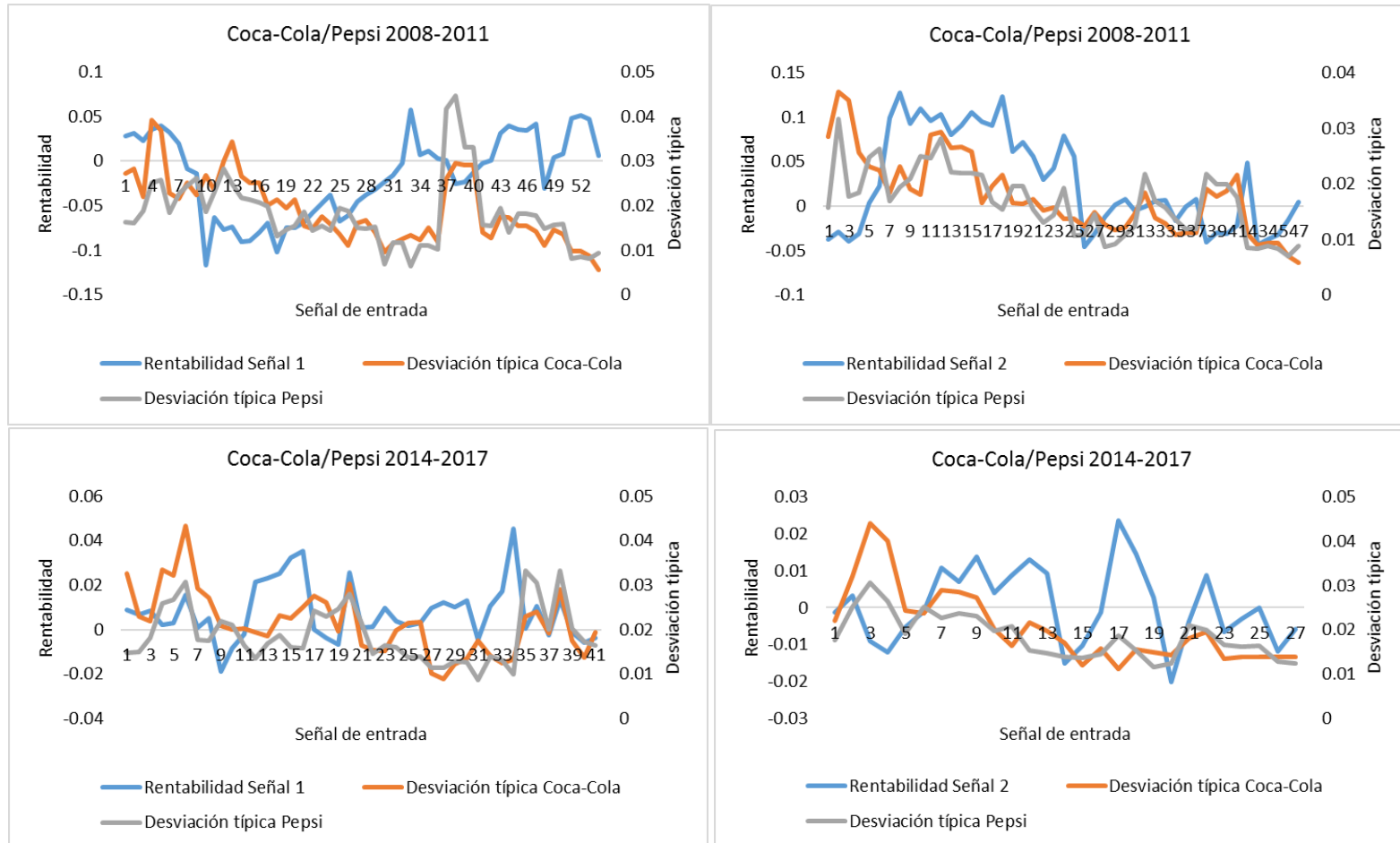
ANEXO 5.2

COMPARACIÓN ENTRE RENTABILIDAD Y DESVIACIÓN TÍPICA DE EXXONMOBIL/CHEVRON



ANEXO 5.3

COMPARACIÓN ENTRE RENTABILIDAD Y DESVIACIÓN TÍPICA DE COCA-COLA/PEPSI



BIBLIOGRAFÍA

- [1] Avellaneda, M. and Lee, J.-H. (2010). Statistical arbitrage in the US equities market. *Quantitative Finance*, 10(7):761–782.
- [2] Bertram, W. K. (2010). Analytic solutions for optimal statistical arbitrage trading. *Physica A: Statistical Mechanics and its Applications*, 389(11):2234–2243.
- [3] Bowen, D., Hutchinson, M. C., and O’Sullivan, N. (2010). High frequency equity pairs trading: transaction costs, speed of execution and patterns in returns. *The Journal of Trading*, 5(3):31–38.
- [4] Chen, H., Chen, S. J., and Li, F. (2012). Empirical investigation of an equity pairs trading strategy. *SSRN Electronic Journal*.
- [5] Do, B. and Faff, R. (2010). Does simple pairs trading still work? *Financial Analysts Journal*, 66(4):83–95.
- [6] Do, B. and Faff, R. (2012). Are pairs trading profits robust to trading costs? *Journal of Financial Research*, 35(2):261–287.
- [7] Do, B., Faff, R., and Hamza, K. (2006). A new approach to modeling and estimation for pairs trading. In *Proceedings of 2006 Financial Management Association European Conference*.
- [8] Elliott, R. J., Van Der Hoek*, John, and Malcolm, W. P. (2005). Pairs trading. *Quantitative Finance*, 5(3):271–276.
- [9] Engelberg, J., Gao, P., and Jagannathan, R. (2009). An anatomy of pairs trading: the role of idiosyncratic news, common information and liquidity. In *Third Singapore International Conference on Finance*.
- [10] Ferreira, L. (2008). New tools for spread trading. *Futures: News, Analysis & Strategies for Futures, Options & Derivatives Traders*, 37(12):38–41.
- [11] Gatev, E., Goetzmann, W. N., and Rouwenhorst, K. G. (1999). Pairs trading: Performance of a relative value arbitrage rule. Working paper, Yale School of Management’s International Center for Finance.
- [12] Gatev, E., Goetzmann, W. N., and Rouwenhorst, K. G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*, 19(3):797–827.
- [13] Huck, N. (2009). Pairs selection and outranking: An application to the S&P 100 index. *European Journal of Operational Research*, 196(2):819–825.
- [14] Huck, N. (2010). Pairs trading and outranking: The multi-step-ahead forecasting case. *European Journal of Operational Research*, 207(3):1702–1716.

- [15] Huck, N. (2015). Pairs trading: does volatility timing matter? *Applied Economics*, pages 1–18.
- [16] Jacobs, H. and Weber, M. (2013). Losing sight of the trees for the forest? Attention shifts and pairs trading. *SSRN Electronic Journal*.
- [17] Jurek, J. W. and Yang, H. (2007). Dynamic portfolio selection in arbitrage. Working paper, Harvard University.
- [18] Krauss, C. and Stübinger, J. (2015). Nonlinear dependence modeling with bivariate copulas: Statistical arbitrage pairs trading on the S&P 100.
- [19] Liew, R. Q. and Wu, Y. (2013). Pairs trading: A copula approach. *Journal of Derivatives & Hedge Funds*, 19(1):12–30.
- [20] Lin, Y.-X., McCrae, M., and Gulati, C. (2006). Loss protection in pairs trading through minimum profit bounds: A cointegration approach. *Journal of Applied Mathematics and Decision Sciences*, 2006(1):1–14.
- [21] Liu, J. and Timmermann, A. (2013). Optimal convergence trade strategies. *Review of Financial Studies*, 26(4):1048–1086.
- [22] Nath, P. (2003). High frequency pairs trading with U.S. treasury securities: Risks and rewards for hedge funds. Working paper, London Business School.
- [23] Papadakis, G. and Wysocki, P. (2007). Pairs trading and accounting information. Working paper, Boston University and MIT.
- [24] Puspaningrum, H. (2012). Pairs trading using cointegration approach. PhD thesis, University of Wollongong.
- [25] Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360.
- [26] Stander, Y., Marais, D., and Botha, I. (2013). Trading strategies with copulas. *Journal of Economic and Financial Sciences*, 6(1):83–107.
- [27] Vidyamurthy, G. (2004). Pairs trading: Quantitative methods and analysis. John Wiley & Sons, Hoboken, N.J.
- [28] Wu, Y., Zou, X., Liew, R.Q., Xie, W. (2014). Pairs trading with copulas.