



Universidad de
San Andrés

UNIVERSIDAD DE SAN ANDRÉS
ESCUELA DE NEGOCIOS
LICENCIATURA EN FINANZAS

*Selección de activos del S&P 500 y el FTSE 100 con
PCA durante la crisis del COVID-19*

Autoras: Micaela Arias (29017), Juana De Rosa (28038)
Tutora: Magdalena Cornejo

Buenos Aires, Argentina, diciembre 2021

Índice

1. Introducción	5
2. Revisión de la literatura.....	8
2.1 PCA en la construcción de portafolios	8
2.2 PCA en la selección de activos	8
3. Metodología	12
3.1. Cálculo de los componentes principales.....	12
3.1.1. <i>¿Sobre qué matriz debería aplicarse PCA?</i>	15
3.1.2. <i>¿Cuántos componentes deberían retenerse?</i>	15
3.2. Construcción de portafolio y selección de activos	16
4. Datos	19
4.1. Datos	19
4.1. Análisis Descriptivo	20
5. Resultados	23
5.1 S&P 500	23
5.1.1 <i>Portafolio PCA para la region estadounidense</i>	23
5.1.2 <i>Selección de activos para la region estadounidense</i>	29
5.2 FTSE 100.....	39
5.1.1 <i>Portafolio PCA para la region britanica</i>	39
5.1.2 <i>Selección de activos para la region britanica</i>	43
6. Conclusiones	49
7. Bibliografía	51
8. Anexos	52

Índice de figuras

1. Análisis descriptivo del S&P 500	21
2. Análisis descriptivo del FTSE 100	22
3. Matriz de Correlaciones de los activos del S&P 500	23
4. Componentes a retener en la segunda ventana para los activos del S&P 500	24
5. Proporción de la varianza explicada por los componentes principales (Activos S&P 500).....	26
6. Primer Componente Principal (S&P 500).....	27
7. Segundo Componente Principal (S&P 500).....	27
8. Portafolio PCA vs S&P 500.....	28
9. Portafolio 50 Activos vs S&P 500	31
10. Portafolio 50 Activos vs Portafolio 74 Activos	34
11. Portafolio 50 Activos vs Portafolio 25 Activos	36
12. Portafolio 50 Activos vs Portafolio 15 Activos	37
13. Portafolios en comparación contruidos a partir de los activos del S&P 500	39
14. Matriz de Correlaciones de los activos del FTSE 100	40
15. Proporción de la varianza explicada por los componentes principales (Activos FTSE 100).....	41
16. Primer Componente Principal (FTSE 100).....	42
17. Portafolio PCA vs FTSE 100.....	42
18. Portafolio 9 Activos vs Portafolio FTSE 100	44
19. Portafolio 9 Activos vs Portafolio 14 Activos	46
20. Portafolios en comparación contruidos a partir de los activos del FTSE 100	48

Índice de tablas

1. Compañías contenidas en el percentil 0.9 (Activos S&P 500).....	30
2. Compañías contenidas en el percentil 0.85 (Activos S&P 500).....	33
3. Ratio de Sharpe para los portafolios contruidos a partir de los activos del S&P 500.....	34
4. Compañías contenidas en el percentil 0.95 (Activos S&P 500).....	35
5. Compañías contenidas en el percentil 0.97 (Activos S&P 500).....	36
6. Compañías contenidas en el percentil 0.03 (Activos S&P 500).....	38
7. Compañías contenidas en el percentil 0.90 (Activos FTSE 100).....	43
8. Compañías contenidas en el percentil 0.85 (Activos FTSE 100).....	45
9. Ratio de Sharpe para los portafolios contruidos a partir de los activos del FTSE 100	46
10. Compañías contenidas en el percentil 0.15 (Activos FTSE 100).....	47

RESUMEN

Esta tesis propone una estrategia basada en el Análisis de Componentes Principales (PCA, por sus siglas en inglés) para la gestión activa de selección de activos en todas las fases de mercado.

Nuestra investigación se centra en la selección de activos individuales dentro del *asset class equity*. El periodo comprendido de estudio irá desde febrero de 2017 hasta febrero de 2021. Esto tiene como objetivo observar el efecto que tuvo la crisis del COVID-19. Para ello, utilizaremos dos mercados de valores representativos, Estados Unidos y el Reino Unido. Utilizaremos como *benchmarks* los índices de S&P 500 y FTSE 100.

En primer lugar, construiremos un portafolio a partir de los componentes principales obtenidos a través de los activos que se encuentran integrados en los índices mencionados. Esto tendrá como objetivo evaluar la efectividad del método de PCA para modelar el comportamiento de un *benchmark*. En segundo lugar, buscaremos observar el efecto de la pandemia en el portafolio construido. A partir de esto, utilizaremos otro criterio de selección de activos para así mitigar el efecto del COVID-19 segmentando en función de las empresas más y menos afectadas por este evento. De esta manera obtendremos un portafolio con una menor cantidad de activos, que tendrá un rendimiento significativamente mejor al del *benchmark*.

1. Introducción

Existen diversas opciones e instrumentos de inversión en los mercados financieros. Los inversores e investigadores cada vez se están dedicando más a estudiar con mayor detenimiento cuáles son los *asset classes* y activos que presentan un rendimiento futuro favorable para que sean componentes de las distintas carteras de inversiones. Sin embargo, siempre existió el debate sobre la importancia relativa de la selección de *asset classes* y la selección de activos individuales. Esto es así porque la creencia popular establece que la asignación de distintas clases de activos es el factor dominante que impulsa el rendimiento relativo en todas las carteras. Es por esto que en la presente tesis existe una motivación para buscar implementar estrategias que bien pueden servir como método de selección activo dentro de un determinado *asset class* o bien para inversores que buscan definir en qué activos invertir individualmente.

Existen varios métodos por los cuales los inversores pueden elegir seleccionar activos. Mientras que algunos inversores utilizan datos económicos para seleccionar sus activos, otros utilizan datos técnicos con la ayuda de gráficos para determinar la eficiencia, y otros se basan en datos fundamentales como ratios financieros (por ejemplo, rentabilidad de una acción, rendimiento de activos, etc.). También existen inversores que utilizan una combinación de información técnica y fundamental para determinar cuáles son los mejores activos. Si bien hay muchas formas en las que uno puede seleccionarlos, aquí proponemos un método que permite reducir la complejidad y la alta dimensionalidad de los datos.

El análisis de componentes principales (PCA, por sus siglas en inglés) es una técnica de análisis multivariado que permite reducir la dimensionalidad de los datos altamente correlacionados. Esto es de gran importancia ya que cuando solo hay unos pocos activos involucrados resulta fácil observar las varianzas y las covarianzas o correlaciones. Pero cuando se quiere incorporar un gran número de activos a un portafolio el análisis no resulta tan sencillo. Entonces, PCA nos proporciona un enfoque alternativo que es capaz de reducir la complejidad: nos devuelve componentes principales que preservan la mayor parte de la información que proviene de esas varianzas y covarianzas o correlaciones y además resulta en un número menor de variables a estudiar.

Es también debido a esto que este enfoque estadístico reduce el problema de la diversificación. En lugar de construir una cartera basada en los activos subyacentes, se pueden tratar los componentes principales como activos de inversión. Además, cuando el número de variables en un conjunto de datos es grande, a menudo ocurre que muchas variables contienen información redundante. Por ello, PCA permite obtener los componentes relevantes dentro de un conjunto. Utilizando esta estrategia el tamaño de la cartera resultante es más significativo desde el punto de vista de la diversificación.

A su vez, la elección aleatoria de activos para agregar a una cartera, incluso cuando alcanzan el número requerido, puede no resultar en la diversificación prometida si los activos elegidos al azar estuvieran más correlacionadas de lo esperado. El análisis de componentes principales viene a solucionar este problema debido a que este método permite transformar las variables originales en nuevas variables incorrelacionadas.

El problema de diversificación se vio claramente expresado durante la crisis financiera de 2008 cuando se observó que no resultaba posible diversificar un portafolio de manera eficaz. Cuando el mercado se vuelve más conectado, el riesgo sistémico es mayor en el sentido de que los choques negativos se propagan más rápida y ampliamente. En los mercados financieros reales las correlaciones se mueven de manera dinámica, a medida que van cambiando junto con las fases de mercado. Es por ello que algunos críticos argumentaron que la crisis financiera de 2008 tuvo como resultado carteras poco diversificadas que además no proporcionaron el control de riesgo esperado cuando más se necesitaba (Allen, 2010). Esta tendencia de mercado también se vio reflejada como consecuencia de la pandemia producto del COVID-19. Esta pandemia tuvo un enorme impacto negativo tanto en la tasa de desempleo como en el Producto Bruto Interno (PBI) de los países, que provocó una gran incertidumbre económica y financiera por la duración y profundidad de la recesión conduciendo así a altas volatilidades en el mercado de valores. Como consecuencia, las grandes bolsas de valores experimentaron una profunda caída de su valor durante el primer trimestre del año, siendo este el desplome histórico más rápido que transcurrió en tan solo un mes (de febrero a marzo de 2020). De esta forma, las tendencias de mercado comenzaron a ser bajistas y las correlaciones tendieron a uno.

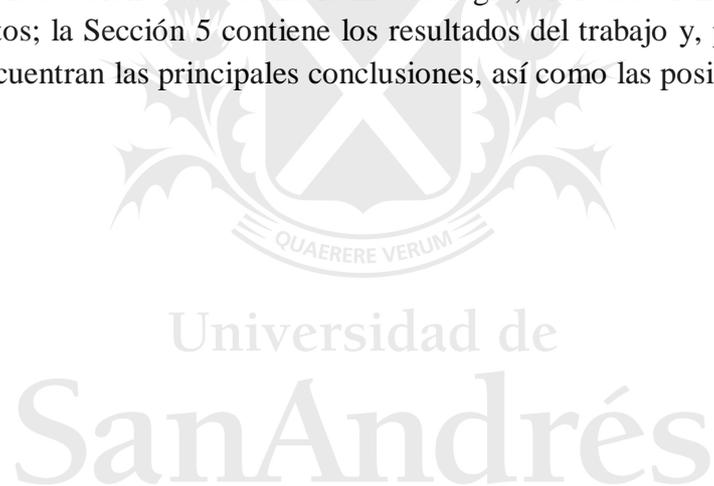
En los mercados financieros reales, los movimientos de los activos suelen estar sincronizados debido a que hay factores de riesgo comunes, lo que suele perjudicar el rendimiento de las carteras en tiempos turbulentos. Por esta razón, el seguimiento de la evolución temporal de la correlación es fundamental dado que esta se profundiza cuando los mercados comienzan a experimentar tendencias bajistas. Es decir, resulta necesario observar esto ya que para hacer posible la diversificación necesitamos que las correlaciones entre los activos sean bajas o negativas. En esta línea, PCA puede servir como estrategia para modelar el comportamiento de un conjunto de activos a partir de los componentes. Sin embargo, es una técnica donde dentro de cada componente principal se encuentran incluidos todos los activos, aunque algunos tengan peso bajo. Es por esto que no solo buscamos modelar un índice, sino que, con el objetivo de evitar las grandes correlaciones entre activos, buscamos establecer un método de selección de activos que haga posible mitigar los efectos de las determinadas crisis. A partir del análisis de componentes principales se puede descubrir y retener la información relevante en el conjunto de datos financieros.

La investigación se enfoca en el uso potencial de PCA para la selección de activos. En particular, en la selección de acciones ya que estas son los instrumentos de inversión más populares en los mercados financieros. Por un lado, buscamos observar si al implementar

PCA resulta posible modelar el comportamiento de un conjunto de activos (índices) a partir de los componentes principales. Por el otro, buscamos diseñar una estrategia de selección de activos que pueda servir en los distintos ciclos de mercado, específicamente cuando están en baja. Es por esto que el periodo de análisis se centra en la crisis generada por el COVID-19.

Si bien nuestro método se aplica en la selección de acciones, de hecho, es bastante general y puede servir para cualquier conjunto de oportunidades de inversión, siempre que se puedan estimar las correlaciones entre los activos. Los resultados podrían ser interesantes para una amplia variedad de profesionales, como administradores de riesgos que desean monitorear la volatilidad y las correlaciones entre activos, administradores de carteras que desean comprender cómo seleccionar activos dentro de una determinada clase e inversores que desean saber cuántos activos se requieren para seguir un índice de referencia.

Esta tesis está estructurada de la siguiente manera. La Sección 2 contiene la revisión de la literatura; en la Sección 3 se detalla la metodología; en la Sección 4 se presenta la selección de datos; la Sección 5 contiene los resultados del trabajo y, por último, en la Sección 6 se encuentran las principales conclusiones, así como las posibles limitaciones del trabajo.



2. Revisión de la literatura

2.1. PCA en la construcción de portafolios

Partovi y Caputo (2004) fueron los primeros autores en proponer la idea de utilizar el método de PCA para analizar el problema de la construcción de portafolios eficientes. Esta idea surge del hecho de que el análisis de los portafolios se simplificaría si los precios de los activos estuvieran incorrelacionados. Los autores argumentan que, si la venta en corto está permitida, cualquier conjunto de activos puede ser transformado en componentes principales. De esta manera, cada componente principal obtenido es tratado como una cartera de inversión individual denominada “portafolio principal”. A partir de dichos portafolios los autores pueden construir una frontera eficiente y concluyen que la estructura de esta se encuentra más relacionada a los portafolios construidos a partir de los componentes principales que a los datos originales.

Por su parte, Attilio Meucci (2010) continuó con la idea de Patovi y Caputo (2004) utilizando el método de PCA para transformar un set de activos correlacionados en un nuevo conjunto de portafolios principales incorrelacionados. El autor escribe sobre la gestión de la diversificación. La diversificación, según él, puede ser definida como “(...) un portafolio está bien diversificado si no está fuertemente expuesto a *shocks* individuales”. De esta forma, el autor propone utilizar la técnica de PCA para expresar la cartera en términos de exposiciones a fuentes de riesgo no correlacionadas, pero establece que PCA puede no resultar manejable en carteras restringidas. Es por esto que propone utilizar un portafolio principal condicional. Así, interpretará las volatilidades como un conjunto de masas probabilísticas.

El trabajo de Meucci (2010) extrajo los componentes principales que determinan la variabilidad de los activos. Estos pueden ser interpretados como portafolios principales representando los riesgos incorrelacionados inherentes a los activos de la cartera. Para que los portafolios estén bien diversificados, su riesgo general tendría que estar distribuido uniformemente a través de estos portafolios principales. Es por esto que el autor introduce un índice de diversificación que representa el número efectivo de apuestas no correlacionadas en una cartera. Es así como, controlando los términos de esta distribución, el administrador puede gestionar eficientemente la diversificación. Entonces, Meucci (2010) propone un enfoque valioso para describir la concentración del riesgo: utiliza una contribución al riesgo para expresar la concentración de volatilidad por componentes y concluye que un portafolio bien distribuido es uno en el que los factores de riesgo se invierten en partes iguales.

2.2. PCA en la selección de activos

Muchas investigaciones discuten métodos para seleccionar con precisión las acciones que presentan un rendimiento futuro favorable para poder incorporarlas dentro de una cartera de inversiones. Sin embargo, el uso de PCA como método de selección de acciones no

está tan detallado en la literatura a pesar de que algunos modelos tradicionales de selección de activos generalmente enfrentan desafíos cuando se trata procesos multidimensionales no lineales.

El uso de PCA resuelve el problema asociado con la elección aleatoria de acciones. Rudin y Morgan (2006) aplicaron PCA para medir la diversificación cuantitativamente, seleccionando activos a partir de los componentes principales. Para ello construyeron un índice de diversificación de portafolios (PDI, por sus siglas en inglés) aplicando PCA. Esto se da ya que los escritores argumentan que la ventaja del enfoque PCA es que permite a un administrador reducir el problema de la diversificación dado que se reducen los datos a un número mucho menor de variables sin perder información relevante. Además, este método evalúa el número efectivo de componentes incorrelacionados dentro de un portafolio. Para ello, los autores comparan las carteras de acciones, ponderadas de igual manera, con el índice S&P 100. De esta forma, observan que un grupo de 40 acciones seleccionadas al azar se encuentra aproximadamente tan diversificado como solo 20 componentes ortogonales. Es por esto que concluyen que este método, además de ayudar a los profesionales a comprender la diversificación, también puede ayudar en la construcción de portafolios, no sólo aplicado a acciones sino con cualquier tipo de *asset class*.

Años después, Fulga y Dedu (2012) aplican PCA para seleccionar los activos de un portafolio y luego el método de clustering para agrupar activos similares, evalúan el riesgo de dichos activos con la medida de *Value at Risk* (VaR) y así construyen la cartera óptima. Este estudio propone utilizar las técnicas de análisis de datos solamente para la información relevante de los datos. El método de PCA, es utilizado, para reducir la dimensionalidad de los datos, así como también, transformar el conjunto de variables correlacionadas en un conjunto de variables incorrelacionadas de la misma dimensión eligiendo las que contienen mayor información y despreciando las que tienen información irrelevante, y entonces, identificar nuevas variables subyacentes significativas. Luego, usan técnicas de clasificación jerárquica para así construir *clusters* compuestos por acciones similares.

Al tener los componentes principales agrupados en diferentes *clusters*, se enfocan en seleccionar activos de mínimo riesgo. En primer lugar, evalúan el resultado de los activos utilizando el retorno logarítmico diario. Dado que en la mayoría de los casos no se conoce la distribución de la variable aleatoria de riesgo, se requiere un método para evaluar o aproximar el VaR. Existen tres enfoques típicos para estimar el VaR: el método analítico o paramétrico, el método de simulación histórica o empírico y el método de simulación de Monte-Carlo. En este caso, ponen foco en el método de simulación histórica y así calculan el Var y eligen el activo con menor VaR para un nivel de probabilidad específico (en este caso 0.99).

Al tener agrupados los activos de mínimo riesgo, se debe optimizar el portafolio utilizando el marco de riesgo medio. Lo cual consiste en encontrar el portafolio de retorno

máximo tal que el VaR de la pérdida de la cartera no exceda cierto límite superior. Así es como construyen entonces la frontera eficiente. Esto, entonces, tiene en consideración tanto la selección de activos como la estimación y optimización del riesgo. Así es como, este enfoque presenta una ventaja importante, ya que garantiza la diversificación de dicha cartera.

Más tarde, Yang et al. (2015) proponen un modelo de selección de acciones que se basa en el análisis de componentes principales. Estos abordan la cuestión de cuántas acciones se necesitan para crear una cartera diversificada y cómo resulta posible identificar qué acciones mantener. Para aplicar PCA utilizan las empresas contenidas dentro del índice australiano ASX200 desde el inicio hasta febrero de 2014. Sin embargo, los autores trabajan con 156 activos y no con 200 ya que retienen aquellos para los cuales hay información para todo el periodo de estudio.

La investigación se divide en dos partes. Primero, Yang et al. (2015) construyen portafolios seleccionando acciones a partir de los componentes principales obtenidos y los comparan con carteras construidas aleatoriamente. Para ello, establecen una regla de detención y un criterio de eliminación que determina cuántos activos se retienen. El criterio de eliminación consiste en seleccionar las acciones con los coeficientes más altos dentro de los componentes principales correspondientes a autovalores inferiores a 1 para descartarlas. Así se eliminan las acciones con las correlaciones más altas con las otras acciones. Luego se repite el procedimiento con los activos restantes hasta que no se consideren necesarias más eliminaciones en función de un criterio de detención. Este criterio consiste en suprimir acciones hasta que los componentes principales de las acciones retenidas tengan autovalores no inferiores a 0,7. Los autores llevan a cabo tres niveles de ciclos de eliminación y retuvieron 72, 32, 15 acciones en cada uno. Al comparar el riesgo y el rendimiento de los portafolios de acciones seleccionadas con los portafolios de acciones seleccionadas aleatoriamente, observan que la cartera de 32 acciones fue ligeramente mejor que la de 15 o 72. Por ello concluyen que 15 acciones no resultan ser suficientes para diversificar una cartera y las 32 acciones seleccionadas por el método implementado explotan todos los beneficios de la diversificación. A su vez, argumentan que los tres números diferentes de acciones seleccionadas explican bien las 156 acciones originales

En la segunda parte, Yang et al. (2015) comparan el portafolio construido a partir del método de selección de acciones, con un criterio de eliminación de 1 y un criterio de detención de 0,7, con el índice ASX200 para todo el periodo de estudio. Para ello dividieron todo el período en siete subperíodos, de dos años cada uno. Esto creó dos conjuntos de acciones seleccionadas basadas en los datos del primer y segundo año respectivamente. Para cada conjunto de acciones seleccionadas, construyeron una cartera con el mismo peso en todas las acciones para cada periodo. A través de esta metodología los autores observan que un portafolio con alrededor de 20 acciones describe bien el índice de 200 acciones y que las acciones seleccionadas se movieron estrechamente en relación con el índice en casi todos los periodos. Sin embargo, esto no fue así para el

periodo cercano a la crisis financiera de 2008. Esto se atribuye al hecho de que el número de acciones necesarias para formar una cartera diversificada no es constante en el tiempo, sino que varía con las condiciones del mercado.

Finalmente, Yang et al. (2015) concluyen que sus resultados indican que para diversificar adecuadamente una cartera, no es necesario incluir las 200 acciones. Además, establecen que para inversores que quieran comprar acciones individuales y replicar la fluctuación del índice, su método de selección de acciones proporciona una forma de hacerlo posible.

En esta línea, Hargreaves y Mani (2015) continúan con el trabajo de Yang et al. (2015) y también proponen que los problemas de elección de acciones se solucionan utilizando el método de PCA, dado que este reduce la dimensionalidad de los datos. Sin embargo, los autores exhiben los factores obtenidos a través de PCA en un mapa perceptivo con el objetivo de visualizar claramente cuáles serían los activos ganadores que deberían ser elegidos.

De igual forma, Hargreaves y Mani (2015) también se enfocan en la bolsa australiana haciendo uso del índice ASX200. Utilizan la técnica de PCA para demostrar una manera óptima de selección de activos ganadores garantizando un aumento en el rendimiento de la inversión (ROI) y así superar al “Australian Stock Market.” Al hacer esto derivan en un rendimiento de la inversión mucho más alto para su portfolio comparándolo con el índice de Australian Securities Exchange (ASX200). Obtienen un ROI aproximadamente 10% más alto en promedio. En este caso reducen de 22 indicadores a 4 variables importantes. Así es como demuestran que, utilizando PCA efectivamente se quedan con los activos ganadores que le dan al inversor un buen rendimiento de la inversión.

San Andrés

3. Metodología

El análisis de componentes principales (PCA, por sus siglas en inglés) fue desarrollado en 1901 por Karl Pearson como complemento de un teorema que se centraba en mecánica. Más tarde fue desarrollado de forma independiente por Harold Hotelling en la década de 1930 (Jolliffe, 1986).

El análisis de componentes principales permite transformar las variables iniciales (que generalmente están correlacionadas) en un número más pequeño de variables (o factores) incorrelacionadas, construidas como combinaciones lineales de las originales. Este método estadístico permite reducir la dimensionalidad de los datos y la alta correlación. En otras palabras, permite reducir un conjunto de datos complejos a una dimensión inferior para revelar la estructura subyacente simplificada. Los componentes principales son ortogonales, es decir, no están correlacionados entre sí y están ordenados secuencialmente de manera que los primeros retienen la mayor parte de la variación presente en todas las variables originales. El primer componente es la combinación lineal con varianza máxima, el segundo es la combinación lineal con varianza máxima en la dirección ortogonal al primer componente principal y así sucesivamente para los demás.

PCA es una de las técnicas más conocidas en el análisis multivariado. Esto se da ya que a través de este método resulta posible observar el comportamiento de las variables y evaluar si éstas presentan algún co-movimiento o una tendencia común. Además, permite identificar las principales fuentes de riesgo y facilita la gestión del riesgo en un portafolio. Del mismo modo, esta técnica también puede servir para ayudar a inversores a formar portafolios óptimos. Su gama de aplicaciones se ha ampliado con la llegada de las computadoras y se ha utilizado en una amplia variedad de áreas durante los últimos 50 años (Alexander, 2008).

3.1. Cálculo de los componentes principales

Considérese Σ , una matriz $n \times n$ de varianzas y covarianzas o correlaciones de R . Donde R representa una matriz $T \times n$ que contiene los datos de n retornos correlacionados (siendo r_{jt} el retorno del activo j en t para $j = 1, \dots, n$ y $t = 1, \dots, T$). A partir de esta matriz Σ , se calcularán los componentes principales. Esta matriz Σ se puede descomponer en sus autovalores (λ) y sus autovectores (W). De esta forma, obtendremos una transformación lineal definida por W que modificará nuestros datos originales de R en un conjunto de variables aleatorias ortogonales que conforman los componentes principales. Esto resultará en la representación de toda la información contenida en la variable R a partir de un número pequeño de variables o factores (componentes principales). Por lo tanto, todos los componentes principales (P) se pueden expresar como el producto de una matriz formada por autovectores (W), multiplicada por el vector R que contiene las variables originales.

$$P = WR \quad (3.1)$$

En R está contenida la información proveniente de los retornos de los activos incluidos en el índice S&P 500 para la región americana y en el índice FTSE 100 para la región europea. Si bien el periodo de análisis resulta ser el mismo para ambos índices R tendrá una dimensión distinta dependiendo de la región que se estudie debido a la diferencia de días hábiles en el año entre ellas. Es decir, en R se encontrarán los datos de los 1006 retornos correspondientes al S&P 500 y los datos de los 1013 retornos correspondientes al FTSE 100 para el periodo elegido.

Reescribiendo, cada componente principal p_j es una combinación lineal de las r_j originales:

$$p_j = w_1 r_1 + w_2 r_2 + \dots + w_n r_n = w_j' r \quad (3.2)$$

Aquí w refiere a los ponderadores o importancia que se le debe asignar a cada una de las series de los activos en cada componente principal, es decir, representa los elementos contenidos en la matriz de autovectores (W). Cabe destacar que al estandarizar las variables evitamos que aquellos retornos cuya escala sea mayor dominen respecto al resto. A través de un ejercicio de maximización se deja que los propios datos indiquen cuáles serán los pesos óptimos.

Los pesos o ponderadores (w_j') se calculan maximizando la varianza de p_j sujeto a que el módulo de del vector w_j' sea 1, es decir, garantizando que los componentes principales sean ortogonales. De esta forma, obtenemos los componentes como nuevas variables incorrelacionadas. Para que esto suceda resulta necesario imponer dicha restricción a la maximización de varianza.

Planteando el problema obtenemos:

$$\max w_j' \Sigma w_j \quad (3.3)$$

$$\text{sujeto a: } w_j' w_j = 1$$

Para resolver el problema, se puede hacer uso de los multiplicadores de Lagrange. A través de este método, podemos maximizar una función de varias variables sujetas a la condición de normalidad.

Entonces, el primer componente se calcula eligiendo w_1 de modo que p_1 tenga la mayor varianza posible, sujeta a la restricción planteada. Entonces, si queremos elegir w_1 , el método habitual para la maximización de una función de varias variables sujeta a restricciones es a través del método de los multiplicadores de Lagrange. El problema consiste entonces en maximizar $w_1' \Sigma w_1$ sujeta a la restricción $w_1' w_1 = 1$.

Así, construimos la función L tal que:

$$L(w_1') = w_1' \Sigma w_1 - \lambda (w_1' w_1 - 1) \quad (3.4)$$

donde λ es el vector de autovalores de Σ .

Y, buscamos el máximo derivando con respecto a w_j e igualando a 0:

$$\frac{\partial L}{\partial w_1} = 2\Sigma w_1 - 2\lambda I w_1 = 0 \quad (3.5)$$

$$(\Sigma - \lambda I) w_1 = 0 \quad (3.6)$$

De esta forma obtenemos un sistema lineal de ecuaciones compuesto por la matriz de correlaciones menos el vector de autovalores que, a su vez, se encuentra multiplicando la matriz identidad que corresponde al primer autovector.

Nótese que para que el sistema tenga una solución distinta de cero, el determinante de la matriz debe ser cero, es decir, debe ser una matriz singular. Por lo tanto:

$$|\Sigma - \lambda I| = 0 \quad (3.7)$$

Como toda matriz de correlaciones, Σ es definida positiva y de orden n , y tendrá n autovalores en general distintos tales que $\lambda_1 > \lambda_2 > \dots > \lambda_n$.

Entonces, el primer componente se calcula eligiendo w_1 de modo que p_1 tenga la mayor varianza posible sujeta a la restricción planteada, dicho componente corresponde al mayor autovector, λ_1 . El segundo componente principal se calcula obteniendo un w_2 de modo que p_2 se encuentre incorrelacionado con p_1 y así sucesivamente.

$$Cov(p_1, p_2) = Cov(w_2' y, w_1' y) = w_2' E [(y - \mu)(y - \mu)'] w_1 = w_2' \Sigma w_1 = 0 \quad (3.8)$$

Es decir, cada nuevo componente que calculamos agrega una restricción adicional que impone incorrelación respecto a los componentes anteriores. Dado que buscamos que los componentes sean ortogonales, se elimina el problema original de la multicolinealidad¹ que suelen presentar las variables financieras, como los retornos de distintos activos. Una vez obtenida la matriz formada por autovectores (W) podemos multiplicarla por el vector R para obtener todos los componentes P .

Utilizando este modelo de factores un administrador de activos puede formar carteras y explorar las características y retornos de estas. A su vez, esta técnica puede ayudar a

¹ El problema de la multicolinealidad sucede cuando las variables explicativas (en nuestro caso, los retornos) se encuentran altamente correlacionados entre sí, por lo que PCA representa una técnica que permite transformar las variables originales para volverlas ortogonales.

construir carteras que con el mínimo riesgo posible devuelven un objetivo determinado y que se encuentren sujetas a restricciones de asignación (Alexander, 2008).

3.1.1. *¿Sobre qué matriz debería aplicarse PCA?*

PCA se basa en la descomposición de autovalores-autovectores de una matriz de correlaciones o de varianzas y covarianzas de cierto conjunto de variables. La elección de la matriz (de varianza-covarianza o de correlaciones) sobre la cual aplicar el método generalmente depende de los datos, sin embargo, para esta tesis, existen algunos problemas asociados con el uso de una matriz de varianzas y covarianzas. Esto se da ya que, si existen grandes diferencias entre las varianzas de las variables, entonces el uso de una matriz de covarianzas resultará en que los componentes principales se encuentren dominados por aquellas variables que poseen mayor varianza. Entonces, dado que no se ilustra correctamente la estructura de correlación de las acciones dentro de los componentes, esto podría impedir obtener información útil para la diversificación. Es por esto por lo que utilizaremos la matriz de correlaciones. De esta forma los datos serán estandarizados previamente a la aplicación de PCA y se le otorgará la misma importancia a todos los retornos.

3.1.2. *¿Cuántos componentes deberían retenerse?*

No existe un número exacto de componentes principales para retener. Este número deberá cambiar a medida que cambian las circunstancias, dado que la variación total de los componentes depende de la correlación entre los rendimientos de las acciones. Durante los períodos en los que los rendimientos de todas las acciones están estrechamente relacionados, se necesitarán menos componentes para lograr un grado adecuado de precisión en la representación de los componentes principales. Mientras que, en los momentos donde el mercado no se encuentra tan conectado, se necesitarán más componentes.

Jolliffe (1986) establece que normalmente deberíamos intentar explicar entre el 70% y el 90% de la variación total. Esto se da ya que, si tratamos de explicar más del 90% de la variación, el modelo puede estar captando "ruido" que no es relevante para las decisiones de asignación a largo plazo. Por otro lado, si explicamos menos de aproximadamente el 70% de la variación, las carteras que se modelan con este marco de referencia podrían presentar riesgos específicos muy grandes.

Para determinar el número de componentes principales a retener varias reglas estadísticas pueden ser aplicadas. Sin embargo, la combinación de estas tendrá como consecuencia un resultado más preciso. Es por esto que en la presente tesis utilizaremos la varianza acumulada de los factores y el criterio de Kaiser como métodos de decisión para definir cuantos componentes principales deben retenerse.

1. La idea del primer criterio es elegir un porcentaje acumulativo de la variación total al que deben contribuir los componentes principales. Como explican Yang

et al. (2015), utilizando una matriz de correlaciones, la varianza porcentual explicada por los primeros m componentes se calcula mediante:

$$\text{Variance}_k = \frac{100}{p} \sum_{k=1}^m l_k \quad (3.9)$$

donde l_k es el autovalor del componente principal k y p resulta ser el número total de variables. Por lo tanto, el número requerido de componentes principales a retener es el valor más pequeño de m para el cual se excede el porcentaje que se elige. Normalmente, se elige un límite entre el 70% y el 90%. En nuestro caso, elegiremos preservar un 80% de la variación de los datos originales. A partir de este umbral será posible observar cuántos componentes principales deben retenerse para cada ventana del periodo de estudio.

2. El criterio de Kaiser se basa en el tamaño de las variaciones de los componentes principales. Tiene como objetivo retener solo aquellos componentes principales cuyas varianzas l_k excedan 1. La idea expresada por Kaiser (1960) es que si todos los elementos de r son ortogonales (están incorrelacionados), entonces los componentes principales son los mismos que las variables originales y todos tienen varianzas unitarias en el caso de una matriz de correlación. Entonces, cualquier componente principal con autovalores menores que 1 contienen menos información que una de las variables originales y, por lo tanto, no vale la pena retenerlo. A través de este método también resultará posible observar cuántos componentes principales deben retenerse para cada ventana del periodo de estudio.

3.2. Construcción del portafolio y selección de activos

Una vez aplicados los dos criterios mencionados anteriormente y habiendo definido cuántos componentes principales deben retenerse, a partir de los *weights* de dichos componentes resulta posible construir un portafolio para cada *benchmark*. Esto se da ya que a través del método de PCA se logra explotar la propia varianza de los datos y a partir de eso resulta posible definir los pesos óptimos. Entonces, si construimos una cartera de acciones asignando la inversión proporcionalmente a los componentes que explican la mayor parte de la varianza de los datos, resultaría posible replicar, aproximadamente, los rendimientos de los determinados índices. Por lo tanto, una vez construidos los determinados portafolios puede evaluarse su rendimiento en ambas ventanas (pre-crisis COVID-19 y COVID-19), así como la efectividad del método propuesto.

El mercado de acciones en 2020 se caracterizó por estar muy conectado y presentar correlaciones entre activos cercanas a uno. Con el objetivo de observar esto, se divide la muestra en dos ventanas: el primer período comprendido irá desde 20/02/2017 al 18/02/2020 y el segundo período irá desde el 19/02/2020 al 19/02/2021. Dado que es en

la segunda ventana donde las tendencias de mercado se vuelven bajistas producto de la pandemia, proponemos un método de selección de activos a partir de los componentes principales. Utilizando PCA resulta posible analizar cómo la pandemia afecta individualmente a las distintas acciones.

Cabe destacar que, para el caso del índice estadounidense, el S&P 500, en el segundo periodo de estudio se utilizara una ventana *rolling*. Esto se da ya que, para poder calcular los componentes principales, la cantidad de días debe ser igual o mayor a la cantidad de observaciones que presenta la variable a estudiar. Por lo tanto, dado que las observaciones dentro de la segunda ventana de análisis conforman un total de 365 días, mientras que, los activos conforman un total de 494 observaciones, resulta necesario hacer uso de esta ventana de media móvil. En cambio, en el caso del índice británico, el FTSE 100, esto no resulta ser necesario ya que la cantidad de días en el segundo periodo de estudio (también 365) es superior a la cantidad de activos (90 observaciones). Para construir la ventana para el S&P 500 primero se realizará PCA para las primeras 1064 observaciones (previas al COVID-19) y luego se continuará avanzando en dicha ventana para poder cubrir los 1430 días totales, iterando 365 veces hasta llegar al número total de componentes necesarios para el análisis. De esta forma, resultara posible establecer un criterio de selección individual de activos para el periodo de la crisis del COVID-19.

Para construir el criterio de selección de activos observaremos la magnitud (en valor absoluto) de cada uno de los elementos de w_j' . Agruparemos a las compañías entre aquellas que poseen los mayores y menores pesos según el primer componente obtenido a través de PCA. Resulta importante mencionar que para esta segunda parte del análisis utilizaremos solamente el primer componente principal dado que es aquel que explica mayor porcentaje de la varianza y puede considerarse como el componente de mercado. En particular, dado que la segunda ventana se caracteriza por ser una de gran volatilidad donde los activos se encuentran más correlacionados, esto significa que el primer componente principal debe explicar un mayor porcentaje de la varianza de los datos. Además, dado que dentro de cada componente principal se encuentran contenidos todos los activos, usar más de un componente podría no derivar en una conclusión significativa en términos de la diversificación.

Entonces, resulta posible construir un portafolio eficiente, *equally weighted* ($\frac{1}{N^2}$), que se encuentra *long* en las compañías menos afectadas según los pesos obtenidos a través del primer componente. Si bien también se podría realizar la venta en corto de aquellas compañías que poseen los menores *weights* según PCA, no es el objetivo de esta tesis y a modo de simplificar en análisis no se tendrá en cuenta, aunque bien podría aplicarse. También, cabe destacar que en la literatura continúa habiendo una creciente discusión sobre cuántas son las acciones necesarias para diversificar un portafolio, por lo que no existe un número fijo de cuantos activos resulta eficiente retener. Es por esto que analizaremos el 10% de las acciones contenidas en cada componente principal y

² Donde N representa los activos elegidos a partir de los *weights* del primer componente principal.

realizaremos un análisis de sensibilidad de dicho porcentaje. Esto tendrá como objetivo observar si los resultados cambian significativamente o no. De esta manera la cartera resultante habrá tenido un rendimiento significativamente mejor que el de los índices, ya que invierte en aquellas empresas que realmente se beneficiaron de la pandemia.



Universidad de
San Andrés

4. Datos

4.1 Datos

Nuestra investigación se centra en la selección de activos individuales dentro del *asset class equity*. Particularmente, utilizaremos dos mercados de valores representativos en regiones distintas, siendo estos, Estados Unidos y el Reino Unido.

Para llevar a cabo el análisis estudiamos los componentes contenidos dentro del índice S&P 500 (US) y el de FTSE 100 (UK). Por un lado, el índice S&P 500 se caracteriza por ser uno de los más importantes y conocidos a nivel mundial y es considerado como una representación real de la economía de Estados Unidos. Captura aproximadamente el 80% de la capitalización del mercado estadounidense y representa la capitalización bursátil de las 500 empresas más grandes. Por otro lado, el FTSE 100 es el índice de las 100 compañías con la capitalización bursátil más grande de la Bolsa de Londres. Aunque existen muchos otros índices que están compuestos por una cantidad mayor de empresas, como el FTSE All-Share index, FTSE 250, FTSE 350, entre otros, este es el indicador bursátil británico más utilizado.

Debido a que el S&P 500 y el FTSE 100 no son índices de rendimiento total, no incluye las ganancias obtenidas de los dividendos pagados por las empresas que los componen. Dado que muchas empresas dentro de estos índices pagan dividendos, resulta necesario incluir esos pagos en el rendimiento general de la inversión para el análisis. Sin embargo, como en el presente trabajo tomamos el rendimiento de cada acción de las distintas empresas de manera individual, logramos simplificar el análisis porque es posible acceder a los precios de cierre ajustados por dividendos

El periodo de estudio comprendido irá desde febrero de 2017 hasta febrero de 2021. Para la aplicación de PCA se distinguen dos periodos: del 20/02/2017 al 18/02/2020 para el período anterior a la crisis del COVID-19 y del 19/02/2020 al 19/02/2021, que es el período donde se asienta la pandemia producto del coronavirus. La distinción entre ambas ventanas se produce debido a que la primera gran caída como consecuencia del COVID-19 que se puede observar es para la región británica y tuvo lugar el 19 de febrero de 2020. Para la región americana, la primera gran caída puede observarse el 23 de marzo de 2020, también dentro de la segunda ventana. Los datos fueron extraídos de de *Yahoo Finance*, en USD para los activos que componen el S&P 500 y en GBP para los activos que componen el FTSE 100, utilizando el precio de cierre ajustado por dividendos. Destacamos que, al aplicar el método de PCA en nuestra investigación, no resulta ser un problema trabajar con datos denominados en diferentes monedas ya que una vez que se aplica el método este mismo re-escala las unidades de medida.

En cuanto a los datos para el análisis, para observar los activos que se encuentran dentro de los índices se hizo uso de las listas de constituyentes del S&P 500 y del FTSE 100 de Wikipedia obtenida en marzo de 2021, descargadas directamente en R Studio.

Ambos índices poseen una frecuencia relativamente alta de activos que fueron agregados o eliminados en el correr del tiempo. Es debido a esto que, con el objetivo de ser consistentes, se utilizarán las acciones que se encuentran listadas al momento del análisis, y que se encuentran disponibles para todo el periodo de estudio. Al ajustar dichos datos, tomando en cuenta lo mencionado, se eliminaron varios activos que no presentaban la cantidad de datos necesarios en el periodo comprendido.

Para el caso del S&P 500 las compañías que no contaban con la totalidad de los datos en el periodo de estudio fueron: Berkshire Hathaway Inc. (BRB.B), Brown-Forman Corporation (BF.B), Organon & Co. (OGN), Otis Elevator Company (OTIS), Ingersoll-Rand (IR), Fox Corp. Class A (FOX.A), Fox Corp. Class B (FOX), Dow Inc. (DOW), Corteva Inc. (CTVA) y Carrier (CARR). Entonces, dado esto, el conjunto de datos finales incluye 495 acciones únicas para el análisis, en vez de un total de 505 acciones. Destacamos que, aunque este índice es llamado el S&P 500, contiene 505 acciones dado que incluye dos clases de acciones de 5 de las empresas que lo componen.

Para el caso del FTSE 100 las acciones que no contaban con la totalidad de los datos en el periodo de estudio fueron: Avast Plc. (AVST.L), Kingfisher Plc. (KGF.L), London Stock Exchange Group Plc. (LSEG.L), M&G Plc. (MNG.L), Pershing Square Holdings Ltd. (PSH.L), RSA Insurance Group Ltd. (RSA.L), Standard Life Aberdeen Plc. (SLA.L), James's Place Plc. (STJ.L) y Unilever Plc. (ULVR.L). Entonces, el conjunto de datos finales incluye 91 acciones únicas para el análisis, en vez de un total de 100 acciones.

A partir de estos datos se obtiene el cálculo del rendimiento de mercado de todos los activos finales para de cada uno de los índices:

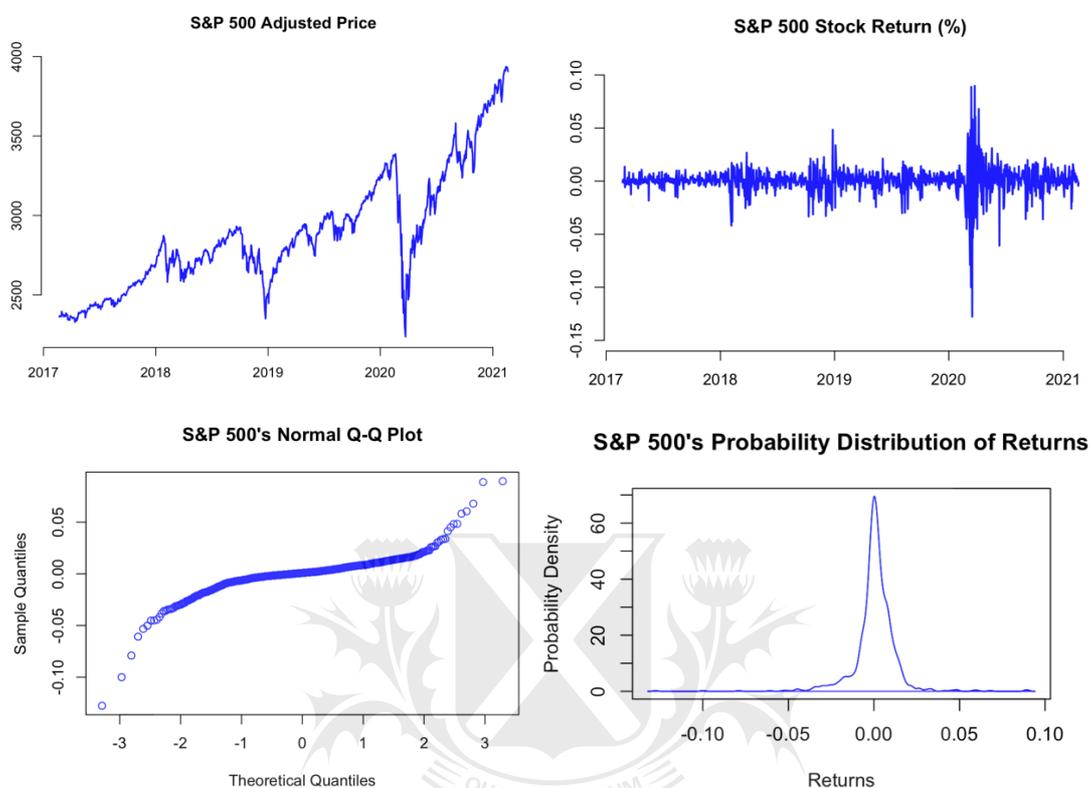
$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (4.1)$$

Una vez calculados los retornos de mercado, entonces, podemos aplicar las diferentes metodologías a los datos.

4.2 Análisis Descriptivo

A partir del análisis estadístico de la distribución de los rendimientos resulta posible determinar muchas características clave del comportamiento de un conjunto de datos.

Figura 1: Análisis descriptivo del S&P 500

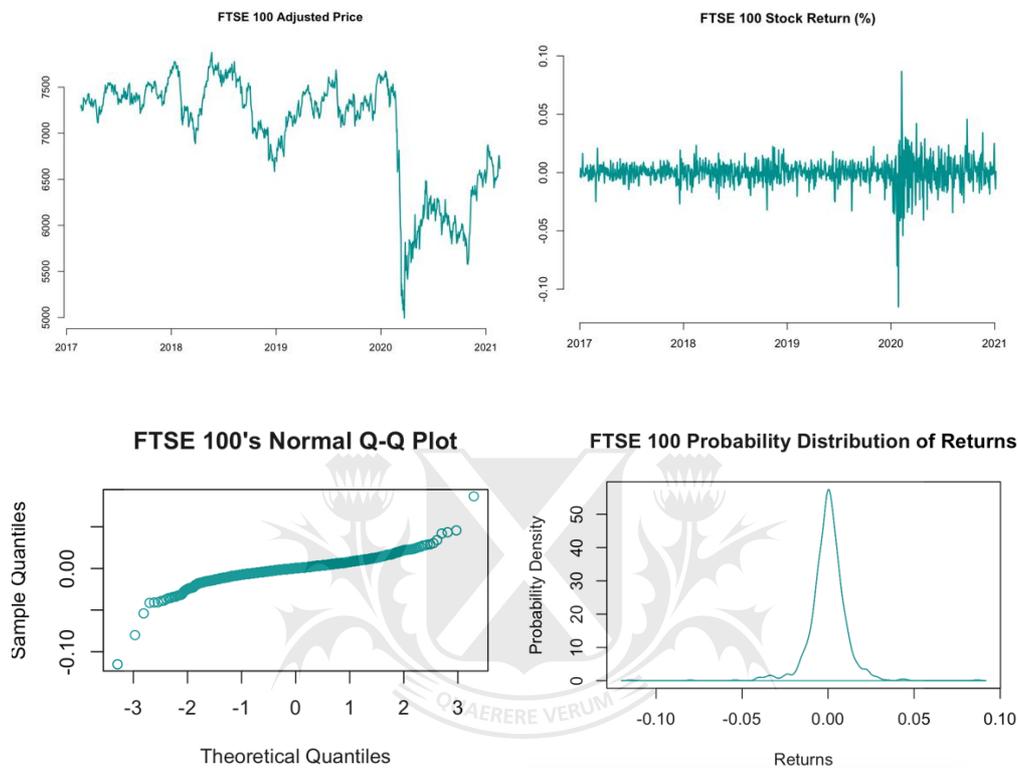


La Figura (1) muestra el análisis descriptivo para el índice del S&P 500. En los dos gráficos superiores queda en evidencia que los picos más importantes se produjeron durante la crisis del COVID-19. A partir del segundo gráfico superior podemos observar un comportamiento típico de retornos, estacionarios en media y varianza que se mantiene relativamente constante para la primera ventana. Sin embargo, luego vemos un gran periodo de volatilidad. También, en el comportamiento de los retornos se logra ver claramente la agrupación de la volatilidad. Es decir, los períodos de baja volatilidad se encuentran juntos y los periodos de alta volatilidad se encuentran juntos, en el otro extremo (Mandelbrot, 1963). Del mismo modo, a partir del QQ plot resulta posible observar la presencia de colas pesadas, a la vez que parecería que los retornos se encuentran sesgados hacia la izquierda. Observando el gráfico de densidad, también queda en evidencia la presencia de estas colas pesadas. Por ello, resulta pertinente realizar el cálculo de la curtosis y el sesgo para los retornos diarios del S&P 500.

La curtosis es una medida del peso combinado de las colas de una distribución en relación con el resto de la distribución, es decir, representa la proporción de la probabilidad total que está en las colas (DeCarlo, 1997). Para el S&P 500 el cálculo de la curtosis indica un valor de 20.32661. Esto indica una distribución de retorno con un exceso de curtosis positiva (una distribución de retorno leptocúrtica) y, por lo tanto, tiene desviaciones extremadamente grandes más frecuentes de la media que una distribución normal. De

igual manera resulta posible calcular el sesgo (una distribución que no es simétrica se llama sesgada). En este caso el sesgo para los retornos del S&P 500 es de -1.129107 . Esto confirma la presencia de una cola pesada en el lado izquierdo de la distribución.

Figura 2: Análisis descriptivo del FTSE 100



La Figura (2) muestra el análisis descriptivo para el índice del FTSE 100. Aquí parecen observarse conclusiones similares a las obtenidas en el S&P 500: se observa que la caída más importante coincide con el periodo del COVID-19 y de igual forma, en el segundo gráfico superior, también resulta posible observar la agrupación de volatilidad. Finalmente, al observar el QQ plot y el gráfico de la densidad también se pueden observar colas pesadas, al mismo tiempo que parecería haber un sesgo hacia la izquierda.

También calculamos la asimetría y la curtosis de los rendimientos diarios del FTSE 100. Estas fueron de -1.236792 y 18.04848 , respectivamente. Una curtosis de esa magnitud significa que la distribución tiene un pico mucho más agudo y colas mucho más pesadas que una distribución normal, aunque el pico es menor para el FTSE 100 que para el S&P 500. De igual forma, conclusiones similares a las anteriores respecto al sesgo pueden hacerse para este índice.

5. Resultados

5.1 S&P 500

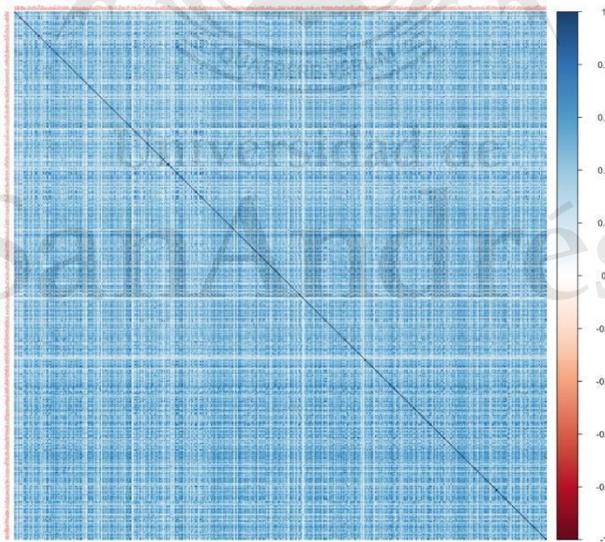
5.1.1 Portafolio PCA para la región estadounidense

En la presente sección investigamos en detalle la efectividad del método de PCA para modelar el comportamiento de un *benchmark*, en particular, el S&P 500.

Para ello, aplicamos PCA, para todo el periodo de estudio, sobre los 495 activos que se encontraban contenidos en el índice a la fecha de análisis. A partir de esto, obtuvimos una matriz P de dimensiones 1006×495 donde en cada columna se observan los componentes principales.

Para un mejor análisis de los datos, comprobamos que los componentes principales resultan ser ortogonales y graficamos la matriz de correlaciones. Dicho gráfico permite tener un mejor entendimiento del co-movimiento de las variables y de la interpretación de las mismas.

Figura 3: Matriz de Correlaciones de los activos del S&P 500

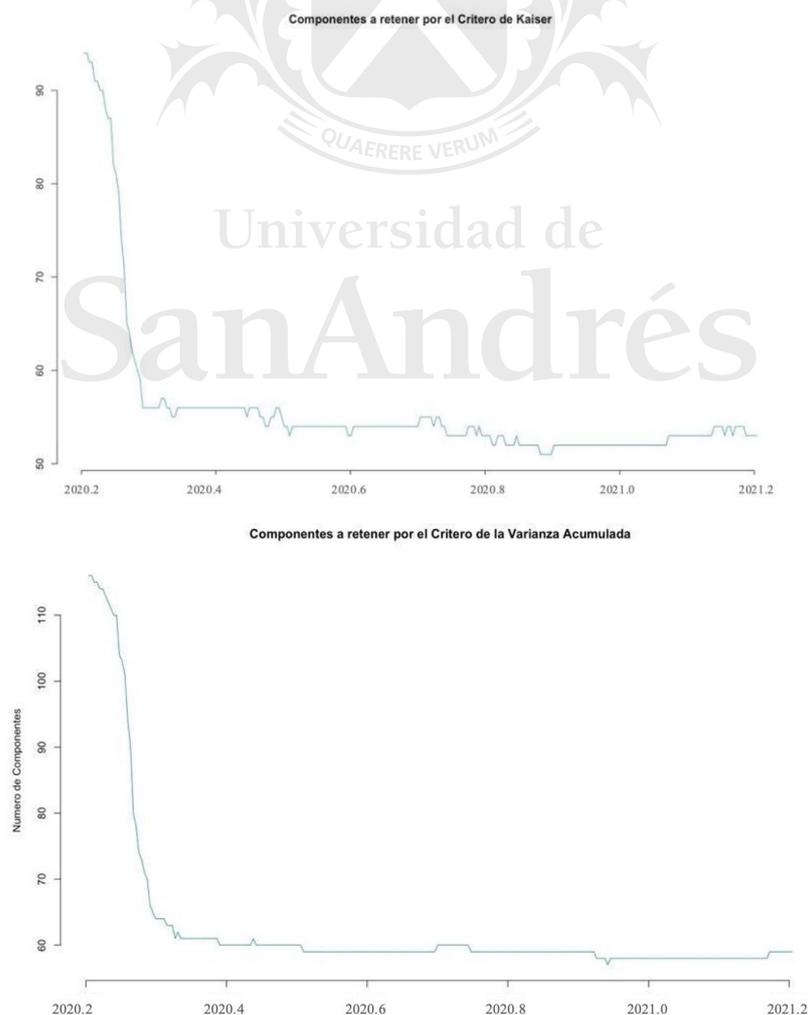


A partir de la Figura (3) podemos observar que predominan los colores azulados que tienen que ver con las correlaciones positivas. Aunque, como puede observarse, el color no resulta ser muy intenso. De cualquier modo, en ningún momento se observa la presencia de colores rojos que indican correlaciones negativas. Es por esto que resulta posible establecer que las variables se mueven conjuntamente y no se comportan de forma inversa. Es decir, se observa una especie de dependencia directa de los activos. Esto podría interpretarse como un primer indicador de que en momentos donde el mercado se encuentre más correlacionado, se necesitará un menor número de componentes para describir el comportamiento del índice.

Por ello, decidimos analizar este supuesto. Para observar la diferencia en los componentes que deben retenerse dependiendo la fase del mercado realizamos PCA sobre las variables para las dos determinadas ventanas, por separado. Dado que, como fue mencionado anteriormente, para poder calcular los componentes principales la cantidad de días debe ser igual o mayor a la cantidad de observaciones que presenta la variable a estudiar, para el segundo periodo fue necesario utilizar una ventana *rolling*.

Los resultados que obtuvimos a partir de este análisis confirmaron nuestras sospechas iniciales; el primer componente principal explica mucha más variación después de la crisis del COVID-19. La varianza explicada por el primer componente fue del 28,22% antes de la crisis y se duplicó a partir del inicio de la pandemia, llegando a un valor de 46,56%. Esto significó que la varianza explicada por el primer componente aumentó en un 61% de un periodo a otro. Como argumentan Yang et al. (2015), la intuición detrás de esto es que después de las crisis financieras, el riesgo de mercado aumenta significativamente en comparación con el riesgo diversificable, por ello hay más volatilidad en los datos para ser capturada por los componentes.

Figura 4: Componentes a retener en la segunda ventana para los activos del S&P 500



La Figura (4) muestra cómo cambia la cantidad de componentes que deberían retenerse según los determinados criterios para el segundo periodo de estudio. En el gráfico superior puede observarse que previo al periodo del COVID-19 aproximadamente 94 componentes eran necesarios para explicar la varianza del índice según el criterio de Kaiser. Sin embargo, a partir del primer mes de 2020 este valor cae drásticamente y oscila entre una cantidad de 50 y 55 para todo el año de pandemia. Conclusiones similares pueden hacerse respecto al gráfico inferior; cuando inicialmente 115 componentes explicaban el comportamiento del índice según el criterio de la varianza acumulada, luego de la pandemia, con tan solo 55 componentes se logra explicar ese mismo comportamiento. Esto hace que uno pueda afirmar que a medida que el mercado se correlaciona, menos componentes serán necesarios para explicar la varianza original de los datos.

Una vez realizado el análisis de las variables, procedemos al cálculo de los componentes principales para todo el periodo de estudio. Aquí analizamos si resulta posible representar adecuadamente toda la información contenida en la variable R a partir de un número pequeño de variables o factores independientemente de las fases de mercado.

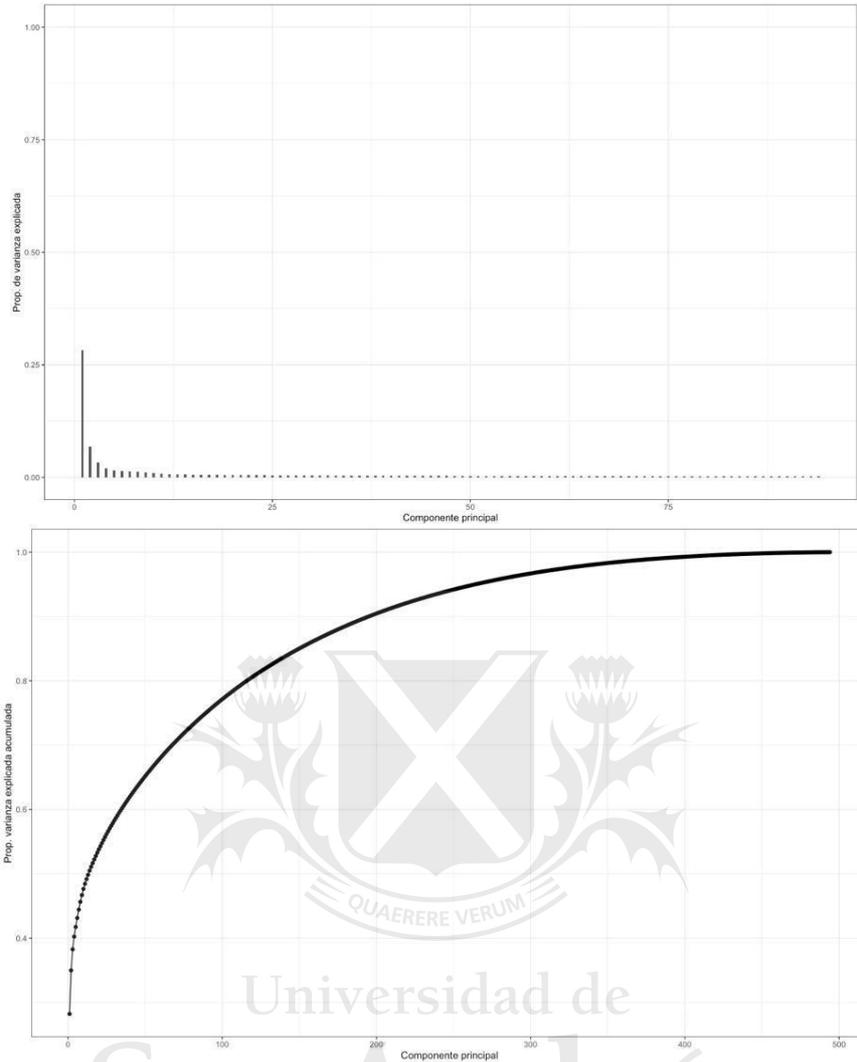
Al estudiar la proporción de la varianza queda en evidencia que los primeros componentes logran capturar una proporción alta de la volatilidad de los datos; para toda la ventana de análisis el primer componente explica un 0,4322 de la varianza. Sin embargo, a medida que se agregan más componentes este número va disminuyendo significativamente. El segundo componente principal tan solo explica un 0,0557 de la variabilidad de los datos.

Por ello, hacemos uso del criterio de Kaiser y de la varianza acumulada para decidir cuántos componentes deben retenerse para la ventana completa. Observamos que a partir del primer criterio resultan ser 55 componentes los que tienen autovalores mayores a 1. Sin embargo, a partir del segundo criterio se puede concluir que son 74 componentes los que explican un 80% de la varianza.

Las dos reglas anteriores sugirieron números significativamente diferentes de componentes para retener y no hay forma de que podamos decir cuál es más precisa. Sin embargo, si un análisis requiere que preservemos la mayor cantidad posible de variación, entonces tal vez resulte más apropiado decidir en función de la variación acumulativa. Si bien observamos que con 55 componentes retenidos por el criterio de Kaiser se explica un 76% de la varianza, decidimos retener 74 componentes para lograr explicar un 80% de los datos.

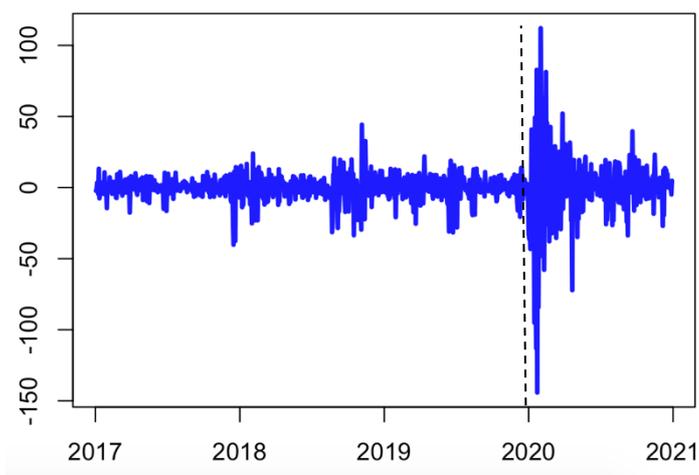
Entonces, el conjunto original de 495 activos se transformó en un sistema principal que incluía 495 componentes principales no correlacionados en los que los primeros 74 componentes principales identifican los principales factores de riesgo de las acciones. Por lo tanto, observamos que sólo reteniendo un 15 % de los componentes principales se puede explicar la proporción de la varianza de toda la serie.

Figura 5: Proporción de la varianza explicada por los componentes principales (Activos S&P 500)



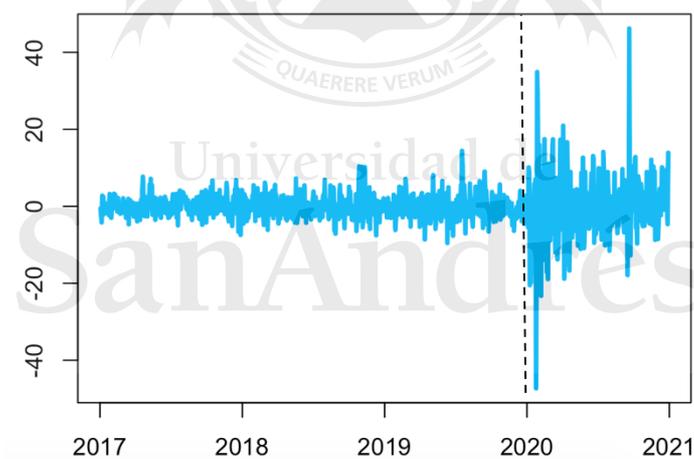
El gráfico superior de la Figura (5) muestra la proporción de la varianza explicada por cada componente, mientras que, el gráfico inferior muestra la proporción de la varianza acumulada explicada. En ambos casos se ve que los primeros componentes principales explican casi toda la varianza del periodo completo. Sin embargo, queda en evidencia que luego del primer componente principal resulta ser muy marginal lo que aproxima el resto y a medida que se consideran más componentes este número continúa decreciendo. Esto es así ya que PCA transforma los activos originales interrelacionados en componentes principales no correlacionados, donde el primer componente explica la mayor parte de la variación y los componentes subsiguientes explican tanto como sea posible la variación restante.

Figura 6: Primer Componente Principal (S&P 500)



Al analizar con detenimiento el primer componente principal podemos observar que este se comporta de una manera casi idéntica a los retornos del S&P 500 (véase Figura (1)). Si bien PCA no respeta unidades de medida debido a la estandarización de los datos, se logra representar perfectamente la estructura de los activos. Es decir, el primer componente está capturando muy bien lo que sucede en toda la serie.

Figura 7: Segundo Componente Principal (S&P 500)



La Figura (7) muestra los resultados obtenidos a partir del segundo componente principal. Resulta posible observar que el comportamiento general de los activos se encuentra representado: inicialmente vemos una serie relativamente constante en media y varianza que luego es seguida por un gran periodo de volatilidad producto de la pandemia. Sin embargo, observando en detalle vemos que los retornos no se comportan de una manera tan similar a los del índice. Esto se debe a que el segundo componente no reúne una proporción muy alta de la varianza explicada.

Habiendo determinado que se retendrán 74 componentes, a partir de los *weights* de estos, resulta posible construir un portafolio con el objetivo de observar si es probable reproducir el comportamiento del S&P 500 en todo el periodo de estudio.

Figura 8: Portafolio PCA vs S&P 500



Observando la Figura (8) se puede establecer que la hipótesis planteada puede comprobarse dado que la cartera formada logra replicar bien el comportamiento del índice. Además, observamos que el portafolio construido no solo logró modelar bien el S&P 500 en su totalidad, sino que, para el primer periodo de estudio, logró superar dicho *benchmark*. Sin embargo, cabe destacar que el rendimiento superior del portafolio que se observa en esta primera ventana podría ser resultado de la inclusión de los dividendos en la matriz R . Recordamos que el S&P 500 no ajusta por dividendos pagados, pero en nuestro análisis asumimos que todos los dividendos se reinvirtieron para todos los componentes.

Si bien el portafolio construido fue capaz de seguir de cerca el movimiento del índice estadounidense, para la segunda ventana de análisis observamos una caída de mayor magnitud al mismo tiempo que vemos que el portafolio no resulta ser el ganador. Esto tiene que ver con el hecho de que, como mencionamos anteriormente, el mercado de valores se vuelve más correlacionado producto de la crisis del COVID-19. Este periodo no puede tomarse como uno normal y debe considerarse con mayor detenimiento debido a que podría considerarse como un *outlier* dentro de la serie histórica del S&P 500. Entonces, dado que cambia la volatilidad de los datos según el periodo, no puede pretenderse utilizar la misma cantidad de componentes. El mayor nivel de varianza explicado por el primer componente principal indica un mayor riesgo sistémico con respecto al riesgo diversificable. Esto sugiere que, teniendo la misma cartera, esta no estaría tan diversificada como en otras ocasiones.

De cualquier forma, resulta posible establecer que los 74 componentes obtenidos y seleccionadas a partir de PCA explican bien las 495 acciones originales. Sin embargo, remarcamos que, aunque la dimensionalidad fue reducida, el portafolio construido sigue siendo una cartera donde dentro se encuentran contenidos los 495 activos originales.

Además, si bien a partir de los componentes se logra comprobar la hipótesis planteada, notamos que para aquellos momentos donde el mercado se encuentra muy correlacionado resulta necesario aplicar otro tipo de análisis para construir un portafolio que, no solo imite el comportamiento de un *benchmark*, sino que resulte ser superior.

5.1.2 Selección de Activos para la región estadounidense

En esta sección proponemos un método para seleccionar activos a través de PCA, enfocándonos en el segundo periodo de estudio. La pregunta central es si resulta posible, en periodos de estrés financiero, construir un portafolio ganador utilizando el primer componente principal como indicador de mercado.

A partir de PCA se pueden agrupar las compañías entre aquellas que fueron más y menos afectadas por la crisis provocada en consecuencia al COVID-19. Como mencionamos anteriormente, observando la magnitud de los autovectores se puede distinguir entre aquellas acciones que se encuentran dentro del percentil superior y que poseen los mayores pesos según el primer componente principal. Entonces, en esta sección observamos si a partir de estas ponderaciones resulta posible construir un portafolio que mitigue la gran caída índice S&P 500 en 2020 producto de la crisis.

Para ello, comenzamos calculando los percentiles del 0.90, para observar los activos con mayores ponderaciones y del 0.10, para poder observar las acciones con las peores ponderaciones. Al analizar dicho percentil superior resulta posible observar las acciones con mayores retornos, en el periodo comprendido por la crisis, según PCA.

Tabla 1: Compañías contenidas en el percentil 0.9 (Activos S&P 500)

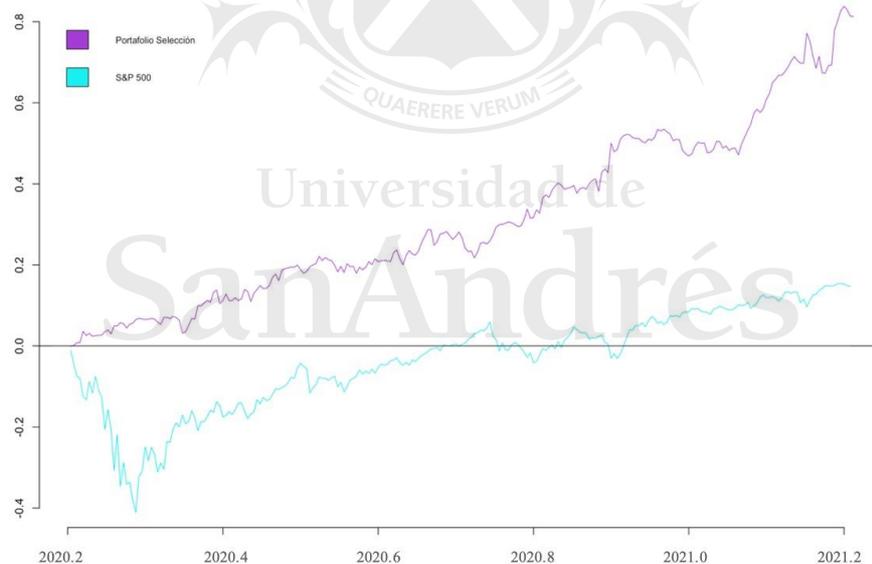
Ticker	Company Name	Industry	Sub-Industry
ABMD	Abiomed	Health Care	Health Care Equipment
AKAM	Akamai Technologies	Information Technology	Internet Services & Infrastructure
ALXN	Alexion Pharmaceuticals	Health Care	Pharmaceutical
AMD	Advanced Micro Devices	Information Technology	Semiconductors
AMZN	Amazon	Consumer Discretionary	Internet & Direct Marketing Retail
ANET	Arista Networks	Information Technology	Communications Equipment
ATVI	Activision Blizzard	Communication Services	Interactive Home Entertainment
BIIB	Biogen	Health Care	Biotechnology
CAG	Conagra Brands	Consumer Staples	Packaged Foods & Meats
CHD	Church & Dwight	Consumer Staples	Household Products
CHRW	C. H. Robinson	Industrials	Air Freight & Logistics
CLX	Clorox	Consumer Staples	Household Products
CMG	Chipotle Mexican Grill	Consumer Discretionary	Restaurants
COG	Cabot Oil & Gas	Energy	Natural gas & Oil
CPB	Campbell Soup	Consumer Staples	Packaged Foods & Meats
CTXS	Citrix Systems	Information Technology	Application Software
DG	Dollar General	Consumer Discretionary	General Merchandise Stores
DLTR	Dollar Tree	Consumer Discretionary	General Merchandise Stores
DPZ	Domino's Pizza	Consumer Discretionary	Restaurants
DVA	DaVita	Health Care	Health Care Facilities
DXCM	DexCom	Health Care	Health Care Equipment
EA	Electronic Arts	Communication Services	Interactive Home Entertainment
EBAY	eBay	Consumer Discretionary	Internet & Direct Marketing Retail
ED	Consolidated Edison	Utilities	Electric Utilities
ENPH	Enphase Energy	Information Technology	Electronic Components
ETSY	Etsy	Consumer Discretionary	Internet & Direct Marketing Retail
GILD	Gilead Sciences	Health Care	Biotechnology
GIS	General Mills	Consumer Staples	Packaged Foods & Meats
HRL	Hormel	Consumer Staples	Packaged Foods & Meats
INCY	Incyte	Health Care	Biotechnology
IPGP	IPG Photonics	Information Technology	Electronic Manufacturing Services
K	Kellogg's	Consumer Staples	Packaged Foods & Meats
KHC	Kraft Heinz	Consumer Staples	Packaged Foods & Meats
KMB	Kimberly-Clark	Consumer Staples	Household Products
KR	Kroger	Consumer Staples	Food Retail
MKC	McCormick & Company	Consumer Staples	Packaged Foods & Meats
MKTX	MarketAxess	Financials	Financial Exchanges & Data
NEM	Newmont	Materials	Gold
NFLX	Netflix	Communication Services	Movies & Entertainment
NLOK	NortonLifeLock	Information Technology	Application Software
PRGO	Perrigo	Health Care	Pharmaceutical
REGN	Regeneron Pharmaceuticals	Health Care	Biotechnology
SJM	JM Smucker	Consumer Staples	Packaged Foods & Meats
TGT	Target	Consumer Discretionary	General Merchandise Stores
TSLA	Tesla	Consumer Discretionary	Automobile Manufacturers
TTWO	Take-Two Interactive	Communication Services	Interactive Home Entertainment
TWTR	Twitter	Communication Services	Interactive Media & Services
VTRS	Viatis	Health Care	Pharmaceuticals
WEC	WEC Energy Group	Utilities	Electric Utilities
WMT	Walmart	Consumer Staples	Hypermarkets & Super Centers

En la Tabla (1) se muestran las compañías que se encuentran contenidas en el percentil 0.90. Es decir, se observan las 50 empresas que presentan el mejor retorno. Se puede notar que, las acciones que se encuentran en el rango superior son aquellas que, no casualmente, fueron las menos afectadas por la crisis del COVID-19. Estas, en su mayoría, pertenecen a industrias de servicios de bienes de consumo, servicios de comunicación, servicios de atención sanitaria, tecnología de la información, y consumo discrecional. Esto tiene sentido ya que estas fueron las industrias menos afectadas, o que se vieron más beneficiadas por la crisis, a raíz de las medidas de cuarentena.

Tanto el consumo de bienes como la tecnología de la información se vieron incrementados por el impulso a las ventas y de aplicaciones o servicios de internet en respuesta a las medidas de la pandemia. La pandemia tuvo implicaciones tanto sanitarias como económicas y en momentos de estrés se consumen en mayor medida aquellos bienes que resultan ser necesarios: claro ejemplo de esto son empresas como Campbell Soup y Walmart. Así mismo, los servicios de comunicación también se vieron beneficiados por las medidas de la cuarentena ya que a raíz de esta se incrementó exponencialmente su consumo, por eso tampoco resulta sorprendente que en el percentil superior se encuentran compañías como Amazon o Netflix. Por otro lado, los servicios de atención sanitaria se vieron beneficiados por el hecho de que fueron los servicios que fueron utilizados para combatir el COVID-19, por ello no es sorprendente que veamos compañías farmacéuticas o de equipos de salud.

A partir de estos resultados resulta posible constatar que utilizando las ponderaciones por PCA se logró agrupar las mejores empresas afectadas, en este caso positivamente, por la crisis del COVID-19, sin conocer previamente sus fundamentos. Entonces, a partir de las 50 compañías contenidas en la Tabla (1) construimos un portafolio *equally weighted*.

Figura 9: Portafolio 50 Activos vs S&P 500



Como se puede observar en la Figura (9), construyendo un portafolio que esté *long* en las compañías conformadas por el 10% superior según los pesos de PCA, se puede mitigar el efecto de la crisis y, del mismo modo, ganarle al índice del S&P 500. El gráfico muestra claramente que la cartera resultante se comportó significativamente mejor que el mercado: dicha cartera en ningún momento tuvo rendimientos negativos y siempre se movió muy por encima del S&P 500. Por lo tanto, utilizando este método de selección se invirtió en aquellas empresas que realmente se vieron beneficiadas por la crisis del COVID-19 y resultó posible reducir el riesgo de la cartera desde el punto de vista de la diversificación al eliminar los activos que se encontraban más correlacionados entre sí.

Luego de realizar una conclusión positiva respecto de la cartera construida, resulta pertinente realizar un análisis de sensibilidad de los resultados obtenidos. A partir de este se puede observar la sensibilidad que tiene la cartera construida a cambios en la cantidad de acciones que la componen. Por ello, realizamos un análisis de sensibilidad sobre los percentiles a utilizar. De esta forma observaremos si, al modificar el porcentaje de compañías elegidas, dichas carteras se encuentran más o menos diversificadas, y si incluyendo o sacando estas acciones, se encontraría una cartera resultante cuyo rendimiento sea aún mayor o, por el contrario, menor.

Para el análisis de sensibilidad realizamos una comparación entre: el percentil 0.85 (constituido por 74 compañías), el percentil 0.95 (constituido por 25 compañías) y el percentil 0.97 (constituido por 15 compañías). Consideramos pertinente realizar una comparación que contenga tanto una mayor cantidad de compañías como una menor cantidad, para así realizar un análisis más preciso y completo.

Se continuó por realizar el cálculo de las compañías que se encuentran contenidas en el percentil 0.85. Es decir, se observan las 74 compañías que presentan el mejor retorno. Dichas acciones se pueden observar en la Tabla (2).

Al observar dicha selección se puede ver claramente que las 50 compañías seleccionadas en la cartera anterior también están incluidas en esta. Pero también, se pueden observar, resaltado en negrita, las 24 compañías que fueron agregadas en esta instancia.

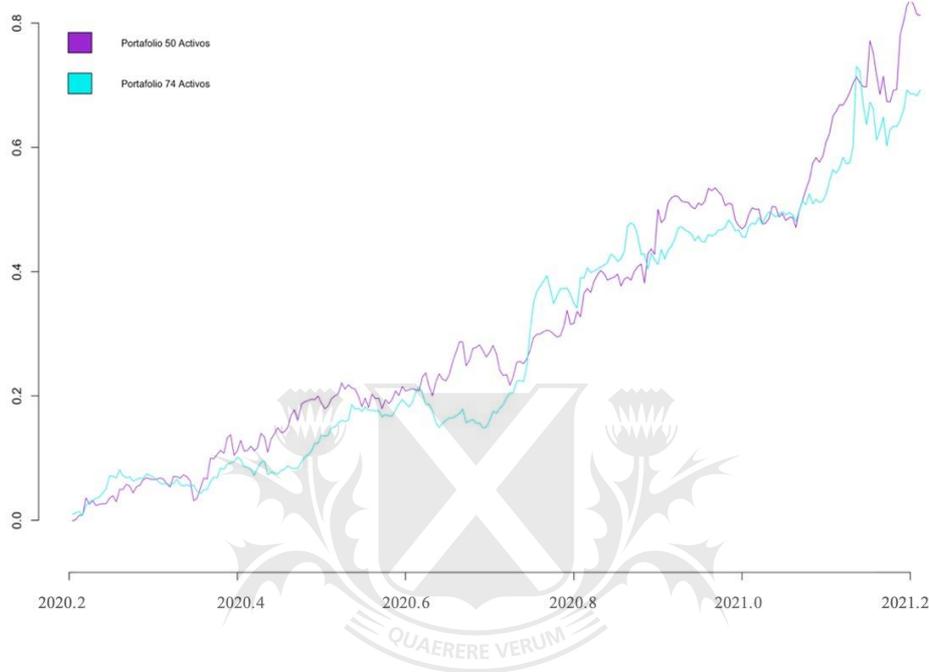
Tabla 2: Compañías contenidas en el percentil 0.85 (Activos S&P 500)

Ticker	Company Name	Industry	Sub-Industry
AAL	American Airlines Group	Industrials	Airlines
ABBV	AbbVie	Health Care	Pharmaceuticals
ABMD	Abiomed	Health Care	Health Care Equipment
AEP	American Electric Power	Utilities	Electric Utilities
AKAM	Akamai Technologies	Information Technology	Internet Services & Infrastructure
ALXN	Alexion Pharmaceuticals	Health Care	Pharmaceutical
AMD	Advanced Micro Devices	Information Technology	Semiconductors
AMZN	Amazon	Consumer Discretionary	Internet & Direct Marketing Retail
ANET	Arista Networks	Information Technology	Communications Equipment
ATVI	Activision Blizzard	Communication Services	Interactive Home Entertainment
BDX	Becton Dickinson	Health Care	Health Care Equipment
BIIB	Biogen	Health Care	Biotechnology
BIO	Bio-Rad Laboratories	Health Care	Life Sciences Tools & Services
BMJ	Bristol Myers Squibb	Health Care	Health Care Distributors
CAG	Conagra Brands	Consumer Staples	Packaged Foods & Meats
COO	Cboe Global Markets	Financials	Financial Exchanges & Data
CHD	Church & Dwight	Consumer Staples	Household Products
CHRW	C. H. Robinson	Industrials	Air Freight & Logistics
CHTR	Charter Communications	Communication Services	Cable & Satellite
CLX	Clorox	Consumer Staples	Household Products
CMG	Chipotle Mexican Grill	Consumer Discretionary	Restaurants
COG	Cabot Oil & Gas	Energy	Natural gas & Oil
CPB	Campbell Soup	Consumer Staples	Packaged Foods & Meats
CTXS	Citrix Systems	Information Technology	Application Software
DG	Dollar General	Consumer Discretionary	General Merchandise Stores
DISCA	Discovery (Series A)	Communication Services	Broadcasting
DLR	Digital Realty Trust	Real Estate	Specialized REITs
DLTR	Dollar Tree	Consumer Discretionary	General Merchandise Stores
DPZ	Domino's Pizza	Consumer Discretionary	Restaurants
DVA	DaVita	Health Care	Health Care Facilities
DXCM	DexCom	Health Care	Health Care Equipment
EA	Electronic Arts	Communication Services	Interactive Home Entertainment
EBAY	eBay	Consumer Discretionary	Internet & Direct Marketing Retail
ED	Consolidated Edison	Utilities	Electric Utilities
ENPH	Enphase Energy	Information Technology	Electronic Components
ETSY	Etsy	Consumer Discretionary	Internet & Direct Marketing Retail
EXR	Extra Space Storage	Real Estate	Specialized REITs
FINTE	Fortinet	Information Technology	Systems Software
GILD	Gilead Sciences	Health Care	Biotechnology
GIS	General Mills	Consumer Staples	Packaged Foods & Meats
HRL	Hormel	Consumer Staples	Packaged Foods & Meats
INCY	Incyte	Health Care	Biotechnology
IPGP	IPG Photonics	Information Technology	Electronic Manufacturing Services
K	Kellogg's	Consumer Staples	Packaged Foods & Meats
KHC	Kraft Heinz	Consumer Staples	Packaged Foods & Meats
KMB	Kimberly-Clark	Consumer Staples	Household Products
KR	Kroger	Consumer Staples	Food Retail
LB	L Brands	Consumer Discretionary	Clothing
LLY	Eli Lilly	Health Care	Pharmaceuticals
LUMN	Lumen Technologies	Communication Services	Alternative Carriers
MKC	McCormick & Company	Consumer Staples	Packaged Foods & Meats
MKTX	MarketAxess	Financials	Financial Exchanges & Data
NEM	Newmont	Materials	Gold
NFLX	Netflix	Communication Services	Movies & Entertainment
NLOK	NortonLifeLock	Information Technology	Application Software
NWL	Newell Brands	Consumer Discretionary	Housewares & Specialties
PRGO	Perrigo	Health Care	Pharmaceutical
PSA	Public Storage	Real Estate	Specialized REITs
REGN	Regeneron Pharmaceuticals	Health Care	Biotechnology
SJM	JM Smucker	Consumer Staples	Packaged Foods & Meats
TGT	Target	Consumer Discretionary	General Merchandise Stores
TSCO	Tractor Supply Company	Consumer Discretionary	Specialty Stores
TSLA	Tesla	Consumer Discretionary	Automobile Manufacturers
TSN	Tyson Foods	Consumer Staples	Packaged Foods & Meats
TTWO	Take-Two Interactive	Communication Services	Interactive Home Entertainment
TWTR	Twitter	Communication Services	Interactive Media & Services
TYL	Tyler Technologies	Information Technology	Application Software
VRTX	Vertex Pharmaceuticals	Health Care	Biotechnology
VTRS	Viatis	Health Care	Pharmaceuticals
VZ	Verizon Communications	Communication Services	Integrated Telecommunication Services
WBA	Walgreens Boots Alliance	Consumer Staples	Drug Retail
WEC	WEC Energy Group	Utilities	Electric Utilities
WMT	Walmart	Consumer Staples	Hypermarkets & Super Centers
WST	West Pharmaceutical Services	Health Care	Health Care Supplies

Analizando la Tabla (2) se pueden llegar a conclusiones similares a las obtenidas con las 50 compañías. Sin embargo, en este caso, notamos que se incluyen compañías que pertenecen a la industria inmobiliaria, industriales (aerolíneas), entre otras.

Entonces, a partir de estos nuevos resultados observamos si resulta más significativo tener en cuenta una mayor cantidad de compañías. Por ello, analizamos si este portafolio supera a la cartera construida previamente (la cual contiene 50 empresas) y así, al índice y puede mitigar aún más la crisis.

Figura 10: Portafolio 50 Activos vs Portafolio 74 Activos



A partir de la Figura (10) podemos concluir que el portafolio compuesto por 50 activos resulta ser superador; este le gana al portafolio compuesto por 74 acciones en la mayor parte del periodo, al mismo tiempo que posee un comportamiento menos volátil. Considerando los resultados obtenidos podemos argumentar que retener más acciones de las óptimas cuando el mercado se encuentra más correlacionado se vuelve redundante.

Para poder realizar una mejor comparación entre ambas carteras y observar realmente cual es mejor en términos de diversificación y rendimiento, se calculó el ratio de Sharpe para ambos portafolios.

Tabla 3: Ratio de Sharpe para los portafolios construidos a partir de los activos del S&P 500

	Ratio de Sharpe
Portafolio 50 compañías	3.28
Portafolio 74 compañías	2.79

El ratio de Sharpe es un ratio que calcula la rentabilidad ajustada según su riesgo y considera el retorno, la *risk free*³ y el desvío de un portafolio. En este caso, se puede

³ La *risk free* utilizada para el cálculo es anualizada a 10 años para la fecha de estudio. Recuperado de: <https://fred.stlouisfed.org/series/DGS10>

observar claramente que el ratio del portfolio construido a partir de 50 compañías es más alto que el construido a partir de las 74 compañías. A partir de esto concluimos que el portfolio de 50 compañías es mejor en términos de rentabilidad ajustada que el portafolio de las 74. Aunque el portafolio construido a partir de 74 compañías tiene una diversificación mayor, esta presenta menos rendimiento ajustado por riesgo. Esto podría atribuirse al hecho de que se agregan compañías que, a pesar de tener buenos fundamentos, se encuentran más correlacionadas entre sí dada la volatilidad del mercado al inicio del 2020.

En línea con lo planteado anteriormente, también se realizó un análisis de sensibilidad para los percentiles 0.95 (constituido por 25 compañías) y el percentil 0.97 (constituido por 15 compañías).

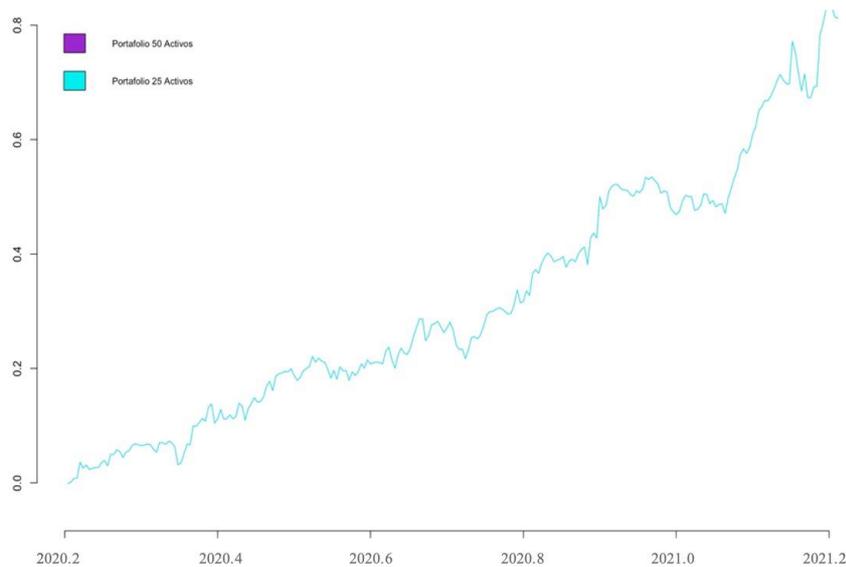
Para el caso del percentil 0.95, se puede ver a continuación en la Tabla (4) las 25 compañías seleccionadas a partir de los autovectores de PCA con sus industrias.

Tabla 4: Compañías contenidas en el percentil 0.95 (Activos S&P 500)

Ticker	Company Name	Industry	Sub-Industry
ABMD	Abiomed	Health Care	Health Care Equipment
ATVI	Activision Blizzard	Communication Services	Interactive Home Entertainment
BIIB	Biogen	Health Care	Biotechnology
CAG	Conagra Brands	Consumer Staples	Packaged Foods & Meats
CHD	Church & Dwight	Consumer Staples	Household Products
CLX	Clorox	Consumer Staples	Household Products
CPB	Campbell Soup	Consumer Staples	Packaged Foods & Meats
CTXS	Citrix Systems	Information Technology	Application Software
DPZ	Domino's Pizza	Consumer Discretionary	Restaurants
DXCM	DexCom	Health Care	Health Care Equipment
EA	Electronic Arts	Communication Services	Interactive Home Entertainment
ENPH	Enphase Energy	Information Technology	Electronic Components
GILD	Gilead Sciences	Health Care	Biotechnology
GIS	General Mills	Consumer Staples	Packaged Foods & Meats
HRL	Hormel	Consumer Staples	Packaged Foods & Meats
INCY	Incyte	Health Care	Biotechnology
K	Kellogg's	Consumer Staples	Packaged Foods & Meats
KR	Kroger	Consumer Staples	Food Retail
NEM	Newmont	Materials	Gold
NFLX	Netflix	Communication Services	Movies & Entertainment
NLOK	NortonLifeLock	Information Technology	Application Software
REGN	Regeneron Pharmaceuticals	Health Care	Biotechnology
SJM	JM Smucker	Consumer Staples	Packaged Foods & Meats
TSLA	Tesla	Consumer Discretionary	Automobile Manufacturers
TTWO	Take-Two Interactive	Communication Services	Interactive Home Entertainment

Nuevamente, a partir de las 25 acciones superiores construimos un portafolio y observamos si con una menor cantidad de compañías resulta posible tener una cartera ganadora respecto al portafolio inicial de 50 activos y a su vez, al S&P 500.

Figura 11: Portafolio 50 Activos vs Portafolio 25 Activos



La Figura (11) muestra que los resultados obtenidos para ambos portafolios son idénticos, dado que ambas carteras se mueven exactamente igual para el segundo periodo de análisis. Entonces, se puede concluir que, incluir o no, 25 compañías más (véase Anexo) resulta redundante ya que ambos portafolios tienen el mismo rendimiento.

Por último, observamos en la Tabla (5) los resultados para el caso del portafolio construido por 15 activos (dentro del percentil 0.97) seleccionados a través de nuestro método.

Tabla 5: Compañías contenidas en el percentil 0.97 (Activos S&P 500)

Ticker	Company Name	Industry	Sub-Industry
ABMD	Abiomed	Health Care	Health Care Equipment
CAG	Conagra Brands	Consumer Staples	Packaged Foods & Meats
CHD	Church & Dwight	Consumer Staples	Household Products
CLX	Clorox	Consumer Staples	Household Products
CPB	Campbell Soup	Consumer Staples	Packaged Foods & Meats
DPZ	Domino's Pizza	Consumer Discretionary	Restaurants
DXCM	DexCom	Health Care	Health Care Equipment
ENPH	Enphase Energy	Information Technology	Electronic Components
GIS	General Mills	Consumer Staples	Packaged Foods & Meats
K	Kellogg's	Consumer Staples	Packaged Foods & Meats
KR	Kroger	Consumer Staples	Food Retail
NEM	Newmont	Materials	Gold
NLOK	NortonLifeLock	Information Technology	Application Software
REGN	Regeneron Pharmaceuticals	Health Care	Biotechnology
SJM	JM Smucker	Consumer Staples	Packaged Foods & Meats

La diferencia que se encuentra entre la cartera construida por las 50 compañías y la resultante puede observarse en el Anexo 1. A partir de estos resultados notamos que, si bien la cantidad de industrias tenidas en cuenta es menor, el método se vuelve más selectivo a la hora de elegir las compañías. En particular, observamos que la industria más seleccionada por el modelo de PCA es la de bienes de consumo, seguida por la del

cuidado de la salud, al mismo tiempo que la ponderación del sector de consumo discrecional se reduce drásticamente. Considerando las restricciones aplicadas en Estados Unidos en el primer trimestre de 2020 resulta posible argumentar que estas 15 empresas se beneficiaron de las nuevas tendencias de consumo producto de la pandemia ya que su esquema de ventas no es cíclico. A diferencia de las empresas de consumo discrecional, las compañías seleccionadas por el modelo son aquellas que venden productos de insumo básicos y necesarios tales como productos del hogar o comida.

Habiendo analizado las compañías dentro del percentil superior, construimos el portafolio de 15 activos y observamos su comportamiento para la segunda ventana de análisis.

Figura 12: Portafolio 50 Activos vs Portafolio 15 Activos



En la Figura (12) notamos que la línea de tendencia del portafolio resulta ser idéntica a la cartera construida a partir de 50 activos y, por lo tanto, también a la de 25 activos. De igual manera, el ratio de Sharpe calculado para estos portafolios resulta ser el mismo. Este resultado sugiere que, para un periodo de estrés financiero, no vale la pena tener más activos de los necesarios: a pesar de que se agreguen compañías estas no contribuyen al rendimiento y por ello retenerlas resulta inútil. Por lo tanto, concluimos que a partir de un portafolio de 15 activos se pueden explotar todos los beneficios de la diversificación. Sin embargo, destacamos que, si bien para la segunda ventana de análisis el resultado de este portafolio es positivo, a medida que las correlaciones de mercado cambian, retener tan pocas compañías podría resultar en un riesgo mayor. Por ello, este es un análisis que debe hacerse considerando el momento y las condiciones de mercado actuales.

Habiendo definido que el portafolio de 15 activos es el ganador, resulta pertinente analizar los *weights* negativos para la misma cantidad de acciones. En la Tabla (6) se observan las acciones con menores retornos, en el periodo de crisis, según PCA.

Tabla 6: Compañías contenidas en el percentil 0.03 (Activos S&P 500)

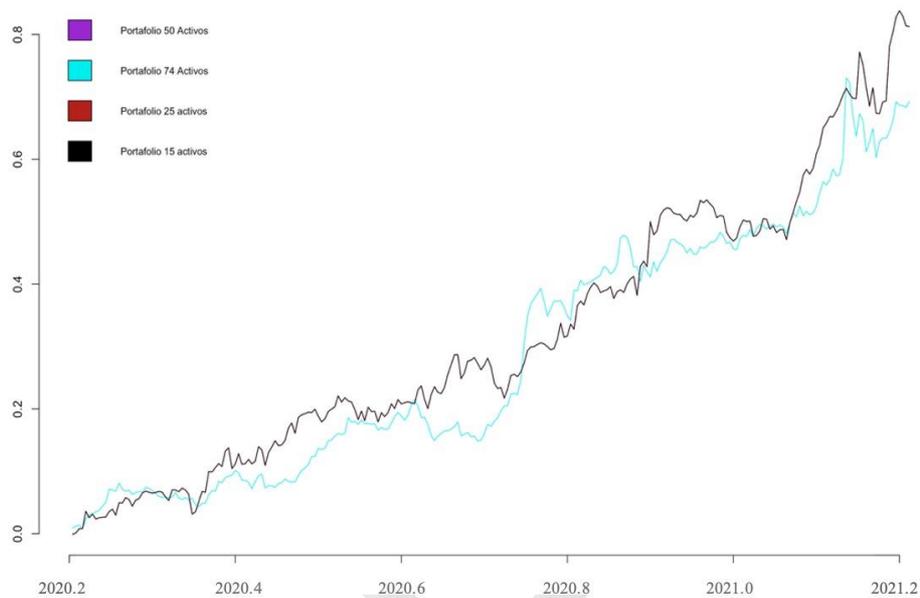
Symbol	Security	GICS Sector	GICS Sub-Industry
AXP	American Express	Financials	Consumer Finance
AMP	Ameriprise Financial	Financials	Asset Management & Custody Banks
AME	Ametek	Industrials	Electrical Components & Equipment
BAC	Bank of America	Financials	Diversified Banks
C	Citigroup	Financials	Diversified Banks
DOV	Dover Corporation	Industrials	Industrial Machinery
ETN	Eaton	Industrials	Electrical Components & Equipment
GL	Globe Life	Financials	Life & Health Insurance
HON	Honeywell	Industrials	Industrial Conglomerates
JPM	JPMorgan Chase	Financials	Diversified Banks
LNC	Lincoln National	Financials	Multi-line Insurance
MET	MetLife	Financials	Life & Health Insurance
MS	Morgan Stanley	Financials	Investment Banking & Brokerage
PH	Parker-Hannifin	Industrials	Industrial Machinery
PRU	Prudential Financial	Financials	Life & Health Insurance

A partir de los resultados queda en evidencia que las acciones que se encuentran en el rango inferior son aquellas que fueron las más afectadas por la crisis del COVID-19. Estas, en su mayoría, pertenecen a industrias de servicios financieros, industriales y de tecnología de la información. Nuevamente, los resultados obtenidos coinciden con el hecho de que estos son los sectores más perjudicados por la crisis a raíz de las medidas de cuarentena.

Al realizar un análisis de las compañías más afectadas por la crisis, se puede observar que la lista está constituida por los bancos más grandes del mundo como, por ejemplo, Citibank, JP Morgan, Bank of America, entre otros. Estos resultados son interesantes debido a que durante el inicio de la pandemia el sector bancario fue uno de los más afectados. Además, al descartar estas compañías se evita una caída grande producto de la correlación o co-movimiento entre todas ellas. Es decir, resulta probable que estas empresas están expuestas a múltiples factores de riesgo y se muevan con el mercado, por lo que excluirlas podría resultar beneficioso. Así también, se observa que empresas dedicadas a la maquinaria industrial forman parte de la lista; esto también tiene sentido ya que dichos sectores se vieron obligados a frenar sus operaciones por la pandemia y por ello se encontraron perjudicados.

En conclusión, se puede decir que, la cartera construida a partir de 15 compañías en el percentil superior resulta ser la mejor en torno al rendimiento ajustado por el riesgo y en términos de selección de activos. En la figura (13) observamos una comparación entre esta cartera y las analizadas anteriormente.

Figura 13: Portafolios en comparación construidos a partir de los activos del S&P 500



De esta forma, concluimos que resulta posible construir un portafolio ganador en periodos de estrés financiero a partir de PCA. A partir del primer componente principal resulta posible agrupar las mejores compañías, menos afectadas por el COVID-19, y construir un portafolio *equally weighted* que se encuentra *long* en las mismas. De este modo, la cartera resultante de 15 activos resulta tener un rendimiento muy superior al del índice de la región americana, el S&P 500.

5.2 FTSE 100

5.2.1 Portafolio PCA para la región británica

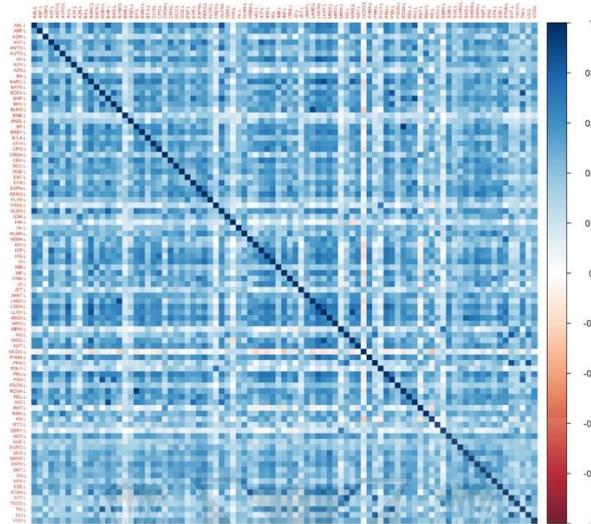
En la presente sección investigamos en detalle la efectividad del método de PCA para modelar el comportamiento de un *benchmark*, en particular, el FTSE 100. Consideramos que este es un análisis pertinente ya que los distintos países lidiaron con la crisis sanitaria y las políticas económicas de maneras distintas. Por lo tanto, en este caso buscamos observar si resulta posible obtener conclusiones significativas para otra región, donde el impacto del COVID -19 fue distinto.

Para ello, aquí también aplicamos PCA, para todo el periodo de estudio, sobre los 91 activos que se encontraban contenidos en el índice a la fecha de análisis. A partir de esto, obtuvimos una matriz P de dimensiones 1013×91 donde en cada columna se observan los componentes principales calculados.

La Figura (14) muestra la matriz de correlaciones entre los activos. Contemplamos que el color que predomina es el azul, aunque este es aún menos intenso que para el S&P 500. Así mismo, al observar el gráfico con detenimiento queda en evidencia que parecería

haber una presencia muy pequeña de colores rojos, que tienen que ver con correlaciones negativas.

Figura 14: Matriz de Correlaciones de los activos del FTSE 100



Estos resultados indican que, aunque los activos dentro de la matriz R se mueven conjuntamente, estos no se encuentran tan correlacionados como los datos para el *benchmark* estadounidense. Sin embargo, como observamos en la Figura (2) el FTSE 100 experimentó una caída de mayor magnitud y, para el periodo de estudio, no logró recuperarse completamente. Esto nos llevará a pensar que el primer componente principal debería explicar una menor proporción de la varianza a la contemplada anteriormente.

Procedemos a analizar la varianza explicada por los componentes principales en las dos distintas ventanas. La varianza explicada por el primer componente fue del 26,02% antes de la crisis y, al igual que en el caso del S&P 500, se duplicó a partir del inicio de la pandemia. Para la segunda ventana la varianza explicada por el primer componente resultó ser de 43,95%. Por lo tanto, estos primeros datos no parecerían indicar una diferencia sustancial respecto al índice de Estados Unidos.

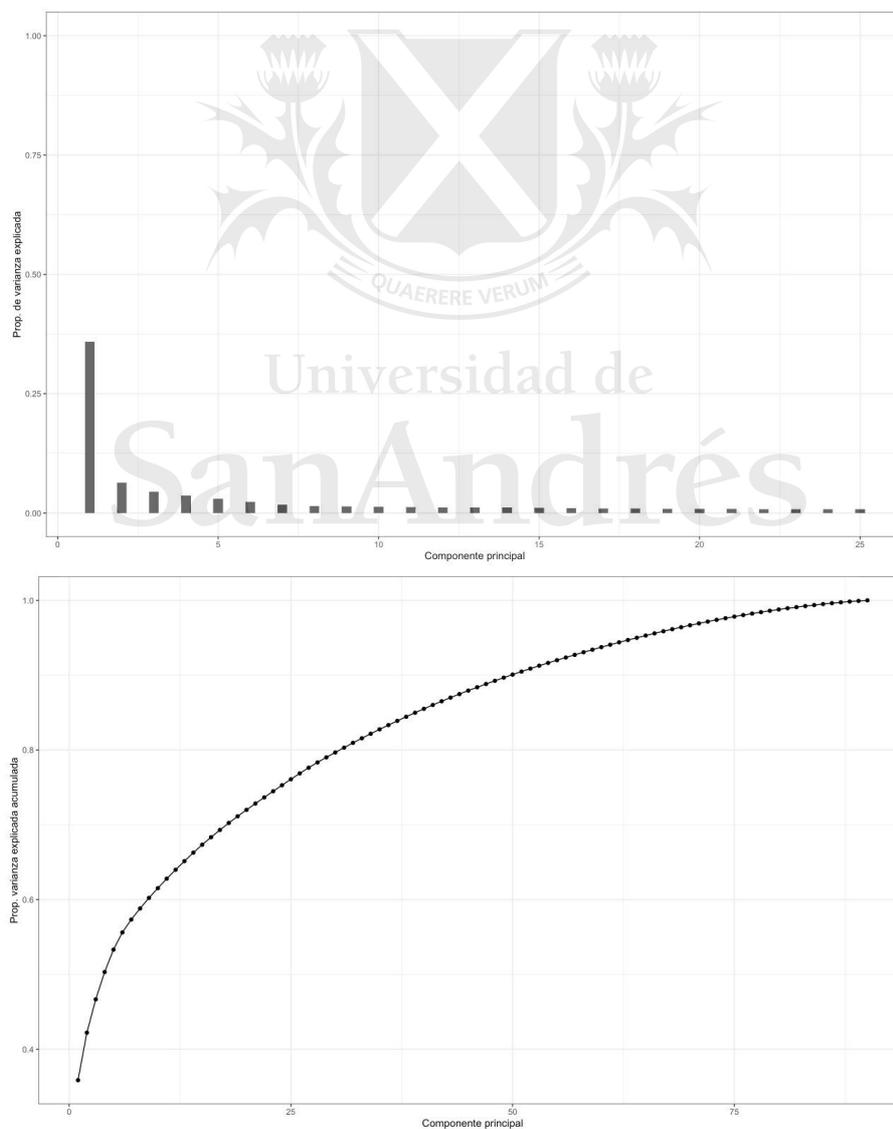
Sin embargo, al estudiar la proporción de la varianza para toda la ventana cambian las conclusiones. Esto se debe a que para todo el periodo de análisis el primer componente explica tan solo un 0,34 de la varianza. Este número parecería indicar una menor correlación entre los activos que componen el FTSE 100.

Utilizando el criterio de Kaiser y de la varianza acumulada decidimos cuántos componentes deben retenerse en la ventana completa, para esta región. Observamos que a partir del primer criterio resultan ser 14 componentes los que tienen autovalores mayores a 1. Sin embargo, a partir del segundo criterio se puede concluir que son 30 componentes los que explican un 80% de la varianza. Esta vez, los datos muestran una diferencia más alta entre las reglas de decisión. Por esto, consideramos que mejores

resultados pueden surgir a partir de una combinación de ambos criterios. Si se retuviesen 14 componentes, como lo indica el criterio de Kaiser, solo se estaría explicando un 66% de la varianza de los datos. Esto podría significar un problema ya que, como mencionamos anteriormente, se podría estar captando ruido de la serie. A su vez, si se retuviese un número mucho menor a 30 componentes se podría estar perdiendo información significativa. Por lo tanto, decidimos retener 25 componentes, de esta forma no se pierde demasiada información y se logra explicar un 75% de la variación de los datos originales.

Entonces, se logró transformar el conjunto original de 91 activos en 25 componentes incorrelacionados que identifican los principales factores de riesgo de las acciones. En este caso, notamos que reteniendo un mayor porcentaje de los componentes principales, un 27%, se puede explicar la proporción de la varianza de la serie completa.

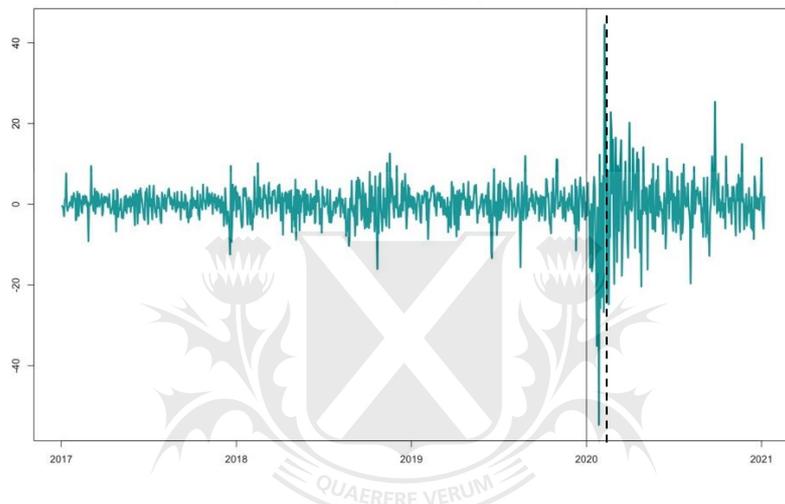
Figura 15: Proporción de la varianza explicada por los componentes principales (Activos FTSE 100)



El gráfico superior de la Figura (15) muestra la proporción de la varianza explicada por cada componente, mientras que, el gráfico inferior muestra la proporción de la varianza acumulada explicada. Luego del primer componente principal, cada componente adicional va aproximando cada vez menos de la varianza de los datos.

A pesar de que el primer componente principal no logra capturar una proporción alta de la volatilidad de los datos, en la Figura (16) observamos que el comportamiento de este resulta ser análogo a los retornos del FTSE 100.

Figura 16: Primer Componente Principal (FTSE 100)



A partir de los *weights* de los 25 componentes retenidos, se puede construir un portafolio con el objetivo de estudiar si es probable reproducir el comportamiento del FTSE 100 en todo el periodo de análisis.

Figura 17: Portafolio PCA vs FTSE 100



Observando la Figura (17) se puede establecer que la hipótesis planteada puede comprobarse también para una región distinta. Observamos que la cartera formada no

solo logra replicar bien el comportamiento del índice en su totalidad, sino que logro superar el *benchmark* ampliamente en todo el periodo estudiado. Además, la caída producto del COVID-19 resultó tener una magnitud mucho menor que la del índice. Esto refleja una conclusión distinta a la planteada anteriormente para la región norte americana, donde el portafolio construido a partir de los componentes principales rindió menos que su *benchmark*. Sin embargo, en este caso también resulta posible que el rendimiento superior del portafolio que se observa en esta primera ventana podría ser resultado de la inclusión de los dividendos en la matriz *R*.

5.2.2 selección de Activos para la región británica

En esta sección proponemos un método para seleccionar activos a través de PCA, enfocándonos en el segundo periodo de estudio, pero considerando una región distinta. De esta forma, buscamos constatar si para este caso, dada la diferencia en los datos, también es posible construir un portafolio ganador utilizando el primer componente principal como indicador de mercado.

Replicaremos toda la metodología de la selección de activos y análisis de sensibilidad para el caso del índice británico. Para ello, comenzamos calculando el percentil del 0.90, para observar los activos con mayores ponderaciones. Al analizar dicho percentil superior resulta posible observar las acciones con mayores retornos, en el periodo comprendido por la crisis, según PCA.

Tabla 7: Compañías contenidas en el percentil 0.90 (Activos FTSE 100)

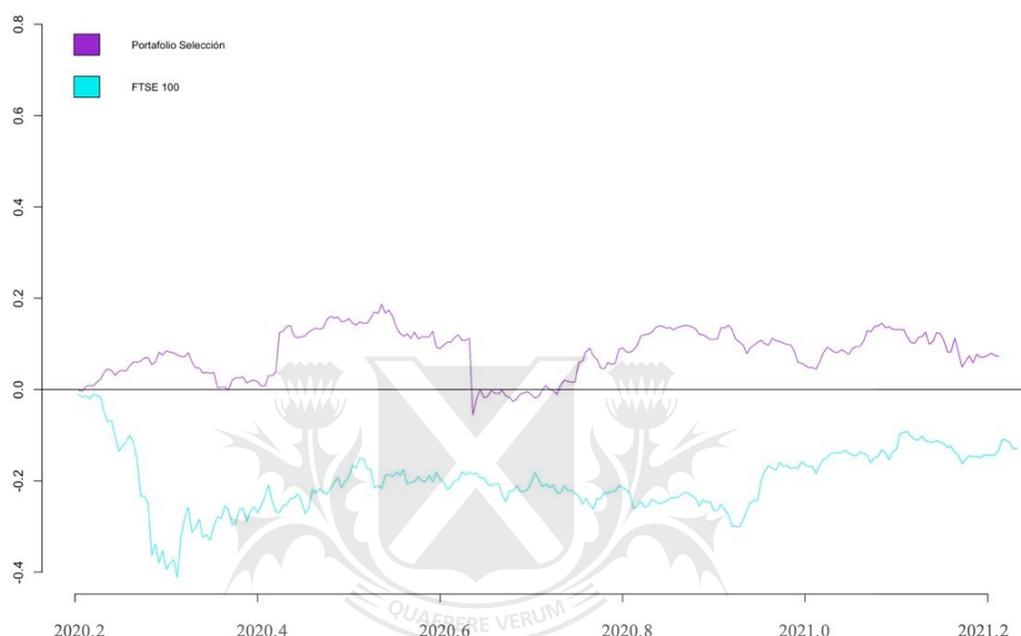
Company	Ticker	Industry
AstraZeneca	AZN	Pharmaceuticals & Biotechnology
B&M	BME	Retailers
Fresnillo	FRES	Mining
Hikma Pharmaceuticals	HIK	Pharmaceuticals & Biotechnology
William Morrison Supermarkets plc	MRW	Retailers
Ocado Group	OCDO	Food & Drug Retailers
Polymetal International	POLY	Precious Metals & Mining
Reckitt	RKT	Household Goods & Home Construction
Sainsbury's	SBRY	Food & Drug Retailers

En la Tabla (7) se muestran las compañías que se encuentran contenidas en el percentil 0.90. Es decir, se observan las 9 empresas que presentan el mejor retorno. Al igual que en el caso del S&P 500, se nota que las acciones contenidas en dicha tabla, no casualmente, fueron las menos afectadas por la crisis del COVID-19. En este caso los activos seleccionados pertenecen a industrias farmacéuticas, de medicamentos y también, al sector de bienes de consumo. Empresas como William Morrison Supermarkets, SainsBury's son claros ejemplo: estas proveen bienes necesarios no de lujo, los cuales vieron un incremento en su demanda. Del mismo modo, AstraZeneca experimentó un gran crecimiento en el 2020 producto de la distribución de la vacuna contra el COVID-19. A su vez, observamos la presencia de industrias de minería o metales preciosos; esto

coincide con el hecho de que, a raíz de la incertidumbre económica, estos sectores estuvieron en auge, sobre todo para el primer trimestre del 2020.

A partir de las compañías contenidas en el percentil superior de los autovectores construimos un portafolio *equally weighted*.

Figura 18: Portafolio 9 activos vs FTSE 100



Universidad de

Como se puede observar en la Figura (14), arribaremos a la misma conclusión que con el índice del S&P 500, construyendo un portafolio que este *long* en las compañías conformadas por el 10% superior según los pesos de PCA, se puede mitigar el efecto de la crisis y, del mismo modo, ganarle al índice del FTSE 100. En el gráfico se observa como claramente la cartera resultante se comportó significativamente mejor que el mercado, aunque para un breve periodo presentó rendimientos negativos. Sin embargo, siempre se movió muy por encima del FTSE 100.

Luego de realizar una conclusión positiva respecto de la cartera construida, resulta pertinente realizar un análisis de sensibilidad de los resultados obtenidos. Para este caso, en comparación con el índice estadounidense, realizamos un análisis de sensibilidad incluyendo más compañías, no menos. Esto fue así porque, desde el punto de vista de la diversificación no resultaría una buena decisión de inversión construir un portafolio con menos de 9 activos, ya que de esta forma se está muy expuesto al riesgo intrínseco de cada compañía.

Para el análisis de sensibilidad realizamos una comparación entre: el percentil 0.85 (constituido por 14 compañías), el percentil 0.80 (constituido por 18 compañías) y el percentil 0.75 (constituido por 23 compañías).

Comenzamos por analizar las compañías el percentil 0.85. En la Tabla (8) se pueden ver presentadas las compañías que lo constituyen.

Tabla 8: Compañías contenidas en el percentil 0.85 (Activos FTSE 100)

Company	Ticker	Industry
Admiral Group	ADM	Nonlife Insurance
AstraZeneca	AZN	Pharmaceuticals & Biotechnology
B&M	BME	Retailers
Fresnillo	FRES	Mining
Hikma Pharmaceuticals	HIK	Pharmaceuticals & Biotechnology
Just Eat Takeaway.com	JET	Online food ordering
William Morrison Supermarkets plc	MRW	Retailers
Ocado Group	OCDO	Food & Drug Retailers
Pennon Group plc	PNN	Water and Waste
Polymetal International	POLY	Precious Metals & Mining
Pearson plc	PSON	Education
Reckitt	RKT	Household Goods & Home Construction
Sainsbury's	SBRY	Food & Drug Retailers
Tesco	TSCO	Food & Drug Retailers

Las empresas que fueron agregadas en esta instancia se encuentran representadas en negrita. Observamos que si bien se agregaron empresas que se beneficiaron por la crisis sanitaria debido al aumento en la demanda por el consumo de bienes necesarios, tales como Just Eat Takeaway.com o los supermercados Tesco, también se agregaron otro tipo de industrias. Por ejemplo, observamos entre la selección compañías de seguros de auto, educación y de agua y residuos. En particular, estas son empresas que, aunque no se encontraban dentro de los sectores más demandados en el 2020, tuvieron muy buenos rendimientos debido a que vieron una oportunidad para readaptar sus negocios. Por ejemplo, Pennon adaptó su modelo de negocios hacia uno más sustentable. También, Admiral Group logró sobrevivir al contexto pandémico y evitó una caída en las ventas de seguro debido a sus políticas implementadas: a través del “Admiral Stay at Home Refund” lograron retener y también atraer nuevos clientes. Del mismo modo, aunque Pearson se vio afectada inicialmente por el cierre de sus centros educativos, logró readaptarse rápidamente a las tendencias de la pandemia ofreciendo cursos online y de esta forma consiguiendo captar un nuevo mercado.

A partir de los 14 activos seleccionados por el método de PCA construimos un nuevo portafolio y analizamos los resultados obtenidos.

Figura 19: Portafolio 9 Activos vs Portafolio 14 Activos



La Figura (15) muestra que el portafolio de 14 activos tiene un mejor rendimiento considerando toda la etapa estudiada, aunque por un corto periodo de tiempo el portafolio de 9 activos lo supera. A pesar de esto, se observa que el portafolio construido a partir de 14 compañías es significativamente mejor: este nunca presenta rendimientos negativos, y comienza a despegarse para los últimos meses del año. Esto quiere decir que las cinco compañías adicionales agregadas poseen una ponderación significativa en la construcción del portafolio. Al mismo tiempo, al contemplar más industrias se logra una mayor diversificación.

Transformando esta interpretación en una observación numérica, calculamos el ratio de Sharpe para ambos portafolios⁴.

Tabla 9: Ratio de Sharpe para los portafolios construidos a partir de los activos del FTSE 100

	Ratio de Sharpe
Portafolio 9 compañías	0.23
Portafolio 14 compañías	1.61

Dado los resultados, se concluye que el portafolio construido a partir de 14 empresas es realmente superador al de 9. Este presenta un rendimiento ajustado por riesgo mucho más alto y es por esta razón que se toma la decisión de elegir este portafolio como el óptimo, o el ganador.

⁴ La *risk free* utilizada para el cálculo es anualizada a 10 años para la fecha de estudio. *Recuperado de:* <https://fred.stlouisfed.org/series/DGS10>

Sin embargo, resta observar que sucede cuando se consideran percentiles más bajos, y por ende, más empresas. Se continuo por realizar el cálculo de las compañías que se encuentran contenidas tanto en el percentil 0.80, como en el percentil 0.75. Es decir, se observan las 18 y 23 compañías que presentan el mejor retorno. Los activos seleccionados pueden verse en el Anexo 2 y Anexo 3, respectivamente.

Los portafolios obtenidos a partir de las 18 y 23 compañías muestran resultados idénticos al portafolio construido a partir de 14 acciones tanto en tendencia como en rendimiento ajustado por riesgo. Si bien se agregan algunas industrias, los resultados sugieren que estas no aportan un rendimiento significativo y que las compañías agregadas tienen una ponderación cercana a cero dentro de los portafolios. Es decir, retener más acciones de las óptimas, cuando el riesgo sistémico aumenta, es redundante. Por ello, para la región británica decidimos que el portafolio de 14 activos es el que posee los resultados más eficientes.

Habiendo definido que el portafolio de 14 activos es el ganador, resulta pertinente analizar los *weights* negativos para la misma cantidad de acciones. Entonces, analizamos el percentil 0.15 para poder observar cuales son las compañías que se vieron más perjudicadas por la crisis. En la Tabla (10) se puede observar el listado de compañías que lo componen.

Tabla 10: Compañías contenidas en el percentil 0.15 (Activos FTSE 100)

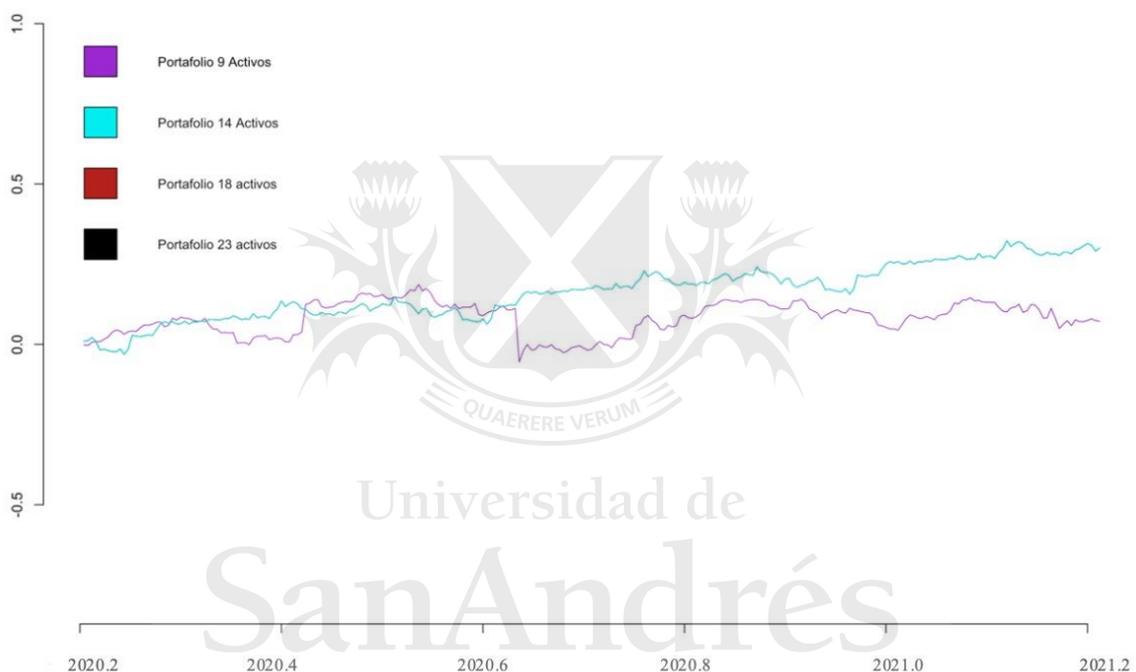
Company	Ticker	Industry
Anglo American plc	AAL	Mining
Aviva plc	AV	Financial Services
Barclays	BARC	Banks
Burberry	BRBY	Personal Goods
Ferguson plc	FERG	Support Services
Glencore	GLEN	Mining
3i	III	Financial Services
Legal & General	LGEN	Life Insurance
Mondi	MNDI	Forestry & Paper
Next plc	NXT	General Retailers
Phoenix Group	PHNX	Life Insurance
Prudential plc	PRU	Life Insurance
Schroders	SDR	Financial Services
Smiths Group	SMIN	General Industrials

Igual que para el S&P 500 observamos que las compañías más afectadas durante el 2020 pertenecen a industrias de servicios financieros, bienes de lujo, industriales, entre otros. La lista está constituida por bancos muy grandes como por empresas de servicios financieros, como lo son Barclays, 3i, Schroders, entre otras. Así también, Burberry, al proveer bienes de lujo, es otro ejemplo de una compañía que se vio afectada por la

pandemia debido al cambio en el patrón de consumo. A diferencia de los resultados obtenidos para el S&P 500, aquí vemos que dentro de las industrias con *weight* más negativo se encuentran los seguros de vida. Esto podría explicarse por el hecho de que, producto de las muertes causadas por el COVID-19, estas compañías tuvieron grandes salidas de capital.

En conclusión, se puede decir que, la cartera construida a partir de 14 compañías en el percentil superior resulta ser la mejor en torno al rendimiento ajustado por el riesgo y en términos de selección. En la figura (13) observamos una comparación entre esta cartera y las analizadas anteriormente.

Figura 20: Portafolios en comparación construidos a partir de los activos del FTSE 100



De esta forma, constatamos que, para periodos de alta volatilidad, el método de selección de activos utilizado a partir de PCA resulta ser eficiente independientemente de la región que se considere. En esta sección también observamos que a partir del primer componente principal resultó posible agrupar las mejores compañías y construir un portafolio eficiente, de modo tal que la cartera resultante de 14 activos mostró tener un rendimiento superior al del FTSE 100.

6. Conclusiones

En el presente trabajo hemos analizado la posibilidad de construir una estrategia basada en PCA para la gestión activa de selección de activos en todas las fases de mercado, utilizando dos mercados de valores representativos: Estados Unidos y el Reino Unido.

Los resultados obtenidos sugieren que el análisis de componentes principales es muy útil para reducir la dimensionalidad de un modelo financiero complejo y encontrar el factor determinante del rendimiento de una cartera. Además, observamos que para ambas regiones las carteras formadas a partir de los componentes principales pudieron replicar bien el comportamiento de los índices, excepto durante el período cercano a la crisis causada por el COVID-19 para el caso del S&P 500. A pesar de esto, podemos concluir que los portafolios construidos a partir de los componentes retenidos explican bien las acciones originales.

El hecho de que el segundo periodo de estudio no logre representarse correctamente se debe a que el mercado de valores se encuentra más concentrado en tiempos de crisis. Por lo tanto, resultó posible comprobar que un portafolio que mantiene la misma posición a lo largo del tiempo se encuentra menos diversificado cuando el mercado se encuentra más correlacionado. La varianza explicada por el primer componente principal es una medida eficaz del nivel de riesgo sistémico y sirvió como un indicador de crisis financiera.

Al utilizar el método de selección de activos a partir del weight del primer componente principal resulta posible argumentar que se podría haber evitado la caída significativa que tuvieron las distintas bolsas del mundo producto del COVID-19. El método de selección empleado muestra que se puede segmentar los activos en función de las empresas más y menos afectadas por este evento y de esta forma construir portafolios que resultan ser ganadores. Este es un hallazgo importante ya en esta tesis proporcionamos un medio para que los inversores seleccionen acciones que darán como resultado la diversificación que no se puede lograr fácilmente en periodos donde el mercado se encuentra conectado. Además, los resultados presentados aquí resultan interesantes debido a que también se evalúa cual es la cantidad de activos que diversifican un portafolio en un periodo de alta volatilidad, y, por el contrario, a partir de cuantas acciones los beneficios de la diversificación prácticamente se agotan.

A pesar de estos aportes, no debemos olvidar las limitaciones de estas conclusiones. En primer lugar, debe notarse que no existe una cantidad única de activos que conformen un portafolio ganador. Esto depende tanto de las condiciones del mercado como del método de selección. Como demostramos, a medida que aumentan las correlaciones entre las acciones en el mercado, como lo hacen en tiempos de euforia y crisis, el número de acciones requeridas disminuye. Esto indica que los resultados que obtuvimos aquí corresponden a un periodo en particular, por lo que el presente método de selección no necesariamente garantiza retornos positivos para aquellos tiempos menos volátiles. Una cartera que tiene la misma cantidad de acciones o los mismos componentes solo puede

ser la mejor combinación para crear una cartera diversificada en un solo momento. Entonces, si bien aquí empleamos un método eficiente de selección de activos para un periodo de alta volatilidad de los datos, no analizamos cómo se comporta este frente a datos normales en varianza. Consideramos que esto abre la puerta a futuras líneas de investigación.

En segundo lugar, aunque demostramos que los portafolios contruidos a partir de PCA tienen rendimientos superiores, en la presente tesis no se realizó ninguna evaluación de riesgo de los activos individuales. Dado que no es aconsejable utilizar una única regla de selección para elegir acciones para mantener en una cartera, recomendamos combinar distintos criterios. Por ejemplo, se podría realizar un análisis fundamental de las compañías obtenidas a partir del criterio de selección, o tal como lo hacen Fulga y Dedu (2012), se podría evaluar el riesgo de los activos con la medida de Value at Risk (VaR) y así agrupar los activos de mínimo riesgo y compararlos con la cartera resultante. De esta forma, futuros trabajos que cuenten con un análisis de riesgo de los activos individuales no solo tendrán en consideración la selección de activos ganadores sino también la estimación y optimización de riesgo.

Los resultados obtenidos en esta tesis podrían ser interesantes para una amplia variedad de profesionales. En particular, demostramos que para los inversores que desean replicar un índice o comprar activos individuales, nuestros dos métodos de construcción de carteras con PCA proporcionan una manera de hacerlo posible.

7. Bibliografía

- Allen, G. C. (2010). *The Risk Parity Approach to Asset Allocation*. [web] <http://www.top1000funds.com/attachments/TheRiskParityApproachtoAssetAllocation2010.pdf>.
- Alexander, C. (2008) *Market risk analysis. Volume II. Practical financial econometrics*. John Wiley & Sons Inc, West Sussex, England, 2008.
- DeCarlo, L. T. (1997) On the meaning and use of Kurtosis. [web] <http://www.columbia.edu/~ld208/psymeth97.pdf>
- Fulga, C. y Dedu, S. (2012) *Mean-Risk Portfolio Optimization with prior PCA-Based stock selection*. Bucharest Academy of Economic Studies.
- Hargreaves, C. A. y Mani, C. K. (2015). *The Selection of Winning Stocks Using Principal Component Analysis*. Business Analytics, Institute of Systems Science, National University of Singapore, Singapore.
- Jolliffe, I.T. (1986). *Principal Component Analysis*, Second Edition. Springer Texts in Statistics, Springer New York, 2002.
- Kaiser, H. F. (1960). *The application of electronic computers to factor analysis*. Educational and Psychological Measurement, Vol 20:141–151.
- Mandelbrot, B. B. (1963). *The variation of certain speculative prices*. Journal of Business 36, 392-417
- Meucci, A. (2010) *Managing Diversification*. Risk, 22, 5, p. 74-79.
- Partovi, M. H. y Caputo, M. (2004) *Principal Portfolios: Recasting the Efficient Frontier*. Economics Bulletin, Vol. 7, No. 3 pp. 1-10.
- Yang, L., Rea, W., y Rea, A. (2015) *Stock Selection with Principal Component Analysis, Working Paper*. Department of Economics and Finance. College of Business and Economics, University of Canterbury, New Zealand.

8. Anexos

Anexo 1: Diferencia en compañías entre el portafolio de 50 y 15 activos (Activos S&P 500)

Symbol	Security	GICS Sector	GICS Sub-Industry
ATVI	Activision Blizzard	Communication Services	Interactive Home Entertainment
AMD	Advanced Micro Devices	Information Technology	Semiconductors
AKAM	Akamai Technologies	Information Technology	Internet Services & Infrastructure
AMZN	Amazon	Consumer Discretionary	Internet & Direct Marketing Retail
ANET	Arista Networks	Information Technology	Communications Equipment
BIIB	Biogen	Health Care	Biotechnology
CHRW	C. H. Robinson	Industrials	Air Freight & Logistics
CMG	Chipotle Mexican Grill	Consumer Discretionary	Restaurants
CTXS	Citrix Systems	Information Technology	Application Software
ED	Consolidated Edison	Utilities	Electric Utilities
DVA	DaVita	Health Care	Health Care Facilities
DG	Dollar General	Consumer Discretionary	General Merchandise Stores
DLTR	Dollar Tree	Consumer Discretionary	General Merchandise Stores
EBAY	eBay	Consumer Discretionary	Internet & Direct Marketing Retail
EA	Electronic Arts	Communication Services	Interactive Home Entertainment
ETSY	Etsy	Consumer Discretionary	Internet & Direct Marketing Retail
GILD	Gilead Sciences	Health Care	Biotechnology
HRL	Hormel	Consumer Staples	Packaged Foods & Meats
INCY	Incyte	Health Care	Biotechnology
IPGP	IPG Photonics	Information Technology	Electronic Manufacturing Services
KMB	Kimberly-Clark	Consumer Staples	Household Products
KHC	Kraft Heinz	Consumer Staples	Packaged Foods & Meats
MKTX	MarketAxess	Financials	Financial Exchanges & Data
MKC	McCormick & Company	Consumer Staples	Packaged Foods & Meats
NFLX	Netflix	Communication Services	Movies & Entertainment
TTWO	Take-Two Interactive	Communication Services	Interactive Home Entertainment
TGT	Target	Consumer Discretionary	General Merchandise Stores
TSLA	Tesla	Consumer Discretionary	Automobile Manufacturers
TWTR	Twitter	Communication Services	Interactive Media & Services
VTRS	Viatis	Health Care	Pharmaceuticals
WMT	Walmart	Consumer Staples	Hypermarkets & Super Centers
WEC	WEC Energy Group	Utilities	Electric Utilities
ALXN	Alexion Pharmaceuticals	Health Care	Pharmaceutical
COG	Cabot Oil & Gas	Energy	Natural gas & Oil
PRGO	Perrigo	Health Care	Pharmaceutical

Anexo 2: Compañías contenidas en el percentil 0.80 (Activos FTSE 100)

Company	Ticker	Industry
Admiral Group	ADM	Nonlife Insurance
AstraZeneca	AZN	Pharmaceuticals & Biotechnology
B&M	BME	Retailers
Bunzl	BNZL	Support Services
Croda International	CRDA	Chemicals
Fresnillo	FRES	Mining
Hikma Pharmaceuticals	HIK	Pharmaceuticals & Biotechnology
Just Eat Takeaway.com	JET	Online food ordering
William Morrison Supermarkets plc	MRW	Retailers
Ocado Group	OCDO	Food & Drug Retailers
Pennon Group plc	PNN	Water and Waste
Polymetal International	POLY	Precious Metals & Mining
Pearson plc	PERSON	Education
Reckitt	RKT	Household Goods & Home Construction
Sainsbury's	SBRY	Food & Drug Retailers
Sage Group	SGE	Software & Computer Services
Tesco	TSCO	Food & Drug Retailers
United Utilities Group plc	UU	Water industry

Anexo 3: Compañías contenidas en el percentil 0.75 (Activos FTSE 100)

Company	Ticker	Industry
Admiral Group	ADM	Nonlife Insurance
Auto Trader Group	AUTO	Media
AstraZeneca	AZN	Pharmaceuticals & Biotechnology
B&M	BME	Retailers
Bunzl	BNZL	Support Services
Croda International	CRDA	Chemicals
Fresnillo	FRES	Mining
Hikma Pharmaceuticals	HIK	Pharmaceuticals & Biotechnology
Imperial Brands	IMB	Tobacco
Just Eat Takeaway.com	JET	Online food ordering
William Morrison Supermarkets plc	MRW	Retailers
Ocado Group	OCDO	Food & Drug Retailers
Pennon Group plc	PNN	Water and Waste
Polymetal International	POLY	Precious Metals & Mining
Pearson plc	PERSON	Education
Reckitt	RKT	Household Goods & Home Construction
Rolls-Royce Holding plc	RR	Aerospace, Defence, Energy, Marine
Rentokil Initial	RTO	Support Services
Sainsbury's	SBRY	Food & Drug Retailers
Sage Group	SGE	Software & Computer Services
Severn Trent	SVT	Gas, Water & Multi-utilities
Tesco	TSCO	Food & Drug Retailers
United Utilities Group plc	UU	Water industry