

PREDICCIÓN Y CAPITALIZACIÓN DE LAS TENDENCIAS DEL MERCADO BURSÁTIL EUROPEO: COMPARATIVA DE MÉTODOS PARAMÉTRICOS Y NO PARAMÉTRICOS

Gael Pereira Goce

Trabajo de investigación 20/001

Master en Banca y Finanzas Cuantitativas

Director: Dr. Jesús Ruiz Andújar

Universidad Complutense de Madrid

Universidad del País Vasco

Universidad de Valencia

Universidad de Castilla-La Mancha

**Predicción y capitalización de las tendencias en el mercado
bursátil europeo: Comparativa de métodos paramétricos y no
paramétricos.**

Gael Pereira Goce

Trabajo de investigación
Máster en Banca y Finanzas Cuantitativas

Director/a: Dr. Jesús Ruíz Andújar

Universidad Complutense de Madrid
Universidad del País Vasco
Universidad de Valencia
Universidad de Castilla-La Mancha

www.finanzascuantitativas.com

Abstract

En este trabajo se trata de encontrar el mejor modelo y la mejor combinación de variables macro-financieras para predecir el signo de las tendencias futuras del mercado bursátil europeo. Se emplearán aproximaciones no paramétricas mediante la estimación de probit dinámicos. Para la selección de las variables explicativas a incluir en dichos modelos se emplea un algoritmo genético guiado por los estadísticos de ajuste “in sample”, y se comparan el ajuste y la capacidad predictiva de los modelos resultantes con los derivados de la aplicación de métodos más simples de selección.

Se plantean también especificaciones paramétricas con las mismas variables seleccionadas para los modelos no paramétricos, de modo que se puedan comparar las predicciones realizadas en base a ambas aproximaciones. Se analiza la capacidad predictiva de los modelos para diferentes retardos de las variables explicativas, y por último se establecen diferentes estrategias de inversión mediante las cuales capitalizar sobre los resultados del estudio.

Se encuentra que el algoritmo genético no es en este caso un mecanismo adecuado para la selección sistemática de modelos. También que los modelos no paramétricos son mejores predictores que los paramétricos, y que los mejores probit dinámicos con una única variable exógena devuelven mejores predicciones que los de esta misma clase pero que incorporan múltiples variables. Se encuentra además que el modelo factorial dinámico predice probabilidades más ajustadas a la realidad que los modelos de regresión clásicos con parámetros cambiantes vía Markov Switching. Por último se concluye que el factor momentum y el “Default Yield Spread” son variables relevantes que aportan capacidad predictiva adicional a los modelos probit dinámicos, al combinarlas con las pendientes de las principales curvas de tipos de interés de referencia.

Contents

1	Introducción	5
2	Revisión de Literatura	8
3	Datos	10
4	Metodología	12
4.1	Criterios de selección	12
4.1.1	In Sample	13
4.1.2	Out of Sample	14
4.1.3	Algoritmo Genético	15
4.2	Modelos no Paramétricos	16
4.3	Modelos Paramétricos	22
4.3.1	Modelos con parámetros cambiantes vía Markov Switching	22
4.3.2	Modelo Factorial Dinámico (MS-DFM)	24
4.4	Selección del umbral óptimo (“Threshold Moving”)	28
5	Comparación de Resultados	31
5.1	Diebold & Mariano	31
5.2	Clark & West	33
6	Estrategias	35
6.1	Estrategia Naive	37
6.2	Momentum	38
6.3	Value	39
7	Conclusiones	40
A	Anexo 1. Justificación Teórica Variables Explicativas.	45
B	Anexo 2. Algoritmo de Bry & Boschan (BBA)	51
C	Anexo 3. Algoritmo Genético (GA)	53

List of Figures

1	Representación Variables	46
2	Transformación Variable Dependiente	52
3	Esquema Algoritmo Genético	55
4	Convergencia GA	70
5	Ajuste Modelos Probit vía GA	71
6	Mejor Probit con Varias Explicativas	71
7	Probabilidades Predichas Markov Switching	72
8	Probabilidades estimadas vía MS-DFM	73
9	Curvas ROC	73
10	Estadísticos de Análisis de Umbrales	74
11	Threshold óptimos	75
12	Rendimiento Acumulado Estrategias Para Mejor Modelo	76

List of Tables

1	Variables Explicativas	45
2	Estadísticos descriptivos Variables Exógenas	47
4	Ajuste Probit Exógenos	56
5	Probit Autorregresivos 1	57
6	Probit Autorregresivos 2	58
7	Probit Autorregresivos 3	59
8	Probit Autorregresivos. Variables seleccionadas vía GA	60
9	Probit con Múltiples Explicativas	61
10	Modelos Markov Switching	62
11	Modelo MS-DFM	63
12	Matrices de Confusión	64
13	Resultados Diebold & Mariano	65
14	Resultados Clark & West	67
15	Estrategias	69

1 Introducción

La predicción de las tendencias del mercado bursátil es un tema de incuestionable relevancia para la literatura financiera, debido al importante papel que éstas juegan en diferentes áreas de la economía financiera. La fase del ciclo en que se encuentra el mercado será un factor determinante para diversas secciones de las finanzas, como pueden ser las decisiones de “asset allocation”, de diversificación de la cartera, o la gestión de riesgos. Pero también será clave en otras disciplinas relacionadas como podría ser la gestión de la economía desde un punto de vista macroeconómico.

Dos de los principales hechos estilizados relativos a las series de rendimientos de los activos financieros son la existencia de correlación negativa entre rendimientos y volatilidad, y el aumento de las correlaciones entre los distintos activos en los periodos de elevada volatilidad. Agregando estos hechos, se puede afirmar que las fases bajistas de mercado coincidirán con periodos de elevada volatilidad, y que en tales circunstancias se experimentarán incrementos en los ratios de correlación entre los diferentes activos. Se podrán distinguir así dos posibles estados o regímenes: el alcista o “bullish”, donde se experimentarán periodos de crecimiento sostenido de los precios de los activos y bajas volatilidad y correlaciones; y el bajista o “bearish”, donde se encontrarán series de rendimientos negativos mayores en valor absoluto que en el estado alternativo, elevada volatilidad, y correlaciones más elevadas entre los activos. Si se lograsen evitar los mercados bajistas y con ello las peores caídas del mercado, se podría limitar el riesgo asumido y el “drawdown” de diferentes estrategias de inversión. Por ello, la predicción de las tendencias del mercado agregado constituirá una de las principales preocupaciones de los inversores y gestores activos de fondos, dado que las estrategias planteadas variarán en función de la fase del ciclo bursátil.

En general, se puede considerar que un mercado es bajista si la caída acumulada desde el precio máximo local anterior sobrepasa el 20%. En caso contrario, si se experimentase una caída de los precios de gran magnitud pero que no llegase a sobrepasar el umbral anterior, se consideraría una corrección de mercado pero se mantendría el mercado alcista. Históricamente se ha experimentado al menos una corrección al año, con una duración media de alrededor de los dos meses. Además, las fases “bearish” tienden a ser más breves en duración pero más bruscas, coincidiendo con los hechos estilizados explicados anteriormente. Para la tendencia alternativa se aplica una definición similar, considerando que el mercado es alcista si el retorno acumulado desde mínimos supera el 20%. Se considera que se ha experimentado un ciclo completo cuando se han superado una fase alcista y una bajista (o viceversa), y se ha iniciado ya la siguiente fase. Cabe destacar que estas definiciones incorporan cierto sesgo de retrospección, ya que requieren conocer lo que pasará en el futuro para definir lo que estaría ocurriendo en el presente.

Los ciclos individuales de mercado se podrían considerar como el horizonte básico para los inversores de corto plazo. Éstos podrán utilizar la predicción de los puntos de inflexión de los ciclos del mercado para especular con el rendimiento que presentarán sus activos en cartera en el futuro inmediato. Sin embargo, también los inversores de largo plazo, aquellos que invierten con un horizonte de planificación de más de un ciclo bursátil, pueden hacer uso de esta información. Un inversor de largo plazo podrá rebalancear su cartera, sobreponderando los activos que mejor se adapten a la tendencia bursátil esperada, para aprovechar los beneficios de los mercados alcistas y evitar o limitar las pérdidas en el tramo bajista del ciclo. También podría aprovechar las fases bajistas para incrementar sus posiciones. Por esto, se puede esperar que la incorporación de una forma sistemática de “market timing” al proceso de inversión permita mejorar el desempeño de diferentes estrategias.

Se pueden destacar dos aspectos clave a la hora de definir un método sistemático para la detección de las tendencias de mercado: la selección de las variables que actúen como indicadores adelantados de los cambios de ciclo, y la elección del modelo que las relacione con el mercado.

En este sentido, en base a la teoría macro-financiera, se pueden encontrar múltiples variables con fun-

damentos que justifican su inclusión como predictores de los puntos de inflexión de los ciclos. Tradicionalmente se ha considerado la curva de tipos de interés y sus componentes como uno de los principales indicadores adelantados a la hora de tratar de predecir las tendencias de mercado. Sin embargo se pueden considerar otras muchas variables macro-financieras con potencial capacidad predictiva. La gran cantidad de posibles indicadores adelantados y las consecuentes distintas combinaciones de éstos que se podrían plantear, llevan a intuir que los algoritmos de selección de modelos podrían facilitar y optimizar la tarea.

En cuanto a los modelos que definen la relación entre las variables, tradicionalmente se han empleado métodos paramétricos de vectores autorregresivos o de cointegración. Se plantean así relaciones lineales y multidireccionales entre las variables, y se trata de predecir el rendimiento o el nivel del mercado en los periodos posteriores. Sin embargo, como se comentaba anteriormente, en algunos estadísticos descriptivos de las series de rendimientos¹ se pueden diferenciar dos estados de la naturaleza, los cuales se pueden asociar a las distintas fases del ciclo bursátil. La relación entre algunas de las variables explicativas consideradas y el rendimiento del mercado podrían ser más complejas que la simple relación lineal planteada en los modelos anteriores. Por ejemplo, el efecto de los cambios en los tipos de interés sobre los rendimientos de las acciones podría ser diferente en función de si el mercado es bajista o alcista. Se podría por tanto plantear modelos que reflejen esta relación no lineal entre las variables incorporando cambios de régimen en los regresores del modelo, y estimando la tendencia del mercado como un parámetro más en el proceso de ajuste a los datos. Para ello se podrían plantear modelos de regresión con relaciones unidireccionales entre las variables, o modelos factoriales dinámicos, ambos con regímenes gobernados por procesos de Markov. La principal alternativa que ha surgido en la literatura a los métodos paramétricos mencionados consiste en emplear modelos de variable dependiente limitada, planteando directamente una relación no lineal y unidireccional entre las variables explicativas retardadas y la tendencia del mercado, y haciendo uso de las definiciones teóricas de los ciclos bursátiles expresadas anteriormente. Será necesario transformar previamente el índice de referencia para que tome valores discretos, siguiendo dichas indicaciones teóricas. En esta segunda alternativa no se obtendría la tendencia subyacente al mercado como un parámetro más del modelo, sino que sería directamente la variable dependiente. En adelante se hará referencia a estos modelos como métodos no paramétricos.

El principal objetivo de este trabajo será encontrar el mejor modelo y la mejor combinación posible de indicadores adelantados para predecir, de forma estática y a un plazo de un mes, la probabilidad de que el mercado europeo experimente una tendencia bajista en un periodo futuro. Para ello se emplearán métodos econométricos y computacionales de diferente índole. Se analizará, para los distintos modelos planteados y para diferentes retardos de las variables exógenas, la capacidad predictiva de ciertos indicadores macro-financieros. Entre ellos se podrían destacar los spread entre los tipos de interés de diferentes economías, la volatilidad implícita en las opciones financieras, el spread entre la rentabilidad de los peores y mejores bonos calificados como de inversión, el cambio en el índice de precios europeo, o el nivel de desempleo de la economía. Los modelos paramétricos empleados serán modelos de regresión con variables exógenas y parámetros cambiantes en el tiempo siguiendo un proceso de Markov, y un modelo factorial dinámico. Esta última clase de modelos se emplea en general para la predicción de ciclos económicos, dado que permite mayor flexibilidad que sus comparables, pero no se ha encontrado literatura en que se haya aplicado al análisis de ciclos bursátiles. El modelo no paramétrico empleado será, siguiendo la literatura previa en el área, un modelo Probit (dinámico) con variables exógenas retardadas(4.2), y extensiones de éste mediante las cuales se trata de tener en cuenta la persistencia de las tendencias bursátiles mediante información intramuestral (4.2)². Los modelos de variable dependiente limitada se emplean para la predicción de los ciclos económi-

¹Para la correlación por ejemplo, se pueden definir dos estados en función del nivel de la volatilidad del mercado.

²Con “información intramuestral” se hace referencia a la evolución pasada de la propia variable dependiente, que podría aportar información valiosa para la predicción de sus tendencias futuras.

cos, y se han trasladado recientemente al ámbito financiero. Para la especificación no paramétrica se aplicará, como se explicará en el apartado de metodología, tanto una modificación previa de la variable dependiente para transformarla en una variable dicotómica (Apéndice B), como un algoritmo genético para seleccionar la mejor especificación posible del modelo (Apéndice C). Una vez comparada la capacidad predictiva de los modelos, se emplearán las mejores especificaciones de cada una de las clases planteadas para tratar de mejorar algunas estrategias de inversión clásicas como pueden ser el “buy and hold” o algunas estrategias de inversión factorial.

El presente trabajo contribuye de diferentes modos a la literatura actual de este área de la economía financiera. Para empezar, se traslada una metodología novedosa, que generalmente no se aplica en este ámbito, como son los modelos factoriales dinámicos con regímenes gobernados por procesos de Markov. Tal método permite identificar los cambios de régimen de forma paramétrica, sin apoyarse en las definiciones teóricas parcialmente arbitrarias que se requieren para establecer modelos no paramétricos y que se reflejan en el planteamiento del algoritmo de Bry & Boschan. Por otro lado, se extienden los hallazgos de la literatura previa de varios modos. [Fernández, Fernández & Sosvilla (2014)] analizan la conexión entre diferentes indicadores macroeconómicos y las tendencias de la bolsa española mediante métodos no paramétricos. Para ello emplean modelos de variable dependiente limitada, aplican un algoritmo genético para seleccionar la combinación de variables que logra un mejor ajuste “in sample”, y lo emplean para predecir. Los autores encuentran que el modelo que maximiza el “fit” a los datos empleados para el proceso de calibración no logra aportar capacidad predictiva para el periodo mantenido para la validación “out of sample” de los modelos. Sin embargo, los autores emplean únicamente variables retardadas en un periodo y configuraciones del modelo probit que no tienen en cuenta la información contenida en el propio índice. En este estudio se tendrán en cuenta retardos adicionales en las variables exógenas de los modelos probit, y se añadirá información extraída de la historia del propio índice de referencia para tratar de mejorar las predicciones. Esto permitirá además contar con configuraciones adicionales a partir de las cuales obtener conclusiones acerca de la idoneidad de emplear el algoritmo genético como mecanismo para seleccionar los indicadores adelantados. Se añade además el factor momentum como método para incorporar dinámica en los modelos. En la literatura previa se emplean habitualmente retardos de la variable dependiente (dicotómica) o del activo subyacente (el índice en este caso). En este trabajo se analiza si al añadir el factor momentum, definido como el rendimiento del índice en el año previo sin tener en cuenta el periodo más reciente, se logra mejorar el ajuste y la capacidad predictiva de los modelos.

Se encuentran varias conclusiones relevantes. Primero, se determina que las variables macro-financieras analizadas son en general predictores de corto y medio plazo. Se concluye que el factor momentum y el “Default Yield Spread” son variables relevantes que aportan capacidad predictiva adicional al combinarlas con las pendientes de las curvas de tipos americana y francesa. También se concluye que el algoritmo genético no es un método de selección de modelos que se pueda emplear de forma recurrente para realizar predicciones sobre las tendencias del mercado, y que los modelos no paramétricos se adaptan mejor al problema que los paramétricos. Se encuentra, además, que las predicciones del modelo factorial dinámico con regresores³ cambiantes en el tiempo en general se ajustan mejor a los datos que los modelos clásicos con parámetros cambiantes vía Markov Switching. Aunque cabe destacar que estas conclusiones se alcanzan considerando la definición de mercado bajista/alcista subyacente al algoritmo de Bry & Boschan, y que la capacidad predictiva del MS-DFM es muy similar a la del modelo para el corto plazo con parámetros cambiantes vía Markov Switching. Finalmente se encuentra que al incorporar un componente de “market timing” se puede mejorar el desempeño de diferentes estrategias de inversión desde el punto de vista del rendimiento ajustado por riesgo.

Lo restante del trabajo se estructura del siguiente modo. A continuación, en la sección 2, se realiza un

³Se permite que la constante de la ecuación del factor varíe entre dos estados.

repasso de la literatura previa en el área de estudio. En la sección 3 se describen las series de datos empleadas en el procedimiento empírico y se explica el fundamento teórico de cada una de ellas. Posteriormente, en 4, se detalla el procedimiento seguido y la metodología empleada para obtener los resultados del trabajo. En la sección 5 se comparan estadísticamente los resultados obtenidos mediante los distintos modelos planteados. Seguidamente, en 6, se establecen y analizan varias estrategias de inversión mediante las cuales capitalizar sobre las predicciones y, por último, en la sección 7 se resumen las principales conclusiones alcanzadas.

2 Revisión de Literatura

La importancia del tema a tratar ha llevado a la aparición de creciente literatura al respecto. Tradicionalmente, para predecir los rendimientos de los mercados o la evolución de otra clase de series temporales relacionadas con variaciones en otras variables, se plantean métodos de Vectores Autorregresivos. Emplear esta clase de métodos econométricos supone asumir que, a priori, la relación entre las variables analizadas puede ser lineal y bidireccional. Por ejemplo, [Laopodis (2006)] examina las interdependencias dinámicas entre la política monetaria y los precios de las acciones americanas empleando métodos de cointegración y de vectores autorregresivos. El autor encuentra que la relación entre variables cambia con el tiempo.

Sin embargo, se encuentra que en general los modelos que plantean relaciones lineales no logran recoger bien las interrelaciones entre los diferentes indicadores adelantados y la evolución del mercado financiero. Se comienza a plantear por tanto modelos no lineales con regresores cambiantes en función del estado subyacente. [Duprey et al. (2015)] emplean modelos con regímenes cambiantes, estimados mediante métodos de Markov Switching y con datos de diferentes países europeos. Generan una cronología de los periodos de estrés financiero sistemático en las diferentes geografías analizadas, generando previamente un indicador con datos de los mercados de acciones, bonos, y forex. Encuentran que los periodos de estrés financiero preceden a las crisis económicas, y que los periodos de crisis tienden a coincidir en los diferentes países europeos. Por otro lado, [Camacho et. al. (2018)] plantean una extensión del modelo factorial dinámico donde se permite que la media de la ecuación del factor común varíe en el tiempo, vía Markov-Switching, para monitorizar la evolución de la economía. Dicho modelo se basa en la suposición de que un conjunto de variables se puede descomponer en un factor común cambiante en el tiempo, y componentes específicos a cada una de las variables consideradas. Se modelizan tanto el factor común como los componentes específicos según un proceso autorregresivo, y se supone que los términos independientes del modelo varían según un proceso de Markov. Las modificaciones que introducen permiten combinar series temporales con diferentes frecuencias y emplear para las predicciones los últimos datos disponibles, aunque éstos sean todavía provisionales. Los autores encuentran que estas modificaciones ayudan a lograr una mejor monitorización de la evolución de la economía en tiempo real, ya que permiten tener en cuenta toda la información disponible. En este trabajo se trasladará el método a la monitorización de los ciclos bursátiles.

La principal alternativa que ha surgido en la literatura financiera frente a los métodos paramétricos consiste en emplear métodos de variable dependiente limitada y, mediante la selección justificada de variables exógenas retardadas a las que se les presupone capacidad predictiva sobre la variable a analizar, tratar de ajustar el mejor modelo posible. Así, [Estrella & Mishkin (1996)] estiman modelos Probit para analizar la capacidad de diferentes indicadores adelantados, y logran predecir los cambios de fase en los ciclos económicos estadounidenses. Para ello aplican previamente el algoritmo de Bry & Boschan (BBA, detallado en B), mediante el cual transforman la serie temporal del PIB en una variable dicotómica y, de este modo, estiman la probabilidad futura de recesión.

Como se ha comentado, la metodología no paramétrica se emplea inicialmente para predecir la probabilidad de experimentar recesiones económicas, pero se ha trasladado al ámbito de las finanzas en trabajos como los de [Fernández, Fernández & Sosvilla (2014)] o [Candelon et. al (2012)]. En estos estudios se con-

trasta la capacidad predictiva de las curvas de tipos de interés sobre los ciclos de los mercados financieros, y se trata de encontrar indicadores adicionales que ayuden a mejorar la precisión de las predicciones que se obtienen al emplear únicamente los tipos de interés. [Fernández, Fernández & Sosvilla (2014)] analizan la conexión entre diferentes indicadores macroeconómicos y las tendencias de la bolsa española mediante métodos no paramétricos. Emplean modelos de variable dependiente limitada, aplican un algoritmo genético para seleccionar la combinación de variables que logra un mejor ajuste “in sample”, y lo emplean para predecir. Encuentran que el modelo que maximiza el “fit” a los datos empleados para la calibración no logra aportar capacidad predictiva sobre la que se obtendría empleando únicamente información procedente de las curvas de tipos. Sin embargo, se podría analizar todavía si el algoritmo genético logra seleccionar modelos con capacidad predictiva para horizontes de predicción mayores, o para diferentes especificaciones de los modelos probit dinámicos en los cuales se incorpore información intramuestral.

Por otro lado, en las últimas décadas se han desarrollado estrategias de inversión basadas en diferentes factores de riesgo, siguiendo los hallazgos de Fama & French. Algunos de los factores a tener en cuenta a la hora de desarrollar una estrategia sistemática de inversión factorial pueden ser el factor momentum o el “value”. En cuanto al primero de ellos, los trabajos de [Jegadeesh & Titman (2001)] y [Asness (1995)] supusieron el punto de partida. Los primeros encuentran evidencia de correlación positiva entre los rendimientos a medio plazo de los activos, y sus rendimientos futuros a corto y medio plazo. Es decir, encuentran momentum en el precio de las acciones. Las que mejor rendimiento presentan en los últimos 3 a 12 meses tienden a mantener rendimientos elevados durante los siguientes 12. Contrastan que a corto y a muy largo plazo las acciones presentan reversión en lugar de momentum, asociando este fenómeno a la iliquidez y al “factor value” respectivamente. Por ello, si se adquiriesen las acciones que peor rendimiento hubiesen presentado en las últimas semanas o en los últimos 5 años, se podría esperar excesos de rendimiento positivos en el corto y medio plazo. De este modo, y como se plantea en el trabajo clásico de [De Bond & Thaler (1985)], se puede emplear el rendimiento de las acciones en los últimos 5 años, sin tener en cuenta el más reciente, para seleccionar una cartera formada por activos “value”. Se podría esperar un rendimiento superior al del mercado asociado a tal cartera, debido a que se estaría asumiendo una prima por riesgo atribuida al factor.

[Asness (1995)] concluye también que la estrategia de reversión a corto plazo se puede explicar en base al “bid-ask spread”, ya que la estrategia es mucho más significativa para las empresas de menor capitalización. Por el contrario, los rendimientos de las estrategias de momentum no se podrían explicar mediante el bid-ask spread, por la iliquidez de las acciones, ni mediante las variables de los modelos de Fama & French. Desde la publicación de estos trabajos, diferentes autores han tratado de dar explicación teórica al exceso de rendimiento de las estrategias fundamentadas en el factor momentum. Sin embargo no hay un consenso claro. Se atribuye en general a la existencia de una prima de riesgo asociada al factor, en línea con la teoría de los mercados eficientes, pero también se encuentran explicaciones basadas en la existencia de sesgos propios de las finanzas conductuales.

Posteriormente, [Asness et. al. (2013)] demuestran que las estrategias basadas en el factor momentum funcionan también en Europa, y que el momentum en el precio es un fenómeno que se da en diferentes clases de activos, y no únicamente en stocks. Por ejemplo, los autores encuentran evidencias de la influencia del factor en los precios de bonos o de índices bursátiles en diferentes regiones del mundo. Por ello, se puede justificar la inclusión del momentum del índice Eurostoxx 50 como variable explicativa en los modelos de “market timing” considerados. Los autores encuentran también correlación negativa entre los rendimientos de las estrategias basadas en los factores “value” y momentum, y que el valor absoluto de esta correlación aumenta durante las fases bajistas. Concluyen de este modo que la combinación de las dos aproximaciones a la inversión factorial permitiría obtener un rendimiento ajustado por riesgo superior al del mercado. Dicha mejora se podría obtener también potencialmente mediante la incorporación de “market timing” a las estrategias, incurriendo además en menores costes de transacción.

3 Datos

En este trabajo se analizará la capacidad predictiva de diferentes variables financieras y macroeconómicas, con frecuencia mensual, para predecir las fases bajistas de los ciclos del índice bursátil europeo de referencia “Eurostoxx 50”. Se emplearán dichas variables para ajustar diferentes clases de modelos econométricos, comparar las predicciones resultantes de éstos, y emplearlas para determinar las ponderaciones asociadas a diferentes clases de activos en varias estrategias de inversión. Se emplearán frecuencias mensuales para poder analizar la capacidad predictiva de los indicadores macroeconómicos, dado que ésta es la periodicidad más elevada para la que se publican. Esto permite también eliminar el ruido propio de los datos de mayor frecuencia. Finalmente la periodicidad mensual es también la que mejor se adapta a las estrategias de inversión factorial que se plantearán al final del estudio.

Se analizará la capacidad predictiva de las variables explicativas retardadas en uno, tres, seis y doce periodos, de modo que se pueda contrastar la capacidad de cada variable o conjunto de variables como predictor de corto (1 mes de retardo), medio (3 y 6 meses) y largo plazo (12 meses). Se emplearán datos obtenidos a través de “Reuters Eikon” que abarcarán, en el mejor de los casos, desde enero de 1987 hasta marzo de 2020. Sin embargo, para ciertas variables la muestra será más corta dado que algunas de las instituciones a las que hacen referencia las series de datos, como puede ser la Zona Euro, no existían al inicio del periodo considerado. Por ello, para realizar el análisis se empleará una muestra de tamaño igual a la mínima longitud de las series de variables independientes individuales. Dado que de entre las consideradas la serie más corta es la de los tipos de interés europeos, se trabajará con las 242 últimas observaciones de cada una de las series de datos. La muestra se corresponderá con el periodo 31/01/2000 - 31/03/2020. De estas observaciones, se empleará el 70% para el ajuste “in sample” de los modelos, y el resto para la validación “out of sample”. Es decir, se emplearán 170 observaciones⁴ para estimar los regresores, y el resto ⁵, para la predicción o validación de los regresores estimados. Dada la descomposición en los conjuntos de estimación y validación se encuentra que, según la definición empleada para la adaptación del algoritmo de Bry & Boschan (BBA), en ambos subconjuntos se experimentan fases tanto alcistas como bajistas. Por ejemplo, el periodo considerado para la estimación “in sample” abarca un periodo de gran estabilidad y crecimiento económico como fue el inicio de la década de los 2000, pero también incluye el mercado bajista de 2008. En el periodo de validación “out of sample” se experimentan dos ciclos completos, con sus respectivas fases alcistas y bajistas, de modo que se podrá contrastar la capacidad predictiva real de los distintos modelos planteados bajo las diferentes condiciones de mercado posibles.

Desde un punto de vista econométrico, muchas de las variables seleccionadas presentan tendencia, por lo que para evitar encontrar relaciones espurias o introducir problemas de multicolinealidad en los modelos, se aplicarán transformaciones sobre las variables exógenas que así lo requieran. La correlación entre las series de variables transformadas es bastante reducida. Por ejemplo, la mediana de las correlaciones de cada variable con el resto en ningún caso se supera el valor de 0.2. Por otro lado, analizando la autocorrelación en las series de variables explicativas transformadas se encuentra que la gran mayoría de ellas presentan autocorrelación para algún retardo reciente. Esto justifica el supuesto que se hará en el modelo factorial dinámico de que los componentes específicos de las variables consideradas se comportan según un proceso autorregresivo.

Como se ha indicado, la variable a explicar será el Eurostoxx 50. Más concretamente, se buscará predecir la probabilidad de que en un determinado periodo se materialice un mercado bajista en el mercado

⁴Tras aplicar el algoritmo de Bry & Boschan se corresponderán con el periodo 31/01/2000 - 30/09/2013. La adaptación a las características del algoritmo lleva a la pérdida de los 6 últimos periodos, lo cual supone que finalmente la muestra “in sample” (70%) sea de 165 observaciones.

⁵71 observaciones (30% de la muestra), correspondientes al periodo 31/10/2013 - 30/09/2019, ya que se elimina el último semestre de la muestra como consecuencia de la aplicación del algoritmo de Bry & Boschan.

europeo. Por tanto, en el procedimiento no paramétrico será necesario definir qué se considera tendencia bajista/alcista. Para ello se empleará una adaptación del procedimiento que Bry & Boschan desarrollan para determinar los periodos recesivos de la economía, que se describirá detalladamente en el anexo (Apéndice B), y se resume de forma esquemática a continuación:

1. Se suavizan los datos aplicando logaritmos a la serie.
2. Con una ventana de 12 meses (los seis previos y los seis posteriores al dato a analizar), se encuentran los máximos y mínimos locales preliminares de la serie a analizar. De este modo, un pico (“peak”) se dará cuando la serie encuentra un máximo local, y un valle (“trough”) se dará cuando en la serie se observe un mínimo local. Es decir, el algoritmo encontrará un pico en el mes “t”, si:

$$p_{t-i} < p_t, \text{ con } i = 1, \dots, 6 \text{ y}$$

$$p_{t+i} < p_t \text{ con } i = 1, \dots, 6$$

y un valle, si:

$$p_{t-i} > p_t, \text{ con } i = 1, \dots, 6 \text{ y}$$

$$p_{t+i} > p_t \text{ con } i = 1, \dots, 6$$

3. Se eliminan los máximos y mínimos locales consecutivos, manteniendo el mayor de los picos o el menor de los valles.
4. Se eliminan las fases cuya duración sea menor que 3 meses.
5. Se fuerza a que el ciclo completo (una fase alcista y una bajista) dure al menos 15 meses.
6. Se eliminan los puntos de inflexión a 3 meses de los extremos de la muestra. Se eliminan los picos y valles en los extremos de la serie que sean menores o mayores, respectivamente, que el primer o último dato de la muestra.

Cabe destacar que la aplicación del BBA supone la pérdida de los primeros y últimos 6 datos de la muestra de la variable dependiente. Esto implica acortar las series de las variables explicativas de los modelos para adaptar las fechas principalmente a la pérdida de las 6 observaciones finales.

Una vez aplicado el algoritmo se encuentra que el 32% de los datos de la muestra corresponden con tendencias bajistas en el mercado europeo. Es decir, en general los periodos alcistas son más frecuentes y de mayor duración que los bajistas, aunque estos últimos son más bruscos. Los resultados coinciden con los hechos estilizados de las series de rendimientos financieros, ya que tradicionalmente se acepta que los mercados bursátiles son en general alcistas y que el crecimiento en el precio de los activos financieros es más sostenido que sus caídas.

Se obtiene la siguiente representación gráfica de los periodos expansivos y recesivos del mercado bursátil:

[Insertar Figura 2]

En cuanto a las variables explicativas, éstas se seleccionan en base a la literatura previa, y se sustentan en criterios teóricos sobre la relación que las fluctuaciones del mercado bursátil europeo pueden tener con los sectores financiero y real de la economía.

Se podrían dividir las variables seleccionadas en: financieras, macroeconómicas reales y macroeconómicas nominales. De las primeras, se tendrán en cuenta dos índices bursátiles diferentes del que se desea

predecir: el Nasdaq norteamericano y el DAX alemán; también la pendiente de las curvas de tipos⁶ de distintas economías correlacionadas con la europea (de Estados Unidos, Alemania, Francia, Italia y la propia Unión Europea), y el nivel de la curva de tipos europea, calculado como la media de los rendimientos de los bonos con vencimientos de 3 meses y 2, 5 y 10 años. Se considera además el “Default Yield Spread”, aproximado por la diferencia entre los rendimientos de los bonos no gubernamentales calificados por Moodys como Baa y Aaa. Adicionalmente, se añaden los precios del petróleo, del oro y el índice VIX.

De entre las posibles variables macroeconómicas reales se analizarán el nivel de desempleo, del índice de construcción y de importaciones desde el exterior de la economía europea.

Por último, de entre las variables macroeconómicas nominales se consideran el índice de precios de consumo europeo (IPC), el tipo de cambio del euro frente al dólar, y el agregado monetario M3 de la zona euro⁷.

Las variables seleccionadas, además de haber sido empleadas en la literatura previa como posibles indicadores adelantados de los puntos de inflexión de los mercados, se pueden justificar en base a la teoría económico-financiera. Esta justificación se expone en el anexo (Apéndice A).

4 Metodología

El objetivo de este trabajo es encontrar la combinación de variables y el mecanismo para combinarlas que permita generar el mejor modelo posible para predecir el signo de las tendencias del mercado, tanto en el ajuste “in sample” como en la predicción “out of sample”. Para seleccionar el mejor modo de combinar las variables se compararán modelos paramétricos y no paramétricos, planteando siempre una relación no lineal entre las variables exógenas y la explicada. Para la selección de variables explicativas se empleará un algoritmo basado en la “selección natural”, y se compararán los resultados obtenidos mediante este método con los que se podrían alcanzar mediante la selección “manual” basada en criterios de ajuste “in sample”.

En esta sección se explican las características y resultados obtenidos mediante los diversos modelos planteados. Se explicarán primero los diferentes criterios empleados para analizar el ajuste de los modelos a la muestra, los estadísticos destinados a determinar su capacidad predictiva en el “out of sample”, y se resumirán las bases del algoritmo genético de selección de variables. Posteriormente se dividirá el apartado en modelos no paramétricos (4.2) y paramétricos (4.3), y finalmente se resumirá el procedimiento y los resultados de la selección del threshold o umbral óptimo (4.4) para cada especificación, obtenidos haciendo uso de la curva ROC (“Receiver Operating Characteristic”).

4.1 Criterios de selección

Como se ha mencionado anteriormente, se distinguirá entre criterios para analizar el ajuste “in sample”, y estadísticos referentes a la capacidad predictiva en el “out of sample”. Es importante realizar esta distinción, ya que los modelos planteados podrían lograr un muy buen ajuste a las características de la muestra “in sample” pero no ser técnicas fiables para realizar predicciones debido, por ejemplo, a sobreparametrización o sobreajuste a las características específicas de un periodo concreto de los datos de la muestra seleccionada para la estimación. Dado que el objetivo es predecir el comportamiento futuro del mercado para poder capitalizar sobre dichas predicciones, se deberá dar peso a la validación “out of sample” aunque la selección de

⁶Se calcula la pendiente de las curvas de tipos como la diferencia entre el rendimiento de los bonos a 2, 5 y 10 años respecto del rendimiento de los bonos de la misma economía pero con vencimiento de 3 meses.

⁷Con la misma intuición económica, se podrían incluir también el M1 y M2, pero en caso de considerar el conjunto de las tres variables se encontraría una elevada correlación. Por ello, se escoge únicamente el M3, que será el agregado con una definición más amplia de “dinero”.

variables se base en el ajuste “in sample”. Esta subsección adquiere especial importancia por el hecho de que se plantean modelos no lineales, lo cual invalida las conclusiones que se puedan alcanzar mediante algunos de los principales estadísticos de ajuste, como puede ser el coeficiente de determinación habitualmente empleado en los modelos de regresión lineal.

Por otro lado, el elevado número de variables con potencial capacidad predictiva consideradas (28 variables exógenas) lleva a que se deban analizar gran cantidad de posibles combinaciones. Se tendrán 2^{28} potenciales especificaciones de los modelos. Para seleccionar la combinación óptima de variables explicativas se empleará un algoritmo genético que seleccionará los indicadores únicamente en base al ajuste de los modelos a los datos, y a la significatividad individual de las variables explicativas. Este algoritmo se resumirá en esta sección y se desarrollará con más detalle en el anexo (Apéndice C).

4.1.1 In Sample

Se emplearán el pseudo- R^2 y el criterio de bondad de ajuste de Akaike (AIC). Estos estadísticos se analizarán solo para los modelos no paramétricos por dos motivos: estos serán los que se emplearán para la selección de variables, y son los que tratan de explicar directamente la evolución de la variable de interés del trabajo (la probabilidad de tendencia bajista).

El pseudo- R^2 para modelos de variable dependiente limitada se puede estimar mediante

$$pseudo - R^2 = 1 - \frac{\ln L_{nr}}{\ln L_0}$$

donde $\ln L_{nr}$ es la log-función de verosimilitud para el modelo estimado sin restringir el valor de ninguno de los regresores del modelo, y se compara con $\ln L_0$, que es la log-función de verosimilitud para un solo regresor. Además se tiene $|\ln L_{nr}| \leq |\ln L_0|$ dado que, en el peor de los casos (si las variables explicativas consideradas no tienen capacidad explicativa), $\frac{\ln L_{nr}}{\ln L_0} = 1 \Rightarrow pseudo - R^2 = 0$.

El estadístico, al igual que el R^2 tradicional, tomará por construcción valores en $[0,1]$. El R^2 se basa en el supuesto de que el modelo subyacente es una regresión lineal⁸.

El pseudo- R^2 no tomará valor 1 si se obtiene un ajuste perfecto, ya que $\Phi(\beta'x) = 1 \Leftrightarrow \beta'x = \pm\infty$. Por ello, el hecho de que el estadístico tome valor 1 indicaría no un ajuste perfecto como es habitual en el coeficiente de determinación para los modelos lineales, sino algún error en el modelo.

Se empleará también el criterio de Akaike como medida de bondad de ajuste de los modelos, y será la métrica principal en la que recaerá la comparativa del ajuste a la muestra de los distintos modelos planteados. Se obtendrá del modo habitual, mediante la fórmula:

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

donde “k” es el número de parámetros en la especificación del modelo, y “lnL” es el valor de la función de verosimilitud en logaritmo para la correspondiente especificación. Se prefieren los modelos con menor valor del criterio, de modo que se penalizan aquellos con elevado número de parámetros, y se buscan los que logren mejor ajuste a los datos. La penalización implica que si por ejemplo se añade una variable sin capacidad explicativa suficiente para justificar el aumentar en una unidad el número de parámetros a estimar,

⁸Si se calculase dicha medida para especificaciones no lineales, como son los modelos Probit, se obtendrían valores más elevados de lo que debería encontrarse, sesgando la interpretación. El estadístico no siempre aumentaría en su valor al implementar mejores modelos no lineales y, en general, no constituiría un criterio de selección fiable. Ésto es así porque en modelos no lineales la suma de los cuadrados explicados más la suma de los cuadrados residuales no igualaría, como sí ocurre en los modelos lineales, a la suma de los cuadrados totales. De este modo, la medida no estaría entre los límites habituales y, en la mayor parte de las ocasiones, no llevaría a seleccionar el modelo correcto.

se preferirá el modelo previo en que no se incluía dicha variable. Se preferirán aquellos modelos que logren una buena estimación, pero que no sean excesivamente complejos.

4.1.2 Out of Sample

Para comparar el ajuste out of sample de los distintos modelos y especificaciones de uno mismo, se emplearán los estadísticos “root mean squared error”, “quadratic probability score”, y “out of sample pseudo- R^2 ”. Estas medidas permitirán determinar fácilmente si las variables presentan o no capacidad predictiva.

La raíz del error cuadrático medio (RMSE) será la raíz cuadrada del segundo momento de las diferencias entre los valores predichos y los observados en la muestra. La medida se obtiene mediante la siguiente fórmula:

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (\hat{P}(R_{t+1} = 1) - R_{t+1})^2}{T}}$$

La “quadratic probability score” (QPS) se conoce también como Brier score, y se puede emplear para datos binarios, como es el caso, o para datos categóricos (siempre y cuando el conjunto de posibles resultados sume uno). Se puede obtener como:

$$QPS = \frac{1}{T} \sum_{t=1}^T (\hat{P}(R_{t+1} = 1) - R_{t+1})^2$$

donde “ T ” es la longitud de la serie mantenida para “out of sample”, $\hat{P}(R_{t+1} = 1)$ es la probabilidad predicha desde el periodo “ t ”, de que se experimente tendencia bajista del mercado en el periodo “ $t + 1$ ”. R_{t+1} es la probabilidad real de experimentar dicha tendencia. Se considerará que la probabilidad real es la obtenida mediante el algoritmo de Bry & Boschan, cuyo procedimiento computacional se expone en el apartado 3. Dado que la variable dependiente es una probabilidad, y por ello toma siempre valores en $[0,1]$, el estadístico tomará valores en $[0,1]$ y se preferirán los modelos con el estadístico QPS más bajo. Es decir, se prefieren aquellos modelos en que $\hat{P}(R_{t+1} = 1) = R_{t+1}$. Los métodos para el cálculo de las medidas RMSE y QPS son muy similares, y los resultados se diferencian básicamente en la escala de los estadísticos. Por ello en este caso se pueden emplear ambas métricas indistintamente, aunque en la literatura previa en el área se emplea habitualmente el QPS.

Por último, el pseudo- R^2 para “out of sample” se puede obtener de forma similar al R^2 habitual:

$$O.S. - pseudo - R^2 = 1 - \frac{\sum_{t=1}^T (R_{t+1} - \hat{P}(R_{t+1} = 1))^2}{\sum_{t=1}^T (R_{t+1} - \bar{R}_{t+1})^2}$$

donde todas las variables, a excepción de \bar{R}_{t+1} , se han definido ya para las métricas anteriores. La variable restante, \bar{R}_{t+1} , será la media de las probabilidades reales hasta la observación “ $t + 1$ ”. El out of sample pseudo- R^2 tiene la misma interpretación que el R^2 habitual, pero en este caso puede tomar valores negativos y mayores que uno. Valores negativos de la métrica implicarían muy baja o nula capacidad predictiva del modelo. Las variables explicativas predecirían peor que la constante por sí sola.

Si la diferencia al cuadrado entre el valor real y el valor predicho es menor que la diferencia al cuadrado entre el dato real y el medio para el periodo predicho (hasta “ t ”), se obtendrán valores positivos del estadístico, ya que el segundo término de la fórmula sería menor que 1. Ésto se interpretará positivamente, ya que indicaría que el modelo logra predecir mejor de lo que lo haría la simple media de los valores observados hasta “ t ”. Si el numerador del segundo término de la fórmula es mayor que el denominador, se obtendrá la conclusión opuesta y un signo negativo del estadístico. Se interpretaría por tanto que el modelo tendría mala capacidad predictiva.

4.1.3 Algoritmo Genético

Además de analizar la capacidad predictiva de las variables seleccionadas consideradas de forma individual, en este trabajo se tratará de encontrar también la mejor combinación de variables en base al ajuste “in sample”. Dado el elevado número de variables consideradas, se empleará para la selección de indicadores adelantados un algoritmo genético que, mediante ensayo y error, y en base al criterio de selección de Akaike, encontrará la combinación de variables que proporcione mejor ajuste a los datos de la muestra.

Al considerarse 28 variables exógenas con potencial capacidad predictiva, se pueden encontrar 2^{28} combinaciones de variables con las cuales predecir la dependiente. El procedimiento se llevará a cabo tanto para los modelos probit con únicamente variables exógenas, como para los probit con información procedente del pasado del propio índice. Se consideran además cuatro retardos diferentes de las variables explicativas para tener en cuenta la posibilidad de que las variables sean indicadores de corto, medio o largo plazo⁹. Es decir, se analizará el ajuste de las 2^{28} posibles configuraciones de modelos probit para 4 posibles retardos de cada variable, y para 2 especificaciones distintas del modelo.

Los algoritmos genéticos son técnicas de optimización basadas en la selección natural que logran superar algunas de las limitaciones propias de los sistemas clásicos. El método evoluciona siguiendo un procedimiento de ensayo y error, buscando maximizar el ajuste “in sample” definido en este caso por el criterio de selección de Akaike. El procedimiento recursivo, explicado en mayor profundidad en el anexo (Apéndice C), se puede resumir en los siguientes pasos:

1. Se genera aleatoriamente una población inicial de modelos a analizar. Ésta consistirá en 20 combinaciones de variables explicativas seleccionadas aleatoriamente.
2. Se analizan las características (estadístico AIC) de cada uno de los modelos de la población.
3. Se mantienen los 2 mejores modelos en base a sus características. Éstos serán los que “se reproduzcan”.
4. Los modelos seleccionados se reproducen generando 20 nuevos modelos con unas variables exógenas seleccionadas estrechamente relacionadas con las que definían a los dos modelos óptimos de la generación anterior. El proceso se conoce como “crossover”, y en él también se introduce aleatoriedad. Se permitirá que se crucen las últimas “i” variables consideradas por cada uno de los dos mejores modelos anteriores.
5. Además del crossover, también se realiza una mutación en las variables seleccionadas por los modelos de la nueva generación. Se podría eliminar una de las variables o añadir otra que no se estaba considerando en los modelos precedentes.
6. La nueva generación de 20 modelos reemplaza a la precedente. Se repite el proceso desde el punto 2. Es decir, los nuevos modelos resultantes de la selección, cruce y mutación de los dos seleccionados en el punto 3, sustituyen a las anteriores 20 combinaciones de variables exógenas consideradas.

Dado que se trata de un procedimiento recursivo, será necesario incorporar criterios de parada. Se definirán dos circunstancias: que el procedimiento encuentre una combinación óptima de variables¹⁰ o que, habiendo probado un elevado número de combinaciones, no logre converger a una combinación óptima.

Una vez seleccionado el conjunto óptimo de variables para una especificación determinada, se eliminan del modelo aquellas que no sean estadísticamente significativas.

⁹Se analizará qué variables retardadas en 1, 3, 6 y 12 meses habría que considerar para obtener el modelo de corto, medio y largo plazo con el mejor ajuste posible a los datos “in sample”.

¹⁰Se considera que se ha alcanzado la combinación óptima cuando, durante 15 generaciones, el modelo que proporciona el mejor ajuste no ha cambiado.

4.2 Modelos no Paramétricos

En esta sección se plantearán las bases teóricas necesarias para la estimación de los modelos no paramétricos, y los principales resultados de la estimación y predicciones asociadas a estos. Se emplearán siempre especificaciones Probit de los modelos de variable dependiente limitada. La sección se dividirá en dos subsecciones: modelos con únicamente variables macro-financieras exógenas retardadas, y modelos que adicionalmente incorporan información procedente del pasado del propio índice de referencia. De este modo se plantea en ambos casos modelos probit dinámicos.

Como se ha explicado previamente, para poder ajustar modelos de variable dependiente limitada primero habrá que transformar la variable de interés (continua) en una dicotómica. En base a la definición de mercado bajista planteada en la sección introductoria (1), se emplea una versión del algoritmo de Bry & Boschan resumida en 3 y explicada en detalle en el anexo (Apéndice B). La variable binaria resultante de dicho procedimiento será la endógena de estos modelos.

Modelos Probit con Variables Exógenas

Los modelos Probit se emplean cuando, como es el caso, se desea obtener una probabilidad de respuesta. Se modeliza $P(y = 1|x) = P(y = 1|x_1, x_2, \dots, x_n)$, donde “x” denota el conjunto de las variables explicativas del modelo. En este caso “x” será la matriz de variables macro-financieras retardadas¹¹ enumeradas en la sección 3.

En los modelos de variable dependiente limitada se consideran funciones no lineales que asumen valores de la variable dependiente estrictamente en $[0, 1]$. Es decir, se establecen modelos de la forma:

$$P(y = 1|x) = G(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n) = G(\beta'x)$$

donde $G(\cdot)$ es una función tal que $0 < G(\beta'x) = G(y^*) < 1 \forall y^* \in \mathbb{R}^2$. Esta clase de especificaciones pueden derivarse a partir de modelos de variable latente. Es decir, se define una ecuación lineal en el espacio de los estados con variable dependiente continua, a la cual se asociará la siguiente expresión:

$$y^* = x\beta + \varepsilon, \text{ con } y = 1_{[y^* > 0]}$$

donde $1_{[y^* > 0]}$ es la función indicatriz. En este caso, el planteamiento del modelo se podría interpretar del siguiente modo: Existirían ciertas variables explicativas que determinarían el nivel de la variable latente y^* . Si dicho nivel es positivo ($y^* > 0$), se experimenta una tendencia bajista que será inobservable en el periodo contemporáneo. La tendencia actual del índice tampoco será una variable observable. En cambio sí se puede encontrar una aproximación al hecho de que las tendencias pasadas del mercado hayan sido alcistas o bajistas, mediante el algoritmo presentado en 3. Es decir, no se observa si la tendencia contemporánea es alcista o bajista, ya que la última observación disponible puede constituir un punto de inflexión, la continuación de una tendencia existente, o los primeros datos de una nueva fase del ciclo. Solo a posteriori se podrá saber si el dato pertenece a una situación de mercado alcista o bajista.

Se modeliza por tanto la tendencia “actual” mediante una variable latente y^* . Partiendo de la ecuación anterior,

$$y = 1_{[y^* > 0]} = 1_{[x\beta + \varepsilon > 0]} = \text{Prob}(x\beta + \varepsilon > 0|x) = \text{Prob}(\varepsilon > -x\beta|x) = G(\beta'x) = G(y^*)$$

¹¹Estos modelos se han empleado generalmente para describir series independientes de datos. Sin embargo, al aplicar estas técnicas a series temporales será necesario tener en cuenta el hecho de que habitualmente se encuentra autocorrelación en los datos. Se pueden plantear varios modos de tener ésto en cuenta. Uno de ellos, que se plantea en la subsección actual, consiste en introducir en el modelo las variables exógenas retardadas. Otra posibilidad, que se considerará en la siguiente subsección, consistiría en introducir retardos de la propia variable dependiente del modelo, o transformaciones del pasado de esta. De tal modo se añadiría información “intramuestral” al modelo, adicionalmente a la “extramuestral” que aportarían las variables exógenas.

Dado que “y” es una función indicatriz, cada una de sus observaciones se considerarán realizaciones de una variable aleatoria con distribución Binaria. Por tanto, la sucesión de dichas observaciones seguirá una distribución Binomial.

En el modelo Probit se supone que la perturbación aleatoria sigue una distribución normal, de modo que “G(.)” será la función de distribución acumulada normal estándar. Es decir, $G(z) = \Phi(z)$ de modo que $\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$. Por ser una función de distribución, “G” devolverá valores en el dominio deseado. Además será una función monótonamente creciente. $G(z) \rightarrow 0$ si $z \rightarrow -\infty$ y $G(z) \rightarrow 1$ si $z \rightarrow \infty$.

Para estimar los valores de los parámetros de esta clase de modelos, por la naturaleza no lineal de $E(y|x)$, se empleará estimación por máxima verosimilitud. Se obtienen así estimadores consistentes de los parámetros. La función a optimizar se obtendrá de:

$$P(Y_t = y_t) = \prod_{y_t=0} (1 - G(\beta'x_i)) * \prod_{y_t=1} G(\beta'x_i) = \prod_{i=1} G(\beta'x_i)^{y_i} (1 - G(\beta'x_i))^{1-y_i}$$

Por tanto, el logaritmo de la función de verosimilitud será:

$$\begin{aligned} \ln L &= \sum_{i=1}^n (1 - y_t) \ln(1 - G(\beta'x_i)) + \sum_{i=1}^n y_t \ln(G(\beta'x_i)) = \\ &= \sum_{y_t=0} \ln(1 - G(\beta'x_i)) + \sum_{y_t=1} \ln(G(\beta'x_i)) = \\ &= \sum_{y_t=0} \ln(1 - \Phi(\beta'x_i)) + \sum_{y_t=1} \ln(\Phi(\beta'x_i)) \end{aligned}$$

Dado que “G” toma valores estrictamente entre cero y uno, la log-verosimilitud estará bien definida para todos los posibles valores de β . Además, la teoría general de la estimación máximo verosímil indica que, bajo condiciones muy generales, para muestras aleatorias la estimación es consistente, asintóticamente normal, y asintóticamente eficiente.

Una vez estimados los parámetros, se podrán obtener los errores estándar asintóticos y, a partir de la matriz de varianzas y covarianzas de los parámetros, se pueden calcular sus estadísticos “t” e intervalos de confianza. Por trabajar asintóticamente, los contrastes “t” para restricciones de variables individuales se compararán con los valores críticos determinados por la distribución normal estándar.

En este apartado se plantean modelos univariantes con las variables exógenas comentadas en 3. Se ajustarán por tanto 28 modelos para cada uno de los cuatro plazos considerados para los retardos de las variables explicativas. En la tabla 4 se muestran los estadísticos de ajuste de los mejores modelos en base a la significatividad de las variables y al ajuste “in sample”. Se reflejan además los estadísticos correspondientes a las estimaciones de los modelos seleccionados por el algoritmo genético.

[Insertar Tabla 4]

Se reflejan los estadísticos correspondientes al nivel de la curva de tipos europea, el spread de las curvas de tipos americana, europea y francesa, las variaciones en el spread entre bonos Baa-Aaa, las variaciones en la volatilidad implícita del mercado, y las variaciones en el IPC. Se eligen estas variables por ser significativas para al menos uno de los retardos considerados, y mantener la significatividad en los modelos con una única variable exógena planteados en la siguiente subsección.

Como se puede observar, el nivel de la curva de tipos europea es la variable que, considerada individualmente, presenta mejor ajuste a los datos para todos los retardos considerados. La variable es también significativa para todos los plazos. Sin embargo su capacidad predictiva es peor que la de los demás modelos

expuestos, por lo que se puede considerar que el buen ajuste de la variable es consecuencia de las características del periodo seleccionado para el ajuste de los parámetros.

En cuanto a los modelos asociados a las restantes variables explicativas, se encuentra que todas son significativas para alguno de los retardos. Pese a ello, el ajuste de los modelos univariantes es pobre tanto en base al pseudo- R^2 como al AIC (a excepción del caso del nivel de los tipos europeos). Además se puede concluir que la capacidad predictiva de estos modelos es nula, ya que presentan un “out of sample pseudo- R^2 ” negativo y estadísticos QPS y RMSE relativamente elevados. Esto lleva a plantear modelos en que se combinen varias explicativas.

En los modelos con múltiples exógenas seleccionadas mediante el algoritmo genético se encuentra que en general el algoritmo emplea las mismas variables que se han comentado individualmente, aunque incluyendo alguna adicional. Al combinar variables explicativas se logra mejorar en gran medida los niveles de los estadísticos de ajuste in sample. Se pasa de obtener pseudo- R^2 que en general no superaban el 0.1 en los univariantes a encontrar ahora valores superiores incluso al 0.60 en algunos de los modelos con las variables seleccionadas mediante el algoritmo genético. Esta mejora en el ajuste “in sample” se aprecia también al comparar los resultados del criterio de Akaike, lo cual se podía esperar dado que como se explica en el anexo (Apéndice C), éste es el estadístico que se establece como referencia en el algoritmo para la selección de modelos. Comparando los resultados del ajuste “in sample” de los modelos con múltiples exógenas para distintos retardos se puede intuir que en general las variables consideradas son indicadores de medio plazo de las tendencias del mercado, dado que los mejores estadísticos de ajuste se encuentran para las variables explicativas retardadas en un trimestre. La misma conclusión se podría alcanzar analizando la significatividad de las distintas variables para los modelos de exógenas consideradas individualmente. Por otro lado, en cuanto a la capacidad predictiva de los modelos seleccionados mediante el algoritmo de selección se encuentra que, al igual que para los modelos con una única variable exógena, la utilidad para predecir las tendencias futuras del mercado sería nula en base a cualquiera de los tres estadísticos que miden la capacidad predictiva en el “out of sample”. Estas conclusiones llevan a añadir dinámica mediante la inclusión de información intramuestral¹² en los modelos.

Modelos Probit con Información intramuestral

Los resultados de los modelos planteados en la subsección anterior muestran que empleando únicamente la información extramuestral aportada por las variables exógenas no se logra establecer modelos con mejor capacidad predictiva que la del paseo aleatorio. Tanto los modelos con una variable exógena como los de múltiples variables explicativas seleccionadas por el algoritmo genético devuelven valores negativos del “out of sample pseudo- R^2 ”, lo cual sugiere nula capacidad de predicción. Los modelos multivariable¹³ estimados sí encuentran un buen ajuste “in sample”, pero mantienen una nula capacidad predictiva. De este modo, los resultados obtenidos coinciden con los encontrados por [Fernández, Fernández & Sosvilla (2014)], aun ampliando los plazos considerados para los retardos de las variables explicativas. Sin embargo, el trabajo de [Candelon et. al (2012)] permite intuir que al introducir dinámica en los modelos añadiendo información intramuestral, es decir, información procedente del pasado de la propia variable transformada para generar la dependiente, se podría mejorar su capacidad predictiva. De este modo se plantean nuevas especificaciones con una única variable exógena, pero con dos regresores adicionales asociados al factor momentum y a un componente autorregresivo del índice de referencia. Adicionalmente se aplicará el algoritmo de selección para tratar de encontrar las variables exógenas que, combinadas con la información intramuestral, optimicen el ajuste in sample.

¹²Con “información intramuestral” se hace referencia a los retardos del propio índice de referencia, o a otras variables que se pueden obtener mediante transformaciones de este. Un ejemplo podría ser el factor “momentum”.

¹³Con “Multivariable” se hace referencia a aquellos modelos en que se incluyen varias exógenas.

Se añadirán dos regresores a los modelos anteriores:

1. Componente autorregresivo: En algunos trabajos de la literatura se propone la estimación del modelo en que se incorpore simplemente la variable dependiente limitada retardada como variable explicativa. Sin embargo, además de que tal especificación del modelo llevará a un sobreajuste a los datos del periodo “in sample”, limitando la capacidad predictiva de los modelos, el hecho de introducir una variable independiente binaria prácticamente idéntica a la dependiente dificultará la convergencia del proceso de optimización. Una alternativa será añadir retardos del precio del índice de referencia (S_{t-k}), en lugar de retardar la propia variable dependiente del modelo. Ésto debería aportar capacidad explicativa y predictiva a los modelos, ya que se encuentra que la serie del Eurostoxx 50 presenta cierto nivel de autocorrelación con los retardos más recientes. Se puede encontrar que existe un componente autorregresivo en la serie del índice Eurostoxx 50 y que, en consecuencia, existirá persistencia en las tendencias si dicho componente es negativo, y reversión si es positivo. Dadas las conclusiones de [Asness (1995), Jegadeesh & Titman (2001)] se puede esperar que los retardos recientes del índice lleven asociados un coeficiente con signo positivo en el modelo, reflejando un factor de reversión. Es decir, al aumentar el nivel del índice aumentaría también la probabilidad de que la tendencia bursátil sea bajista en el periodo siguiente. La especificación resultante de incorporar dinámica al modelo probit mediante la introducción del proceso AR(k) correspondiente a la variable Eurostoxx 50 será:

$$P(y_t = 1|x) = \pi_t = G(\alpha + x_{t-k}\beta_k + S_{t-k}\phi_k) \text{ donde } k = 1, 3, 6, 12.$$

Donde S_{t-k} es el retardo del nivel del índice. Una especificación alternativa podría ser:

$$P(y_t = 1|x) = \pi_t = G(\alpha + \sum_{j=1}^k x_{t-j}\beta_j + \sum_{j=1}^k S_{t-j}\phi_j) \text{ donde } j = 1, 3, 6, 12.$$

Sin embargo, dadas la cantidad de variables explicativas consideradas y el número de retardos a analizar (4), el número de especificaciones aumentaría de forma exponencial hasta alcanzar magnitudes inabarcables. Habría un total de 2^{112} combinaciones posibles de variables.

2. Factor Momentum: Adicionalmente se introducirá una variable que refleje el momentum del índice de referencia y, con ello, la persistencia de las tendencias. Según el estudio de [Asness et. al. (2013)], no solo los stocks individuales presentan momentum, sino que dicho factor está presente en los bonos, en commodities o en índices. Asness concluye también que el momentum es especialmente marcado en los activos financieros europeos. Por tanto, el introducir el momentum del índice como variable explicativa en los modelos será un añadido novedoso que debería mejorar tanto el ajuste a la muestra como la capacidad predictiva de éstos. Cabe destacar que, como se explicaba en secciones anteriores (2), los activos financieros presentan en general reversión a la media a largo y corto plazo, pero presentan momentum en el medio plazo. Por ello, se introducirá el rendimiento generado por el índice de referencia en el año anterior a la observación de la variable dependiente, pero sin considerar el rendimiento del activo en el mes inmediatamente previo. El modelo planteado será por tanto el siguiente:

$$\pi_t = G(\alpha + \phi_{t-k}S_{t-k} + \gamma_{t-k} \sum_{j=t-12-k}^{t-2-k} \Delta S_j + x_{t-k}\beta_k) \text{ donde } k = 1, 3, 6, 12.$$

Tras estimar los modelos probit dinámicos con una única variable exógena pero añadiendo los factores explicados en esta sección, se encuentran grandes mejoras en todos los estadísticos considerados. En las

tablas 5 y siguientes se recogen los resultados para las variables exógenas comentadas en la subsección anterior.

[Insertar Tabla 5]

Se observa que las variables seleccionadas mantienen la significatividad para algunos de los retardos considerados. En base a esto se puede concluir que el nivel y el spread de las curvas de tipos son indicadores de medio plazo, a excepción de la curva de tipos francesa que se puede considerar predictor de largo plazo. También la inflación se define claramente como predictor de medio plazo, mientras que el “Default Yield Spread” tiene cierta influencia tanto en el corto como en el medio plazo. En cuanto al momentum, se encuentra que presenta significatividad individual para todos los retardos considerados, coincidiendo con los resultados obtenidos en la literatura previa¹⁴. Analizando el signo del coeficiente asociado a este factor se observa una relación inversa entre el valor de la variable y la probabilidad de experimentar un mercado recesivo en el siguiente periodo. Se puede concluir por tanto que el momentum es una variable relevante a la hora de tratar de predecir el signo de las tendencias futuras del mercado, y que refleja la persistencia de las tendencias del activo en cuestión. Por otro lado, se encuentra que también el componente autorregresivo es significativo para todos los plazos, y que el signo que lleva asociado es positivo. Se establece por tanto una relación directa entre la variable y la probabilidad estimada de mercado bajista.

Analizando el ajuste “in sample” de los modelos se aprecia una mejora importante en el valor de ambos estadísticos considerados. El pseudo- R^2 por ejemplo pasa de no superar el umbral de 0.1 en los modelos con únicamente variables exógenas, a obtenerse consistentemente valores superiores al 0.2. En este sentido, se encuentra nuevamente que la pendiente de la curva de tipos de interés estadounidense es la variable cuyo modelo asociado logra un mejor ajuste “in sample”, al obtener un pseudo- $R^2 = 0.4975$ para el medio plazo. También el nivel de la curva de tipos europea y la inflación destacan por el buen ajuste que alcanzan, ya que obtienen respectivamente pseudo- R^2 de 0.33 y de 0.34 para las especificaciones de medio plazo. Se obtiene una relación inversa entre los spread europeo y francés y la probabilidad de recesión. En cambio, se encuentra una relación directa entre dicha probabilidad y el “Default Yield Spread”, la variación en el VIX o la inflación. Estos resultados coinciden con las bases teóricas planteadas en el anexo (Apéndice A). Sin embargo se encuentra una relación directa también con el spread de la curva de tipos americana, lo cual no se esperaba. Ésto se podría explicar por la diferente reacción de los bancos centrales americano y europeo a la crisis económica y bursátil desatada tras la quiebra de “Lehman Brothers”, ya que la FED reacciona de forma más rápida y contundente en la reducción de los tipos de interés, mientras que el BCE reacciona de forma más conservadora y mantiene los tipos bajos durante un periodo más prolongado. Además, en 2012 se experimentan problemas económicos en la Unión Europea que no necesariamente habrían afectado al ciclo bursátil americano. De este modo el signo que acompaña al spread de la curva de tipos estadounidenses se podría asociar a las características intrínsecas del periodo analizado.

En cuanto a la predicción “out of sample”, se encuentra que los modelos de una única variable exógena presentan capacidad predictiva sobre las tendencias del mercado cuando se añade información intramuestral y que, como se podía esperar, las mejores predicciones se asocian a los modelos de medio plazo. Las variaciones en el IPC (out of sample pseudo- $R^2 = 0.2939$ en el medio plazo) o el VIX (o.s. pseudo- $R^2 = 0.2411$ en el medio plazo) y el ratio Baa-Aaa (o.s. pseudo- $R^2 = 0.2691$ en el medio plazo) presentan capacidad predictiva al considerar cada variable en conjunto únicamente con el factor momentum y el componente autorregresivo. Se puede concluir de este modo que el componente autorregresivo, el factor momentum y las variables comentadas son indicadores adelantados muy relevantes a tener en cuenta, adicionalmente a

¹⁴[Griffin et. al. (2003)] demuestran que las estrategias que basan la selección de activos en el momentum de las acciones devuelven rendimientos positivos durante los siguientes 12 meses, alcanzando los máximos rendimientos entre los 3 y 6 meses de haber formado la cartera.

las curvas de tipos, a la hora de plantear modelos para predecir las tendencias futuras del mercado bursátil europeo.

Los modelos con múltiples variables explicativas seleccionadas mediante el algoritmo genético nuevamente emplean las 7 variables analizadas, y devuelven un muy buen ajuste “in sample”. Los resultados se exponen en la tabla 8:

[Insertar Tabla 8]

En este sentido, al añadir los nuevos factores se logra mejorar el ajuste a la muestra, alcanzándose incluso un pseudo- R^2 de 0.70 en el modelo de medio plazo. Sin embargo, cuando se analiza la capacidad predictiva de estos modelos se observa que solo el de largo plazo presenta un valor positivo del “out of sample pseudo- R^2 ”. Es decir, los modelos de corto y medio plazo no permitirían predecir las tendencias del mercado. En cuanto al de largo plazo, aunque se obtienen estadísticos positivos para las medidas de validación “out of sample”, es destacable el hecho de que sea éste el único modelo con varias exógenas seleccionadas vía “GA” que presente capacidad predictiva, ya que se ha concluido en base a los modelos univariantes de ambas subsecciones que las variables son en general predictores de corto/medio plazo. Se observa además que el algoritmo selecciona por ejemplo las pendientes de la ETTI americana para los plazos de 2 y 5 años¹⁵, y éstas llevan asociados coeficientes con signos opuestos, lo cual no es coherente con la teoría financiera. El resultado podría considerarse “data mining”, y en base a los 8 modelos con múltiples variables exógenas se puede concluir que el método de selección de modelos mediante algoritmos genéticos no es adecuado para realizar predicción, ya que se sobreajusta a la muestra.

En cuanto a la convergencia del procedimiento, como se aprecia en la figura 4, se encuentra que el algoritmo converge rápidamente al modelo óptimo, en un número de generaciones similar al de la subsección anterior. El procedimiento alcanza la combinación óptima de variables según el criterio de Akaike en un número reducido de generaciones. Pese al gran número de posibles combinaciones de variables, para las especificaciones de medio plazo es suficiente con el paso de 7 generaciones para alcanzar la combinación óptima. Para los plazos restantes, el número de generaciones que pasan hasta lograr la convergencia es, en el peor de los casos, de 15. Ésto indica que el coste computacional del procedimiento es bastante reducido. Aunque no es un mecanismo adecuado para el problema de selección planteado, el algoritmo genético sí es un buen método de optimización dada la rápida convergencia y el ajuste “in sample” de los modelos seleccionados.

[Insertar Figura 4]

Dado que se ha concluido que el algoritmo no se puede emplear como procedimiento para identificar los factores que conjuntamente conformen un modelo que presente buen ajuste tanto “in sample” como “out of sample”, se estimará un modelo con múltiples variables exógenas seleccionando, de entre las que se han comentado en estas subsecciones, las que mantienen la significatividad individual al emplearlas de forma conjunta. Los próximos modelos estarán conformados por el factor momentum y el componente autorregresivo, por la pendiente de las curvas de tipos de interés americana y francesa, y por el “Default Yield Spread”. Estas serán las variables que se emplearán para plantear y comparar los modelos paramétricos y no paramétricos con múltiples explicativas. Cabe destacar que el índice VIX, la inflación y el nivel y

¹⁵Es decir, la diferencia entre las rentabilidades de los bonos americanos con vencimientos a 5 y 2 años menos la rentabilidad de los bonos del mismo país pero con rendimiento a 3 meses.

pendiente de la ETTI europea pierden la significatividad individual al combinarlas con las demás variables consideradas.

Como se puede observar en la tabla 9, todas las variables mantenidas son significativas para varios de los plazos considerados, se obtiene un ajuste “in sample” mejor que el de los modelos univariantes, aunque peor que el resultante de la selección del algoritmo genético, y las variables en conjunto presentan capacidad predictiva en el medio plazo. Además, los signos de los regresores son coherentes con los de los modelos con una única variable exógena.

[Insertar Tabla 9]

4.3 Modelos Paramétricos

En este apartado se plantean modelos no lineales en los cuales la variable de referencia es la tasa de variación del índice bursátil europeo Eurostoxx 50, y donde la probabilidad de tendencia bajista se estima como un parámetro más de los modelos. De este modo no es necesario realizar transformaciones de la variable de interés en base a criterios teóricos (más allá de estacionarizar el índice). Se plantearán dos posibles especificaciones de modelos paramétricos: modelos con variables exógenas y cambios de régimen en los coeficientes y en la varianza de las perturbaciones, y un modelo factorial dinámico cuya media seguirá un proceso de Markov. En ambos casos se considerarán, además de la variable de interés, las exógenas seleccionadas para los modelos multivariantes anteriores (Baa-Aaa, spread americano y spread francés), el factor momentum del índice, y el componente autorregresivo.

4.3.1 Modelos con parámetros cambiantes vía Markov Switching

En este apartado se planteará un primer método paramétrico a través del cual extraer las probabilidades predichas de tendencia bajista en el mercado europeo. Se estimarán simplemente regresiones con las variables independientes definidas para los últimos modelos probit con múltiples explicativas de la sección anterior, y se permitirán cambios de régimen en los regresores de los modelos, así como en las matrices de covarianzas. Se emplearán el factor momentum y el componente autorregresivo, la pendiente de las curvas de tipos americana y francesa, y el “Default Yield Spread”.

Se plantean dos posibles escenarios: uno alcista, con signo positivo esperado del término independiente y baja volatilidad en el mercado, y otro con $\mu_{s_t} < 0$, de mayor magnitud en valor absoluto y elevada volatilidad. Es decir, se ajusta el siguiente modelo:

$$r_t = \mu_{s_t} + \delta_{s_t, t-k} P_{t-k} + \gamma_{s_t, t-k} \sum_{t-12-k}^{t-2-k} r_t + \sum_{i=1}^n \beta_{i, s_t} x_{i, t-k} + \varepsilon_t$$

con $\varepsilon_t \sim^{iid} N(0, \Omega_{s_t})$, y donde P_t es el nivel del índice de referencia, r_t es la tasa de variación logarítmica de éste, x_t es una matriz que recoge el valor de las variables explicativas consideradas en el momento “t”, y $s_t = 0, 1$ son los dos posibles estados de la naturaleza considerados. De este modo se plantea un modelo no lineal donde se permite que los regresores y la varianza de las perturbaciones varíen al unísono entre dos posibles regímenes. El segundo y tercer término de la ecuación reflejarán respectivamente el componente autorregresivo y el factor momentum.

El modelo se estimará mediante la maximización de la función de verosimilitud, haciendo uso de la información que se va incorporando en cada momento del tiempo para actualizar los parámetros estimados

del modelo. En este caso, el parámetro de interés serán las probabilidades de experimentar una fase recesiva en periodo siguiente, que vendrá dada por:

$$\widehat{\xi}_{t+1|t} = E(\xi_{t+1}|s_t = i) = \begin{pmatrix} p_{i,1} \\ p_{i,2} \end{pmatrix}$$

con $j=1,2$ y donde $p_{i,j} = p(S_t = j|\Theta_{t-1})$ son las probabilidades de que la observación t-ésima sea generada por el régimen “j”, condicionada a la información conocida hasta t-1.

Se puede definir la matriz de probabilidades de transición como:

$$P = \begin{pmatrix} P_{1,1} & 1 - P_{2,2} \\ 1 - P_{1,1} & P_{2,2} \end{pmatrix}$$

Además se supondrá que las innovaciones siguen una distribución Gaussiana. De este modo se pueden estimar las probabilidades condicionales de forma recursiva, mediante las siguientes ecuaciones:

$$\widehat{\xi}_{t|t} = \frac{\widehat{\xi}_{t|t-1} * \eta_t}{1'(\widehat{\xi}_{t|t-1} * \eta_t)}$$

y

$$\widehat{\xi}_{t+1|t} = \widehat{P}\widehat{\xi}_{t|t}$$

donde η_t es la función de densidad de la normal estándar dada por las perturbaciones ε_t , y $\widehat{\xi}_{t|t}$ son las probabilidades filtradas de que la observación t-ésima venga generada por un régimen determinado.

Para que los resultados sean comparables a los obtenidos mediante los modelos no paramétricos se definen nuevamente 4 retardos posibles para las variables independientes: $k = 1, 3, 6$ y 12 .

En la tabla 10 se recogen los parámetros resultantes de la estimación de los modelos. Los únicos parámetros significativos para todos los plazos considerados son la varianza de los modelos y las probabilidades de transición. Se encuentra que el procedimiento distingue entre un estado de la naturaleza de volatilidad relativamente baja y muy persistente, y otro menos volátil y ligeramente menos duradero. El hecho de que las variables exógenas no sean en general significativas indica que, aunque son indicadores relevantes para predecir las tendencias del mercado, no lo serían para predecir los rendimientos futuros aun teniendo en cuenta los cambios de régimen.

Sin embargo, si se analiza el modelo de corto plazo se encuentra que el momentum del índice influye de forma asimétrica sobre los rendimientos futuros en función de si se encuentra en uno u otro estado de la naturaleza. Se encuentra también que el componente autorregresivo es solo significativo en los modelos de corto y largo plazo, y que su efecto es significativo solo en el régimen poco volátil. Las variables restantes son solo significativas para el modelo de corto plazo, a excepción del spread de la curva de tipos de interés francesa que, coincidiendo con los resultados obtenidos en los modelos no paramétricos, se encuentra que es significativa para el modelo de largo plazo. Solo lo es para el régimen alcista, donde influye sobre los rendimientos del índice del mismo modo que lo hacía sobre las tendencias en las estimaciones de los modelos no paramétricos. Se puede concluir por tanto que los efectos de las variables consideradas sobre los rendimientos del índice europeo son asimétricos, y que el signo de las tendencias subyacentes al mercado serán importantes para determinar dichos efectos.

[Insertar Tabla 10]

En la figura 7 se grafican las probabilidades de mercado bajista predichas a un periodo para los diferentes retardos analizados de las exógenas. Se añaden las probabilidades “reales” obtenidas mediante el algoritmo de Bry&Boschan, y las probabilidades suavizadas ($\widehat{\xi}_{r|T}$) resultantes de aplicar el algoritmo de Kim una vez conocidos todos los datos de la muestra. Se observa que las predicciones son muy similares para los distintos plazos considerados, lo cual es consecuencia directa de la falta de significatividad de las variables explicativas. La excepción es el modelo de corto plazo, para el cual sí se habían encontrado varias variables explicativas significativas. También se aprecia que las predicciones no difieren en gran medida de las probabilidades suavizadas, aunque sí son más bruscas.

[Insertar Figura 7]

En cuanto a la capacidad predictiva de los modelos, como se puede apreciar en la figura 11, las predicciones realizadas en base a los modelos con variables exógenas y parámetros cambiantes en el tiempo son claramente inferiores a las obtenidas mediante los mejores probit.

4.3.2 Modelo Factorial Dinámico (MS-DFM)

El modelo se apoya en la relación entre las variables inobservables, pertenecientes al espacio de los estados, y el conjunto de variables observables consideradas. Permite recoger los comovimientos de las variables y las asimetrías que surgen como consecuencia de la variación de los ciclos bursátiles. La principal dificultad para la estimación de modelos factoriales es el hecho de incluir factores inobservables, que se pueden tratar como variables estado. Los principales supuestos subyacentes en el modelo son la descomposición lineal entre factor común y componentes específicos a las variables, y que los errores se distribuyen según una normal. Además del comovimiento de las variables, el modelo incorporará los cambios de régimen mediante regresores cambiantes en el tiempo vía Markov Switching. La principal ventaja de esta clase de modelos es la flexibilidad que permiten, y la adaptabilidad a los nuevos datos que se incorporen. Permitirán a los agentes económicos o a los inversores producir y actualizar en tiempo real los cambios en la actividad económica o en las variables financieras de interés. Con el modelo factorial dinámico se podrán tener en cuenta en la misma estimación indicadores publicados para distintas frecuencias, variables que se hagan públicas con retraso, o incluso aquellas que se publiquen de forma provisional para posteriormente actualizar sus valores.

Una vez definido el modelo en forma del espacio de los estados, se podrán predecir tanto las variables observables como las inobservables. Se empleará el filtro de Kalman para la estimación, haciendo uso de la función de verosimilitud del modelo. El filtro se descompone en un proceso de dos etapas: Primero se producen estimaciones de las variables de estado inobservables y del grado de incertidumbre asociada a éstas; posteriormente, una vez observada la materialización de la siguiente medida, se actualizan las predicciones mediante una media ponderada atribuyendo más peso a las estimaciones con menor grado de incertidumbre. De este modo se genera un algoritmo recursivo donde se emplea la medida de las variables observables, el estado calculado previamente, y la matriz de incertidumbre.

El modelo se definirá mediante la descomposición de las variables en factores comunes y específicos a cada variable, de la siguiente forma:

$$y_t = \Lambda f_t + u_t$$

donde y_t es un vector de N-indicadores relacionados con el estado de la economía. En este caso se tendrán en cuenta 5 variables: El propio índice de referencia (Eurostoxx50), el momentum del índice, la pendiente de las curvas de tipos de interés estadounidense y francesa¹⁶, y la variación en el “Default Yield Spread”.

¹⁶Definidas como la diferencia entre el rendimiento de los bonos a dos años y a 3 meses.

Por tanto, y_t será un vector de $N = 5$ variables. No se incluyen retardos de las variables dado que se tiene en cuenta el pasado de éstas mediante las ecuaciones siguientes.

$f_t = (f_{1,t}, \dots, f_{r,t})$ será el componente idiosincrásico, común a las variables consideradas en el vector y_t , y u_t será el componente específico a cada variable. En este caso se considerará un único factor común, y por tanto $r = 1$. La modelización del comportamiento de las variables según la ecuación anterior lleva a suponer que los comovimientos en las series se derivan directa y únicamente del factor común, de modo que $f_t \perp u_t$. Por último, el término “ Λ ” será una matriz de “factor loadings” de tamaño “ $N \times r$ ”, donde “ N ” es el número de variables y “ r ” el de factores comunes considerados.

Por otro lado, el factor común se considera dinámico, y tanto su evolución como la de los componentes específicos se modelizan mediante procesos autorregresivos con cambio de régimen en la media del factor. Éste seguirá la siguiente ecuación:

$$f_t = \mu_{s_t} + \phi_1 f_{t-1} + \dots + \phi_p f_{t-p} + a_t$$

con $a_t \sim^{iid} N(0, \Sigma_a)$ y $\Sigma_a \perp s_t$. Se supondrá también que f_t evoluciona según un proceso autorregresivo de primer orden, de modo que $p = 1$. La asimetría del ciclo se tendrá en cuenta mediante la variable de estado s_t , que regirá el comportamiento de los factores comunes y afectará directamente al término independiente. Se tendrá:

$$s_t = 0 \rightarrow \text{Alcista}$$

$$s_t = 1 \rightarrow \text{Bajista}$$

Se establece que la variable evoluciona según una cadena de Markov de primer orden, de modo que:

$$p(s_t = j | s_{t-1} = i, s_{t-2} = h, \dots, s_{t-n} = k) = p(s_t = j | s_{t-1} = i) = p_{i,j}$$

Por último, la ecuación que modeliza la evolución de los componentes específicos será:

$$u_t = (u_{1,t}, u_{2,t}, \dots, u_{N,t})' = \Psi_1 u_{t-1} + \dots + \Psi_p u_{t-p} + \varepsilon_t$$

donde $\varepsilon_t \sim^{iid} N(0, \sigma_{s_t}^2)$, Ψ_j y $\text{var}(\varepsilon_t) = \sigma_{s_t}^2 = \text{diag}(\sigma_{1,s_t}^2, \dots, \sigma_{N,s_t}^2)$ son matrices diagonales de tamaño “ $N \times N$ ”, con $q = 1$. Esto tiene su fundamento en el hecho comentado en la sección 3, donde se había determinado la existencia de un componente autorregresivo en las series temporales de la mayoría de variables exógenas consideradas. Cabe destacar que al plantear una matriz de covarianzas diagonal no se permite la existencia de correlación entre los distintos componentes específicos u_t .

Bajo el supuesto de que existe un único factor común que determina los movimientos compartidos por las variables, se puede representar el modelo en el espacio de los estados como:

$$y_t = Hx_t \tag{1}$$

$$x_t = M_{s_t} + Fx_{t-1} + Ge_t \tag{2}$$

con $u_t \sim N(0, R)$, $e_t \sim N(0, Q)$, $x_t = \begin{bmatrix} f_t \\ u_t \end{bmatrix}$ y $e_t = \begin{bmatrix} a_t \\ \varepsilon_t \end{bmatrix}$.

La primera ecuación se conoce como ecuación de observación, y modela las variables observables como una función lineal del factor común y de los componentes específicos. La segunda ecuación se conoce como ecuación de estado, y define la evolución en el tiempo de los componentes de la ecuación de observación. Los parámetros del modelo se estimarán de forma conjunta y, por ello, no se pueden identificar las matrices de forma independiente.

Matricialmente se puede expresar la ecuación de estado como:

$$\begin{bmatrix} f_t \\ u_t \end{bmatrix} = \begin{bmatrix} M_{s_t} \\ 0_{N,1} \end{bmatrix} + \begin{bmatrix} F_1 & 0_{1,N} \\ 0_{N,1} & \Psi_1 \end{bmatrix} \begin{bmatrix} f_{t-1} \\ u_{t-1} \end{bmatrix} + I_{N+1} * \begin{bmatrix} a_t \\ \varepsilon_t \end{bmatrix}$$

y la de observación como:

$$y_t = H * x_t = \begin{bmatrix} \Lambda_{N,1} & I_{N,N} \end{bmatrix} \begin{bmatrix} f_t \\ u_t \end{bmatrix}$$

con $\Lambda = (\lambda_1, \dots, \lambda_N)'$.

Como se comentaba anteriormente, los parámetros de los modelos definidos en el espacio de los estados se pueden estimar mediante el procedimiento recursivo definido por el filtro de Kalman. Éste permite obtener las predicciones óptimas de las variables de estado y de las observables, siempre y cuando se planteen relaciones lineales en la ecuación (1). Partiendo del modelo general definido por las ecuaciones (1) y (2), asumiendo normalidad, y denotando los posibles estados de la naturaleza (tendencia alcista/bajista) por “i” y “j”, se pueden expresar (1) y (2) como

$$f(f_t | f_{t-1}, Y_{t-1}) = N(F f_{t-1}, G Q G')$$

$$f(y_t | f_t, Y_{t-1}) = N(M_{s_t} + H f_t, R)$$

con $Y_{t-1} = (y_{t-1}, y_{t-2}, \dots)$.

Se pueden descomponer estas ecuaciones en función de la información sobre la que se condiciona. Para empezar, condicionando sobre Y_{t-1} (es decir, la densidad de la variable estado condicionando sobre toda la información pasada), se tendría:

$$f(f_t | Y_{t-1}) = N(f_{t|t-1}, P_{t|t-1}) \quad (3)$$

donde $P_{t|t-1}$ es la matriz de covarianzas de las predicciones del factor un periodo hacia delante ($f_{t|t-1}$). Las ecuaciones que corresponderán a dicha función serán:

$$f_{t|t-1}^{i,j} = M_{s_t} + F f_{t-1|t-1}^{i,j}$$

$$P_{t|t-1}^{i,j} = F P_{t-1|t-1}^{i,j} F' + G Q G'$$

Condicionando sobre Y_t (densidad de la variable estado condicional a la información actual y pasada) se obtienen las funciones que permiten actualizar de forma óptima las predicciones de la variable estado cuando se obtiene la nueva información sobre las variables observables. Por tanto se puede definir la función de actualización:

$$f(f_t | Y_t) = N(f_{t|t}, P_{t|t}) \quad (4)$$

que se corresponderá con:

$$f_{t|t}^{i,j} = f_{t|t-1}^{i,j} + P_{t|t-1}^{i,j} H' F_t^{-1} u_t^{i,j},$$

$$P_{t|t}^{i,j} = P_{t|t-1}^{i,j} - P_{t|t-1}^{i,j} H' F_t^{-1} H P_{t|t-1}^{i,j}$$

con $u_t^{i,j} = y_t^{i,j} - y_{t|t-1}^{i,j}$. Cabe destacar que las ecuaciones de predicción dependen de las ecuaciones de actualización definidas en (4).

También será necesario obtener la verosimilitud del proceso de la ecuación de observación dado el pasado de las propias variables observables. Se tiene:

$$f(y_t|Y_{t-1}) = N(y_{t|t-1}, F_t) \quad (5)$$

que se corresponde con

$$y_{t|t-1}^{i,j} = H f_{t|t-1}^{i,j},$$

$$F_{t|t-1} = H P_{t|t-1} H' + R$$

de modo que la función de verosimilitud depende también de las ecuaciones de predicción. Además, $F_t = E(u_t u_t')$ es la varianza teórica de las perturbaciones de la ecuación (6).

La fórmula anterior se puede denotar como función de verosimilitud, dado que

$$f(y_1, y_2, \dots) = \prod_{t=1}^T f(y_t|Y_{t-1}) \quad (6)$$

es la función de verosimilitud conjunta.

El procedimiento a seguir para la estimación será el siguiente:

1. Se obtienen $f_{2|1}$ y $P_{2|1}$ a partir de las condiciones iniciales.
2. Se emplean $f_{2|1}$ y $P_{2|1}$ para obtener $y_{2|1}$ y $F_{2|1}$ mediante las ecuaciones de verosimilitud en (5).
3. Se emplean las estimaciones de los dos pasos anteriores para obtener $f_{2|2}$ y $P_{2|2}$ a través de las ecuaciones de actualización definidas por (4).
4. Se generan $f_{3|2}$ y $P_{3|2}$ empleando las ecuaciones de predicción (3).
5. Se obtienen $y_{3|2}$ y $F_{3|2}$ mediante la función de verosimilitud (5), y partiendo de las estimaciones del paso anterior.
6. Se aplican las ecuaciones de actualización sobre los resultados de los dos pasos anteriores.
7. Se continua iterando hasta $t = T$. En tal punto se podrá obtener la verosimilitud total mediante la expresión (6) o la función

$$\log L(\theta) = -\frac{NT}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log |F_t| - \frac{1}{2} \sum_{t=1}^T u_t' F_t u_t$$

donde u_t son los errores de las predicciones a un periodo y F_t es su matriz de covarianzas.

Siguiendo todo el procedimiento explicado, se obtienen los siguientes valores de los parámetros:

[Insertar Tabla 11]

En este caso lo más destacable es el hecho de que los cambios de régimen son significativos, pasando de un régimen en el que la media es positiva y cercana a cero, a otro en el cual la media de la ecuación del factor común pasa a ser negativa y de mayor magnitud en valor absoluto. De este modo se puede asociar el estado “0” a la fase alcista del ciclo bursátil, y el estado “1” a la fase bajista. Se observa además que la tendencia expansiva es muy persistente, mientras que la bajista lo es en menor medida. Se puede concluir por tanto que la estimación es coherente con lo esperado en base a los hechos estilizados de los rendimientos financieros.

El factor común afecta positivamente a todas las variables consideradas excepto al spread Baa-Aaa, lo cual es coherente con la teoría. Se encuentra una correlación de alrededor de 0.6 entre el factor común y la

variación logarítmica del índice de referencia, y de 0.32 con el momentum. Además, como se puede observar en la figura 8, el factor subyacente está muy correlacionado (negativamente) con la probabilidad predicha de tendencia bajista. De este modo se puede identificar con la tendencia subyacente a la bolsa europea, aunque no se puede interpretar como una probabilidad al no tomar valores en el dominio pertinente.

Se encuentra por tanto que existe un factor común a las variables consideradas que ejerce una influencia significativa sobre la evolución de éstas, y que se puede modelizar mediante un proceso autorregresivo de primer orden con dos posibles regímenes: uno alcista muy persistente y uno bajista que lo es en menor medida.

En la figura 8 se recogen las probabilidades filtradas estimadas mediante el modelo factorial dinámico con componentes cambiantes en el tiempo. Se observa que los periodos en que el modelo predice tendencia bajista coinciden en general con los obtenidos mediante el algoritmo de Bry & Boschan, aunque son menos persistentes dado que el procedimiento no exige la duración mínima de 3 meses de las fases. Para comparar las predicciones de los modelos se obtendrán las probabilidades estimadas un periodo hacia delante para el periodo “out of sample”, partiendo de las probabilidades filtradas expuestas, de modo muy similar al realizado para los modelos de la subsección anterior.

[Insertar Figura 8]

Comparando la capacidad predictiva del modelo se puede concluir, por ejemplo en base a la métrica QPS que se recoge en la figura 11, que esta especificación es inferior a los modelos no paramétricos, pero mejor en general a la de los modelos con regímenes cambiantes vía Markov Switching planteados en la subsección anterior. Se ahondará en la comparación de las predicciones en la sección 5.

4.4 Selección del umbral óptimo (“Threshold Moving”)

Uno de los principales determinantes de la rentabilidad obtenida por las estrategias de inversión que incorporan un componente de “market timing” será la selección del umbral o threshold a partir del cual se considera que la probabilidad predicha es suficientemente elevada como para definir una tendencia bursátil bajista. Es decir, se buscará seleccionar un valor de las probabilidades predichas por los modelos a partir del cual se determinaría que la bolsa europea pasa de una fase a otra, considerando que existen únicamente dos posibles estados subyacentes. Generalmente se emplea como threshold un valor estándar de 0.5. Sin embargo, éste podría no ser óptimo si existiese asimetría en la serie de rendimientos, o si el coste de cometer uno de los errores de clasificación fuese mayor que el del alternativo.

Como se ha mencionado previamente en 3, el número de ocasiones en que se determina mediante el algoritmo de Bry & Boschan que la bolsa experimenta una fase bajista es inferior al número de periodos en que se concluye que es alcista. Existirá por tanto cierto grado de asimetría en los datos de la variable endógena de los modelos. Por otro lado los periodos recesivos en los activos financieros tienden a ser más cortos y bruscos que aquellos en que se experimentan rendimientos positivos. Es decir, habitualmente se materializan tendencias alcistas con rendimientos positivos de magnitud relativamente reducida, frente a periodos bajistas menos frecuentes pero con valores superiores de los rendimientos en términos absolutos. Esta diferencia será más marcada todavía en las estrategias factoriales, donde históricamente se encuentra que los rendimientos en determinados periodos bajistas pueden ser lo suficientemente negativos para eliminar la revalorización experimentada durante las fases alcistas anteriores. De este modo, se puede considerar que una estrategia de inversión será más eficaz si logra evitar los periodos de grandes caídas en los precios aun a costa de abandonar el mercado en algunos periodos alcistas incorrectamente identificados. En el largo plazo

sería mejor cometer el error del tipo 2, “abandonar el mercado cuando éste es alcista”, que el de tipo 1, “mantenerse en el mercado cuando éste es bajista”.

En consecuencia, fijar el threshold predeterminado de 0.5 no sería eficiente. Se empleará un procedimiento conocido como “threshold moving” para encontrar el umbral óptimo a partir del cual considerar que la probabilidad estimada es lo suficientemente elevada para definir el régimen subyacente al mercado como bajista. Se basará en el análisis de la curva ROC (“Receiver Operating Characteristic”) y consistirá en el diagnóstico, para diferentes thresholds, de las predicciones realizadas partiendo de la estimación de cada modelo. Para cada threshold considerado se calculan las medidas “True Positive Rate” (TPR) y “False Positive Rate” (FPR) mediante las siguientes fórmulas:

$$TPR = \frac{TP}{P} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos} + \text{Falsos Negativos}} = 1 - FNR$$

$$FPR = \frac{FP}{N} = \frac{\text{Falsos Positivos}}{\text{Verdaderos Negativos} + \text{Falsos Positivos}}$$

La curva ROC será una representación de la TPR frente a la FPR para cada umbral considerado, donde la bisectriz del gráfico indica una capacidad predictiva nula. La diagonal se conoce como línea de no discriminación, y representa las predicciones que se obtendrían mediante la prueba aleatoria o “random guess”, que podría representarse por ejemplo mediante el lanzamiento de una moneda perfectamente equilibrada. Se graficarán las tasas relativas a cada uno de los thresholds analizados, y se buscará el punto de la curva más cercano al vértice superior izquierdo del gráfico. Dicho punto representará una clasificación perfecta de los datos predichos. Se preferirá el threshold que obtenga una mayor tasa de verdaderos positivos, pero teniendo en cuenta también la proporción de falsos positivos del modelo y threshold. Se busca así maximizar la TPR, pero minimizando la FPR. Los resultados se exponen en la figura 9.

[Insertar Figura 9]

En la primera figura, cada curva representa un modelo y cada punto, un threshold. Los umbrales considerados varían entre 0.05 y 0.95, con paso de tamaño 0.05. De este modo, los puntos más cercanos al cero en el eje “x” se corresponden con umbrales bajos. Se encuentra que las curvas de algunos de los modelos permiten discriminar de forma adecuada el estado subyacente, mientras que otros no logran mejorar el resultado del “random guess”. En la figura tridimensional se puede observar que los primeros, que se corresponden con los probit con una única variable exógena, logran acercarse a valores elevados de la TPR manteniendo realizaciones reducidas de la FPR. Los últimos modelos no logran en general valores elevados de TPR, manteniéndose cerca de la bisectriz.

Una vez calculados los datos de la curva, se busca seleccionar para cada modelo el threshold que más se acerca al vértice superior izquierdo. Para ello se pueden emplear la “Media Geométrica” (G-Mean) o el estadístico “J de Youden”. Éstos se obtienen según:

$$G - Mean = \sqrt{TPR * (1 - FPR)}$$

$$J - statistic = Sensitivity - Specificity = TPR + (1 - FPR) - 1 = TPR - FPR$$

y aunque el procedimiento para calcular los estadísticos es diferente, los resultados tienden a coincidir. Se preferirá el threshold que obtenga mayor valor de los estadísticos, ya que será el umbral más cercano a la clasificación perfecta. Los resultados se plantean en la figura 10.

[Insertar Figura 10]

Se encuentra que los *threshold* óptimos varían para los diferentes modelos, principalmente en función de si son paramétricos o no paramétricos. Claramente los primeros (no paramétricos y con una única variable exógena) son los que logran mayores valores de ambos estadísticos para alguno de los *thresholds* considerados. Los umbrales óptimos para cada modelo, resultado del análisis de los estadísticos, se recogen en la figura 11.

[Insertar Figura 11]

En general, los modelos paramétricos requieren la fijación de un umbral muy inferior al de los no paramétricos. Esto se puede atribuir al hecho de que, como se concluirá en la sección siguiente (5), los modelos no paramétricos presentan mayor capacidad de discriminación entre estados de la naturaleza, y como se observa en los distintos gráficos de predicciones, las probabilidades predichas por los modelos paramétricos son más extremas que las predichas por los no paramétricos. Como se mencionaba anteriormente, se puede relacionar el nivel del *threshold* obtenido con la capacidad predictiva del modelo y, en consecuencia, con el error de tipo 1 que se comete. Se encuentra una correlación de -0.76 entre el estadístico QPS y el nivel del *threshold* determinado para cada modelo. De este modo se puede afirmar que a mayor magnitud del estadístico QPS (es decir, cuanto menor sea la capacidad del modelo para discriminar entre estados de la naturaleza), menor será el nivel del umbral óptimo definido.

Por último, una vez determinados los *threshold* para los distintos modelos planteados, se pueden obtener las matrices de confusión para el periodo “out of sample”. De este modo se podrá cuantificar la capacidad predictiva de los modelos en función del porcentaje de periodos en que cada modelo logra identificar correctamente el régimen subyacente (considerando como siempre los resultados del algoritmo de Bry&Boschan como los estados “reales”). Los resultados para algunas de las principales especificaciones se plantean en la tabla 12.

[Insertar Tabla 12]

El mejor modelo probit con una única variable exógena es el que mayor porcentaje de las observaciones del conjunto de validación logra identificar correctamente.

El modelo con variables seleccionadas mediante el algoritmo genético logra un porcentaje de periodos correctamente predichos similar al del modelo factorial dinámico, pero este último es el que en más ocasiones comete error de tipo 1 (“mantenerse en el mercado cuando es bajista”). Los probit (a excepción del seleccionado mediante algoritmo genético) logran limitarlo en gran medida. Por otro lado, los modelos con parámetros cambiantes vía Markov Switching y el mejor modelo Probit con múltiples explicativas no logran superar el 70% de acierto total en las clasificaciones en el periodo.

Estos resultados se podrían atribuir al hecho de que se toman las tendencias determinadas por el BBA como las reales, mientras que en los métodos paramétricos no se tiene en cuenta la definición teórica de mercado bajista ni la duración requerida para éste. De ser este el motivo, las estrategias definidas mediante la inclusión del componente de “market timing” deberían devolver mayores rendimientos ajustados por riesgo para los modelos paramétricos, en comparación con los no paramétricos. Como se analizará en la sección 6, éste no es el caso.

En base a estas tablas se podría concluir que los probit con una única variable exógena serían los mejores predictores, seguidos del seleccionado mediante el algoritmo genético y del modelo factorial dinámico. Los modelos con parámetros cambiantes vía Markov Switching y el probit con múltiples variables serían los peores predictores según este criterio, aunque lograrían limitar el error de tipo 1 en mayor medida de la que lo hace el modelo factorial.

El procedimiento seguido en la sección se podría resumir del siguiente modo:

1. Predecir las probabilidades en el “out of sample” para cada uno de los modelos seleccionados en base a los criterios de selección planteados en 4.1.
2. Para cada threshold considerado (se analizan todos entre 0.5 y 0.95 con pasos de tamaño 0.05), se convierten las probabilidades en clases o estados, se evalúan las clases frente a las reales, y se selecciona el mejor threshold.
3. Dado que se analizan varios modelos, y por ello varias predicciones para el “out of sample”, se selecciona el threshold óptimo para cada modelo en base a los estadísticos planteados.
4. Se analiza la capacidad de discriminación de cada modelo con su correspondiente threshold óptimo mediante la matriz de confusión.
5. Se empleará el umbral seleccionado para cada modelo para definir las estrategias posteriores.

5 Comparación de Resultados

En esta sección se comparará estadísticamente la capacidad predictiva de los diferentes modelos estimados. Para ello se emplearán los tests de Diebold & Mariano y de Clark & West. El primero se empleará para comparar los modelos de una misma categoría. Por ejemplo, se comparará la capacidad predictiva de los diferentes modelos probit univariantes. El segundo test se empleará para comparar las distintas clases de modelo planteadas. Se hará esta distinción por el sesgo que aparece en el contraste de Diebold & Mariano cuando se comparan modelos anidados, y que se logra solventar mediante el segundo método planteado.

Para los modelos probit univariantes se seleccionarán los 10 mejores en base a su capacidad predictiva, según el valor de su QPS, de entre los modelos para los cuales las variables exógenas son significativas y que tienen buen ajuste “in sample”. Se ordenarán de menor a mayor valor del estadístico, de modo que los primeros serán los mejores predictores. Adicionalmente se tendrán en cuenta el modelo probit con múltiples variables explicativas seleccionadas mediante el algoritmo genético, los 4 probit con exógenas seleccionadas según su significatividad individual y ajuste in sample (9), los 4 modelos con regresores cambiantes en el tiempo en base a un proceso de Markov (10), y el modelo factorial dinámico con cambios de régimen (11).

5.1 Diebold & Mariano

El primer test considerado, introducido en [Diebold & Mariano (1995)], permite comparar estadísticamente la precisión de las estimaciones de dos modelos. El test parte de los residuos de las predicciones de los modelos, es decir, de:

$$e_1 = y_t - \hat{y}_{a,t}$$

$$e_2 = y_t - \hat{y}_{b,t}$$

Se define la función de pérdida lineal de los residuos de las predicciones al cuadrado:

$$d_t = e_1^2 - e_2^2$$

Se obtienen la media muestral de esta serie, y su coeficiente de autocorrelación de orden “k”. Se calcula:

$$\bar{d} = \frac{1}{T} \sum_{t=1}^T d_t$$

$$\gamma_k = \frac{1}{T} \sum_{t=k+1}^T (d_t - \bar{d})(d_{t-k} - \bar{d})$$

Por ejemplo, γ_0 hará referencia a la varianza muestral de la serie temporal de la función de pérdida. El estadístico de Diebold & Mariano dependerá de un último parámetro “h”, que reflejará el número de periodos a predecir. Es decir, “h” reflejará si las predicciones son estáticas o dinámicas. En este caso se realiza predicción estática, ya que se incorporan los nuevos datos cuando éstos se conocen, de modo que se fijará $h = 1$. Así, para $h > 1$,

$$DM = \frac{\bar{d}}{\sqrt{\frac{\gamma_0 + 2 \sum_{k=1}^{h-1} \gamma_k}{T}}} \sim N(0, 1)$$

Las hipótesis nula y alternativa del contraste serán:

$$\begin{cases} H_0 : E(d) = 0 \\ H_1 : E(d) \neq 0 \end{cases}$$

y, para el caso considerado donde las predicciones son estáticas ($h=1$), el estadístico de contraste será

$$DM = \frac{\bar{d}}{\sqrt{\frac{\gamma_0}{T}}} \sim N(0, 1)$$

De este modo se contrastará si, para un nivel de significatividad determinado ($\alpha = 5\%$), se puede rechazar la hipótesis nula de que las predicciones obtenidas mediante dos modelos son iguales. Cabe destacar que el test no se puede emplear para comparar modelos anidados, ya que el hecho de que las especificaciones tengan diferente número de regresores introduce un sesgo en la distribución. Los resultados del contraste se recogen en la tabla 13.

[Insertar Tabla 13]

En general, en los mejores modelos univariantes se puede rechazar la hipótesis nula de que las predicciones de los modelos sean estadísticamente iguales. Por ejemplo, comparando el modelo para la inflación retardada en un trimestre frente al del VIX retardado en un periodo, se obtiene un p-valor de 0.0262, inferior al nivel de significatividad fijado, de modo que se puede rechazar la hipótesis nula. Tal hipótesis no se puede rechazar, en general, cuando los estadísticos QPS son similares. Se debe destacar que todos los modelos probit con una única variable exógena como explicativa incluyen también los factores “momentum” y “reversión”. Es decir, son los resultantes de la estimación correspondiente a lo planteado en la subsección 5. Sin embargo no se puede encontrar un patrón relacionado con el plazo considerado para los retardos de las variables explicativas, a excepción del hecho de que no se ha seleccionado ningún probit de largo plazo

con una única variable exógena. Por ello, como se podía esperar, los indicadores considerados reflejan la evolución de factores macro-financieros diferentes.

No se puede rechazar que la capacidad para predecir las tendencias del mercado europeo de los probit con múltiples variables exógenas para distintos retardos sean estadísticamente iguales entre sí. El menor p-valor obtenido al comparar estos modelos se da cuando se enfrentan las predicciones resultantes del modelo multivariable con explicativas retardadas en 2 trimestres, con el modelo para un retardo de un año. Se obtiene un p-valor de 0.0603, mayor que el umbral fijado para el nivel de significatividad, de modo que no se puede rechazar la hipótesis nula de igualdad en las predicciones.

En cuanto a los modelos paramétricos con regresores cambiantes en el tiempo, se puede concluir que las predicciones realizadas por el modelo con variables retardadas en un solo periodo devuelve predicciones estadísticamente diferentes de las de los demás modelos de esta clase. Se encuentra que, por ejemplo, al comparar el modelo de corto plazo con el que incorpora explicativas retardadas en un trimestre, el p-valor del test es de 0.0002, claramente inferior al límite del 5%. Se puede por tanto rechazar la hipótesis nula del contraste. La conclusión es coherente con el análisis planteado en la subsección 4.3.1, ya que el modelo de corto plazo era el único para el cual se encontraban varias variables exógenas significativas para explicar las variaciones en el rendimiento del índice. De este modo, en base a los resultados del contraste y a los estadísticos QPS expuestos en 11, se puede afirmar que el modelo paramétrico de corto plazo es mejor predictor de las probabilidades de tendencia bajista en el mercado europeo que los modelos para los demás plazos. Se puede concluir que las variables analizadas tienen cierta capacidad para predecir los rendimientos del índice en el corto plazo, pero no en los demás considerados.

5.2 Clark & West

El segundo contraste a realizar, siguiendo a [Clark & West (2005)], se define también una función de pérdida lineal de los residuos al cuadrado, cuyo valor reflejará la diferencia entre el valor real de la serie y la predicción realizada por los modelos “a” y “b”. Es decir, se basa también en el análisis de los residuos de las predicciones:

$$e_1 = F(y_{a,t}) = (y_t - \hat{y}_{a,t})^2$$

$$e_2 = F(y_{b,t}) = (y_t - \hat{y}_{b,t})^2$$

Las precisión de la estimación se basa en si la diferencia entre las funciones de pérdida es nula. Es decir,

$$d_t = e_1 - e_2$$

Se tienen en cuenta de este modo la serie real y las dos series predichas. Clark & West demuestran que cuando el modelo “b” anida al “a”, el hecho de que uno de los modelos tenga más predictores que el otro generará un sesgo y, por ello, será necesario introducir un ajuste que vuelva a centrar el estadístico.

$$d_t - adj_t = F(y_{a,t}) - (F(y_{b,t}) - adj_t) = (y_t - \hat{y}_{a,t})^2 - (y_t - \hat{y}_{b,t})^2 + (\hat{y}_{a,t} - \hat{y}_{b,t})^2$$

Una vez realizado el ajuste, el estadístico finalmente obtenido se podrá comparar asintóticamente con el valor de la distribución normal estándar. Por último, se realiza una regresión de mínimos cuadrados ajustada por heterocedasticidad mediante el procedimiento de Newey-West, donde se regresa la diferencia de errores al cuadrado ajustados frente a únicamente el término independiente. Es decir, se plantea un modelo lineal sin variables explicativas. El estadístico del test de Clark & West será el t-estadístico del término independiente de la regresión de Newey-West.

Una vez realizado el procedimiento, se podrá contrastar si las predicciones son significativamente diferentes. Las hipótesis del contraste serán las siguientes:

$$\begin{cases} H_0 : MSE_a \leq MSE_b \\ H_1 : MSE_a > MSE_b \end{cases}$$

con estadístico $CW_{stat} \sim N(0, 1)$.

Se contrastará si, para un nivel de significatividad $\alpha = 5\%$, se puede rechazar la hipótesis nula de que las predicciones obtenidas mediante uno de los modelos sean significativamente mejores que las del otro. En este caso se empleará el contraste para comparar la capacidad predictiva de las diferentes clases de modelos planteadas, dado que las diferentes especificaciones de una misma categoría ya se han analizado en el subapartado anterior. Los resultados de los contrastes se pueden encontrar en la tabla 14, donde el modelo “a” será la fila, y el “b” la columna. De este modo se contrasta si las predicciones del modelo de la fila son significativamente mejores que las del modelo de la columna.

[Insertar Tabla 14]

Comparando los modelos univariados con los probit con múltiples exógenas planteados se encuentra que en todos los casos se puede rechazar la hipótesis nula de que el modelo con variables seleccionadas mediante algoritmo genético sea mejor predictor que cualquiera de los modelos univariados. Por ejemplo, al comparar el modelo seleccionado por el algoritmo genético con el modelo univariable de medio plazo (retardo de un trimestre) para la inflación, se obtiene un p-valor de 0.0013, lo cual es un dato inferior al umbral previamente fijado en el 5%. De este modo se podría rechazar la hipótesis nula de que las predicciones del modelo seleccionado por el algoritmo sean significativamente mejores o iguales a las del modelo para la inflación. Se puede concluir que la complejidad adicional que introduce el algoritmo genético no se compensa con una mayor capacidad predictiva. Esta misma conclusión se encuentra para todos los demás modelos no paramétricos multivariados al compararlos con los 6 mejores probit univariados¹⁷. Sin embargo, se encuentra que a partir del sexto mejor modelo con una única variable exógena, para los probit con varias exógenas seleccionadas manualmente en base a su significatividad individual y al ajuste “in sample”¹⁸, no se puede rechazar la hipótesis nula de que éstos tengan mayor o igual capacidad predictiva. Por ejemplo, al comparar el modelo multivariable de corto plazo (“Multivariable, t-1”) con la pendiente de “USA” para el retardo de un trimestre, se obtiene un p-valor de 0.058, ligeramente superior al nivel de significatividad de 0.05 que se emplea como límite en todo el estudio. En estos casos se puede concluir, combinando los resultados del contraste con los estadísticos QPS de los modelos, que los que combinan distintas exógenas devuelven predicciones significativamente mejores. Cabe remarcar que habiendo analizado los modelos para 28 variables exógenas y 4 plazos diferentes de cada una de ellas, solo 6 de estos modelos univariados son significativamente mejores predictores que los de múltiples variables exógenas para el medio plazo, por lo que se puede considerar que la combinación de las variables momentum, spread americano y francés, spread entre los bonos calificados como Baa y Aaa y el componente autorregresivo, permite estimar modelos con una capacidad predictiva elevada en el medio plazo.

A excepción del modelo de largo plazo, no se puede rechazar la hipótesis nula al comparar entre sí los modelos con varias exógenas para distintos plazos. Ésto coincide con los resultados del test anterior, de modo que se puede afirmar que la capacidad para predecir el signo de las tendencias de los modelos para el corto y medio plazo es similar, y significativamente mejor que la del modelo de largo plazo. Sin embargo se puede

¹⁷Con una única variable exógena, pero manteniendo los factores reversión y momentum.

¹⁸En las tablas, “Multivariable, t-k”

rechazar la hipótesis nula de que los modelos con múltiples variables exógenas seleccionadas “manualmente” tengan mejor o igual capacidad predictiva que los de variables seleccionadas mediante el algoritmo genético. Al comparar el modelo probit multivariable de corto plazo con el seleccionado por el algoritmo, se alcanza un p-valor de $0.0043 < 0.05$, de modo que se puede rechazar H_0 , y concluir que el modelo multivariable no es significativamente mejor o igual predictor que el seleccionado por el algoritmo genético. Combinando estas conclusiones con los estadísticos QPS asociados a estos modelos, se podría concluir que el modelo seleccionado por el algoritmo genético para el largo plazo sería mejor predictor que los demás probit con que se lo compara.

Comparando estos modelos con los paramétricos, se concluye que los no paramétricos son claramente mejores predictores, ya que se puede rechazar siempre la hipótesis nula de que los modelos con regresores cambiantes vía markov switching sean significativamente mejores predictores que los modelos probit con múltiples variables explicativas. Si se comparan las predicciones para un retardo de un trimestre en ambas clases de modelos, se encuentra un p-valor de $7.58e^{-05} < 0.05$, de modo que se puede rechazar que el modelo Markov Switching sea mejor o igual predictor que el probit dinámico multivariable.

Por último, en la comparación entre las diferentes clases de modelos paramétricos analizadas, no se puede rechazar que el modelo factorial dinámico sea significativamente mejor predictor que los demás modelos paramétricos planteados, a excepción del modelo de corto plazo. Al comparar el MS-DFM con el MS para el retardo de un trimestre, se encuentra un p-valor de $0.2078 > 0.05$, de modo que no se puede rechazar la hipótesis nula de que el modelo factorial dinámico sea significativamente mejor o igual predictor que el modelo Markov Switching de medio plazo. Combinado con el análisis del QPS de los diferentes modelos paramétricos y con la comparación de las matrices de confusión de estos, se puede afirmar que el modelo factorial es significativamente mejor que los modelos markov switching de medio y largo plazo, pero que su capacidad predictiva es similar a la del modelo de corto plazo.

En resumen, se puede concluir que los mejores modelos univariantes son significativamente mejores especificaciones para la predicción de los puntos de inflexión de los ciclos del mercado europeo que todas las demás especificaciones planteadas. Sin embargo se pueden encontrar modelos con múltiples variables explicativas comparables a los de variables exógenas individuales. Para encontrar tales modelos no es de utilidad el algoritmo genético considerado, ya que aunque se encuentra que el algoritmo presenta una capacidad predictiva estadísticamente superior a la de los modelos probit con exógenas seleccionadas “manualmente”, los resultados se podrían achacar a “data mining” al ser positivos solo para el modelo de largo plazo. Por último, los modelos no paramétricos son significativamente mejores mecanismos de predicción que los paramétricos. Pese a ello, el modelo factorial dinámico con cambios de régimen en base a un proceso de Markov, además de permitir la monitorización del mercado en tiempo real y mayor flexibilidad con las frecuencias de las variables explicativas, devuelve mejores predicciones que los modelos paramétricos de regresión habituales con parámetros cambiantes vía Markov Switching, aun siendo la especificación que comete error de tipo 1 en mayor medida.

6 Estrategias

En esta sección se definirán y analizarán diferentes estrategias de inversión, a las cuales se incorporará un componente de “market timing” mediante las predicciones de los modelos estimados previamente. Se analizará si con ello se puede mejorar su rendimiento ajustado por riesgo para el periodo “out of sample”. El principal objetivo que se persigue al incorporar las predicciones de las tendencias del mercado será limitar las pérdidas potenciales cuando se experimentan tendencias recesivas, que se derivan de mantener las estrategias definidas para mercados alcistas. Ésto será muy relevante por ejemplo en las estrategias basadas en el factor momentum, donde el “drawdown” potencial puede llegar a eliminar el exceso de rendimiento generado por

la estrategia durante los periodos alcistas.

Se analizarán tres estrategias: El “buy and hold” del índice pero abandonando el mercado en los periodos en que se predice signo bajista, una estrategia de inversión basada en la evolución del factor momentum pero abandonándola también en los periodos bajistas, y una estrategia basada en el “factor value” incorporando el “market timing” del mismo modo que en las anteriores. Cuando se predice una tendencia recesiva, se pasará a invertir en bonos europeos a 3 meses.

Para definir las estrategias factoriales se seleccionarán los activos en los que invertir de entre el universo de inversión que proporciona el índice europeo Eurostoxx 600. Para evitar anomalías en los datos, se eliminan aquellos activos que presenten una historia de cotización inferior a los 200 meses. Además, para evitar los activos más ilíquidos, se eliminan también los stocks por encima del percentil 95 cuando se los ordena según su volatilidad en el periodo. De este modo se obtiene un conjunto de potenciales inversiones de alrededor de 300 activos.

Se compararán los rendimientos de las estrategias, su asimetría, su máxima caída en un periodo, y los ratios de Sharpe¹⁹ y Sortino²⁰ asociados, con la estrategia básica de “buy and hold” del índice Eurostoxx 50. También se compararán los rendimientos de las estrategias factoriales con componente de market timing frente a los resultados que se obtendrían de las mismas estrategias si no se tuviesen en cuenta las tendencias del mercado agregado. Los ratios de Sharpe se calcularán en comparación con el tipo europeo a 3 meses. Sin embargo, para calcular el ratio de Sortino se tomará el índice Eurostoxx 50 como referencia, de modo que se pueda comparar de forma más directa el desempeño de las estrategias frente al “buy and hold”.

En la figura 12 se pueden observar los rendimientos acumulados en el periodo “out of sample” de las diferentes estrategias planteadas, empleando las probabilidades predichas por el mejor modelo probit con una única variable exógena.

[Insertar Figura 12]

El índice no ha devuelto una rentabilidad acumulada elevada en el periodo de 6 años considerado, ya que se experimentan varias fases bajistas. Ha sido sin embargo un periodo en el que la estrategia de momentum ha permitido obtener rendimientos relativamente elevados, sin experimentar grandes caídas. Esta estrategia se caracteriza por generar rendimientos muy negativos en determinados periodos históricos, que tienden a coincidir con las mayores caídas del mercado en agregado, eliminando las revalorizaciones de los periodos anteriores. Un ejemplo de dichos periodos es el que define el mercado bajista de 2008. Es por ello que se puede argumentar que incorporar “market timing” a la estrategia podría limitar el “drawdown” potencial y que, aun pudiendo limitar parte del recorrido alcista de la estrategia, se podría mejorar el rendimiento ajustado por riesgo. Por otro lado se observa que en el “out of sample” la estrategia “value” ha presentado periodos de rendimientos negativos sostenidos y que en general no ha devuelto un rendimiento superior al que se obtendría mediante la inversión indexada en el periodo analizado.

Los principales supuestos considerados en la sección serán unos costes de transacción de 20 puntos básicos cada vez que se lleven a cabo cambios en la composición de la cartera, y que los activos son perfectamente divisibles. Por otro lado, para generar las carteras asociadas a las estrategias de momentum y “value” que se plantean a continuación, se ponderarán los activos en base a su precio. De este modo se

¹⁹El ratio de Sharpe se define como $SR(r, \theta) = \frac{\mu_r - \theta}{\sigma_r}$, donde $\mu_r = E(r)$ es el rendimiento esperado de la estrategia, θ es la rentabilidad del activo de referencia (en este caso, el bono europeo a 3 meses), y σ_r es la volatilidad de la cartera.

²⁰El ratio de Sortino es un caso particular de las medidas Kappa definidas como $\kappa_m(r, \theta) = \frac{\mu_r - \theta}{m \sqrt{LPM_m(r, \theta)}}$. El ratio de Sortino será aquel que fija $m = 2$, de modo que emplea como medida de ajuste por riesgo el “Lower Partial Moment” (LPM) en lugar de la volatilidad. De este modo se tiene en cuenta la desviación típica de únicamente los rendimientos negativos de la cartera (downside risk). Se define el LPM como $LPM_m(r, \theta) = E_f[\max(\theta - r, 0)^m] = \int_{\theta}^{\infty} (r - \theta)^m f(r) dr$.

invertirá una proporción mayor en los activos de mayor valor. Esta decisión se toma por conservadurismo y para compensar posibles infravaloraciones respecto a la realidad en los costes de transacción, ya que en la literatura previa se ha comprobado que el efecto de los factores es más marcado en los stocks de menor capitalización.

6.1 Estrategia Naive

La estrategia consiste en, cada mes, tomar posiciones en el índice de referencia europeo o en el activo libre de riesgo en función de si la tendencia predicha para dicho periodo es alcista o bajista. Es decir, el rendimiento de la cartera será:

$$R_t = w_t * \Delta B_t + (1 - w_t) * \Delta S_t$$

donde ΔB_t es el cambio en el precio del bono europeo con vencimiento a 3 meses y $\Delta S_t = \text{rendimiento índice}$ es la variación en el precio del índice de referencia. Por otro lado,

$$w_t = 1 \text{ si } \hat{P}_{t-1}(R_t = 1) \geq \text{Threshold}$$

$$w_t = 0 \text{ si } \hat{P}_{t-1}(R_t = 1) < \text{Threshold}$$

siendo $\hat{P}_{t-1}(R_t = 1)$ la probabilidad de mercado bajista predicha a un mes vista, obtenida en base a los mejores modelos planteados en las secciones anteriores. El activo libre de riesgo será el bono a 3 meses, y los umbrales serán los que se determinan para cada modelo en la sección 4.4. Así, si por ejemplo la probabilidad predicha para el siguiente periodo es del 70% y el threshold se define en el 40%, se pasaría a invertir en el activo libre de riesgo. Si por el contrario la probabilidad predicha estuviese por debajo del umbral, se invertiría en el índice.

Los resultados para cada modelo planteado se recogen en la tabla 15.

[Insertar Tabla 15]

La estrategia devuelve mayor rendimiento que el índice de referencia siempre que se emplean las probabilidades predichas mediante modelos probit. Por ejemplo, al emplear el modelo de “market timing” de medio plazo para el “default yield spread” se obtiene un rendimiento logarítmico acumulado en el periodo “out of sample” de 0.4876, claramente superior al 0.1593 devuelto por la inversión indexada al Eurostoxx 50. Sin embargo, y coincidiendo con los resultados obtenidos en la comparación de la capacidad predictiva de los modelos, se observa que el rendimiento acumulado en el periodo para las estrategias “naive” que incorporan “market timing” mediante métodos paramétricos de predicción no logran en general batir al índice. Es destacable el rendimiento acumulado obtenido por la estrategia basada en el modelo factorial dinámico, ya que se genera un rendimiento logarítmico acumulado de únicamente 0.0562, muy inferior al 0.1593 del índice.

Todos los modelos, a excepción del factorial dinámico, son superiores al índice cuando se analiza el rendimiento ajustado por riesgo en el periodo de validación. Por ejemplo, el modelo de medio plazo para la variable Baa-Aaa genera un ratio de Sharpe de 0.5274, mientras que el asociado al índice de referencia es de únicamente 0.1933. Cuando se analiza el ratio de Sortino, se encuentra que todos los modelos probit son superiores al índice si se ajusta el rendimiento únicamente por la volatilidad de los rendimientos negativos. Sin embargo los modelos paramétricos no siempre logran mejorar el rendimiento del Eurostoxx 50 cuando

se emplea el “downside risk” como medida para el ajuste por riesgo. El modelo factorial dinámico genera un ratio de Sortino (tomando el Eurostoxx 50 como referencia) de -0.101.

De este modo se puede afirmar que las estrategias de “market timing” logran mejorar el rendimiento y reducir el riesgo asumido, sobre todo cuando se toman las decisiones de entrada/salida del mercado en base a métodos no paramétricos. Adicionalmente, en la tabla 15 se recoge también la asimetría de los rendimientos de las distintas estrategias en el periodo. Se puede apreciar que, a excepción nuevamente del modelo factorial dinámico, todos los demás modelos de “market timing” permiten establecer estrategias convexas aun partiendo de una serie de rendimientos con asimetría negativa. La asimetría en los rendimientos logarítmicos pasa del -0.3052 del índice, a por ejemplo un 0.5952 cuando se incorpora “market timing” mediante el modelo de medio plazo para el “default yield spread”. En general se logra además reducir la máxima caída en más de un 35%, de modo que, como se esperaba, al introducir market timing se logra abandonar el mercado en los periodos en que se materializan los rendimientos más negativos.

6.2 Momentum

La segunda estrategia considerada consistirá en seleccionar aquellas acciones europeas con mejor rendimiento en el año anterior (sin incluir el rendimiento del mes más reciente), e invertir en ellos si se predice que la tendencia del mercado será alcista en el siguiente periodo. En caso de predecir tendencia bajista, la cartera para el mes siguiente estará compuesta únicamente por bonos europeos a 3 meses. Los activos se mantendrán en cartera durante un mes y, a su finalización, se rebalanceará el portfolio siguiendo el mismo criterio. Es decir, se busca invertir cada mes en los ganadores europeos del año anterior ²¹, y se sigue la estrategia mientras se encuentren perspectivas alcistas para el mercado. La definición de la estrategia se basa en los trabajos de [Asness (1995), Asness et. al. (2013), Jegadeesh & Titman (2001)], cuyas aportaciones a la teoría de la inversión factorial se resumen en la sección 2. Se puede expresar el proceso del siguiente modo:

$$R_t = w_t * \Delta B_t + (1 - w_t) * \Delta S_t$$

donde R_t es el rendimiento de la cartera en el periodo “t”, ΔB_t es el cambio en el precio del bono europeo con vencimiento a 3 meses, y $\Delta S_t = Q_2(\Delta S(12, 2))$ es la variación en el precio de los stocks cuyo rendimiento en el año previo (sin incluir el mes más reciente) se encuentra entre los dos deciles superiores de los rendimientos de los activos del Eurostoxx 600 en dicho periodo. Por otro lado, nuevamente:

$$w_t = 1 \text{ si } \hat{P}_{t-1}(R_t = 1) \geq \text{Threshold}$$

$$w_t = 0 \text{ si } \hat{P}_{t-1}(R_t = 1) < \text{Threshold}$$

siendo $\hat{P}_{t-1}(R_t = 1)$ la probabilidad estimada mediante los modelos de “market timing” de que el mercado sea bajista en el mes siguiente. Se comparará el desempeño de esta estrategia con el que se obtendría de invertir siempre en los activos con mejor momentum (es decir, sin “incorporar market” timing a la estrategia), y con el rendimiento pasivo del Eurostoxx 50 en el periodo.

El exceso de rendimiento sobre el del mercado de referencia obtenido por las estrategias basadas en el momentum se explica en general en base a la teoría de los mercados eficientes, pero también existen explicaciones basadas en las finanzas conductuales. En la literatura no existe consenso sobre cual de las dos es la más realista.

²¹ Activos cuyo rendimiento se encuentra entre los primeros dos deciles en los 12 meses previos, sin incluir el mes más reciente.

Analizando la evolución temporal de la estrategia basada en el factor momentum, se observa que ésta presenta periodos en que el rendimiento es consistentemente negativo. Estos periodos coinciden con momentos en que la tendencia general del mercado es bajista, como se puede apreciar en el gráfico 12a. También la estrategia “value” presenta rendimientos negativos de forma consistente en los periodos bajistas.

[Insertar Figura 12a]

Se podría esperar por tanto que la rentabilidad de las estrategias basadas en el factor será función en parte de la capacidad del inversor para detectar los cambios en las tendencias bursátiles, y con ello los periodos en que la estrategia devolverá rendimientos negativos. Esta explicación justifica el incluir un componente de “market timing”, ya que si se logra predecir el signo de la tendencia del mercado de referencia, se recogería en la estrategia el exceso de rentabilidad esperada asociada a los periodos alcistas, logrando evitar los “momentum crashes” o periodos en que la estrategia deja de funcionar.

En la tabla 15 se recogen los resultados para el periodo “out of sample” de la estrategia basada en el factor momentum. Se incluye también la que no incorpora “market timing”, a modo de comparación.

Se encuentra que en general el incorporar “market timing” a la estrategia de momentum no permite mejorar el rendimiento acumulado en el periodo. El modelo probit dinámico multivariable para explicativas retardadas en un trimestre devuelve por ejemplo un rendimiento logarítmico acumulado en el periodo de validación de 0.298, lo cual es inferior al rendimiento de la estrategia sin incorporar “market timing” (0.4046), aunque superior al rendimiento del índice. La excepción son los mejores modelos probit con una única explicativa.

Cuando se analiza el ratio de Sharpe se observa que para todos los modelos se logra superar al índice, aunque no en todos los casos se puede mejorar la estrategia de momentum “básica”. Los probit dinámicos univariados sí logran en general mejorar el ratio de Sharpe obtenido por la estrategia base. Por ejemplo, el modelo para la inflación retardada en un trimestre obtiene un ratio de 0.4053, superior al 0.3281 que genera la estrategia básica de “momentum”.

Nuevamente se encuentra que la predicción de las tendencias del mercado agregado permiten pasar de carteras con una distribución de rendimientos con asimetría negativa (-0.1951 de la estrategia de momentum base) a estrategias con valores positivos del estadístico (0.9502 para el modelo de medio plazo de la inflación). Este cambio en el signo de la asimetría podría indicar que el hecho de que no se mejore el rendimiento ajustado por riesgo podría deberse simplemente a que, como se observa en las figuras 12a y 12, no se habrían experimentado grandes rendimientos negativos de la estrategia en el periodo “out of sample”. Comparando el periodo del “out of sample” con el mercado bajista de 2008, se encuentra que la peor caída de la estrategia es un 57% mayor que la del periodo de validación²². Por último, analizando los peores periodos de la estrategia en cuanto a rentabilidad, se observa que con la introducción de “market timing” mediante los modelos no paramétricos se logra reducir también la máxima caída en más del 50%, en el 60% de los casos planteados. Sin embargo los modelos paramétricos tampoco logran en general ganancias en este aspecto.

6.3 Value

Finalmente se plantea una estrategia basada en la selección de activos bursátiles europeos en base a su exposición al factor de riesgo “Value”. [Asness et. al. (2013)] analizan la eficacia de las estrategias de momentum y “value investing” en diferentes áreas geográficas y en distintas clases de activos. Destacan que las estrategias que combinan la aproximación “value” y la de momentum logran rendimientos positivos sostenidos y estables (aunque de magnitud reducida), y llevan asociados mejores ratios de Sharpe. Logran carteras menos

²² Algo similar sucede con las estrategias “Value” y “Naive”.

volátiles a lo largo del tiempo. Los autores basan la estrategia en la existencia de una relación negativa entre el “value” y el momentum, que se acentuaría en las fases bajistas del mercado. Por ello, el combinar las estrategias permitirá limitar las pérdidas potenciales en mercados bajistas, pero implicaría enormes costes de transacción. Se podría lograr lo propio mediante el “market timing” sistemático.

Por otro lado, tanto [Jegadeesh & Titman (2001)] como [De Bond & Thaler (1985)] concluyen que las estrategias de reversión de más largo plazo devuelven una prima en el rendimiento que se puede explicar mediante la exposición al “factor value”. De tal manera se podrían seleccionar los activos que peor rendimiento hayan presentado en los últimos 5 años, sin tener en cuenta el más reciente, para replicar el rendimiento que se obtendría si se generase una cartera en la que se incluyesen los activos con mayor ratio “Value to Book”.

En este apartado, la estrategia planteada consiste en seleccionar los activos que se encuentren por encima del percentil 80 de peores rendimientos acumulados en los 5 años previos sin tener en cuenta el más reciente. Dicho procedimiento permitiría replicar el rendimiento de una estrategia “Value” basada en la selección de activos con el mayor “Value to Book” posible. Además, se incorporará el componente de “market timing” del mismo modo que se hacía en la estrategia de momentum. Es decir, abandonando la estrategia en los periodos bajistas del mercado.

Los resultados se reflejan en la tabla 15. Las estrategias que incorporan “market timing” logran batir en general tanto al índice como a la estrategia base cuando las predicciones se realizan mediante modelos probit dinámicos. Sin embargo, cuando los modelos empleados son paramétricos, el rendimiento de la estrategia es muy reducido. El modelo factorial dinámico por ejemplo, genera un rendimiento logarítmico acumulado de 0.0367, inferior al 0.1677 que se obtendría con la estrategia “value” sin “market timing”, e inferior también al rendimiento de 0.1593 devuelto por el índice. Se encuentra también que, según los ratios de Sortino y Sharpe, las estrategias “Value” con “market timing” presentan mejor rendimiento ajustado por riesgo que el índice de referencia cuando se emplean modelos probit. El modelo de la inflación a medio plazo por ejemplo, devuelve un ratio de Sharpe de 0.3764 y un Sortino de 0.0801, superiores a los asociados a la estrategia “value” base, que serán respectivamente de 0.1901 y de 0.0217. Se llega a la conclusión opuesta (según el ratio de Sortino) para los modelos paramétricos. Por otro lado, en este caso no se puede afirmar que en general se logre cambiar el signo de la asimetría en la distribución de los rendimientos de las estrategias. Se encuentra sin embargo que, aunque en menor medida que para las estrategias anteriores, sí se logra reducir en la gran mayoría de los casos el máximo rendimiento negativo.

7 Conclusiones

El principal objetivo de este trabajo ha sido la selección de variables macro-financieras y la comparación de diferentes modelos para tratar de predecir las tendencias futuras del mercado bursátil europeo, planteando relaciones no lineales entre las variables consideradas. Para seleccionar los factores se ha analizado la capacidad de un algoritmo genético basado en el ajuste “in sample” de los modelos para determinar cual es la combinación óptima de indicadores adelantados. Se han comparado además la capacidad predictiva de diferentes modelos paramétricos y no paramétricos, determinando los umbrales óptimos asociados a cada uno de ellos, y comprobando si se pueden definir estrategias de inversión para capitalizar sobre sus predicciones.

Se encuentra que el algoritmo genético de selección de modelos no es un método que se pueda emplear de forma sistemática para realizar predicciones cuando el problema planteado presenta un componente elevado de aleatoriedad, ya que el algoritmo sobreajusta a los datos “in sample”.

Por otro lado, y coincidiendo con la literatura previa, se obtiene que los modelos no paramétricos con una única variable explicativa no logran predecir adecuadamente las tendencias. Esta conclusión cambia cuando a la variable exógena de cada modelo se añade información recogida en el pasado de la serie del índice de referencia, como pueden ser el factor momentum o un componente autorregresivo. En este caso se encuentra

que tanto el ajuste “in sample” de los modelos como su capacidad predictiva mejoran significativamente. En consecuencia se puede concluir que los indicadores del momentum y el componente autorregresivo del índice son variables a tener en cuenta a la hora de tratar de predecir las tendencias futuras del mercado. A estos factores se pueden añadir el “Default yield spread” y las pendientes de las curvas de tipos americana y francesa para definir modelos dinámicos con múltiples variables exógenas y capacidad predictiva en el “out of sample”, ya que se comprueba que éstas son significativas para diferentes plazos. Además, las variaciones en el índice VIX y el spread de la curva de tipos europea permiten establecer también modelos probit dinámicos con una única variable exógena y elevada capacidad predictiva, aunque estos indicadores pierden la significatividad individual al combinarlos con los mencionados anteriormente.

También se compara estadísticamente la capacidad predictiva de los modelos paramétricos y no paramétricos planteados. Se encuentra que los modelos no paramétricos son en este sentido significativamente mejores que los paramétricos, y que el modelo factorial dinámico es en general mejor predictor que los modelos clásicos de regresores cambiantes en el tiempo vía Markov Switching. La excepción a esto sería el modelo de corto plazo, para el cual se puede concluir que ambas clases de modelos paramétricos tienen una capacidad predictiva similar. Se comprueba que los mejores modelos con una única variable exógena (pero incorporando los factores de momentum y reversión) son preferibles frente a los modelos con varios factores exógenos macro-financieros, aunque estos últimos presentan una capacidad predictiva comparable en el medio plazo. Además se encuentra que los modelos paramétricos cometen error de tipo 1 en la clasificación de las tendencias futuras en mayor medida de lo que lo hacen los no paramétricos.

Es destacable también el hecho de que los threshold óptimos encontrados para los modelos paramétricos son mucho menores en nivel que los umbrales asociados a los modelos no paramétricos. Esto se puede atribuir al hecho de que los modelos paramétricos predicen probabilidades más extremas y que, en consecuencia, se puede definir un umbral bajo a partir del cual considerar que el modelo predice un mercado bajista.

Por último, se concluye que los modelos dinámicos de elección binaria son una herramienta útil para predecir el signo de los movimientos futuros del mercado. Podrán ser empleados por los inversores para plantear diferentes estrategias con componente de “market timing”, ya que se encuentra que este añadido permite superar a la estrategia de “buy and hold” tanto en rendimiento acumulado como en las medidas de “performance” asociadas. Esto no es tan evidente en las estrategias factoriales, aunque se encuentran resultados positivos en los estadísticos descriptivos de los rendimientos de las estrategias. Se encuentra por ejemplo que el componente de “market timing” permite establecer estrategias convexas a partir de planteamientos como el “buy and hold” y la inversión en momentum, que serían estrategias cóncavas por sí solas. También se logra mejorar los ratios de Sharpe de las estrategias “value”, o reducir las caídas máximas de las estrategias planteadas.

References

- [Acosta & Fernández (2007)] Acosta, E., Fernández, F., 2007. Model selection via genetic algorithms illustrated with cross-country growth data. *Empirical Economics*
- [Akgül et al. (2015)] Akgül, I., Bildirici, M., Özdemir, S., 2015. Evaluating the non-linear linkage between gold prices and stock market index using Markov Switching bayesian VAR models. *Procedia - Social and Behavioral Sciences*, Vol. 210, pp. 408-415.
- [Asness (1995)] Asness, C., 1995. The power of past stock returns to explain future stock returns. SSRN.
- [Asness et. al. (2013)] Asness, C., Moskowitz, T., Pedersen, L., 2015. Value and momentum everywhere. *The Journal of Finance*, Vol 68, No. 3.
- [Camacho et. al. (2018)] Camacho, M., Pérez-Quiros, G., Poncela, P., 2018. Markov Switching dynamic factor models in real time. *International Journal of Forecasting*, Vol. 34, pp. 598-611.
- [Candelon et. al. (2008)] Candelon, B., Piplack, J., Straetmans, S., 2008. On measuring synchronization of bulls and bears: The case of East Asia. *Journal of Banking & Finance*, Vol. 32, pp. 1022-1035.
- [Candelon et. al (2012)] Candelon, B., Ahmed, J., Straetmans, S., 2012. Predicting and capitalizing on stock market bears in the U.S. Maastricht University.
- [Chauvet & Piger (2008)] Chauvet, M., Piger, J., 2008. A comparison of the real time performance of business cycle dating methods. *Journal of Business and Economic Statistics*, Vol. 26, pp. 42-49.
- [Chen & Liang (2007)] Chen, Y., Liang, B., 2007. Do market timing hedge funds time the market? *Journal of Financial and Quantitative Analysis*, Vol. 42, No. 4, pp. 827-856.
- [Clark & West (2005)] Clark, T., West, K., 2005. Using out of sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of Econometrics*, Vol. 135, pp. 155-186.
- [Cremers (2000)] Cremers, M., 2000. Stock return predictability: A bayesian model selection perspective. Stern School of Business.
- [De Bond & Thaler (1985)] De Bond, F., Thaler, R., 1985. Does the stock market overreact? *The Journal of Finance*, Vol XL, No. 3.
- [Diebold & Mariano (1995)] Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. *Journal of Business and Economic Research*, Vol. 13, pp. 253-265.

- [Do, Faff & Veeraraghavan (2009)] Do, V., Faff, R., Veeraraghavan, M., 2009. Do Australian hedge fund managers possess timing abilities?, *Applied Financial Economics*, Vol. 19, pp. 27-38.
- [Duprey et al. (2015)] Duprey, T., Klaus, B., Peltonen, T., 2015. Dating systemic financial stress episodes in the EU countries. ECB working paper.
- [Estrella & Mishkin (1996)] Estrella, A., Mishkin, F., 1996. Predicting U.S. recessions: Financial variables as leading indicators. Federal Reserve Bank of New York research paper No. 9609.
- [Fernández & Sosvilla (1997)] Fernández, F., Sosvilla, S., García, M., 1997. Using nearest neighbour predictors to forecast the spanish stock market. *Investigaciones Económicas*, Vol. XXI. pp. 75-91.
- [Fernández, Fernández & Sosvilla (2014)] Fernández, A., Fernández, F., Sosvilla, S., 2014. The term structure of interest rates as predictor of stock returns: Evidence for the Ibex 35 during a bear market. *International Review of Economics and Finance*, Vol. 31, pp. 21-33.
- [Griffin et. al. (2003)] Griffin, J., Ji, S., Spencer, J., 2003. Momentum investing and business cycle risk: Evidence from pole to pole. *The Journal of Finance*, Vol 58, No. 6.
- [Hood & Malik (2013)] Hood, M., Malik, F., 2013. Is gold the best hedge and a safe haven under changing stock market volatility? *Review of Financial Economics*, Vol 22, pp. 47-52.
- [Ibbotson & Kaplan (2001)] Ibbotson, R., Kaplan, P., 2001. Does asset allocation policy explain 40, 90 or 100 percent of performance? *Financial Analysts Journal*, Vol 56.
- [Jegadeesh & Titman (2001)] Jegadeesh, N., Titman, S., 2001. Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of Finance*, Vol. 56, No. 2, pp. 699-720.
- [Laopodis (2006)] Laopodis, N., 2006. Dynamic linkages between monetary policy and the stock market. *Journal of Business & Economic Research*, Vol. 4, No. 12.
- [McDermott & Baur (2009)] McDermott, T., Baur, D., 2009. Is gold a safe haven? *International evidence*. *IIIS*, No. 310.
- [Papadamou & Stephanides (2007)] Papadamou, S., Stephanides, G., 2007. Improving technical trading systems by using a new Matlab-based genetic algorithm procedure. *Mathematical And Computer Modelling*, Vol 26, pp. 189-197.
- [Pesaran & Timmermann (1995)] Pesaran, M., Timmermann, A., 1995. Predictability of stock returns: Robustness and economic significance. *The Journal of Finance*, Vol. 50, No. 4. pp. 1201-1228.

[Rodríguez & Sosvilla (2006)]

Rodríguez, P., Sosvilla, S., 2006. Understanding and forecasting stock price changes. Is it possible? FEDEA.

[Stock & Watson (1989)]

Stock, J., Watson, M., 1989. New indexes of coincident and leading economic indicators. National Bureau of Economic Research. pp. 351-409.

[Osinga et. al. (2016)]

Osinga, A., Schauten, M., Zwinkels, R., 2016. The factor timing ability of hedge fund managers.

A Anexo 1. Justificación Teórica Variables Explicativas.

En este primer anexo se expone la justificación teórica tras las variables explicativas empleadas a lo largo del trabajo como posibles indicadores adelantados de los puntos de inflexión de los ciclos bursátiles europeos. Las variables en cuestión son las enumeradas en la sección 3:

Table 1: Variables Explicativas

Componentes Curva de Tipos	Variables Macro-Financieras
Euro Zona nivel ²³	Nasdaq
Euro Zona pdte. (10YT - 3m) ²⁴	Dax
Euro Zona pdte. (5YT - 3m)	TC Euro/Dollar
Euro Zona pdte. (2YT - 3m)	Brent (petróleo)
Estados Unidos pdte. (10YT - 3m)	Oro spot
Estados Unidos pdte. (5YT - 3m)	Ratio vol. Put, Call Dax ²⁵
Estados Unidos pdte. (2YT - 3m)	Índice VIX Europeo
Alemania pdte. (10YT - 3m)	Ratio bonos Baa-Aaa USA
Alemania pdte. (5YT - 3m)	M3 ²⁶
Alemania pdte. (2YT - 3m)	IPC Euro Zona
Francia pdte. (10YT - 3m)	Construcción Euro Zona
Francia pdte. (5YT - 3m)	Importaciones Euro Zona
Francia pdte. (2YT - 3m)	Desempleo Euro Zona
Italia pdte. (10YT - 3m)	
Italia pdte. (5YT - 3m)	
Italia pdte. (2YT - 3m)	

Tras aplicar las pertinentes transformaciones para estacionarizar las series temporales, se tienen las siguientes representaciones gráficas:

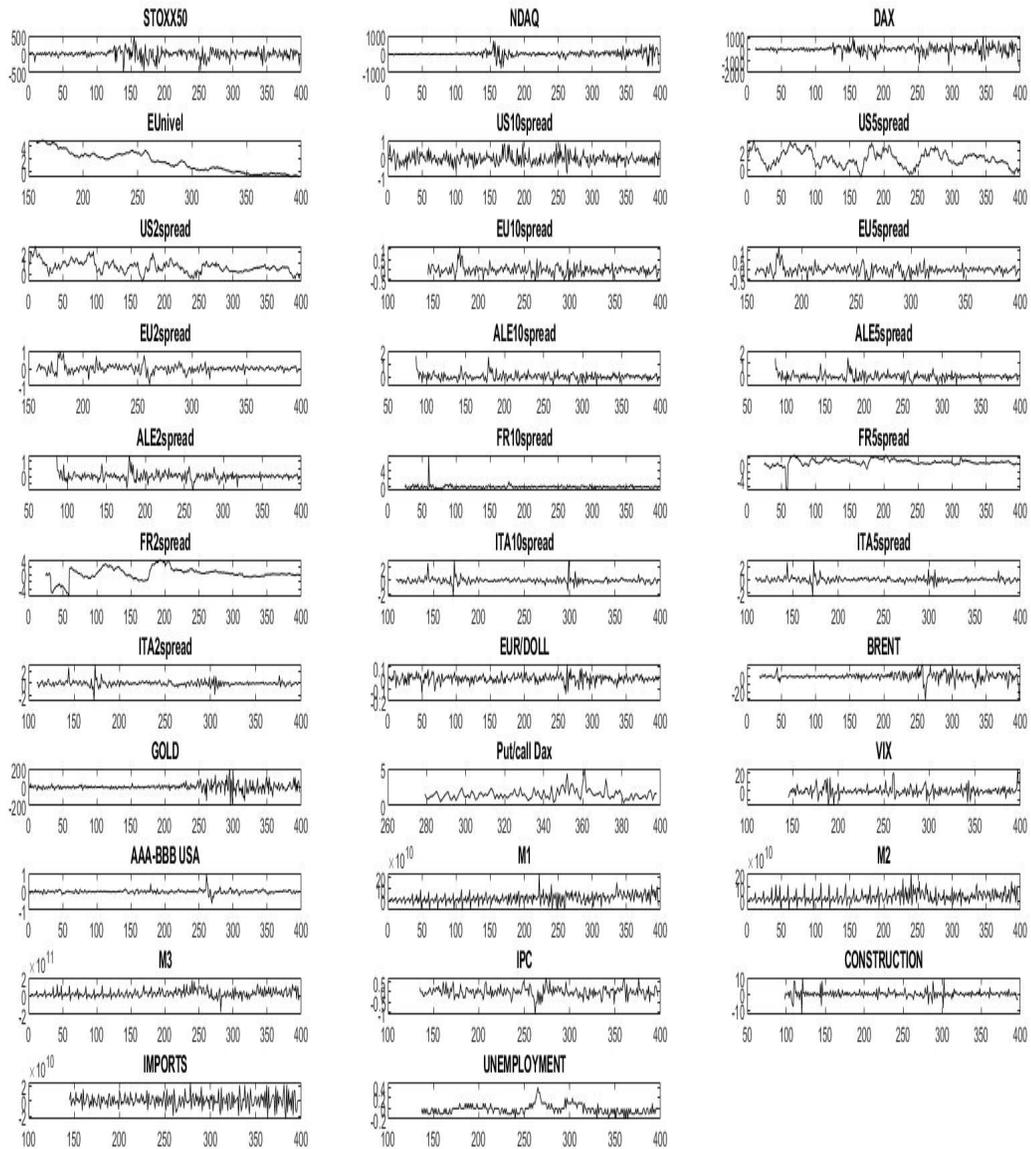
²³El nivel se calcula como la media de los rendimientos de los bonos europeos para los 4 plazos considerados.

²⁴Las pendientes se calculan como la diferencia entre la rentabilidad de los bonos a "x" años y la rentabilidad a 3 meses, con x = 2,5,10.

²⁵Finalmente se elimina por la longitud de la muestra.

²⁶También se analizan M1 y M2, pero no se incluyen en los modelos por la elevada correlación entre las tres variables.

Figure 1: Representación Variables



Nota: Aplicando el contraste aumentado de Dickey-Fuller, se puede rechazar la hipótesis nula de que las variables presenten raíz unitaria. Se trabaja con variables estacionarias.

En la tabla 2 se pueden encontrar los principales estadísticos descriptivos asociados a las variables exógenas:

Table 2: Estadísticos descriptivos Variables Exógenas

Variable	T	Media	Mediana	Std.	K	Skew.	JB	ADF	Δ
Stoxx50	398	5.65	5.6498	120.19	5.4237	-0.6487	0.001***	0.001***	Si
Nasdaq	398	19.74	19.735	168.63	8.7797	-0.4810	0.001***	0.001***	Si
Dax	387	24.38	24.376	342.42	5.6952	-0.7049	0.001***	0.001***	Si
EUnivel	242	1.55	1.5503	1.5952	1.971	0.3179	0.0054***	0.0084***	No
USA10Yspread	398	-0.0033	-0.0033	0.282	3.5979	0.4609	0.0016***	0.001***	Si
USA5Yspread	398	1.1969	1.1969	0.8534	2.3564	0.0076	0.0336***	0.0463**	No
USA2Yspread	398	0.6202	0.6202	0.5250	3.5924	0.4918	0.0012***	0.0040***	No
EU10Yspread	255	-0.0063	-0.0063	0.2318	5.7905	0.6679	0.001***	0.001***	Si
EU5Yspread	242	-0.0078	-0.0077	0.2452	5.9087	0.6449	0.001***	0.001***	Si
EU2Yspread	242	-0.0048	-0.0048	0.2359	6.4009	0.4008	0.001***	0.001***	Si
Ale10Yspread	313	0.0023	0.0023	0.2736	9.0472	1.3359	0.001***	0.001***	Si
Ale5Yspread	313	0.0036	0.0036	0.2788	7.4924	1.1687	0.001***	0.001***	Si
Ale2Yspread	313	0.0036	0.0036	0.2637	8.6867	1.0846	0.001***	0.001***	Si
Fra10Yspread	374	0.0013	0.0013	0.4800	173.09	10.873	0.001***	0.001***	Si
Fra5Yspread	374	0.5823	0.5823	0.9563	30.167	-3.5811	0.001***	0.001***	No
Fra2Yspread	374	0.2546	0.2546	1.726	4.9004	-0.7188	0.001***	0.0033***	No
Ita10Yspread	290	0.0024	0.0024	0.4717	19.52	1.9028	0.001***	0.001***	Si
Ita5Yspread	290	0.0024	0.0024	0.4577	16.752	1.3009	0.001***	0.001***	Si
Ita2Yspread	290	0.0019	0.0019	0.4624	16.503	0.7811	0.001***	0.001***	Si
Eur/Doll	398	8.79e-06	8.79e-06	0.0354	4.5779	-0.4318	0.001***	0.001***	Si
BRENT	398	0.0566	0.0566	4.8263	9.6056	-1.0499	0.001***	0.001***	Si
Gold	398	2.8288	2.8288	41.479	8.2349	-0.0242	0.001***	0.001***	Si
P/C ratio ²⁷	120	1.5681	1.4581	0.7294	8.2101	1.7758	0.001***	0.0029***	No

²⁷Finalmente no se considera el ratio entre los volúmenes de las opciones put y call dado que la longitud de la muestra no permite distinguir entre submuestras para el “in y out of sample”.

Variable	T	Media	Mediana	Std.	K	Skew.	JB	ADF	Δ
VIX	362	0.0897	0.0897	4.3794	8.4166	1.112	0.001***	0.001***	Si
Baa-Aaa	397	-0.0013	-0.0013	0.10056	28.625	1.5933	0.001***	0.001***	Si
M1	396	2.06e+10	2.06e+10	4.08e+10	5.1678	0.4873	0.001***	0.001***	Si
M2	396	2.63e+10	2.63e+10	4.08e+10	4.292	0.8549	0.001***	0.001***	Si
M3	396	2.77e+10	2.77e+10	3.94e+10	4.8737	0.4122	0.001***	0.001***	Si
IPC	263	8.44e-19	8.44e-19	0.2445	4.6439	-0.2167	0.001***	0.001***	Si
Construcción	300	-0.0433	-0.0433	2.3447	10.301	-0.0012	0.001***	0.001***	No
Importaciones	253	4.24e+08	4.24e+08	9.26e+09	2.9472	0.0865	0.5	0.001***	Si
Desempleo	261	-0.0123	-0.0123	0.0899	5.1824	0.8794	0.001**	0.0301**	Si

Notas: Se recogen los principales estadísticos descriptivos de las variables consideradas. La columna “T” indica el número de observaciones en la serie de cada variable. La columna JB recoge el p-valor para el resultado del test de normalidad de Jarque Bera. Un p-valor bajo implica que se puede rechazar que la distribución de la variable sea una normal. La columna del ADF recoge el resultado del test aumentado de Dickey-Fuller. Un p-valor bajo implica que se puede rechazar la hipótesis nula de raíz unitaria en la serie de datos de la variable. Es decir, la variable sería estacionaria. Como es habitual, ***,** y * indican significatividad estadística al 1%, 5% y 10% respectivamente. La columna de Δ indica si se ha tomado primeras diferencias de la variable original para obtener una variable estacionaria.

- *Índices:* Dada la globalización de la economía en las últimas décadas, las bolsas internacionales están actualmente muy relacionadas. Los avances tecnológicos han permitido eliminar las barreras que los inversores podían tener para invertir en instrumentos ajenos a sus economías de residencia. Cabe analizar si los rendimientos de las acciones de las principales empresas de los países económicamente más vinculados a la Zona Euro permiten predecir las tendencias de los títulos de las principales empresas en este área. Se tendrán en cuenta los índices NASDAQ (estadounidense) y DAX (alemán).
 - ΔIPC : La variación del índice de precios de consumo (IPC) estará, teóricamente, estrechamente correlacionada con los rendimientos de las acciones. Existen distintas teorías que tratan de explicar la relación negativa observada históricamente entre la inflación y el rendimiento de los activos financieros. Destaca la hipótesis “proxy” de Fama²⁸, que argumenta que la relación entre las variables analizadas (Inflación y rendimientos de los stocks) evidencia las relaciones negativa entre tasa de inflación y actividad económica esperada, y positiva entre actividad esperada y rendimiento de los activos. De este modo, a mayor inflación, menor rendimiento futuro de los activos, pero por vía de las interrelaciones con la actividad esperada.
- Otra explicación posible es la teoría de la variabilidad, que defiende una relación causal directa entre inflación y rendimiento de los activos, basada en que la inflación afectaría negativamente a las expectativas de los agentes económicos, haciendo disminuir la actividad económica.
- $\Delta Desempleo$: La tasa de desempleo influirá también negativamente sobre los precios de las acciones. Menor número de empleados implica que la población general tendrá menores ingresos y, en consecuencia, los beneficios y los rendimientos de las acciones de las empresas serán menores.

²⁸Existen otras teorías relativas a esta relación, algunas de las cuales defienden la existencia de una interconexión asimétrica dependiente del contexto, representado por el estado de la economía.

- *Construcción*: El nivel de producción en el sector de la construcción se comportará de forma opuesta. La variable es especialmente interesante por el efecto de causación que la burbuja de los 2000 tuvo sobre la crisis de 2008. Sin embargo, por la evolución de la economía en los años recientes, se puede esperar que la variable presente capacidad de ajuste in sample, pero que su capacidad predictiva sea limitada.
- *ΔImportaciones*: Las importaciones desde el extranjero influirían a priori de forma positiva sobre los precios de los stocks. El aumento de las importaciones de productos procedentes de economías extranjeras indicará que en Europa se demandan productos (o cantidades de productos) que no se pueden producir en la economía doméstica. Un aumento de las importaciones indicará que la demanda de las economías europeas está creciendo a un nivel superior al de la oferta.
- *ΔTC*: El tipo de cambio del euro respecto al dólar provocará dos efectos sobre las acciones. Si el euro se deprecia, la demanda de productos europeos desde el extranjero aumentará, dado que los productos producidos en la Euro Zona se abaratarían relativamente para los consumidores americanos. Los beneficios de las empresas europeas crecerían. También se produciría un impacto en los costes de las empresas locales al aumentar relativamente el coste de las importaciones de los inputs necesarios para los procesos productivos. Se podría esperar que el efecto del aumento de la demanda extranjera domine sobre el aumento de los costes de las importaciones, ya que el aumento en los costes se podría trasladar al precio del producto final.
- *Curvas de tipos*: Los tipos de interés se relacionan con el precio de las acciones de varios modos. Para empezar, serán la tasa en la cual se basan los modelos de valoración estándar para activos financieros. Los beneficios futuros de una empresa se descuentan a la tasa “libre de riesgo” cuando se calcula el valor actual de una compañía. Por ello, las variaciones en los tipos de referencia tendrán una influencia directa en la valoración de las empresas. Las variaciones en el tipo de interés de una economía afectarán también al coste de financiación de las empresas, y por tanto, indirectamente, a su valor. La variación en el tipo de interés libre de riesgo también influye en el valor de las empresas por el efecto del coste de oportunidad. Todas estas relaciones implican un efecto negativo de la curva de tipos sobre el mercado de acciones. Pero la curva de tipos también recogerá las expectativas que los inversores tienen con respecto a la evolución futura de la economía. El hecho de que la rentabilidad de los bonos a largo plazo sea menor o igual que la de los activos de corto plazo indicará que los inversores esperan una recesión en algún periodo futuro. Se puede emplear la pendiente entre distintos vencimientos de la curva como predictores de tendencias bajistas.

Los tres primeros efectos mencionados se podría reflejar mediante el nivel de la curva de tipos. El cuarto y último efecto lo recogería la pendiente. La capacidad predictiva de la pendiente de la ETTI se ha contrastado en numerosos trabajos previos. En este caso se emplearán diferentes curvas, por las características heterogéneas de las economías de la Zona Euro. Se consideran las curvas de Alemania y Francia, por ser las principales potencias europeas, y la de Italia, en representación de los ciclos de las economías mediterráneas²⁹. Se emplea la curva del conjunto de la eurozona y, en base a las conclusiones de [Fernández, Fernández & Sosvilla (2014)], que contrastan que la ETTI estadounidense tiene capacidad predictiva sobre las tendencias del mercado español, se puede esperar que presente también cierta capacidad predictiva sobre las tendencias de las acciones europeas en su conjunto. Se considera por tanto también la ETTI de Estados Unidos.

²⁹Como se comprueba durante la crisis de 2012, las economías mediterráneas se comportan de forma diferente a los países continentales.

- $\Delta M3$: El agregado monetario refleja la cantidad de liquidez en la economía europea. Al aplicarse medidas expansivas de política monetaria debería aumentar el nivel de consumo. Por tanto, incrementos en el nivel de la variable M3 llevarían a aumentos en los beneficios esperados de las empresas y, con ellos, a rendimientos crecientes. Las medidas de política monetaria inciden también en las expectativas del mercado.
- “*Default Yield Spread*”: La variación en la diferencia entre los rendimientos de los bonos no gubernamentales Baa-Aaa recogería el cambio en el riesgo de “default” percibido en el mercado de bonos. Si los inversores de estos mercados se vuelven más aversos al riesgo, exigirán un rendimiento mayor a los bonos menos seguros. La demanda de bonos calificados Baa caerá. Con ella lo hará el precio y, por tanto, su “yield” aumentará. En las mismas circunstancias, la demanda de los bonos más seguros aumentaría (“Flight to Quality”), y por ello el yield de los bonos calificados Aaa caerá o se mantendrá. Las primas por riesgo que se asignan a los activos de “equity” también dependerán de la aversión al riesgo de los inversores. Se exigirán mayores primas cuando aumente la aversión general del mercado³⁰. La variable será indicador de las condiciones más generales de los negocios que emiten los bonos de referencia. Por ello, debería poder capturar las variaciones de largo plazo en el ciclo bursátil.
- ΔOro : El oro es uno de los activos refugio de referencia. Es de esperar que su rendimiento esté relacionado negativamente con el del mercado. [McDermott & Baur (2009)] encuentran que para el mercado europeo, canadiense y estadounidense, el oro se comporta como un activo refugio, y que los beneficios de la inversión en oro son más destacables cuando se dan sucesos con probabilidades de ocurrencia menores del 1%. Sin embargo, los autores encuentran que el oro es un activo refugio débil. Es decir, se correlaciona negativamente con el mercado cuando las caídas de éste son ligeras, pero bajo condiciones extremas la correlación pasa a ser positiva. Se podrá esperar por tanto una relación negativa pero asimétrica entre el rendimiento del oro y la tendencia del mercado.
- $\Delta Brent$: El petróleo es uno de los principales inputs de la industria de las economías desarrolladas, y por ello el aumento en su precio debería influir negativamente sobre los beneficios esperados. Las crisis de los años 70 son ejemplos de recesiones causadas por variaciones en los precios del petróleo que generaron caídas en los índices bursátiles mundiales.
- ΔVIX : El VIX reflejará la volatilidad implícita en las opciones de un mercado. Uno de los hechos estilizados de los rendimientos de las acciones es el conocido como efecto leverage. Éste hace referencia al hecho de que la llegada de noticias negativas al mercado ocasiona aumentos en la volatilidad. La llegada de noticias positivas generaría reducciones de volatilidad y rendimientos positivos. El efecto sería además asimétrico. Es decir, la llegada de malas noticias generaría un aumento de la volatilidad y un efecto negativo sobre los rendimientos de mayor magnitud que el efecto positivo que generarían las noticias favorables.

En periodos de elevada volatilidad (y por tanto cuando el VIX toma valores elevados), se esperará experimentar rendimientos negativos en los mercados. [Hood & Malik (2013)] concluyen tras evaluar la capacidad de diferentes metales preciosos como activos de cobertura, que el índice VIX se puede considerar una cobertura fuerte, a diferencia del oro. De este modo, se puede esperar una relación positiva entre la variación en el índice VIX y la probabilidad de que se experimente una tendencia bajista.

³⁰Algunos autores plantean también una relación inversa entre la evolución de esta variable y la variación en la liquidez en el mercado. Cuando aumenta este spread, aumenta la aversión y, con ello, se reduce la liquidez.

B Anexo 2. Algoritmo de Bry & Boschan (BBA)

El algoritmo de Bry & Boschan se desarrolla originalmente para detectar los ciclos económicos y, en particular, para identificar las expansiones y recesiones de la economía en base a la definición de la “National Bureau of Economic Research” (NBER). El algoritmo, mediante la búsqueda de mínimos y máximos locales, encuentra patrones en las series temporales y, en base a ciertas reglas predeterminadas, determina si una sucesión de datos constituye un periodo alcista o bajista.

Originalmente, Bry & Boschan definen el algoritmo de modo que se siguen los siguientes pasos:

1. Se suavizan los datos. Se eliminan “outliers”, y se aplican logaritmos a la serie.
2. Con una ventana de 12 meses (los seis previos y los seis posteriores al dato a analizar), se encuentran los máximos y mínimos locales preliminares de la serie a analizar. De este modo, un pico (“peak”) se dará cuando la serie encuentra un máximo local, y un valle (“trough”) se dará cuando en la serie se observe un mínimo local. Es decir, el algoritmo encontrará un pico en el mes “t” si:

$$p_{t-i} < p_t, \text{ con } i = 1, \dots, 6 \text{ y}$$

$$p_{t+i} < p_t \text{ con } i = 1, \dots, 6$$

y un valle si:

$$p_{t-i} > p_t, \text{ con } i = 1, \dots, 6 \text{ y}$$

$$p_{t+i} > p_t \text{ con } i = 1, \dots, 6$$

3. Se eliminan los máximos y mínimos locales consecutivos, manteniendo el mayor de los picos o el menor de los valles.
4. Se eliminan las fases cuya duración sea menor que 5 meses, en base a la definición de recesión planteada por el NBER.
5. Se fuerza a que el ciclo completo (una fase expansiva y una recesiva) dure al menos 15 meses, también en base a la definición de recesión aceptada por el NBER.
6. Elimina los puntos de inflexión a 6 meses de los extremos de la muestra. Se eliminan los picos y valles en los extremos de la serie que sean menores o mayores, respectivamente, que el primer o último dato de la muestra.

Sin embargo, en éste trabajo se trata con activos financieros como variable dependiente. Sus características intrínsecas, así como las de las tendencias del mercado que se desean captar con el procedimiento, difieren de las que originalmente buscaban reflejar Bry & Boschan. Por ello se modificará el algoritmo para adaptarlo a la definición de mercado bajista mayoritariamente aceptada, y a los métodos empleados en la literatura previa.

Pese a la falta de un consenso claro, en general se podría considerar que un mercado es bajista o “bearish” si la caída acumulada supera el 20%. Si la caída es mayor al 10 pero no supera el 20%, se suele considerar que se ha producido simplemente una corrección del mercado. Es decir, no se llega a abandonar el mercado alcista. Para un mercado “bullish” se aplica una definición equivalente, considerando que el mercado es alcista si el crecimiento acumulado supera el 20%.

Históricamente se experimentan al menos una corrección al año, y su duración ronda los dos meses. Por tanto, en el algoritmo se restringirá un mercado bajista a aquel periodo entre un máximo y un mínimo local

que dure, al menos, 3 meses³¹. De este modo, se modifica el punto 4 eliminando aquellas fases cuya duración sea inferior a 3 meses, en lugar de 5. Se definirá también un mercado alcista como aquel periodo entre un mínimo y un máximo local, con una duración mínima igual a la del mercado bajista. Estas definiciones de mercados alcista y bajista implica que se mantiene el punto 3.

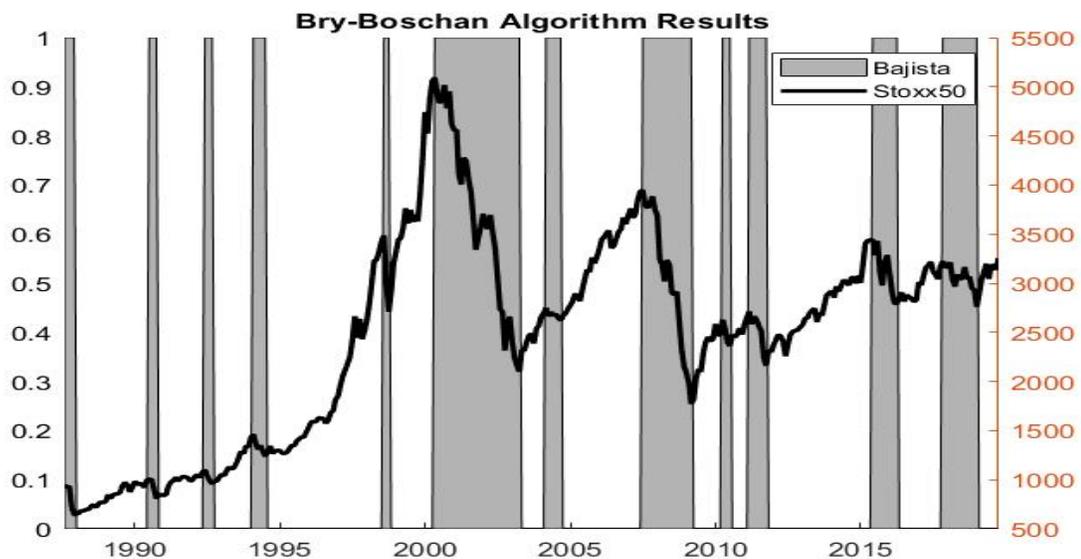
Al igual que en la literatura previa, se mantiene la ventana de seis meses. Se eliminan los picos/valles cercanos a los extremos de la muestra, pero en este caso se considera una distancia a dichos extremos del mismo tamaño que la duración mínima considerada para un mercado bajista (3 meses). Además, no se eliminan los “outliers” dado que son características esenciales de los periodos “bullish” y “bearish”.

Dado que se emplea una ventana de 6 meses para determinar los puntos de inflexión (picos/valles), la variable dependiente perderá necesariamente 12 observaciones al aplicar el algoritmo y transformar en discreta la variable continua original.

Por otro lado, es necesario destacar que la naturaleza del método introduce un sesgo de “retrospectiva” en la definición de las tendencias. Es decir, para concluir si la tendencia bursátil que subyace a una observación es alcista o bajista es necesario conocer si el dato a analizar se sitúa entre máximos y mínimos locales, o viceversa. Se requiere conocer por tanto lo que ocurre en el mercado en los meses posteriores a la observación analizada. Por ello, dado que los modelos estimados no cuentan con la ventaja que se incorpora en el método de Bry & Boschan, las probabilidades que se obtengan serán ruidosas y será necesario determinar un umbral a partir del cual se considera que la probabilidad de mercado bajista es lo suficientemente elevada. Para cada modelo se define mediante la curva ROC (4.4) un threshold diferente que permita transformar las variables predichas (contínuas en el dominio [0,1]) en variables dicotómicas. Es decir, se busca definir un límite que permita pasar de probabilidades predichas, continuas, a probabilidades “teóricas”, discretas.

Los resultados de aplicar el procedimiento se observan en la siguiente figura:

Figure 2: Transformación Variable Dependiente



Nota: Resultados de la transformación de la variable dependiente en una variable dicotómica mediante el algoritmo de Bry&Boschan. Para volver a la sección correspondiente a la figura: 3.

³¹Ésto históricamente garantiza que se captura una caída/subida del mercado de un 20%.

C Anexo 3. Algoritmo Genético (GA)

En general, se pueden distinguir dos procedimientos para modelizar la selección de indicadores: En primer lugar se podría aplicar un procedimiento conocido como modelización densa, que consistiría en considerar factores de riesgo obtenidos mediante métodos de descomposición espectral, como podría ser la descomposición en componentes principales. Este procedimiento por sí mismo no permitiría realizar análisis estructural ni predicciones. Es decir, no permitiría encontrar una relación entre variables económico-financieras para realizar seguimiento, a no ser que se logre identificar los factores subyacentes con alguna variable medible. La alternativa que se empleará en este trabajo será definir relaciones económicas teóricas y, partiendo de éstas, tratar de encontrar la combinación que mejor replique los factores de riesgo en su conjunto. Para garantizar que se obtiene la mejor configuración posible de los modelos dadas las variables explicativas disponibles, se puede recurrir a algoritmos de optimización.

Los algoritmos genéticos son, siguiendo la definición de [Acosta & Fernández (2007)], técnicas de optimización basadas en la selección natural que tratan de sobreponerse a las limitaciones habituales que se encuentran los métodos clásicos. Entre estas limitaciones se pueden destacar la no continuidad o diferenciabilidad de la función de pérdida y la convergencia a óptimos locales. Estos algoritmos logran solventar tales dificultades mediante el concepto de ensayo y error y, sobre todo, mediante la introducción de aleatoriedad en el proceso. Esto último garantiza que el procedimiento abandona los posibles óptimos locales y, en este caso, que se tengan en cuenta todas las variables en el proceso de optimización. El concepto clave del procedimiento será la supervivencia. Sobreviven los individuos con las mejores capacidades para adaptarse al problema planteado, y se reproducen dando lugar a una nueva generación de individuos con características relacionadas con las suyas. La segunda generación de individuos sustituirá a sus predecesores, repitiendo el proceso. Se plantea de este modo un proceso recursivo de selección, cruce o reproducción y mutación que, como se explicará a continuación, durará hasta que se alcanza la configuración genética correspondiente al mejor individuo posible en base a un criterio de selección y a un criterio de parada.

Los algoritmos genéticos se basan en el proceso darwiniano de selección natural. Según esta teoría, sobreviven y se reproducen aquellos individuos cuyos genes se asocian a las mejores capacidades para mantener a la especie. En biología, los individuos o cromosomas estarán formados por genes, que serán dígitos binarios que definen sus características y, en consecuencia, sus probabilidades de supervivencia. En este caso, los cromosomas serán los diferentes modelos, los genes harán referencia a las variables consideradas, y las características de los modelos que determinan las “probabilidades de supervivencia” serán simplemente el ajuste de los modelos a los datos.

En la búsqueda del mejor modelo posible, los individuos o cromosomas serán, para el algoritmo, los distintos conjuntos de variables consideradas. Se mantendrán aquellos que presenten un mejor “fit” a la variable endógena. Es decir, se buscará la combinación de variables que devuelva el mejor valor de un determinado criterio de selección. En este caso la selección de variables se fundamentará en el ajuste de los modelos a la submuestra mantenida para el ajuste “in sample” de los parámetros, y el criterio de bondad de ajuste que se emplea es el estadístico de Akaike. El motivo por el que se basa la selección de modelos en este criterio es el hecho de que, como se explica en 4.1.1, el AIC penaliza a los modelos por el número de variables que incorporan. Claramente, el modelo que minimiza el sumatorio de los cuadrados de los errores será el que seleccione todas las variables. Dado que el objetivo del trabajo es predecir el comportamiento futuro del mercado, habrá que seleccionar una función de pérdida que penalice la complejidad del modelo, buscando evitar la sobreparametrización.

El procedimiento recursivo que seguirá el algoritmo de selección será el siguiente:

1. Se genera aleatoriamente una población inicial de modelos a analizar. Ésta consistirá en un número determinado de combinaciones de variables explicativas. En este caso se considerarán 20 combina-

ciones de variables en cada periodo, en base a los resultados de [Papadamou & Stephanides (2007)], que concluyen que poblaciones de dicho tamaño logran el mejor balance entre coste computacional y precisión del resultado. Es decir, se genera una población de 20 individuos (modelos, o combinaciones de variables explicativas) que se diferenciarán en función de las variables que emplean para tratar de explicar la variable dependiente. Sin embargo, no se restringe a que los modelos generados sean siempre distintos entre sí. Los modelos se representarán como vectores de longitud 28 (igual al número de variables exógenas consideradas), formados por dígitos binarios. Así, cada individuo se representa del siguiente modo: $individuo_i = [x_1, \dots, x_k]$ donde $x_i = 0, 1$ y $k = 28$. Se podría considerar que los cromosomas iniciales seleccionados aleatoriamente actúan como semillas de las que parte el procedimiento de búsqueda y que, por introducir aleatoriedad en la selección, permiten reducir significativamente el rango de combinaciones de variables a analizar.

2. Se analizan las características de cada uno de los individuos de la población. Es decir, se ajustan a los datos “in sample” los 20 modelos considerados en la primera fase, y se obtiene el estadístico de Akaike para cada uno de ellos.
3. Se mantienen los 2 mejores cromosomas en base a sus características. Éstos serán los que, emulando el proceso darwiniano de selección natural, se reproduzcan. Según la teoría de la selección natural, sobreviven y se reproducen aquellos individuos cuyos genes se asocian a las mejores capacidades para mantener a la especie. En el problema planteado, se seleccionan los dos modelos que contengan aquellas variables que lleven a un menor valor del estadístico de Akaike.
4. Los individuos seleccionados se reproducen generando 20 hijos con unos genes estrechamente relacionados con los suyos. En el problema econométrico, se cruzan los modelos seleccionados dando lugar a 20 combinaciones de variables explicativas relacionadas con las combinaciones que seleccionaban los dos modelos que habían superado la criba de la fase 3. El proceso se conoce como crossover, y en él se introduce aleatoriedad. Se permitirá que se crucen los últimos “i” genes³² o variables consideradas por cada uno de los dos mejores modelos anteriores, según el AIC.
5. Además del cruce de los modelos, también se realiza una mutación en los genes de los “hijos”. Es decir, se considerará que con una probabilidad del 40% fijada de antemano³³, podría variar alguna de las variables explicativas seleccionadas tras el “crossover” de los 2 modelos previos. De este modo se introducen shocks aleatorios sobre los cromosomas. Se podría eliminar una de las variables o añadir otra que no se estaba considerando en los modelos precedentes. Así, se logra que las variables que podrían no haberse considerado en las combinaciones iniciales (aleatorias), se tengan en cuenta en posteriores generaciones. Se consigue que el algoritmo no converja a un mínimo local de la función de pérdida. Por otro lado, el introducir una probabilidad de mutación demasiado elevada podría dañar las buenas estructuras o combinaciones de genes. Por ello se selecciona una probabilidad de mutación relativamente baja, del 40%. Cabe destacar que además del 40% de probabilidad de que se produzca mutación en alguno de los 28 elementos definitorios del nuevo modelo respecto a lo que se había obtenido con el cruce de los dos óptimos de la fase anterior, se introducirá aleatoriedad en la propia

³²Se divide el vector binario que representa cada cromosoma en dos subvectores, uno de los cuales tendrá tamaño “i”, y el otro, 1 - “i”. Los dos subvectores de la derecha (es decir, los subvectores de tamaño “i” correspondientes uno a cada “padre”) se intercambian entre los vectores originales, obteniéndose así dos nuevos cromosomas.

³³[Papadamou & Stephanides (2007)] concluyen también que la probabilidad de mutación no influye de forma decisiva en la eficiencia computacional del proceso, ni en la precisión de la solución. Los autores remarcan que probabilidades del 40 o 50% arrojan resultados equiparables.

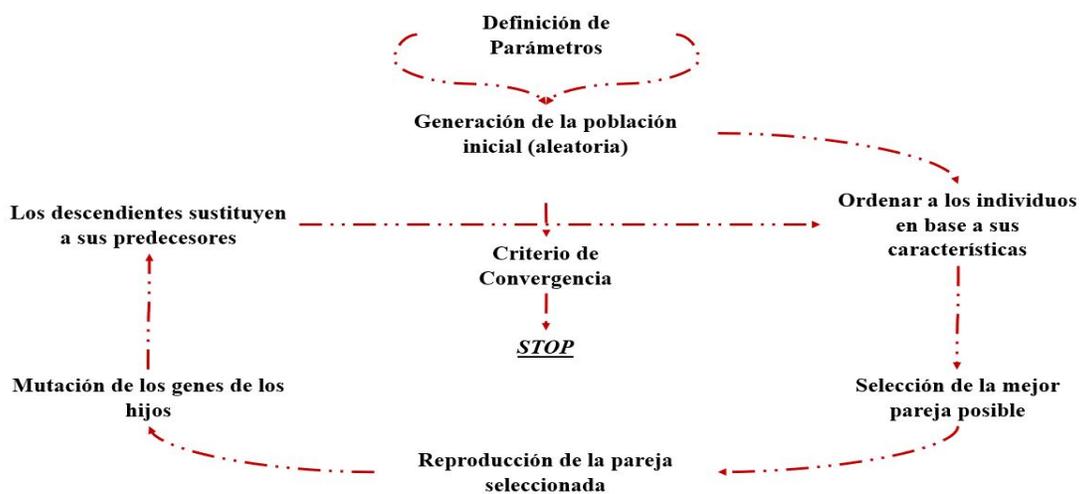
mutación al permitir, mediante la simulación de una variable uniforme estándar, que varíe cualquiera de los genes o variables seleccionadas que definen el nuevo modelo³⁴.

- La nueva generación de 20 “hijos” reemplaza a la generación precedente. Se repite el proceso desde el punto 2. Es decir, los nuevos modelos resultantes de la selección, cruce y mutación de los dos individuos seleccionados en el punto 3, sustituyen a las anteriores 20 combinaciones de variables consideradas.

Dado que se trata de un procedimiento recursivo, será necesario incorporar criterios de parada. Se definirán dos circunstancias, que son las que habitualmente se emplean en la literatura: que el procedimiento encuentre una combinación óptima de variables o que, habiendo probado un elevado número de combinaciones, no logre converger a una combinación óptima. El primer criterio de parada será que durante 15 generaciones o iteraciones no se logre una combinación de variables nueva que mejore el “fit” de los dos modelos seleccionados en la generación anterior. El otro criterio será simplemente que no se superen las 1000 iteraciones, el cual en general (y en particular en este estudio) será un número suficientemente elevado para la convergencia.

El procedimiento se resume en el siguiente esquema:

Figure 3: Esquema Algoritmo Genético



Nota: Para volver a la sección correspondiente a la figura: 4.1

Una vez seleccionado el conjunto óptimo de variables para una especificación determinada se eliminan del modelo aquellas que no sean estadísticamente significativas. Cabe esperar que en la especificación óptima el número variables no significativas sea relativamente reducido ya que, de no serlo, el efecto de la variable sobre el estadístico AIC no debería ser positivo (por la penalización que el criterio impone al número de parámetros).

³⁴Una variante simplificadora del proceso consistiría en considerar que únicamente pueda variar uno de los 28 elementos que definen el cromosoma, determinado arbitrariamente.

Anexo 4. Gráficos y Tablas.

Tablas

Table 4: Ajuste Probit Exógenos

(a) Univariables							(b) Univariables						
$P_{t-k}(S_t = 1) = \Phi(\alpha + \beta x_{t-k})$							$P_{t-k}(S_t = 1) = \Phi(\alpha + \beta x_{t-k})$						
		In Sample		Out Of Sample					In Sample		Out Of Sample		
Variable	k	R^2	AIC	RMSE	QPS	pseudo- R^2	Variable	k	R^2	AIC	RMSE	QPS	pseudo- R^2
EU Nivel	1	0.2660	169.3	0.5884	0.3462	-0.5448	VIX	1	0.0187	225.02	0.4895	0.2396	-0.0692
	3	0.333	152.37	0.5967	0.3561	-0.589		3	0.0067	224.94	0.4934	0.2434	-0.0864
	6	0.3098	154.32	0.5955	0.3547	-0.5826		6	0.0012	221.52	0.4880	0.2382	-0.0629
	12	0.1828	174.03	0.5806	0.3370	-0.5041		12	0.0044	211.16	0.4871	0.2373	-0.0589
US 5Y spread	1	0.0215	224.36	0.4886	0.2388	-0.0655	Baa-Aaa	1	0.0344	221.48	0.4741	0.2247	-0.0029
	3	0.0118	223.82	0.4895	0.2396	-0.0691		3	0.0245	220.99	0.4797	0.2301	-0.0267
	6	0.0011	221.56	0.4898	0.2399	-0.0704		6	0.0003	221.74	0.4883	0.2385	-0.0642
	12	0.0791	195.61	0.5013	0.2513	-0.1213		12	2.44e-05	212.07	0.4857	0.2359	-0.0527
EU 5Y spread	1	0.0025	228.67	0.4934	0.2435	-0.0865	IPC	1	0.0033	228.73	0.4909	0.2410	-0.0755
	3	0.0019	226.02	0.4899	0.2400	-0.0712		3	0.0146	223.19	0.4829	0.2332	-0.0407
	6	0.0163	218.24	0.4898	0.2399	-0.0706		6	0.0317	214.89	0.4993	0.2493	-0.1127
	12	0.0072	210.57	0.4852	0.2355	-0.0508		12	0.0003	212.01	0.4854	0.2356	-0.0514
Fra 5Y spread	1	0.0137	226.14	0.5025	0.2525	-0.1269	G.A.	1	0.6539	97.943	0.6005	0.3605	-0.6089
	3	0.0064	225.02	0.495	0.2450	-0.0934		3	0.6468	90.554	0.6089	0.3707	-0.6543
	6	0.0075	220.15	0.4885	0.2386	-0.065		6	0.5705	107.54	0.6125	0.3751	-0.6739
	12	0.0749	196.48	0.4822	0.2325	-0.0376		12	0.4581	128.76	0.5998	0.3598	-0.6057

Nota: La fórmula correspondiente al modelo con variables seleccionadas mediante el algoritmo genético, ya que se emplean múltiples variables exógenas, será la siguiente:

$$P_{t-k}(S_t = 1) = \Phi\left(\alpha + \sum_{i=1}^n \beta x_{i,t-1}\right)$$

Para volver a la sección correspondiente a las tablas: 4.2

Table 5: Probit Autorregresivos 1

$$P_{t-k}(S_t = 1) = \Phi(\alpha + \delta_{t-k}S_{t-k} + \gamma_{t-k}\sum_{i=12-k}^{t-2-k}\Delta S_t + \beta x_{t-k})$$

		In Sample				Out of Sample				
Variable	k	α	δ	γ	β	R^2	AIC	RMSE	QPS	pseudo- R^2
EU Nivel	1	-0.6175 (0.3692)	-0.0007 (0.0378**)	0.5135 (0.4731)	1.1337 (3.31e-07***)	0.2899	167.93	0.6005	0.3606	-0.609
	3	-1.9726 (0.0072***)	-0.0002 (0.6038)	0.4330 (0.5577)	1.0133 (7.27e-06***)	0.3347	156	0.5992	0.3591	-0.6024
	6	-3.6795 (2.64e-06***)	0.0008 (0.034**)	-0.8398 (0.2590)	0.5159 (0.0114**)	0.3307	153.78	0.5454	0.2975	-0.3276
	12	-4.6965 (5.02e-09***)	0.0016 (1.74e-05***)	-2.1128 (0.0062***)	-0.1915 (0.3481)	0.2771	158.42	0.5157	0.2659	-0.1867
US 5Y spread	1	-6.7779 (1.25e-10***)	0.0018 (7.56e-10***)	-1.6327 (0.0088***)	1.066 (6.75e-08***)	0.3154	162.19	0.4687	0.2197	0.0196
	3	-12.446 (1.65e-10***)	0.0035 (5.58e-10***)	-1.9809 (0.0098***)	1.7739 (1.56e-09***)	0.4975	119.78	0.465	0.2162	0.0351
	6	-9.4742 (1.33e-09***)	0.0027 (1.81e-09***)	-1.8598 (0.0098***)	1.0442 (8.86e-06***)	0.4199	134.34	0.4356	0.1897	0.1534
	12	-3.9047 (6.21e-07***)	0.0013 (2.04e-08***)	-1.7826 (0.0063***)	-0.1159 (0.4672)	0.2757	158.71	0.4861	0.2363	-0.0544
EU 5Y spread	1	-3.1244 (2.43e-08***)	0.001 (4.21e-08***)	-1.9064 (0.0009***)	-0.8053 (0.0359**)	0.1645	196.17	0.4359	0.19	0.1522
	3	-3.1244 (2.43e-08***)	0.001 (4.21e-08***)	-1.9064 (0.0009***)	-0.8053 (0.0359**)	0.1645	196.17	0.4359	0.19	0.1522
	6	-3.9052 (4.09e-10***)	0.0013 (9.06e-10***)	-1.5955 (0.0073***)	-0.3001 (0.4539)	0.2263	180.11	0.4129	0.1705	0.2391
	12	-4.653 (2.07e-11***)	0.0015 (6.91e-11***)	-1.896 (0.0031***)	0.2376 (0.5926)	0.3012	160.2	0.4174	0.1742	0.2225

Table 6: Probit Autorregresivos 2

$$P_{t-k}(S_t = 1) = \Phi(\alpha + \delta_{t-k}S_{t-k} + \gamma_{t-k}\sum_{i=12-k}^{t-2-k}\Delta S_i + \beta x_{t-k})$$

		In Sample				Out of Sample				
Variable	k	α	δ	γ	β	R^2	AIC	RMSE	QPS	pseudo- R^2
FR 5Y spread	1	-2.7304 (2.29e-06***)	0.0009 (3.69e-07***)	-1.6886 (0.0028***)	-0.0862 (0.5886)	0.1464	200.23	0.4403	0.1938	0.135
	3	-3.9203 (1.80e-09***)	0.0012 (6.15e-10***)	-1.5735 (0.0081***)	0.0763 (0.6423)	0.2248	180.44	0.4087	0.1670	0.2547
	6	-4.9039 (3.59e-11***)	0.0015 (2.04e-11***)	-1.9601 (0.0023***)	0.1184 (0.4909)	0.3020	160.02	0.415	0.1722	0.2314
	12	-3.7893 (4.43e-08***)	0.0013 (2.72e-09***)	-1.7327 (0.0071***)	-0.5016 (0.0086***)	0.3074	152.11	0.4831	0.2333	-0.0413
VIX	1	-2.942 (3.27e-08***)	0.0009 (4.83e-08***)	-1.8778 (0.0010***)	0.0452 (0.02321**)	0.1687	195.22	0.4283	0.1835	0.1813
	3	-3.8223 (1.66e-10***)	0.0012 (3.48e-10***)	-1.6134 (0.0066***)	0.0301 (0.1268)	0.2345	178.28	0.4124	0.1701	0.2411
	6	-4.7133 (4.82e-12***)	0.0015 (1.73e-11***)	-1.9435 (0.0023***)	0.0185 (0.3599)	0.3037	159.64	0.413	0.1705	0.239
	12	-4.2857 (5.38e-11***)	0.0013 (1.86e-10***)	-1.6585 (0.0092***)	-0.0199 (0.3354)	0.2775	158.33	0.4828	0.2331	-0.0400
Baa-Aaa	1	-2.7934 (1.53e-07***)	0.0009 (2.46e-07***)	-1.7611 (0.0023***)	2.0785 (0.0205**)	0.1726	194.35	0.4197	0.1762	0.2138
	3	-3.7524 (4.39e-10***)	0.0012 (9.89e-10***)	-1.5181 (0.0113***)	1.5327 (0.0671*)	0.2395	177.17	0.4047	0.1638	0.2691
	6	-4.7657 (6.2e-12***)	0.0015 (2.23e-11***)	-1.9163 (0.0025***)	-0.5111 (0.5342)	0.3015	160.13	0.4194	0.1759	0.2151
	12	-4.2855 (4.15e-11***)	0.0013 (1.40e-10***)	-1.6914 (0.0074***)	-0.3380 (0.6772)	0.2740	159.05	0.482	0.2323	-0.0367

Table 7: Probit Autorregresivos 3

$$P_{t-k}(S_t = 1) = \Phi(\alpha + \delta_{t-k}S_{t-k} + \gamma_{t-k} \sum_{i=12-k}^{t-2-k} \Delta S_i + \beta x_{t-k})$$

		In Sample					Out of Sample			
Variable	k	α	δ	γ	β	R^2	AIC	RMSE	QPS	pseudo- R^2
IPC	1	-2.8566 (7.09e-08***)	0.0009 (1.11e-07***)	-1.8568 (0.0016***)	0.3999 (0.3654)	0.1489	199.68	0.4330	0.1875	0.1631
	3	-3.869 (2.14e-10***)	0.0012 (4.78e-10***)	-1.9243 (0.0022***)	0.8758 (0.0724*)	0.2402	177	0.3978	0.1582	0.2939
	6	-5.0341 (6.04e-12***)	0.0016 (2.25e-11***)	-2.5687 (0.0003***)	1.6817 (0.0033***)	0.3442	150.82	0.4445	0.1976	0.1184
	12	-4.2753 (3.90e-11***)	0.0013 (1.33e-10***)	-1.7466 (0.0084***)	0.0862 (0.8625)	0.2733	159.2	0.4835	0.2338	-0.0431

Nota: La columna “k” se refiere al número de periodos en que se retardan las variables explicativas. Bajo la estimación de cada regresor se recoge entre paréntesis su p-valor asociado. Como siempre, *, ** y *** representan significatividad estadística al 10, 5 y 1% respectivamente. El R^2 refleja el pseudo- R^2 empleado para valorar el ajuste de los modelos de variable dependiente limitada, y se obtiene mediante un ratio de funciones de verosimilitud. Los demás estadísticos contemplados son los analizados en la sección de metodología (4.1).

Para volver a la sección correspondiente a la tabla: 4.2

Table 8: Probit Autorregresivos. Variables seleccionadas vía GA

(a) t-1				(b) t-3			
$P_{t-1}(S_t = 1) = \Phi(\alpha + \delta_{t-1}S_{t-1} + \gamma_{t-1}\sum_{i=1}^{t-2-1}\Delta S_t + \sum_{i=1}^n \beta x_{i,t-1})$				$P_{t-3}(S_t = 1) = \Phi(\alpha + \delta_{t-3}S_{t-3} + \gamma_{t-3}\sum_{i=1}^{t-2-3}\Delta S_t + \sum_{i=1}^n \beta x_{i,t-3})$			
	Variable	β	p-valor		Variable	β	p-valor
In Sample	Intercept	-4.6662	2.85e-09***	In Sample	Intercept	-17.008	1.09e-06***
	EU nivel	1.6186	1.53e-09***		Stoxx50 Ret	0.0031	0.0013***
	US 5Y spread	2.7792	2.78e-07***		EU Nivel	2.0426	0.0007***
	US 2Y spread	-3.5563	0.0001***		US 5Y Spread	6.1849	1.42e-07***
	Fra 5Y Spread	-1.266	9.87e-06***		US 2Y Spread	-8.3553	6.05e-06***
	Baa-Aaa	3.7184	0.0019***		Fra 2Y Spread	-0.6382	0.0128**
	~	~	~		Baa-Aaa	5.5984	0.0075***
	AIC	113.84			AIC	78.802	
	R ²	0.5478			R ²	0.7087	
Out of Sample	RMSE	0.6012		Out of Sample	RMSE	0.6433	
	QPS	0.3615			QPS	0.4139	
	pseudo-R ²	-0.6130			pseudo-R ²	-0.8469	

(c) t-6				(d) t-12			
$P_{t-6}(S_t = 1) = \Phi(\alpha + \delta_{t-6}S_{t-6} + \gamma_{t-6}\sum_{i=1}^{t-2-6}\Delta S_t + \sum_{i=1}^n \beta x_{i,t-6})$				$P_{t-12}(S_t = 1) = \Phi(\alpha + \delta_{t-12}S_{t-12} + \gamma_{t-12}\sum_{i=1}^{t-2-12}\Delta S_t + \sum_{i=1}^n \beta x_{i,t-12})$			
	Variable	β	p-valor		Variable	β	p-valor
In Sample	Intercept	-16.623	6.42e-09**	In Sample	Intercept	-8.4407	1.12e-08***
	Stoxx50 Ret	0.0047	1.00e-08***		Stoxx50 Ret	0.0029	7.08e-09***
	Momentum	-2.2024	0.0103**		US 5Y spread	2.0599	3.94e-05***
	US 5Y spread	4.0908	2.61e-08***		US 2Y spread	-3.9833	1.87e-05***
	US 2Y spread	-4.9278	4.32e-06***		FRA 5Y spread	-2.7669	1.25e-05***
	FRA 10Y spread	3.0611	0.0021***		FRA 2Y spread	1.1712	2.47e-05***
	AIC	99.064			AIC	117.49	
	R ²	0.6002			R ²	0.4930	
Out of Sample	RMSE	0.5806		Out of Sample	RMSE	0.4389	
	QPS	0.3371			QPS	0.1927	
	pseudo-R ²	-0.5041			pseudo-R ²	0.1402	

Table 9: Probit con Múltiples Explicativas

$$P_{t-k}(S_t = 1) = \Phi(\alpha + \delta_{t-k}S_{t-k} + \gamma_{t-k}\sum_{l=1}^{t-k} \Delta S_l + \sum_{l=1}^n \beta_{x_{l,t-k}})$$

k	In Sample							Out Of Sample			
	α	δ	γ	$\beta_{spread\ USA}$	$\beta_{spread\ FRA}$	$\beta_{Ban-Aaa}$	R^2	AIC	RMSE	QPS	O.S. pseudo- R^2
1	-6.8517 (1.10e-09***)	0.0019 (2.26e-09***)	-1.9301 (0.0041***)	1.266 (7.03e-09***)	-0.3802 (0.0608*)	3.3103 (0.0032***)	0.3781	152.07	0.4782	0.2286	-0.0203
3	-13.204 (9.79e-10***)	0.0037 (1.74e-09***)	-2.2617 (0.0053***)	1.9456 (1.78e-09***)	-0.0386 (0.8727)	3.305 (0.0125***)	0.5338	115.69	0.4649	0.2161	0.0355
6	-9.4298 (3.27e-09***)	0.0027 (2.46e-09***)	-1.8295 (0.012**)	1.0437 (8.79e-06***)	-0.0412 (0.8434)	-0.3556 (0.6700)	0.4208	138.13	0.438	0.1918	0.144
12	-3.8103 (2.58e-06***)	0.0013 (4.09e-08***)	-1.7064 (0.0106**)	0.0042 (0.9801)	-0.5099 (0.012**)	-0.4507 (0.5627)	0.3090	155.78	0.4828	0.2331	-0.0400

Nota: Se seleccionan las variables individualmente significativas para alguno de los plazos, y con un ajuste “in sample” elevado. Se eliminan aquellas que, en combinación con las demás, dejan de ser individualmente significativas. Bajo la estimación de cada regresor se recoge entre paréntesis su p-valor asociado. Como siempre, *, ** y *** representan significatividad estadística al 10, 5 y 1% respectivamente.

Para volver a la sección correspondiente a la tabla: 4.2

Table 10: Modelos Markov Switching

$$r_t = \mu_{S_t} + \delta_{S_t,t-k} S_{t-k} + \gamma_{S_t,t-k} \sum_{i=1}^{t-2-k} \Delta S_t + \sum_{i=1}^t \beta_{S_t,t-k} x_{i,t-k} + \varepsilon_{S_t,t}$$

k	Intercept		S_{t-k}		Momentum		US spread 5Y		Fra spread 5Y		Baa-Aaa		$\sigma_{\varepsilon, S_t=0}$		$\sigma_{\varepsilon, S_t=1}$			
	$\mu_{S_t=0}$	$\mu_{S_t=1}$	$\delta_{S_t=0}$	$\delta_{S_t=1}$	$\gamma_{S_t=0}$	$\gamma_{S_t=1}$	$\beta_{S_t=0}$	$\beta_{S_t=1}$	$\beta_{S_t=0}$	$\beta_{S_t=1}$	$\beta_{S_t=0}$	$\beta_{S_t=1}$	$\sigma_{\varepsilon, S_t=0}$	$\sigma_{\varepsilon, S_t=1}$	Prob	$S_t = 0$	$S_t = 1$	
1	0.0867 (4.08e-04***)	0.0517 (0.2858)	-2.33e-05 (0.0029***)	-1.84e-05 (0.1409)	-0.0485 (0.0011***)	0.1531 (3.68e-06***)	-0.0117 (2.86e-04***)	-0.0076 (0.0182**)	0.0069 (0.0988*)	-0.0021 (0.8318)	0.0256 (0.2868)	-0.0085 (0.8204)	0.0274 (0***)	0.0506 (0***)	$S_t = 0$	0.9263	$S_t = 1$	0.149
3	0.0384 (0.11229)	0.0497 (0.2302)	0.0010 (0.2034)	-1.41e-05 (0.1814)	-0.0012 (0.9575)	0.0094 (0.7700)	-0.0020 (0.7897)	-0.0090 (0.3694)	-0.0045 (0.7250)	-0.0688 (0.6686)	0.0381 (0.3223)	-9.48e-06 (0.0555*)	0.0257 (0***)	0.0558 (0***)	$S_t = 0$	0.9599	$S_t = 1$	0.0494
6	0.0370 (0.2551)	0.0579 (0.1316)	-8.1e-06 (0.3430)	-1.82e-05 (0.0578*)	-0.0054 (0.7635)	0.0288 (0.1722)	-0.0047 (0.3261)	-0.0045 (0.6303)	0.0058 (0.2066)	-0.0027 (0.7472)	-0.0065 (0.9618)	0.0064 (0.9027)	0.0257 (4.60e-05***)	0.0565 (4.45e-09***)	$S_t = 0$	0.9611	$S_t = 1$	0.0456
12	0.0429 (0.4858)	0.0191 (0.1657)	-1.93e-05 (0.0180**)	-6.17e-06 (0.4989)	7.59e-04 (0.9831)	0.0231 (0.1967)	-0.0086 (0.1978)	0.0027 (0.5300)	0.0295 (0.0031***)	0.0030 (0.5226)	0.0027 (0.9357)	0.0038 (0.9288)	0.0250 (0***)	0.0532 (0***)	$S_t = 0$	0.9620	$S_t = 1$	0.0383

Nota: En este caso no interesan los estadísticos de ajuste del modelo dado que no se está estimando directamente la variación en las probabilidades de recesión, sino que se ajustan las tasas de variación del índice, obteniendo las probabilidades de tendencia bajista en el periodo siguiente como un parámetro más del modelo. Bajo la estimación de cada regresor se recoge entre paréntesis su p-valor asociado. Como siempre, *, ** y *** representan significatividad estadística al 10, 5 y 1% respectivamente.

Para volver a la sección correspondiente a la tabla: 4.3.1

Table 11: Modelo MS-DFM

	Coefficiente	p-valor		Coefficiente	p-valor
λ_1	0.0882	0***	ϕ_1	0.3608	0***
λ_2	-0.0146	0.5272	σ_1	0.376	0***
λ_3	0.0528	0.0125**	σ_2	0.716	0***
λ_4	0.1928	0***	σ_3	0.7191	0***
λ_5	-0.0365	0***	σ_4	1.3741	0***
ψ_1	0.1022	0***	σ_5	0.4123	0***
ψ_2	0.4853	0***	μ_0	0.0852	0.0046***
ψ_3	0.4872	0***	μ_1	-6.5338	0***
ψ_4	0.483	0***	p_{00}	0.9929	0***
ψ_5	0.3137	0***	p_{11}	0.8456	0***

Recordando que se puede expresar la ecuación de estado como: $\begin{bmatrix} f_t \\ u_t \end{bmatrix} = \begin{bmatrix} M_{S_t} \\ 0_{N,1} \end{bmatrix} + \begin{bmatrix} F_1 & 0_{1,N} \\ 0_{N,1} & \Psi_1 \end{bmatrix} \begin{bmatrix} f_{t-1} \\ u_{t-1} \end{bmatrix} + I_{N+1} * \begin{bmatrix} a_t \\ \varepsilon_t \end{bmatrix}$

y la de observación como: $y_t = H * x_t = \begin{bmatrix} \Lambda_{N,1} & I_{N,N} \end{bmatrix} \begin{bmatrix} f_t \\ u_t \end{bmatrix}$, con $\Lambda = (\lambda_1, \dots, \lambda_N)'$.

Nota: Para volver a la sección correspondiente a la figura: 4.3.2

Table 12: Matrices de Confusión

	True Condition Positive	True Condition Negative	True Condition Positive	True Condition Negative
Predicted Condition Positive	25 (34.72%)	11 (15.28%)	18 (25.00%)	10 (13.89%)
Predicted Condition Negative	1 (1.39%)	35 (48.61%)	8 (11.11%)	36 (50.00%)
Porcentaje total correctamente predicho (ACC): 83.33%				
(a) Mejor probit con una sola explicativa, threshold óptimo (0.45)				
	True Condition Positive	True Condition Negative	True Condition Positive	True Condition Negative
Predicted Condition Positive	23 (31.94%)	20 (27.78%)	20 (27.78%)	19 (26.39%)
Predicted Condition Negative	3 (4.17%)	26 (36.11%)	6 (8.33%)	27 (37.50%)
Porcentaje total correctamente predicho (ACC): 68.05%				
(c) Mejor probit con múltiples variables explicativas, threshold óptimo (0.4)				
	True Condition Positive	True Condition Negative	True Condition Positive	True Condition Negative
Predicted Condition Positive	17 (23.61%)	22 (30.56%)	11 (15.28%)	4 (5.55%)
Predicted Condition Negative	9 (12.50%)	24 (33.33%)	15 (20.83%)	42 (58.83%)
Porcentaje total correctamente predicho (ACC): 56.94%				
(e) Modelo Markov Switching t-3, threshold óptimo (0.15)				
	True Condition Positive	True Condition Negative	True Condition Positive	True Condition Negative
Predicted Condition Positive	17 (23.61%)	22 (30.56%)	11 (15.28%)	4 (5.55%)
Predicted Condition Negative	9 (12.50%)	24 (33.33%)	15 (20.83%)	42 (58.83%)
Porcentaje total correctamente predicho (ACC): 56.94%				
(f) Modelo MS-DFM, threshold óptimo (0.175)				

Nota: Se presentan las matrices de confusión asociadas a los mejores modelos considerados de cada clase. Se expresan los valores absolutos de las clasificaciones obtenidas, así como el porcentaje sobre el total de observaciones en el periodo "out of sample".

Para volver a la sección correspondiente a la tabla: 4.4

Table 13: Resultados Diebold & Mariano

	"IPC M4 t-3"	"Baa-Aaa M4 t-3"	"EU5spread M4 t-3"	"Baa-Aaa M4 t-1"	"VIX M4 t-1"	"US5spread M4 t-6"	"EU5spread M4 t-1"	"IPC M4 t-6"	"US5spread M4 t-3"	"US5spread M4 t-1"
"IPC M4 t-3"	NaN	0.2727	0.0628*	0.0745*	0.0262**	0.0877*	0.0078***	0.0544*	0.0116**	0.0042***
"Baa-Aaa M4 t-3"	-	NaN	0.0326**	0.0611*	0.0103**	0.0969*	0.0007***	0.0682*	0.0116**	0.0036***
"EU5spread M4 t-3"	-	-	NaN	0.2496	0.0702*	0.1718	0.0124**	0.1150	0.0282**	0.0099***
"Baa-Aaa M4 t-1"	-	-	-	NaN	0.1869	0.2499	0.0159**	0.1700	0.0496**	0.0063***
"VIX M4 t-1"	-	-	-	-	NaN	0.3818	0.2078	0.2704	0.1001	0.0295***
"US5spread M4 t-6"	-	-	-	-	-	NaN	0.4945	0.3644	0.1584	0.1138
"EU5spread M4 t-1"	-	-	-	-	-	-	NaN	0.3707	0.1503	0.0558*
"IPC M4 t-6"	-	-	-	-	-	-	-	NaN	0.3053	0.2542
"US5spread M4 t-3"	-	-	-	-	-	-	-	-	NaN	0.4297
"US5spread M4 t-1"	-	-	-	-	-	-	-	-	-	NaN
"GenAlgorithm"	-	-	-	-	-	-	-	-	-	-
"Multivariable, t-1"	-	-	-	-	-	-	-	-	-	-
"Multivariable, t-3"	-	-	-	-	-	-	-	-	-	-
"Multivariable, t-6"	-	-	-	-	-	-	-	-	-	-
"Multivariable, t-12"	-	-	-	-	-	-	-	-	-	-
"MS, t-1"	-	-	-	-	-	-	-	-	-	-
"MS, t-3"	-	-	-	-	-	-	-	-	-	-
"MS, t-6"	-	-	-	-	-	-	-	-	-	-
"MS, t-12"	-	-	-	-	-	-	-	-	-	-
"MS-DFM"	-	-	-	-	-	-	-	-	-	-

	"GenAlgorithm"	"Multivariable, t-1"	"Multivariable, t-3"	"Multivariable, t-6"	"Multivariable, t-12"	"MS, t-1"	"MS, t-3"	"MS, t-6"	"MS, t-12"	"MS-DFM"
"IPC M4 t-3"	0.1166	0.0052***	0.0238**	0.0739*	5.49e-06***	0.0301**	0.00013***	0.0015***	0.0004***	0.0213**
"Baa-Aaa M4 t-3"	0.1516	0.0044***	0.0228**	0.0795*	6.72e-05***	0.0332**	0.0014***	0.0016***	0.0004***	0.0262**
"EUSpread M4 t-3"	0.2119	0.0107**	0.0487**	0.1461	0.0001***	0.0585*	0.0027***	0.0031***	0.0009***	0.0418**
"Baa-Aaa M4 t-1"	0.2812	0.0075***	0.0717*	0.2169	0.0031***	0.0845*	0.0049***	0.0058***	0.0019***	0.0595*
"VIX M4 t-1"	0.3725	0.0291**	0.1263	0.3436	0.0060***	0.1392	0.0095***	0.0108**	0.0040***	0.0827*
"USSpread M4 t-6"	0.4599	0.0704*	0.1742	0.0030***	0.0515*	0.2183	0.0149**	0.0196**	0.0078***	0.1591
"EUSpread M4 t-1"	0.4615	0.0512*	0.1773	0.4634	0.0163**	0.1857	0.0118**	0.0150**	0.0049***	0.124
"IPC M4 t-6"	0.4342	0.2025	0.3186	0.3999	0.0491**	0.2942	0.0156**	0.0327**	0.0089***	0.1884
"USSpread M4 t-3"	0.2639	0.2409	0.4947	0.1772	0.3211	0.4835	0.1344	0.1615	0.1031	0.3771
"USSpread M4 t-1"	0.2025	0.1818	0.4365	0.1305	0.3424	0.4399	0.1271	0.1512	0.0945*	0.3972
"GenAlgorithm"	NaN	0.1728	0.2822	0.4885	0.0538*	0.1223	0.0003***	0.0003***	0.0002***	0.0596*
"Multivariable, t-1"	-	NaN	0.2545	0.0806*	0.4532	0.3587	0.2024	0.2406	0.1657	0.4891
"Multivariable, t-3"	-	-	NaN	0.193	0.3342	0.4855	0.1481	0.1763	0.1160	0.3840
"Multivariable, t-6"	-	-	-	NaN	0.0603*	0.2401	0.018**	0.0239**	0.0097***	0.1735
"Multivariable, t-12"	-	-	-	-	NaN	0.2573	0.131	0.1938	0.0819*	0.4623
"MS, t-1"	-	-	-	-	-	NaN	0.0002***	0.0005***	0.0003***	0.2465
"MS, t-3"	-	-	-	-	-	-	NaN	0.2145	0.247	0.0848*
"MS, t-6"	-	-	-	-	-	-	-	NaN	0.0481**	0.1475
"MS, t-12"	-	-	-	-	-	-	-	-	NaN	0.0755*
"MS-DFM"	-	-	-	-	-	-	-	-	-	NaN

Nota: En esta tabla y en la anterior se contrasta si la capacidad predictiva de los modelos de la fila y de la columna es estadísticamente distinta. En la tabla se recogen los p-valores asociados los tests realizados. De este modo, para un nivel de significatividad del 5%, si el valor en la tabla es menor que 0.05, se podrá afirmar que existe suficiente evidencia para rechazar la hipótesis nula y que, por tanto, la capacidad predictiva de los modelos es significativamente diferente. M2 hace referencia a la especificación probit con únicamente variables explicativas exógenas. M4 corresponde al modelo con variables exógenas, componente autorregresivo (reversión), y factor momentum como explicativas. "t-k" hace referencia al retardo que se considera para las variables explicativas. Los modelos "Multivariables" son aquellos probit que incorporan varias variables exógenas, además del factor momentum y el componente autorregresivo. Como siempre, *, ** y *** representan significatividad estadística al 10, 5 y 1% respectivamente.

Para volver a la sección correspondiente a la tabla: 5.1

Table 14: Resultados Clark & West

	"IPC M4 t-3"	"Baa-Aaa M4 t-3"	"EU5spread M4 t-3"	"Baa-Aaa M4 t-1"	"VIX M4 t-1"	"US5spread M4 t-6"	"IPC M4 t-6"	"US5spread M4 t-3"	"US5spread M4 t-1"
"IPC M4 t-3"	NaN	-	-	-	-	-	-	-	-
"Baa-Aaa M4 t-3"	0.0739*	NaN	-	-	-	-	-	-	-
"EU5spread M4 t-3"	0.0061***	0.0290**	NaN	-	-	-	-	-	-
"Baa-Aaa M4 t-1"	0.0132**	0.0361**	0.1178	NaN	-	-	-	-	-
"VIX M4 t-1"	0.0008***	0.0004***	0.0057***	0.0417**	NaN	-	-	-	-
"US5spread M4 t-6"	0.0058***	0.0190**	0.0206**	0.0446**	0.0392**	NaN	-	-	-
"EU5spread M4 t-1"	0.0008**	0.0004***	0.0049***	0.0023***	0.0287**	0.0539*	-	-	-
"IPC M4 t-6"	0.0046***	0.0099***	0.0141**	0.0223**	0.0306**	0.0295**	NaN	-	-
"US5spread M4 t-3"	0.0003***	0.0001***	0.0003***	0.0003***	0.0005***	0.0012***	0.0006***	NaN	-
"US5spread M4 t-1"	0.0001***	4.07e-05***	0.0001***	0.0001***	0.0003***	0.0029***	0.0019***	0.0948*	NaN
"GenAlgorithm"	0.0013***	0.0013***	0.0023***	0.0024***	0.0038***	0.0034***	0.0342**	0.0005***	0.0059***
"Multivariable, t-1"	0.0003***	0.0001***	0.0003***	0.0005***	0.0008***	0.0009***	0.0013***	0.0587*	0.1377
"Multivariable, t-3"	0.0008**	0.0005***	0.0010***	0.0008***	0.0013***	0.0012***	0.0010***	0.3683	0.0588*
"Multivariable, t-6"	0.0049***	0.0158**	0.0174**	0.0388**	0.0335**	0.0086***	0.0787*	0.2112	0.2405
"Multivariable, t-12"	1.27e-05***	0.0001***	0.0001***	0.0007***	0.001***	0.0014***	0.0042***	0.0013***	0.0042***
"MS, t-1"	0.0020***	0.0039***	0.0055***	0.0068***	0.0088**	0.0059***	0.0133**	0.0038***	0.0174**
"MS, t-3"	0.0001***	0.0001***	0.0019***	0.0001***	0.0001***	0.0001***	0.0003***	0.0001***	0.0002***
"MS, t-6"	7.93e-05***	0.0001***	0.0002***	0.0001***	0.0002***	8.80e-05***	0.0006***	5.41e-05***	0.0002***
"MS, t-12"	3.14e-05***	3.27e-05***	4.81e-05***	2.47e-05***	4.23e-05***	4.29e-05***	0.0001***	6.43e-05***	0.0002***
"MS-DFM"	0.0016***	0.0021***	0.0027***	0.0031***	0.0044***	0.0021***	0.0049***	0.0030***	0.0079***

	"GenAlgorithm"	"Multivariable, t-1"	"Multivariable, t-3"	"Multivariable, t-6"	"Multivariable, t-12"	"MS, t-1"	"MS, t-3"	"MS, t-6"	"MS, t-12"	"MS-DFM"
"IPC M4 t-3"	-	-	-	-	-	-	-	-	-	-
"Baa-Aaa M4 t-3"	-	-	-	-	-	-	-	-	-	-
"EU5spread M4 t-3"	-	-	-	-	-	-	-	-	-	-
"Baa-Aaa M4 t-1"	-	-	-	-	-	-	-	-	-	-
"VIX M4 t-1"	-	-	-	-	-	-	-	-	-	-
"US5spread M4 t-6"	-	-	-	-	-	-	-	-	-	-
"EU5spread M4 t-1"	-	-	-	-	-	-	-	-	-	-
"IPC M4 t-6"	-	-	-	-	-	-	-	-	-	-
"US5spread M4 t-3"	-	-	-	-	-	-	-	-	-	-
"US5spread M4 t-1"	-	-	-	-	-	-	-	-	-	-
"GenAlgorithm"	NaN	-	-	-	-	-	-	-	-	-
"Multivariable, t-1"	0.0043***	NaN	-	-	-	-	-	-	-	-
"Multivariable, t-3"	0.0340**	0.1281	NaN	-	-	-	-	-	-	-
"Multivariable, t-6"	0.0244**	0.3267	0.1989	NaN	-	-	-	-	-	-
"Multivariable, t-12"	0.0027***	0.0029***	0.0008***	0.0016***	NaN	-	-	-	-	-
"MS, t-1"	0.0138**	0.0107**	0.0025***	0.0064***	0.0768*	NaN	-	-	-	-
"MS, t-3"	0.0002***	0.0002***	7.58e-05***	0.0001***	0.0043***	0.0002***	NaN	-	-	-
"MS, t-6"	0.0001***	0.0001***	3.67e-05***	0.0001***	0.006***	0.0007***	0.4212	NaN	-	-
"MS, t-12"	0.0004***	9.36e-05***	3.85e-05***	4.89e-05***	0.0015***	0.0013***	0.1922	0.0464**	NaN	-
"MS-DFM"	0.0029***	0.0059***	0.0022***	0.0022***	0.0194**	0.0383**	0.2078	0.1357	0.1542	NaN

Nota: Se contrasta si la capacidad predictiva del modelo de cada fila es estadísticamente superior a la del modelo de cada columna. Se recogen los p-valores de cada test, y si se obtiene un valor menor de 0.05 se podrá afirmar que existe evidencia suficiente respaldando rechazar la hipótesis nula. Es decir, para un $p < 0.05$, se puede concluir que el modelo de la columna es mejor predictor que el de la fila. M2 hace referencia a la especificación `probit` con únicamente variables explicativas exógenas. M4 corresponde al modelo con variables exógenas, componente autorregresivo, y factor momentum como explicativas. "t-k" hace referencia al retardo que se considera para las variables explicativas. Los modelos "Multivariables" son aquellos `probit` que incorporan varias variables exógenas, además del factor momentum y el componente autorregresivo. Como siempre, *, ** y *** representan significatividad estadística al 10, 5 y 1% respectivamente.

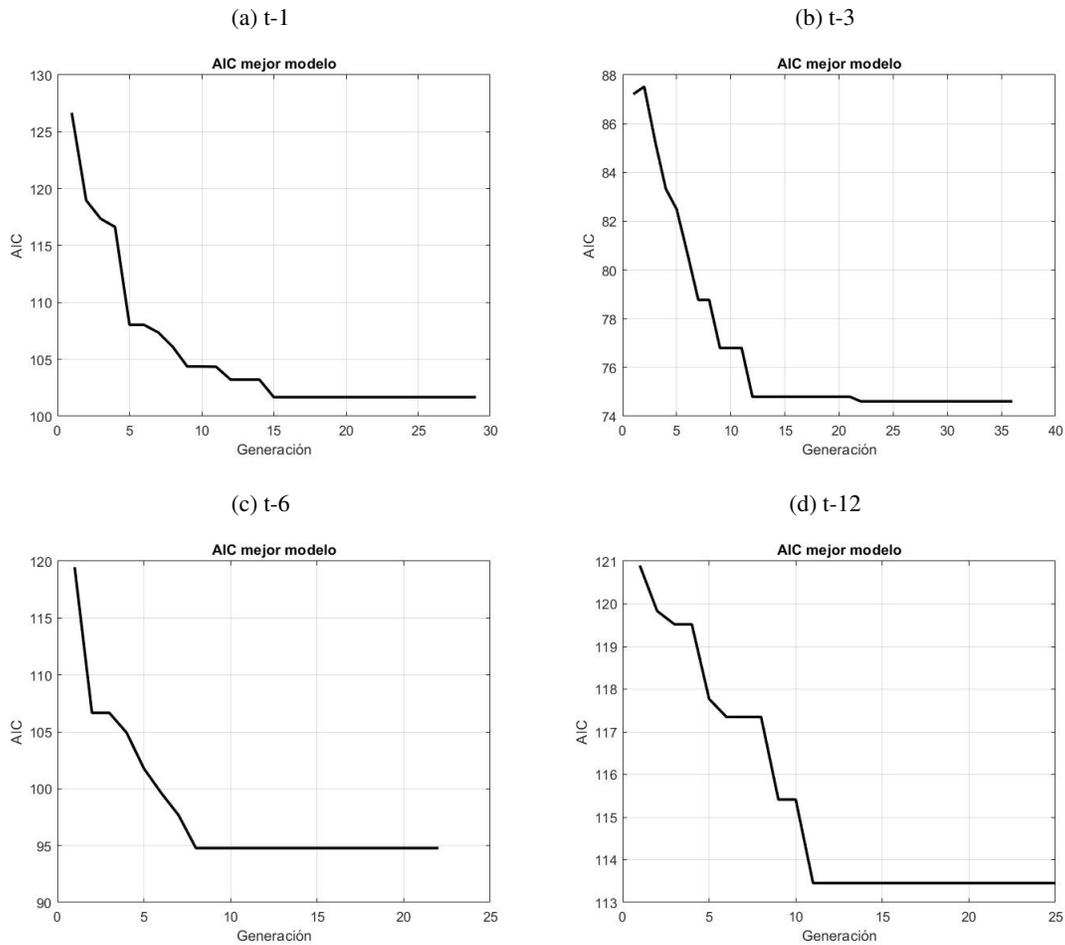
Para volver a la sección correspondiente a la tabla: 5.2

Table 15: Estrategias

	Rdto.	Sharpe	Sortino	Skew.	Máx. Caída	Rdto.	Sharpe	Sortino	Skew.	Máx. Caída	Rdto.	Sharpe	Sortino	Skew.	Máx. Caída	
Stoxx 50	0.1593	0.1933	~	-0.3052	-0.0985	0.1593	0.1933	-	-0.3052	-0.0985	0.1593	0.1933	-	-0.3052	-0.0985	
~	~	~	~	~	~	Mom.	0.4046	0.3281	0.1819	-0.1951	0.1677	0.1901	0.0217	-0.1778	-0.0736	
"IPC M4 t-3"	0.4847	0.5645	0.3481	1.1583	-0.0326	⇒	0.2555	0.4053	0.0996	0.9502	⇒	0.2543	0.3764	0.0801	0.1489	-0.0645
Baa - Aaa M4 t-3"	0.4876	0.5274	0.3536	0.5952	-0.0502	⇒	0.4610	0.4716	0.2449	0.7055	⇒	0.2736	0.3431	0.0866	0.1046	-0.0645
"EU5spread M4 t-3"	0.4288	0.4506	0.3	0.2545	-0.0612	⇒	0.4899	0.4814	0.2651	0.6715	⇒	0.4048	0.4133	0.1665	0.0544	-0.0645
"Baa - Aaa M4 t-1"	0.4281	0.4794	0.2843	0.5880	-0.0502	⇒	0.2632	0.332	0.0936	0.1420	⇒	0.2896	0.3217	0.0915	-0.0604	-0.0731
"VIX M4 t-1"	0.1978	0.3597	0.0340	0.2144	-0.0632	⇒	0.4160	0.4771	0.2087	1.0811	⇒	0.2582	0.4262	0.0821	0.6418	-0.0559
"USSpread M4 t-6"	0.2624	0.4609	0.0849	0.2402	-0.0557	⇒	0.0804	0.2536	-0.0241	0.3386	⇒	-0.0593	0.1861	-0.1078	-1.0185	-0.0731
"EU5spread M4 t-1"	0.2238	0.3677	0.0573	0.3083	-0.0577	⇒	0.1933	0.3258	0.0516	0.938	⇒	0.1514	0.2810	0.0195	0.0729	-0.0666
"IPC M4 t-6"	0.4557	0.5202	0.3092	0.6016	-0.0502	⇒	0.4011	0.4898	0.1949	1.3398	⇒	0.0227	0.2301	-0.0530	0.1718	-0.0653
"USSpread M4 t-3"	0.2359	0.3832	0.0676	0.1078	-0.0612	⇒	0.3365	0.4382	0.1541	1.0879	⇒	0.2782	0.3805	0.0989	0.0593	-0.0645
"USSpread M4 t-1"	0.1928	0.3395	0.0310	0.1002	-0.0612	⇒	0.1563	0.3066	0.0243	-0.3371	⇒	0.3704	0.3771	0.1501	-0.2055	-0.0731
"GenAlgorithm"	0.3604	0.3975	0.2219	0.1309	-0.0612	⇒	0.2383	0.3092	0.0772	0.2323	⇒	0.1495	0.2554	0.0163	-0.2589	-0.0731
"Multivariable,t-1"	0.2293	0.3910	0.0614	0.1623	-0.0612	⇒	0.0325	0.2464	-0.056	-0.3744	⇒	0.2484	0.3286	0.0765	-0.1833	-0.0731
"Multivariable,t-3"	0.2433	0.4142	0.0719	0.1878	-0.0612	⇒	0.298	0.428	0.1258	1.2151	⇒	0.258	0.374	0.0855	0.0661	-0.0645
"Multivariable,t-6"	0.2714	0.4733	0.0923	0.1822	-0.0557	⇒	0.1026	0.2768	-0.0102	0.3643	⇒	-0.034	0.2187	-0.0947	-0.9487	-0.0731
"Multivariable,t-12"	0.2530	0.4920	0.0789	1.5789	-0.0326	⇒	0.2673	0.4419	0.0993	1.6676	⇒	0.0364	0.291	-0.0481	0.6605	-0.0653
"MS,t-1"	0.1000	0.3234	-0.0176	0.4968	-0.0557	⇒	0.1268	0.2676	0.0035	0.4174	⇒	0.0871	0.2821	-0.0202	-0.0118	-0.0731
"MS,t-3"	0.1385	0.3394	0.0153	0.6773	-0.0557	⇒	0.2167	0.3341	0.0594	0.2576	⇒	0.0404	0.2290	-0.0448	-0.1156	-0.0731
"MS,t-6"	0.2169	0.3656	0.0534	0.7466	-0.0557	⇒	0.2542	0.3467	0.0869	0.1976	⇒	0.0834	0.2485	-0.0224	-0.1103	-0.0731
"MS,t-12"	0.2427	0.3534	0.0809	0.4758	-0.0557	⇒	0.2787	0.3569	0.1012	0.1517	⇒	0.016	0.1998	-0.0573	-0.2277	-0.0731
"MS-DFM"	0.0562	0.1839	-0.101	-4659	-0.1005	⇒	0.2597	0.3089	0.0959	0.3146	⇒	0.0367	0.1784	-0.0459	-0.1171	-0.0736

Gráficos

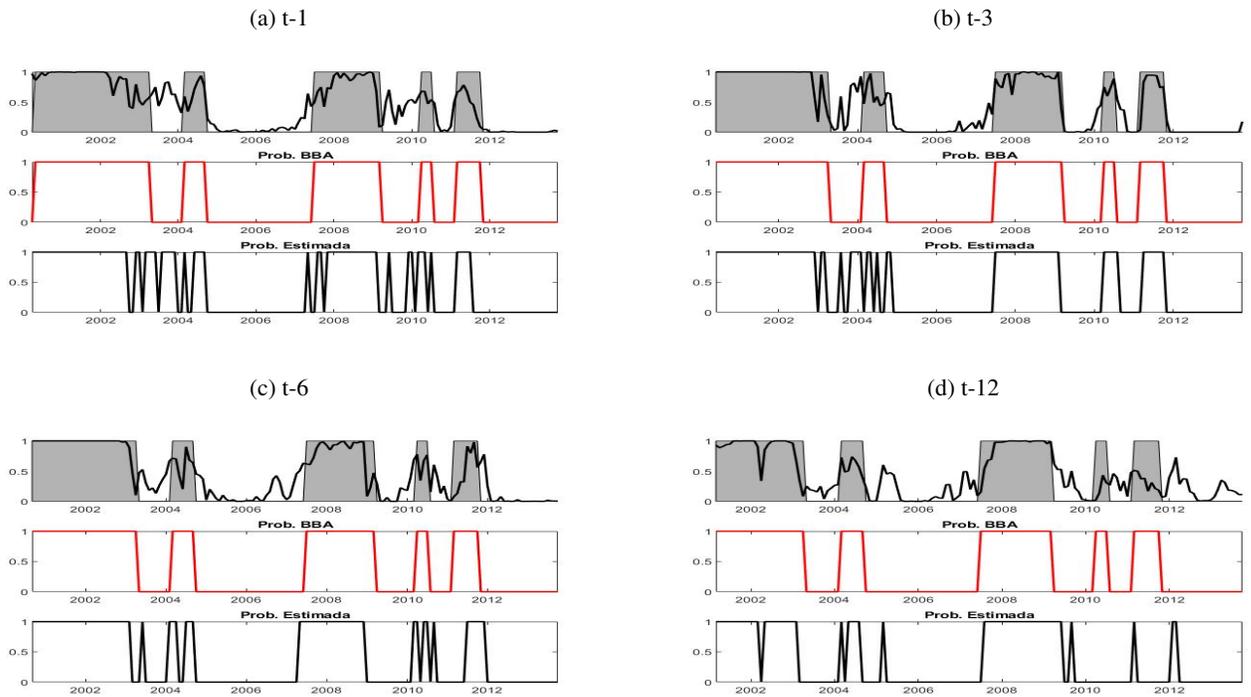
Figure 4: Convergencia GA



Nota: Se reflejan los estadísticos de Akaike correspondientes al mejor modelo (en base a este criterio) de cada generación del algoritmo genético, hasta que se alcanza la convergencia. El primer gráfico refleja la convergencia cuando las variables explicativas analizadas están retardadas en 1 periodo, y el segundo, tercer y cuarto gráfico reflejan la convergencia para las explicativas retardadas en 3, 6 y 12 periodos respectivamente.

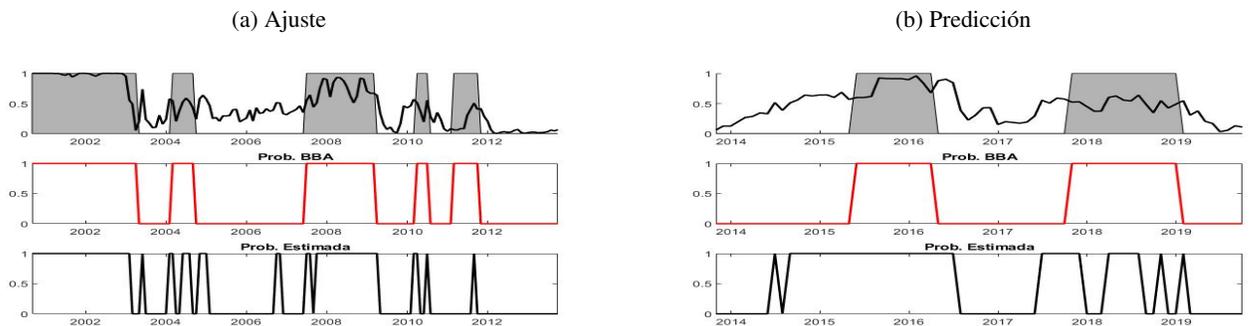
Para volver a la sección correspondiente a la figura: 4.2

Figure 5: Ajuste Modelos Probit vía GA



Nota: Se reflejan las probabilidades de recesión estimadas para el periodo “in sample” para los modelos de variable dependiente limitada con múltiples exógenas seleccionadas mediante el algoritmo genético. Las zonas sombreadas reflejan la probabilidad “real” de tendencia bajista para cada periodo, obtenida mediante el algoritmo de Bry & Boschan.

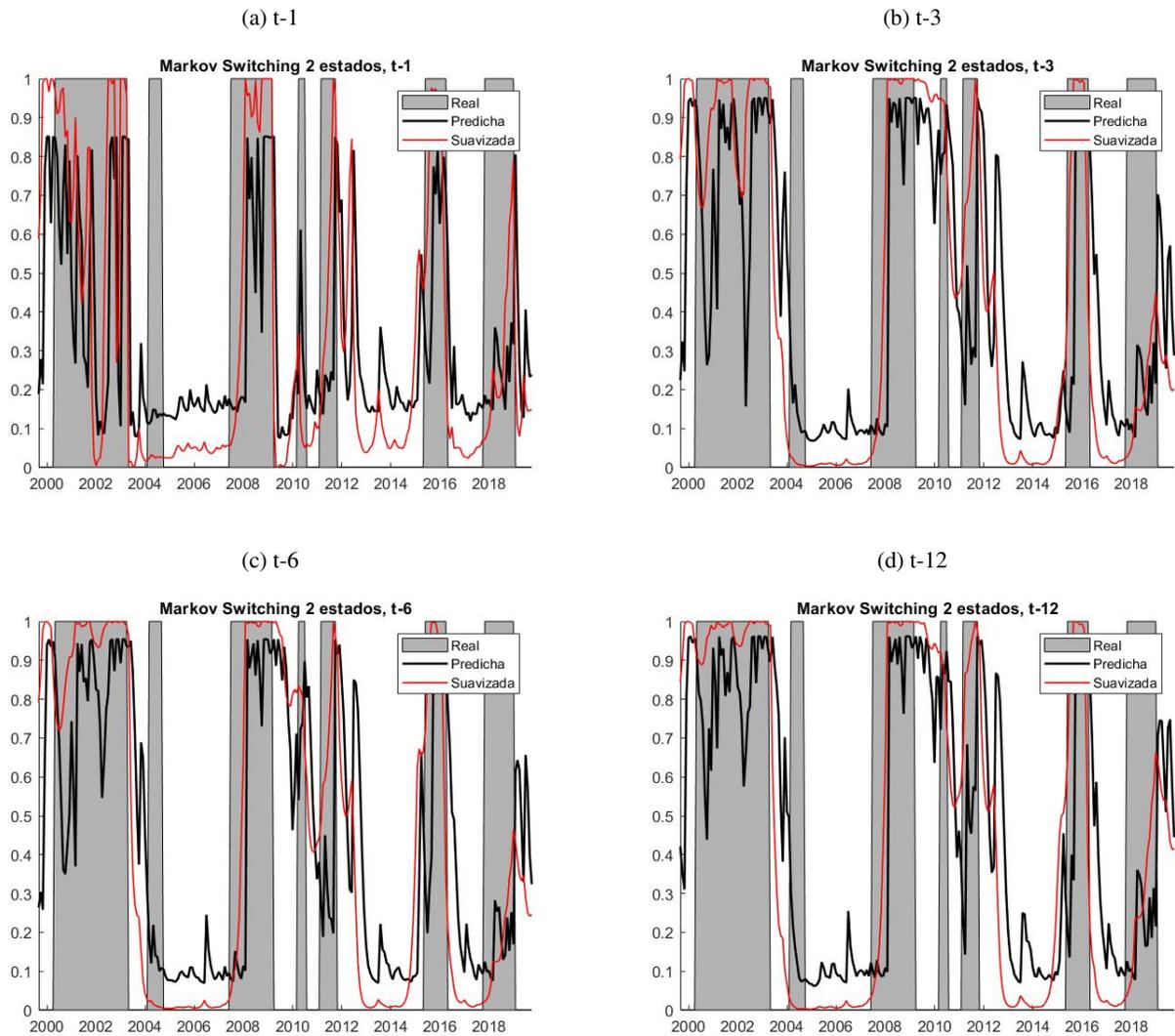
Figure 6: Mejor Probit con Varias Explicativas



Nota: Se reflejan el ajuste (a) y las predicciones (b) correspondientes al modelo probit con múltiples variables explicativas retardadas en 6 periodos, y que incluye el componente autorregresivo, el factor momentum, y el “Default Yield Spread”, junto con las pendientes de las curvas de tipos americana y francesa. En la tercera gráfica de cada figura se recoge la probabilidad estimada para un threshold estándar de 0.5 aplicado a las probabilidades correspondientes al modelo. Las zonas sombreadas reflejan las probabilidades “reales” de tendencia recesiva obtenidas mediante el algoritmo de Bry & Boschan.

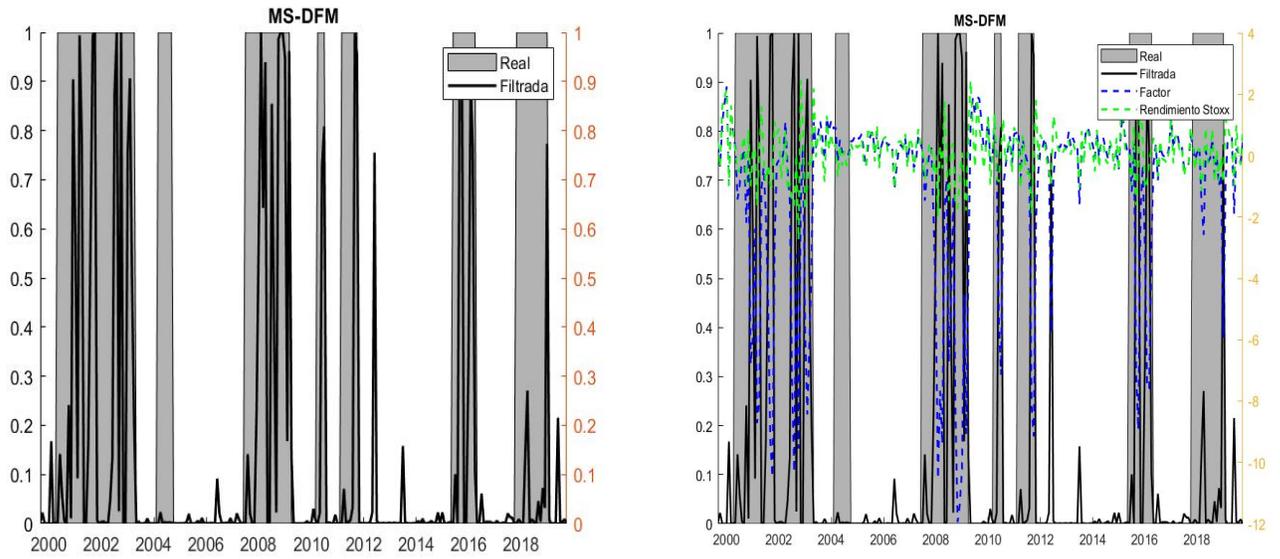
Para volver a la sección correspondiente a las figuras: 4.2

Figure 7: Probabilidades Predichas Markov Switching



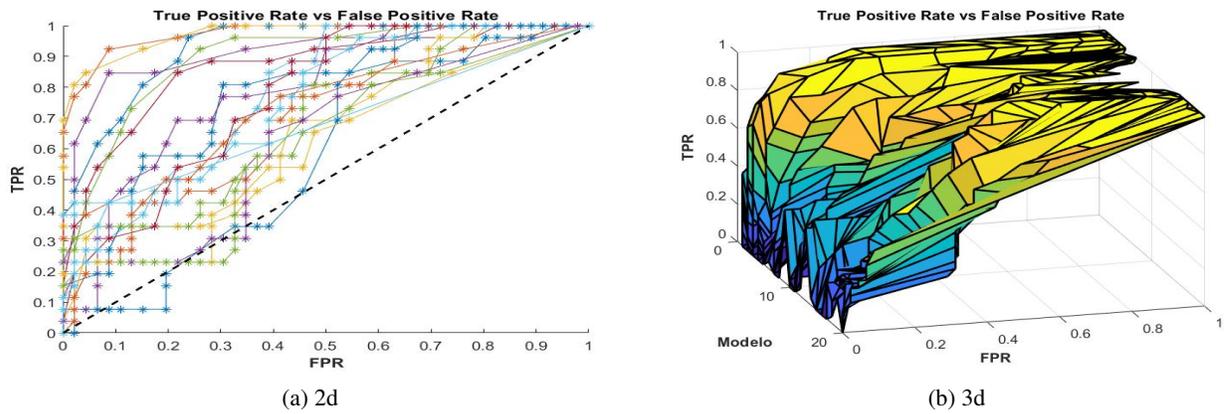
Nota: Se reflejan las probabilidades de recesión suavizadas (en rojo) y predichas (en negro) para los modelos con parámetros cambiantes en el tiempo vía Markov Switching. Las zonas sombreadas reflejan la probabilidad “real” de tendencia bajista para cada periodo, obtenida mediante el algoritmo de Bry & Boschan.

Figure 8: Probabilidades estimadas vía MS-DFM



Nota: Se reflejan las probabilidades de recesión estimadas para el modelo factorial dinámico. Las zonas sombreadas reflejan la probabilidad “real” de tendencia bajista para cada periodo, obtenida mediante el algoritmo de Bry & Boschan.

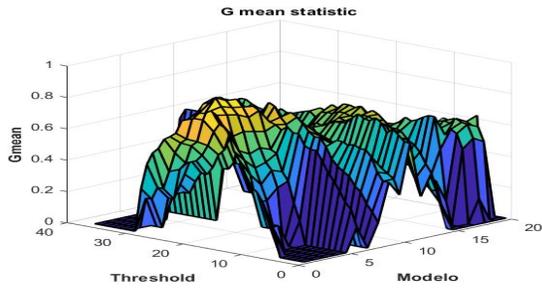
Figure 9: Curvas ROC



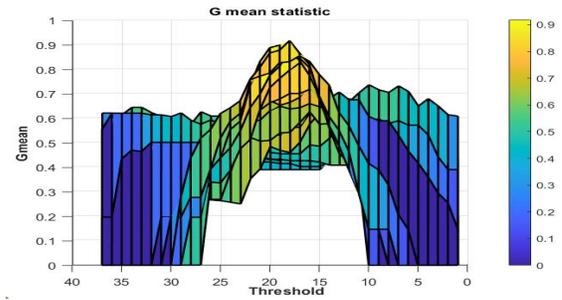
Nota: Se grafica el ratio TPR frente al FPR para los 20 modelos considerados y para cada threshold analizado (del 0.5 al 0.95 con pasos de tamaño 0.05).

Para volver a la sección correspondiente: 4.4

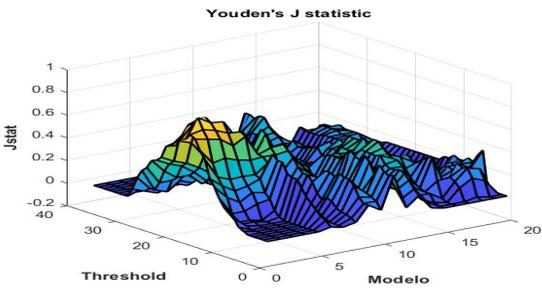
Figure 10: Estadísticos de Análisis de Umbrales



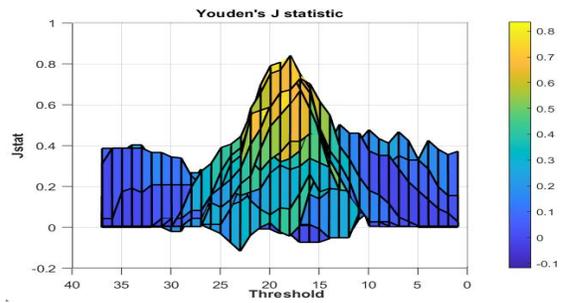
(a) G-Mean



(b) G-Mean Lateral



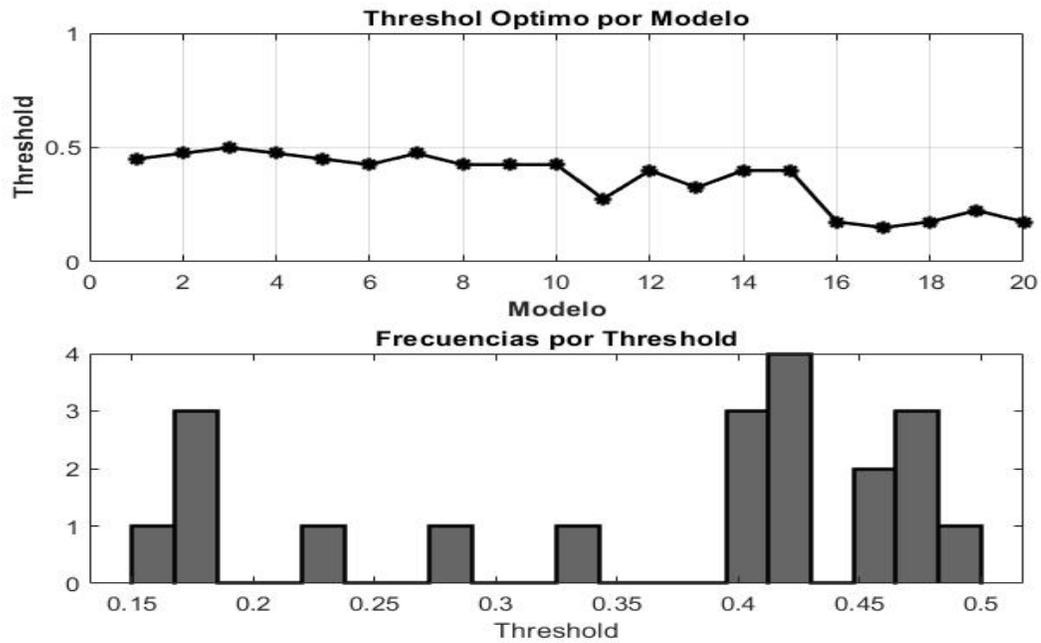
(c) J Statistic



(d) J-Statistic Lateral

Nota: Se representan los estadísticos G-mean y J de Youden para cada modelo y para cada threshold considerados. Se preferirá, para cada modelo, el umbral que aporte un mayor valor de los estadísticos.
Para volver a la sección correspondiente: 4.4

Figure 11: Threshold óptimos



	"IPC M4 t-3"	"Baa-Aaa USA M4 t-3"	"EU5spread M4 t-3"	"AAA-BBB USA M4 t-1"	"VIX M4 t-1"
Threshold	0.45	0.475	0.5	0.475	0.45
QPS	0.1582	0.1638	0.1705	0.1762	0.1835

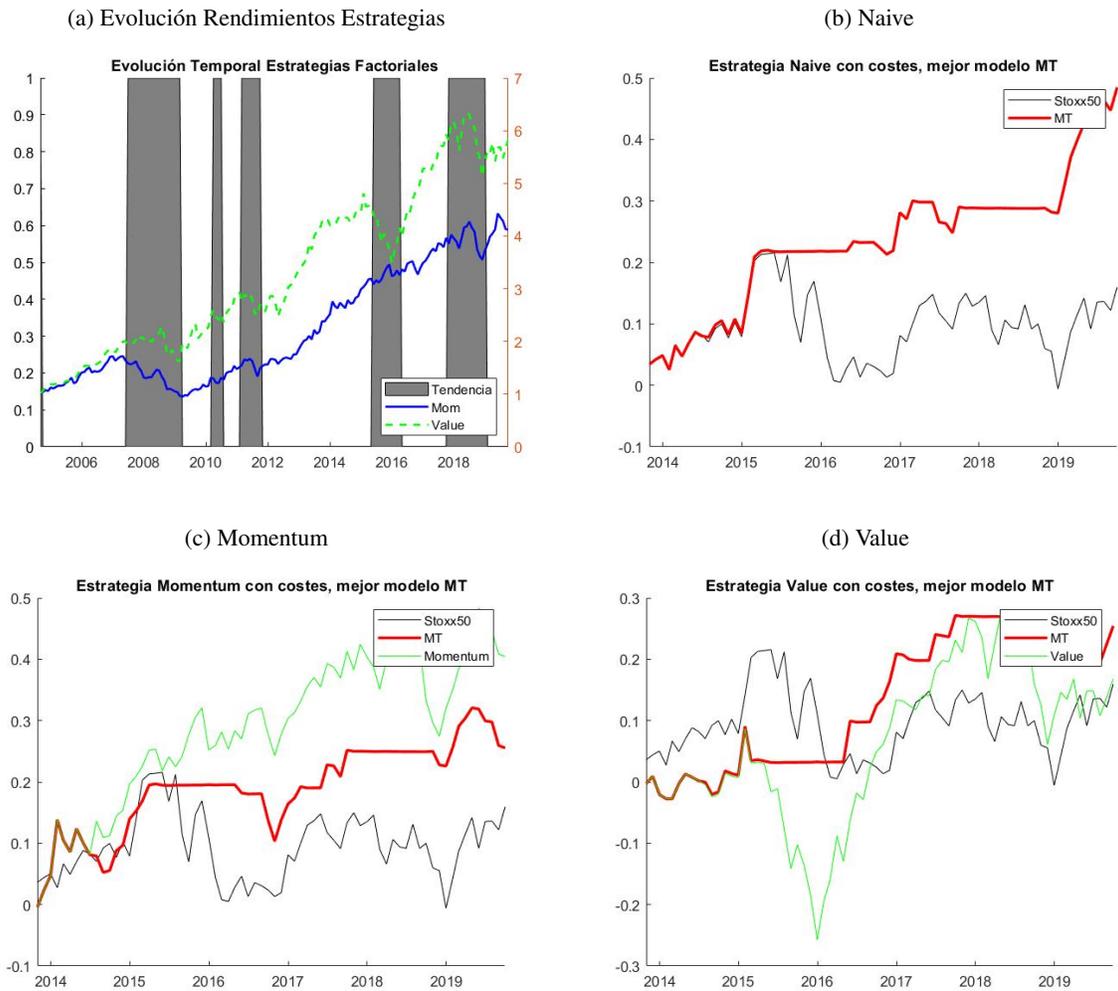
	"US5spread M4 t-6"	"EU5spread M4 t-1"	"IPC M4 t-6"	"US5spread M4 t-3"	"US5spread M4 t-1"
Threshold	0.425	0.475	0.425	0.425	0.425
QPS	0.1897	0.1899	0.1976	0.2162	0.2197

	"GenAlgorithm"	"Multivariable, t-1"	"Multivariable, t-3"	"Multivariable, t-6"	"Multivariable, t-12"
Threshold	0.275	0.4	0.325	0.4	0.4
QPS	0.1927	0.2286	0.2161	0.1918	0.2331

	"MS, t-1"	"MS, t-3"	"MS, t-6"	"MS, t-12"	"MS-DFM"
Threshold	0.175	0.15	0.175	0.225	0.175
QPS	0.2146	0.2683	0.2594	0.2751	0.2298

Nota: Se recogen los threshold óptimos para cada modelo seleccionados mediante el análisis de las curvas ROC. También los estadísticos QPS de los modelos a comparar.

Figure 12: Rendimiento Acumulado Estrategias Para Mejor Modelo



Nota: En la primera figura se representa la evolución de los rendimientos acumulados de las estrategias asociadas al factor momentum y al HmL de Fama & French (conocido como factor “Value”) para un periodo mayor que el del “out of sample”. Se puede observar en particular cómo el momentum presenta rendimientos negativos sostenidos durante el mercado bajista de 2008, o como la estrategia “value” presenta caídas bruscas durante los mercados bajistas. En los siguientes gráficos se comparan los rendimientos acumulados en el periodo de validación de las estrategias planteadas que incorporan un componente de market timing con los rendimientos acumulados del Eurostoxx y de las estrategias factoriales.

Para volver a la sección correspondiente a la figura: 6.2