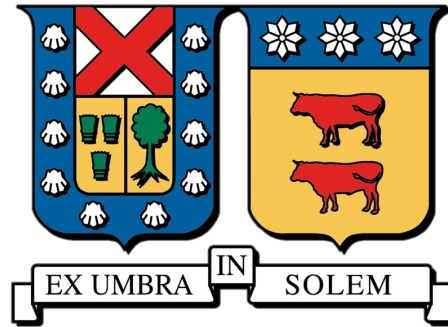




**UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INGENIERÍA COMERCIAL**

**COMPARACIÓN DE MODELOS DE VALORIZACIÓN DE ACTIVOS
USANDO MÉTODOS DE MACHINE LEARNING**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO COMERCIAL

AUTOR

BASTIÁN ALBERTO PEZO VERA

PROFESOR GUÍA

MARCELO VILLENA

PROFESOR CO-REFERENTE

PABLO ISLA

SANTIAGO DE CHILE, FECHA 16 DE ENERO DE 2025

Tabla de Contenidos

1	Resumen	4
2	Introducción	5
3	Problema de Investigación	7
4	Objetivos	10
4.1	Objetivo General	10
4.2	Objetivos Específicos	10
5	Marco Teórico	11
5.1	Valorización de Activos	11
5.1.1	Modelo de valorización de Activos Financieros (CAPM)	12
5.1.2	Modelo de Tres Factores de Fama & French	14
5.1.3	Modelo de Cinco Factores de Fama & French	18
5.1.4	Incorporación del Factor Momentum al Modelo de Cinco Factores	20
5.2	Métodos Estadísticos	21
5.2.1	Mínimos Cuadrados Ordinarios (OLS)	22
5.2.2	Regresión Ridge	27
5.2.3	Regresión LASSO	30
5.2.4	Regresión Elastic Net	31
5.2.5	Árboles de Decisión	34
6	Metodología	38

6.1	Datos y división de la muestra	38
6.2	Hiperparametrización	39
6.3	Evaluación del rendimiento de los métodos/modelos	41
7	Resultados	42
8	Conclusión	49

Lista de Tablas

1	Hiperparámetros para ElasticNet y Random Forest	41
2	Valores R^2 por Sector, Método y Modelo	43
3	Tiempos de ejecución en minutos por Sector, Método y Modelo	44

Lista de Figuras

1	División de la muestra para dos ventanas consecutivas.	39
2	Boxplots de MSE de prueba agrupados por modelo, sector, para todas las ventanas	47
3	Market Betas a través del tiempo para las distintas combinaciones de método, modelo, sector.	55
4	Gráficos de densidad del MSE de prueba para las distintas combinaciones de método, modelo, sector.	60



A mis padres, quienes siempre confiaron en mí.

1 Resumen

Este estudio evalúa la precisión predictiva de diversos modelos de valorización de activos, incluyendo el CAPM y los modelos multifactoriales de Fama-French, en comparación con técnicas de machine learning como Elastic Net y Random Forest. En un contexto de alta volatilidad y complejidad de mercado, la investigación adopta una metodología de rolling windows para validar los pronósticos fuera de muestra, aplicados en múltiples sectores económicos. Los resultados muestran que los métodos de regularización como Elastic Net ofrecen un equilibrio óptimo entre ajuste y generalización, especialmente en sectores con menor volatilidad, mientras que Random Forest enfrenta desafíos de sobreajuste en entornos predominantemente lineales. Este análisis sugiere que la combinación de enfoques tradicionales y machine learning en la valorización de activos mejora la capacidad de los modelos para capturar relaciones complejas y no lineales, proporcionando así herramientas más robustas para la gestión de carteras. La investigación contribuye a la literatura financiera demostrando que, aunque los modelos tradicionales siguen siendo relevantes, el machine learning complementa estos enfoques al ofrecer una alternativa flexible para enfrentar la dinámica cambiante de los mercados financieros.

2 Introducción

La valorización de activos ha sido un tema central en la teoría y práctica financiera, dado su papel crucial en la toma de decisiones de inversión, la gestión de riesgos y la asignación eficiente de recursos. Desde los primeros desarrollos del Modelo de Valoración de Activos Financieros (CAPM) de Sharpe (1964) y Lintner (1965), hasta los modelos multifactoriales de Fama y French (1993), el campo de la valoración de activos ha evolucionado con el objetivo de capturar de manera más precisa las relaciones riesgo-retorno. Estos modelos, aunque fundamentales, han mostrado limitaciones al ser aplicados en mercados financieros contemporáneos, donde factores como la alta dimensionalidad de los datos, la volatilidad y los cambios estructurales en la economía desafían las suposiciones tradicionales.

En las últimas dos décadas, el desarrollo de nuevas tecnologías y la disponibilidad de grandes volúmenes de datos han facilitado la incorporación de técnicas de machine learning en la valorización de activos. Estas metodologías permiten una mayor flexibilidad en la modelización, al captar patrones no lineales y relaciones complejas entre variables que suelen escapar a los enfoques econométricos tradicionales. A pesar de su potencial, la implementación de machine learning en finanzas enfrenta desafíos específicos, tales como el riesgo de sobreajuste y la interpretación de los resultados, lo que exige un balance entre precisión predictiva y robustez del modelo.

Este estudio se centra en evaluar el rendimiento predictivo de modelos tradicionales de valorización de activos, como el CAPM y los modelos de Fama-French de tres y cinco factores, y compararlos con métodos de machine learning, específicamente Elastic Net y Random Forest. Se adopta una metodología de rolling windows para validar el poder de predicción out-of-sample de estos modelos en distintos sectores económicos, lo cual permite evaluar la capacidad de cada

metodología para adaptarse a la evolución del mercado a lo largo del tiempo. La hipótesis central de esta investigación es que la combinación de enfoques tradicionales y modernos puede proporcionar una herramienta más robusta y adaptable para la valorización de activos en contextos volátiles y heterogéneos.

Este trabajo contribuye a la literatura de valorización de activos al ofrecer una comparación empírica entre técnicas de machine learning y modelos financieros tradicionales en un entorno controlado de validación cruzada. Los resultados de este análisis pueden ser de utilidad para gestores de portafolio y analistas financieros, quienes enfrentan el reto de balancear precisión y generalización al implementar modelos predictivos en la toma de decisiones de inversión. En definitiva, este estudio busca aportar evidencia sobre el potencial de machine learning para complementar los modelos tradicionales de valorización de activos, al tiempo que destaca las limitaciones y desafíos de aplicar estas técnicas avanzadas en el ámbito financiero.

3 Problema de Investigación

La valorización de activos ha sido un tema central en la teoría financiera y económica desde mediados del siglo XX, con implicaciones profundas para la asignación de recursos, la gestión de riesgos y la toma de decisiones de inversión (Markowitz, 1952). El desafío fundamental radica en determinar el valor justo de un activo en un entorno de mercado caracterizado por la incertidumbre y la información asimétrica (Fama, 1970). Este problema ha evolucionado desde los primeros intentos de formalización, como el Modelo de Valoración de Activos Financieros (CAPM) desarrollado por Sharpe (1964) y Lintner (1965), hasta modelos multifactoriales más sofisticados como los propuestos por Fama y French (1993).

La necesidad de pronosticar los retornos se ha convertido en un aspecto fundamental de la valorización de activos, ya que proporciona una base para la toma de decisiones de inversión y la gestión de riesgos (Campbell y Thompson, 2008). La capacidad de estimar con precisión los retornos futuros no solo es crucial para los inversores que buscan optimizar sus carteras, sino también para las empresas que necesitan evaluar el costo de capital para sus decisiones de inversión y financiamiento (Fama y French, 2004). Además, los pronósticos de retornos desempeñan un papel vital en la evaluación de la eficiencia del mercado y en la identificación de posibles anomalías (Cochrane, 2011). Sin embargo, la tarea de pronosticar retornos es notoriamente difícil debido a la naturaleza estocástica de los mercados financieros, la presencia de ruido en los datos y la posibilidad de cambios estructurales en la economía (Goyal y Welch, 2008). Esta dificultad ha llevado al desarrollo de una variedad de técnicas de pronóstico, desde modelos econométricos tradicionales hasta enfoques más recientes basados en aprendizaje automático, cada uno con sus propias fortalezas y limitaciones (Gu *et al.*, 2020).

En este contexto, los modelos de valorización de activos como el CAPM y los modelos multifactoriales de Fama-French se han convertido en herramientas fundamentales para pronosticar retornos esperados. Estos modelos típicamente se implementan mediante un proceso de dos etapas (Fama y French, 2015). En la primera etapa, se estiman los coeficientes (betas) para cada activo utilizando regresiones de series temporales. Estos betas representan la sensibilidad de los retornos del activo a varios factores de riesgo, como el exceso de retorno del mercado en el CAPM, o factores adicionales como tamaño y valor en los modelos de Fama-French. En la segunda etapa, estos betas estimados se utilizan en regresiones de sección transversal para determinar las primas de riesgo asociadas con cada factor. Finalmente, los retornos esperados se pronostican combinando los betas estimados con las primas de riesgo (Cochrane, 2011).

Sin embargo, este enfoque enfrenta una limitación significativa conocida como overfitting o sobreajuste in-sample (Harvey *et al.*, 2016). El overfitting ocurre cuando un modelo estadístico se ajusta demasiado bien a los datos de entrenamiento, capturando no solo las relaciones subyacentes verdaderas, sino también el ruido aleatorio presente en la muestra específica. En el contexto de la valorización de activos, esto significa que los modelos pueden parecer tener un excelente desempeño cuando se prueban en los mismos datos utilizados para su estimación (in-sample), pero fallar en generalizar a nuevos datos (out-of-sample) (McLean y Pontiff, 2016).

El problema del overfitting se agrava particularmente en los modelos multifactoriales más complejos. A medida que se añaden más factores al modelo, aumenta su flexibilidad para ajustarse a peculiaridades específicas de la muestra de datos utilizada, potencialmente capturando relaciones espurias que no se mantienen fuera de la muestra (Gu *et al.*, 2020). Esto plantea un dilema fundamental en la modelización financiera: por un lado, se busca capturar la complejidad de los mercados financieros mediante la inclusión de múltiples factores; por otro lado, esta misma complejidad

aumenta el riesgo de overfitting y reduce la robustez y generalización del modelo.

Para abordar estas limitaciones, la investigación reciente en valorización de activos ha enfatizado la importancia de la validación out-of-sample y la necesidad de técnicas más robustas para la selección y evaluación de modelos (McLean y Pontiff, 2016). Esto ha llevado al desarrollo de enfoques como la validación cruzada, el uso de períodos de espera (holdout periods), y la implementación de técnicas de regularización para controlar la complejidad del modelo (Gu *et al.*, 2020).

En este contexto, surge la necesidad de explorar métodos alternativos que puedan abordar eficazmente el problema del overfitting mientras mantienen la capacidad de capturar las complejas relaciones en los mercados financieros. Los métodos de machine learning, con su capacidad para manejar grandes cantidades de datos y relaciones no lineales, se presentan como una alternativa prometedora (Gu *et al.*, 2020).

El problema de investigación de este trabajo se centrará en comparar el rendimiento out-of-sample de diversos modelos de valorización de activos, abordando la limitación crítica del overfitting presente en los enfoques tradicionales.

4 Objetivos

4.1 Objetivo General

Comparar el poder predictivo out-of-sample de los modelos de valorización de activos CAPM, Fama-French de tres factores (FF3) y Fama-French de cinco factores (FF5), utilizando tres métodos de estimación distintos: Mínimos Cuadrados Ordinarios (OLS), Elastic Net y Random Forest, para determinar qué combinación de modelo y método de estimación ofrece los pronósticos más precisos y robustos de los retornos esperados en el mercado financiero.

4.2 Objetivos Específicos

1. Recopilar y preparar un conjunto de datos de retornos para 11 sectores de la economía estadounidense desde el año 2000 hasta 2022.
2. Implementar los modelos CAPM, Fama-French de tres factores (FF3) y Fama-French de cinco factores (FF5) utilizando tres métodos de estimación: Mínimos Cuadrados Ordinarios (OLS), Elastic Net Regression y Random Forest.
3. Aplicar una metodología de rolling windows para evaluar el poder predictivo out-of-sample de cada combinación de modelo y método de estimación a lo largo del tiempo.
4. Realizar el proceso de hyperparameter tuning para los métodos Elastic Net y Random Forest en cada ventana de estimación, optimizando su rendimiento.
5. Calcular y comparar medidas de ajuste y estadísticos relevantes entre los diferentes modelos y métodos de estimación.

5 Marco Teórico

5.1 Valorización de Activos

Los fundamentos de la teoría de valorización de activos fueron establecidos por Markowitz (1952), quien introdujo el concepto de diversificación de carteras y la frontera eficiente. Markowitz demostró cómo los inversores podían optimizar sus carteras al considerar no solo el rendimiento esperado de los activos individuales, sino también las correlaciones entre ellos. Este trabajo pionero estableció el marco para comprender el equilibrio entre riesgo y retorno, sentando las bases para los modelos posteriores de valorización de activos.

Basándose en la teoría de carteras de Markowitz, Tobin (1958) introdujo el concepto de cartera "super-eficiente", que más tarde se conocería como la cartera de mercado. El trabajo de Tobin proporcionó un puente crucial entre la selección de activos individuales y el equilibrio del mercado, avanzando significativamente en el desarrollo de la teoría de valorización de activos.

La intuición económica detrás de la valorización de activos se basa en el principio de que, en mercados de capitales eficientes, el precio de un activo debe reflejar sus flujos de efectivo futuros, descontados a una tasa que compense a los inversores por el valor temporal del dinero y los riesgos asociados. Este principio se resume en la ecuación fundamental de valorización:

$$P_0 = \sum_{t=1}^{\infty} \frac{E(CF_t)}{(1+r)^t}, \quad (1)$$

donde P_0 es el precio actual del activo, $E(CF_t)$ es el flujo de efectivo esperado en el tiempo t , y r es la tasa de descuento apropiada.

El desafío en la valorización de activos radica en determinar la tasa de descuento adecuada,

la cual debe reflejar el riesgo inherente al activo. Aquí es donde los modelos de valorización de activos entran en juego, al intentar cuantificar la relación entre riesgo y retorno esperado. Como señala Cochrane (2009), la pregunta central de la valorización de activos es: "¿Cuál es el valor de este pago, o de esta serie de pagos?". La búsqueda de una respuesta a esta pregunta ha impulsado décadas de investigación teórica y empírica, conduciendo a avances significativos en nuestra comprensión de los mercados financieros y la relación entre riesgo y retorno.

5.1.1 Modelo de valorización de Activos Financieros (CAPM)

El Modelo de valorización de Activos Financieros (CAPM, por sus siglas en inglés), desarrollado independientemente por Sharpe (1964) y Lintner (1965), representó un avance significativo en la formalización de la relación riesgo-retorno. Basándose en el trabajo previo de Markowitz (1952) sobre selección de carteras, el CAPM proporcionó un marco sencillo pero poderoso para la valorización de activos individuales en el contexto de carteras bien diversificadas.

La idea central del CAPM es que, en un equilibrio de mercado, el retorno esperado de un activo debe ser proporcional a su contribución al riesgo general de una cartera bien diversificada. Esta contribución se mide mediante el "beta" del activo, que cuantifica su sensibilidad a los movimientos del mercado en su conjunto. Por lo tanto, el CAPM reduce el complejo mundo de la valorización de activos a un modelo de un solo factor, donde la prima de riesgo del mercado es el único determinante de los retornos esperados más allá de la tasa libre de riesgo.

Formalmente, el CAPM se expresa como:

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f], \quad (2)$$

donde $E(R_i)$ es el retorno esperado del activo i , R_f es la tasa libre de riesgo, β_i es el beta del activo i , y $E(R_m)$ es el retorno esperado de la cartera de mercado. El término $[E(R_m) - R_f]$ representa la prima de riesgo del mercado.

El coeficiente beta, β_i , que captura la sensibilidad del activo a los movimientos del mercado, se calcula como:

$$\beta_i = \frac{\text{Cov}(R_i, R_m)}{\text{Var}(R_m)}, \quad (3)$$

donde $\text{Cov}(R_i, R_m)$ es la covarianza entre el retorno del activo y el retorno del mercado, y $\text{Var}(R_m)$ es la varianza del retorno del mercado.

A pesar de su elegancia teórica, las pruebas empíricas del CAPM revelaron varias anomalías que el modelo no podía explicar adecuadamente. Estas anomalías, o patrones en los retornos promedio que no se explicaban únicamente por la beta del mercado, desafiaron la suficiencia del CAPM de un solo factor y estimularon la investigación de factores explicativos adicionales.

Uno de los primeros y más destacados desafíos al CAPM provino de Banz (1981), quien documentó el "efecto tamaño". Banz encontró que las empresas más pequeñas, medidas por capitalización de mercado, tendían a tener retornos promedio más altos que las empresas más grandes, incluso después de controlar por la beta del mercado. Esta prima de tamaño no era explicada por el CAPM, sugiriendo que el tamaño podría ser un proxy de algún otro factor de riesgo no capturado por la beta del mercado.

Posteriormente, Rosenberg *et al.* (1985) identificaron el "efecto valor", mostrando que las empresas con altos ratios de valor en libros a valor de mercado (acciones de valor) tendían a superar a aquellas con bajos ratios (acciones de crecimiento), nuevamente después de controlar por la beta del mercado. Esta prima de valor fue corroborada además por Fama y French (1992),

quienes demostraron que la combinación de tamaño y ratio de valor en libros a valor de mercado proporcionaba una poderosa explicación para la sección transversal de los retornos promedio de las acciones.

Otras anomalías emergieron, como el efecto momentum documentado por Jegadeesh y Titman (1993), donde las acciones que habían tenido un buen desempeño en el pasado reciente continuaban superando al mercado en el futuro cercano, y la prima de rentabilidad identificada por Novy-Marx (2013), donde las empresas más rentables obtenían retornos más altos que las menos rentables.

Estos hallazgos empíricos plantearon desafíos significativos al CAPM. Como señalan Fama y French (2004), el fracaso del CAPM para explicar estos patrones en los retornos promedio podría deberse a verdaderas ineficiencias del mercado o a deficiencias en el modelo mismo, específicamente la suposición de que la beta del mercado es suficiente para describir la relación riesgo-retorno.

En respuesta a estos desafíos, los investigadores comenzaron a explorar modelos multifactoriales que pudieran explicar estas anomalías. Basándose en la Teoría de valorización por Arbitraje de Ross (1976), que proporcionó un marco teórico para los modelos de valorización de activos multifactoriales, diversos investigadores propusieron y probaron diferentes combinaciones de factores.

5.1.2 Modelo de Tres Factores de Fama & French

En el contexto de la creciente evidencia empírica que cuestionaba la suficiencia del CAPM de un solo factor y la búsqueda de un modelo de valorización de activos más robusto, Fama y French (1993) introdujeron su modelo de tres factores. Al incorporar los factores de tamaño y valor

junto con la beta de mercado, Fama y French buscaban capturar las anomalías más prominentes observadas en los retornos de las acciones y proporcionar una descripción más completa y precisa de los retornos promedio.

El modelo de tres factores de Fama-French (FF3) representa un avance significativo en la teoría de valorización de activos, abordando directamente las limitaciones del CAPM al integrar factores adicionales que reflejan dimensiones importantes del riesgo que no estaban siendo capturadas. Este modelo postula que el retorno esperado de una cartera en exceso de la tasa libre de riesgo puede explicarse por la exposición de la cartera a tres factores clave:

El factor de riesgo de mercado ($E(R_m) - R_f$): Como en el CAPM, este factor captura la prima de riesgo asociada con el mercado en su conjunto. Representa el retorno adicional que los inversores esperan por asumir el riesgo sistemático inherente a la inversión en acciones en lugar de activos libres de riesgo.

El factor de tamaño (SMB, Small Minus Big): Este factor captura el rendimiento adicional que históricamente han ofrecido las empresas de pequeña capitalización en comparación con las de gran capitalización. La justificación económica detrás del factor de tamaño radica en que las empresas más pequeñas suelen ser más riesgosas debido a su menor diversificación, mayores costos de financiamiento, menor liquidez y mayor sensibilidad a condiciones económicas adversas. Por lo tanto, los inversores exigen una prima por asumir este riesgo adicional.

El factor de valor (HML, High Minus Low): Este factor refleja la prima asociada con invertir en empresas con altos ratios de valor en libros a valor de mercado (acciones de valor) en comparación con aquellas con bajos ratios (acciones de crecimiento). Las acciones de valor suelen representar empresas que están subvaloradas por el mercado o enfrentan dificultades financieras, lo que implica un mayor riesgo. La teoría sugiere que los inversores requieren una prima por invertir

en estas empresas, ya que enfrentan incertidumbres mayores en términos de ganancias futuras y estabilidad financiera.

Formalmente, el modelo FF3 se expresa como:

$$E(R_i) - R_f = \beta_i[E(R_m) - R_f] + s_i E(\text{SMB}) + h_i E(\text{HML}), \quad (4)$$

donde $E(R_i)$ es el retorno esperado del activo i , R_f es la tasa libre de riesgo, β_i es la sensibilidad del activo al factor de mercado, s_i es la sensibilidad al factor de tamaño, y h_i es la sensibilidad al factor de valor.

La construcción de los factores SMB y HML es fundamental para el modelo. El factor SMB se calcula tomando la diferencia entre los retornos promedio de carteras de empresas pequeñas y grandes. Específicamente, se forman carteras agrupando empresas según su capitalización de mercado y luego se calcula el retorno diferencial. De manera similar, el factor HML se construye tomando la diferencia entre los retornos promedio de carteras de empresas con altos y bajos ratios de valor en libros a valor de mercado. Estas carteras se forman clasificando las empresas según sus ratios y calculando el retorno diferencial entre las de alto y bajo ratio.

La inclusión de estos factores adicionales permite al modelo FF3 capturar mejor las variaciones en los retornos de las acciones que no son explicadas por el riesgo de mercado general. Los factores de tamaño y valor representan riesgos adicionales que los inversores consideran al tomar decisiones de inversión, y por los cuales exigen una compensación en forma de retornos esperados más altos.

La justificación teórica para estos factores se basa en la idea de que las empresas pequeñas y de valor están sujetas a riesgos sistémicos específicos que no pueden diversificarse. Por ejemplo,

las empresas pequeñas pueden ser más sensibles a cambios en las condiciones económicas, tener acceso limitado al crédito y enfrentar mayores costos operativos relativos. Las empresas de valor pueden estar en industrias en declive o enfrentando problemas financieros, lo que las hace más vulnerables en períodos de recesión.

Empíricamente, el modelo FF3 ha demostrado un poder explicativo superior en comparación con el CAPM. Fama y French (1996) proporcionan evidencia de que el modelo de tres factores captura una porción significativa de las variaciones en los retornos promedio de las acciones a través de diferentes períodos y mercados. Esto ha llevado a una amplia adopción del modelo tanto en la investigación académica como en aplicaciones prácticas en la gestión de carteras y evaluación del desempeño.

Sin embargo, el modelo FF3 no está exento de críticas. Algunos investigadores argumentan que los factores de tamaño y valor pueden no representar verdaderos riesgos sistémicos, sino que podrían ser el resultado de ineficiencias del mercado o comportamiento de los inversores. Otros han propuesto la inclusión de factores adicionales para capturar otras anomalías observadas, lo que ha llevado al desarrollo de modelos multifactoriales más complejos.

En resumen, el modelo de tres factores de Fama-French amplía el marco de valorización de activos al incorporar factores adicionales que reflejan dimensiones importantes del riesgo no capturadas por el CAPM. Al considerar los riesgos asociados con el tamaño y el valor, el modelo proporciona una herramienta más robusta para entender y predecir los retornos esperados de las acciones en un contexto de cartera bien diversificada.

5.1.3 Modelo de Cinco Factores de Fama & French

Para mejorar la capacidad explicativa del modelo de tres factores (FF3) y capturar anomalías adicionales observadas en los retornos de las acciones, Fama y French (2015) extendieron su modelo incorporando dos factores adicionales relacionados con la rentabilidad y las políticas de inversión de las empresas. El modelo de cinco factores (FF5) se expresa matemáticamente como:

$$E(R_i) - R_f = \beta_i[E(R_m) - R_f] + s_i E(\text{SMB}) + h_i E(\text{HML}) + r_i E(\text{RMW}) + c_i E(\text{CMA}), \quad (5)$$

donde:

- $E(\text{RMW})$ es la prima esperada del factor de rentabilidad (*Robust Minus Weak*).
- $E(\text{CMA})$ es la prima esperada del factor de inversión (*Conservative Minus Aggressive*).

El factor RMW captura la diferencia en los retornos entre empresas con alta y baja rentabilidad operativa. Se define como:

$$\text{RMW} = \text{Retornos de carteras con alta rentabilidad} - \text{Retornos de carteras con baja rentabilidad}. \quad (6)$$

La rentabilidad operativa se mide comúnmente como el ingreso operativo dividido por el patrimonio contable. Este factor se basa en la observación empírica de que las empresas más rentables tienden a obtener retornos superiores. La justificación es que las empresas con alta rentabilidad generan mayores flujos de efectivo y tienen mayor capacidad para soportar condiciones económicas adversas, lo que reduce su riesgo sistemático. Sin embargo, si el mercado no precifica adecuadamente esta menor percepción de riesgo, los inversores pueden obtener una prima por invertir

en empresas más rentables.

El factor CMA mide la diferencia en los retornos entre empresas con políticas de inversión conservadoras y agresivas. Se calcula como:

$$\text{CMA} = \text{Retornos de carteras con inversión conservadora} - \text{Retornos de carteras con inversión agresiva.} \quad (7)$$

La tasa de inversión se determina generalmente como el crecimiento en los activos totales dividido por los activos totales del período anterior. Las empresas que invierten agresivamente pueden estar persiguiendo proyectos de menor rentabilidad marginal o enfrentando mayores riesgos financieros, lo que puede conducir a retornos futuros más bajos. Por el contrario, las empresas con políticas de inversión conservadoras tienden a ser más selectivas, enfocándose en proyectos con mayores retornos ajustados por riesgo. Este factor refleja cómo las decisiones de inversión de las empresas afectan su perfil de riesgo y, por ende, los retornos esperados de sus acciones.

La inclusión de los factores RMW y CMA en el modelo FF5 se fundamenta en evidencia empírica que muestra que la rentabilidad y las políticas de inversión son determinantes significativos de los retornos de las acciones. Estos factores capturan dimensiones del riesgo sistemático no consideradas en el modelo FF3:

- **Rentabilidad (RMW):** Las empresas con alta rentabilidad pueden estar expuestas a menores riesgos financieros y operativos, lo que afecta su sensibilidad a factores macroeconómicos. Los inversores exigen una prima por invertir en empresas menos rentables debido al mayor riesgo asociado.
- **Inversión (CMA):** Las tasas de inversión reflejan las expectativas de crecimiento y el opti-

mismo de la gerencia. Las empresas que invierten agresivamente pueden ser más vulnerables a fluctuaciones económicas, aumentando su riesgo sistémico. Los inversores requieren una compensación por asumir el riesgo adicional asociado con empresas que tienen altas tasas de inversión.

Mientras que el modelo FF3 considera los factores de mercado, tamaño y valor para explicar los retornos esperados, el modelo FF5 amplía este enfoque al integrar RMW y CMA. Estos nuevos factores permiten capturar anomalías adicionales que el modelo de tres factores no logra explicar completamente, proporcionando una comprensión más integral de las fuentes de riesgo que influyen en los retornos de las acciones.

5.1.4 Incorporación del Factor Momentum al Modelo de Cinco Factores

Además de los factores considerados en el modelo de cinco factores de Fama y French, la literatura financiera ha identificado el *factor momentum* como un determinante significativo de los retornos accionarios. Este factor captura la tendencia de que las acciones que han tenido un buen desempeño en el pasado reciente continúen superando al mercado, y viceversa para las acciones con bajo desempeño (Jegadeesh y Titman, 1993).

El modelo extendido que incluye el factor momentum se expresa matemáticamente como:

$$E(R_i) - R_f = \beta_i[E(R_m) - R_f] + s_i E(\text{SMB}) + h_i E(\text{HML}) + r_i E(\text{RMW}) + c_i E(\text{CMA}) + m_i E(\text{MOM}), \quad (8)$$

donde:

- $E(\text{MOM})$ es la prima esperada del factor momentum.

El factor MOM se calcula como:

$$\text{MOM} = \text{Retornos de carteras de acciones ganadoras} - \text{Retornos de carteras de acciones perdedoras.} \quad (9)$$

Este cálculo se basa en formar carteras según el desempeño pasado de las acciones, generalmente utilizando horizontes de 3 a 12 meses. Las acciones ganadoras son aquellas que han obtenido altos retornos en el período reciente, mientras que las perdedoras han obtenido bajos retornos.

La inclusión del factor momentum en el modelo FF5+M permite capturar anomalías adicionales no explicadas por los otros factores. Carhart (1997) demostró que el momentum es un factor persistente en la predicción de los retornos accionarios y propuso un modelo de cuatro factores que incluye este elemento. Aunque Fama y French inicialmente se mostraron escépticos sobre la inclusión del momentum debido a su naturaleza transitoria (Fama y French, 1996), estudios posteriores han confirmado su relevancia en distintos mercados y períodos (Moskowitz *et al.*, 2012).

5.2 Métodos Estadísticos

La investigación empírica de valorización de activos se ha basado tradicionalmente en técnicas estadísticas relativamente simples, principalmente la regresión por mínimos cuadrados ordinarios (OLS), para estimar y probar modelos de valorización de activos. Sin embargo, como señalan Gu *et al.* (2020), la creciente complejidad de los mercados financieros, la proliferación de potenciales factores de valorización y la disponibilidad de grandes conjuntos de datos han motivado a los investigadores a explorar enfoques estadísticos más sofisticados.

Esta sección examina una gama de métodos estadísticos empleados en la valorización em-

pírica moderna de activos, pasando de las técnicas clásicas a los algoritmos avanzados de aprendizaje automático. Siguiendo el marco establecido por Gu *et al.* (2020), comenzamos con OLS como punto de referencia, luego discutimos métodos de regresión penalizada como la regresión ridge, lasso y elastic net que ayudan a abordar la alta dimensionalidad y el potencial sobreajuste. También cubrimos enfoques no lineales como árboles de regresión y bosques aleatorios que pueden capturar interacciones complejas entre variables.

5.2.1 Mínimos Cuadrados Ordinarios (OLS)

La regresión por Mínimos Cuadrados Ordinarios (OLS, por sus siglas en inglés) es un método estadístico fundamental ampliamente utilizado en la valorización de activos para estimar los parámetros de modelos lineales de factores (Cochrane, 2009). Este método es apreciado por su simplicidad y por las sólidas propiedades estadísticas de los estimadores que produce bajo ciertas condiciones. Aunque a menudo se introduce en el contexto de la regresión lineal simple, OLS es particularmente potente cuando se aplica a múltiples variables, lo que lo hace adecuado para analizar modelos complejos de valorización de activos.

En el contexto multivariado, el modelo de regresión lineal puede ser expresado como:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + \epsilon_i, \quad i = 1, 2, \dots, n \quad (10)$$

donde:

- y_i es la variable dependiente (por ejemplo, retornos de activos).
- x_{ji} son las variables independientes (por ejemplo, retornos de factores).

- β_j son los coeficientes a estimar.
- ϵ_i es el término de error aleatorio.

El objetivo del método OLS es encontrar los valores de los coeficientes β_j que minimicen la suma de los cuadrados de los residuos (diferencias entre los valores observados y los predichos por el modelo):

$$\min_{\beta} S(\beta) = \min_{\beta} \left\{ \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ji} \right)^2 \right\} \quad (11)$$

Para simplificar la notación y facilitar la manipulación matemática, es común expresar el modelo en forma matricial:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon \quad (12)$$

donde:

- $\mathbf{y}$ es el vector $n \times 1$ de observaciones de la variable dependiente.
- $\mathbf{X}$ es la matriz $n \times (p + 1)$ de variables independientes (incluyendo una columna de unos para el término intercepto).
- β es el vector $(p + 1) \times 1$ de coeficientes a estimar.
- ϵ es el vector $n \times 1$ de términos de error.

El problema de minimización de OLS se convierte entonces en:

$$\min_{\beta} S(\beta) = \min_{\beta} (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta) \quad (13)$$

Para encontrar el vector de coeficientes $\hat{\beta}$ que minimiza $S(\beta)$, tomamos la derivada parcial respecto

a β y la igualamos a cero:

$$\frac{\partial S(\beta)}{\partial \beta} = -2\mathbf{X}'(\mathbf{y} - \mathbf{X}\beta) = 0 \quad (14)$$

Resolviendo para β , obtenemos las ecuaciones normales:

$$\mathbf{X}'\mathbf{X}\beta = \mathbf{X}'\mathbf{y} \quad (15)$$

Siempre que $\mathbf{X}'\mathbf{X}$ sea invertible, la solución para β está dada por:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (16)$$

Este estimador posee varias propiedades deseables bajo los supuestos clásicos del modelo de regresión lineal, conocidos como los supuestos de Gauss-Markov:

1. **Linealidad en los parámetros:** El modelo es lineal en los coeficientes β .
2. **Exogeneidad:** Los errores ϵ_i tienen esperanza cero y son no correlacionados con las variables independientes, es decir, $E[\epsilon_i|\mathbf{X}] = 0$.
3. **Homoscedasticidad:** Los errores tienen varianza constante, $\text{Var}(\epsilon_i|\mathbf{X}) = \sigma^2$.
4. **No autocorrelación:** Los errores ϵ_i son no correlacionados entre sí, $\text{Cov}(\epsilon_i, \epsilon_j|\mathbf{X}) = 0$ para $i \neq j$.
5. **No multicolinealidad perfecta:** Las variables independientes no están perfectamente correlacionadas, lo que asegura que $\mathbf{X}'\mathbf{X}$ es invertible.

Bajo estos supuestos, el estimador OLS es el *estimador lineal insesgado de mínima varianza* (BLUE, por sus siglas en inglés), según el teorema de Gauss-Markov.

La matriz de varianza-covarianza del estimador $\hat{\beta}$ está dada por:

$$\text{Var}(\hat{\beta}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \quad (17)$$

donde σ^2 es la varianza del término de error, que puede estimarse como:

$$\hat{\sigma}^2 = \frac{1}{n - p - 1} (\mathbf{y} - \mathbf{X}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\beta}) \quad (18)$$

Esta estimación es crucial para construir intervalos de confianza y realizar pruebas de hipótesis sobre los coeficientes estimados.

En aplicaciones de valorización de activos, OLS es utilizado para estimar simultáneamente las sensibilidades (betas) de los retornos de activos respecto a múltiples factores de riesgo. Por ejemplo, al probar el modelo de cinco factores de Fama-French, el modelo de regresión sería:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{iM}(R_{Mt} - R_{ft}) + s_i\text{SMB}_t + h_i\text{HML}_t + r_i\text{RMW}_t + c_i\text{CMA}_t + \epsilon_{it} \quad (19)$$

donde:

- R_{it} es el retorno del activo i en el tiempo t .
- R_{ft} es la tasa libre de riesgo en el tiempo t .
- R_{Mt} es el retorno del mercado.
- $\text{SMB}_t, \text{HML}_t, \text{RMW}_t, \text{CMA}_t$ son los factores adicionales de Fama-French.

- α_i es el término de rendimiento anormal.
- $\beta_{iM}, s_i, h_i, r_i, c_i$ son las sensibilidades del activo i a cada factor.
- ϵ_{it} es el término de error.

Este enfoque permite evaluar el poder explicativo incremental de cada factor mientras se controlan los efectos de los demás. Sin embargo, a medida que aumenta el número de factores potenciales, surgen varios desafíos estadísticos y econométricos que pueden afectar la fiabilidad de los estimadores OLS.

Uno de los principales desafíos es la **multicolinealidad**, que ocurre cuando las variables independientes están altamente correlacionadas entre sí. Matemáticamente, esto implica que la matriz $\mathbf{X}'\mathbf{X}$ es casi singular, lo que puede llevar a estimaciones inestables de $\hat{\beta}$ y a varianzas infladas de los estimadores. Esto dificulta la interpretación individual de los coeficientes y puede reducir la potencia de las pruebas estadísticas.

Otro problema es el **sobreajuste**. Al incluir demasiados factores, el modelo puede capturar el ruido específico de la muestra en lugar de la verdadera relación subyacente, lo que resulta en un rendimiento predictivo pobre fuera de la muestra. Esto se manifiesta matemáticamente en un aumento de la varianza de los estimadores y una disminución de su capacidad para generalizar.

Además, existe el riesgo de identificar **correlaciones espurias** cuando se consideran muchos factores. Estadísticamente, al realizar múltiples pruebas, aumenta la probabilidad de obtener resultados significativos por azar, lo que es conocido como el problema de *minería de datos* o *p-hacking* (Harvey *et al.*, 2016).

Estos desafíos han llevado a los investigadores a explorar técnicas de regularización y métodos más avanzados. Las técnicas de regularización, como la regresión ridge y el LASSO, intro-

ducen penalizaciones en la función de minimización, modificando el criterio de OLS para controlar la complejidad del modelo y mejorar su capacidad predictiva.

Kozak *et al.* (2020) demuestran el poder de estas técnicas de regularización en la valorización de activos, mostrando cómo pueden ser utilizadas para construir modelos de factores parsimoniosos a partir de un gran conjunto de características. Su trabajo destaca el potencial de la regresión ridge, LASSO y métodos relacionados para cerrar la brecha entre los modelos de factores tradicionales y la vasta gama de anomalías documentadas en la literatura.

A pesar de sus limitaciones, utilizaremos el método de OLS como punto de referencia a lo largo de este trabajo para ilustrar los potenciales beneficios de métodos más complejos. Al comparar los resultados obtenidos con OLS frente a los obtenidos con técnicas avanzadas, podremos evaluar de manera más efectiva las mejoras en términos de estabilidad del modelo y rendimiento predictivo. Esto nos permitirá comprender mejor en qué contextos las técnicas tradicionales son suficientes y cuándo es necesario recurrir a métodos más sofisticados para capturar las dinámicas subyacentes en la valorización de activos.

5.2.2 Regresión Ridge

La **Regresión Ridge**, también conocida como regresión de crestas, es una técnica de regularización que modifica el método de Mínimos Cuadrados Ordinarios (OLS) para abordar algunos de sus problemas inherentes, como la multicolinealidad y el sobreajuste (Hoerl y Kennard, 1970). Al introducir un término de penalización en la función de minimización, Ridge Regression restringe los coeficientes estimados, reduciendo su varianza y mejorando la capacidad predictiva del modelo, especialmente cuando se trabaja con un gran número de variables correlacionadas. La Regresión Ridge modifica el criterio de minimización al añadir un término de penalización pro-

porcional al cuadrado de la norma de los coeficientes:

$$\min_{\beta} S_{\text{Ridge}}(\beta) = \min_{\beta} \{(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta) + \lambda\beta'\beta\} \quad (20)$$

donde $\lambda \geq 0$ es el **parámetro de regularización** que controla la magnitud de la penalización. Cuando $\lambda = 0$, el criterio se reduce al de OLS. A medida que λ aumenta, la penalización sobre los coeficientes se vuelve más fuerte. El término de penalización $\lambda\beta'\beta$ tiene varias implicaciones importantes:

- **Reducción de la varianza:** Al penalizar los coeficientes grandes, se reduce la varianza de los estimadores, lo que puede mejorar la capacidad predictiva del modelo en nuevas muestras.
- **Control de la multicolinealidad:** La penalización hace que los coeficientes de variables correlacionadas se reduzcan, mitigando los efectos negativos de la multicolinealidad.
- **Sesgo-introducción:** A cambio de una varianza reducida, se introduce un sesgo en los estimadores. Sin embargo, este trade-off puede resultar en un error cuadrático medio (MSE) total más bajo.

La solución para β en la Regresión Ridge se obtiene al resolver:

$$\frac{\partial S_{\text{Ridge}}(\beta)}{\partial \beta} = -2\mathbf{X}'(\mathbf{y} - \mathbf{X}\beta) + 2\lambda\beta = 0 \quad (21)$$

Reorganizando términos, tenemos las **ecuaciones de normalidad modificadas**:

$$(\mathbf{X}'\mathbf{X} + \lambda\mathbf{I})\beta = \mathbf{X}'\mathbf{y} \quad (22)$$

donde $\mathbf{I}$ es la matriz identidad de dimensión $(p + 1) \times (p + 1)$. La solución para β es entonces:

$$\hat{\beta}_{\text{Ridge}} = (\mathbf{X}'\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}'\mathbf{y} \quad (23)$$

Es importante notar que, a diferencia del método OLS, la matriz $(\mathbf{X}'\mathbf{X} + \lambda\mathbf{I})$ es invertible incluso cuando $\mathbf{X}'\mathbf{X}$ no lo es (por ejemplo, cuando hay multicolinealidad perfecta o $p > n$), siempre que $\lambda > 0$.

La Regresión Ridge modifica la estimación de los coeficientes al equilibrar el ajuste del modelo y la complejidad de los coeficientes. Matemáticamente, esto se traduce en un **sesgo hacia cero** de los coeficientes estimados. Este sesgo es controlado por el parámetro λ :

- **Valores pequeños de λ :** El modelo se acerca al OLS, con coeficientes menos penalizados y potencialmente mayor varianza.
- **Valores grandes de λ :** Los coeficientes se acercan a cero, lo que puede llevar a un modelo demasiado simplificado si se penaliza en exceso.

La elección óptima de λ es crucial y generalmente se determina mediante métodos de validación cruzada, buscando el valor que minimiza el error de predicción en datos no vistos.

Desde una perspectiva geométrica, la Regresión Ridge busca soluciones dentro de una esfera (o hiperesfera en dimensiones superiores) definida por el término de penalización, en contraste con el espacio completo de soluciones de OLS. Esto impone restricciones en la magnitud de los coeficientes, evitando soluciones con coeficientes excesivamente grandes que pueden resultar de la multicolinealidad.

5.2.3 Regresión LASSO

La **Regresión LASSO** (Least Absolute Shrinkage and Selection Operator) es una técnica de regularización y selección de variables que mejora el modelo de Mínimos Cuadrados Ordinarios (OLS) al introducir una penalización basada en la norma $L1$ de los coeficientes (Tibshirani, 1996). Esta metodología es especialmente útil cuando se trabaja con conjuntos de datos que contienen un gran número de variables explicativas, algunas de las cuales pueden ser irrelevantes o redundantes. La Regresión LASSO modifica el criterio de minimización de OLS al añadir un término de penalización proporcional a la suma de los valores absolutos de los coeficientes:

$$\min_{\beta} S_{\text{LASSO}}(\beta) = \min_{\beta} \left\{ \frac{1}{2n} |\mathbf{y} - \mathbf{X}\beta|_2^2 + \lambda |\beta|_1 \right\} \quad (24)$$

donde:

- $|\mathbf{y} - \mathbf{X}\beta|_2^2$ es la suma de los cuadrados de los residuos (norma $L2$).
- $|\beta|_1 = \sum_j |\beta_j|$ es la norma $L1$ de los coeficientes.
- $\lambda \geq 0$ es el parámetro de regularización que controla la magnitud de la penalización.

La principal distinción entre LASSO y otras técnicas de regularización, como la Regresión Ridge, radica en el uso de la norma $L1$ en el término de penalización. Esto tiene varias implicaciones clave:

1. **Selección de Variables Integrada:** Debido a la naturaleza de la penalización $L1$, LASSO tiene la propiedad de forzar algunos coeficientes exactamente a cero cuando λ es suficientemente grande. Esto implica que efectúa una selección de variables automática, eliminando aquellas que no contribuyen significativamente al modelo.

2. **Sparsity (Esparsidad):** El modelo resultante es *esparso*, es decir, contiene un subconjunto reducido de variables explicativas con coeficientes distintos de cero. Esto mejora la interpretabilidad y puede revelar relaciones subyacentes más claras.
3. **Trade-off Sesgo-Varianza:** Al igual que otras técnicas de regularización, LASSO introduce un sesgo en los estimadores a cambio de una reducción en la varianza. La penalización controla este trade-off, y la elección adecuada de λ es crucial para optimizar el desempeño del modelo.

La penalización $L1$ en LASSO impone una restricción que puede ser visualizada geométricamente como un rombo (o hipercubo en dimensiones superiores) en el espacio de los coeficientes. La solución de LASSO se encuentra en la intersección de este rombo con las elipses de nivel de la función de pérdida (suma de cuadrados de residuos).

Debido a la forma angular del rombo, es más probable que la solución óptima se encuentre en un vértice, lo que corresponde a coeficientes exactamente cero para algunas variables. Esto contrasta con la penalización $L2$ de Ridge, que tiene una restricción circular y, por lo tanto, tiende a producir coeficientes pequeños pero generalmente no exactamente cero.

5.2.4 Regresión Elastic Net

La **Regresión Elastic Net** es una técnica de regularización que combina las penalizaciones de la norma $L1$ (utilizada en LASSO) y la norma $L2$ (utilizada en Ridge), con el objetivo de aprovechar las ventajas de ambos métodos (Zou y Hastie, 2005). Esta metodología es particularmente útil en situaciones donde el número de variables independientes es grande y existe una alta correlación entre ellas.

La función objetivo de Elastic Net introduce una penalización que es una combinación convexa de las penalizaciones $L1$ y $L2$:

$$\min_{\beta} S_{\text{EN}}(\beta) = \min_{\beta} \left\{ \frac{1}{2n} |\mathbf{y} - \mathbf{X}\beta|_2^2 + \lambda \left(\alpha |\beta|_1 + \frac{1}{2} (1 - \alpha) |\beta|_2^2 \right) \right\} \quad (25)$$

donde:

- $|\mathbf{y} - \mathbf{X}\beta|_2^2 = (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta)$ es la suma de los cuadrados de los residuos.
- $|\beta|_1 = \sum_j |\beta_j|$ es la norma $L1$ de los coeficientes.
- $|\beta|_2^2 = \sum_j \beta_j^2$ es el cuadrado de la norma $L2$ de los coeficientes.
- $\lambda \geq 0$ es el parámetro de regularización que controla la magnitud total de la penalización.
- $\alpha \in [0, 1]$ es el parámetro que determina el balance entre las penalizaciones $L1$ y $L2$:
 - Si $\alpha = 1$, Elastic Net se reduce a LASSO.
 - Si $\alpha = 0$, Elastic Net se convierte en Ridge Regression.

La inclusión de ambos términos de penalización permite que Elastic Net realice selección de variables (a través de la penalización $L1$) y maneje problemas de multicolinealidad (gracias a la penalización $L2$).

Para resolver el problema de minimización, se pueden utilizar métodos numéricos como el *descenso por coordenadas cíclicas*. Debido a que la función objetivo es convexa pero no diferenciable en todos los puntos (por la presencia del término $L1$), no es posible obtener una solución analítica cerrada.

La derivada subgradiente de la función objetivo respecto a β se utiliza para encontrar el mínimo. Para cada coeficiente β_j , la condición óptima viene dada por:

$$-\frac{1}{n}\mathbf{x}'_j(\mathbf{y} - \mathbf{X}\beta) + \lambda(\alpha \cdot \text{sgn}(\beta_j) + (1 - \alpha)\beta_j) = 0 \quad (26)$$

donde $\mathbf{x}_j$ es la j -ésima columna de la matriz $\mathbf{X}$ y $\text{sgn}(\beta_j)$ es la función signo. Este sistema de ecuaciones debe resolverse iterativamente para todos los coeficientes β_j . El proceso general implica:

1. Estandarizar las variables para que tengan media cero y varianza unitaria, lo que evita que la penalización sea afectada por la escala de las variables.
2. Inicializar los coeficientes β (por ejemplo, en cero).
3. Actualizar cada coeficiente β_j manteniendo los demás fijos, utilizando una regla de actualización derivada de la condición óptima.

La regla de actualización para cada coeficiente β_j en Elastic Net es:

$$\beta_j \leftarrow \frac{S\left(\frac{1}{n}\mathbf{x}'_j(\mathbf{y} - \mathbf{X}_{-j}\beta_{-j}), \lambda\alpha\right)}{1 + \lambda(1 - \alpha)} \quad (27)$$

donde:

- $S(z, \gamma)$ es la función de *soft-thresholding* definida como $S(z, \gamma) = \text{sgn}(z) \cdot \max(|z| - \gamma, 0)$.
- $\mathbf{X}_{-j}$ es la matriz $\mathbf{X}$ excluyendo la columna j .
- β_{-j} es el vector de coeficientes excluyendo β_j .

Este proceso se repite hasta que la convergencia es alcanzada, es decir, hasta que los cambios en los coeficientes entre iteraciones son menores que un umbral predefinido.

La elección de los parámetros λ y α es crítica y generalmente se realiza mediante validación cruzada. El parámetro λ controla el grado total de regularización, mientras que α ajusta el balance entre las penalizaciones $L1$ y $L2$. Valores intermedios de α (por ejemplo, $\alpha = 0.5$) permiten obtener un compromiso entre la selección de variables y la estabilización de los coeficientes.

5.2.5 Árboles de Decisión

Los **Árboles de Decisión** son modelos predictivos que utilizan una estructura de árbol para representar una serie de decisiones y sus posibles consecuencias (Breiman *et al.*, 1984). Se utilizan tanto para problemas de clasificación como de regresión, y son especialmente útiles por su interpretabilidad y simplicidad.

El objetivo principal en un árbol de decisión es dividir iterativamente el espacio de características en regiones homogéneas en términos de la variable objetivo. Esto se logra mediante la selección de puntos de corte en las variables independientes que maximizan la pureza de las particiones resultantes.

Para formalizar el proceso de particionado, consideremos un conjunto de datos $(\mathbf{x}_i, y_i)_{i=1}^n$, donde $\mathbf{x}_i \in \mathbb{R}^p$ es el vector de características y y_i es la variable objetivo. El árbol se construye recursivamente seleccionando en cada nodo la variable j y el punto de corte s que resuelven el siguiente problema de optimización:

$$\min_{j,s} \left\{ \min_{c_1} \sum_{x_i^j \leq s} L(y_i, c_1) + \min_{c_2} \sum_{x_i^j > s} L(y_i, c_2) \right\} \quad (28)$$

donde:

- x_i^j es el valor de la j -ésima característica para el i -ésimo individuo.
- $L(y_i, c)$ es una función de pérdida que mide la discrepancia entre el valor real y_i y la predicción constante c en la partición.

En problemas de clasificación, comúnmente se utilizan medidas de impureza como el índice de Gini o la entropía para la función de pérdida:

1. Índice de Gini:

$$G(t) = \sum_{k=1}^K p_k(t)[1 - p_k(t)] \quad (29)$$

2. Entropía:

$$H(t) = - \sum_{k=1}^K p_k(t) \log p_k(t) \quad (30)$$

donde $p_k(t)$ es la proporción de observaciones de la clase k en el nodo t , y K es el número total de clases.

En problemas de regresión, se suele utilizar el error cuadrático medio como función de pérdida:

$$L(y_i, c) = (y_i - c)^2 \quad (31)$$

El árbol se expande recursivamente hasta que se cumple un criterio de parada, como una profundidad máxima, un número mínimo de muestras en un nodo o una mejora mínima en la reducción de la impureza.

Aunque los árboles de decisión son fáciles de interpretar y pueden capturar relaciones no lineales y de interacción entre variables, presentan ciertas limitaciones:

- **Varianza Alta:** Los árboles tienden a sobreajustar los datos, lo que conduce a una alta varianza y pobre generalización en conjuntos de prueba.
- **Inestabilidad:** Pequeñas variaciones en los datos pueden resultar en árboles significativamente diferentes.

Para mitigar estos problemas, se utilizan métodos de ensamblado como el **Random Forest**. El Random Forest es un método de aprendizaje ensemble que construye múltiples árboles de decisión y combina sus predicciones para mejorar la precisión y la robustez del modelo (Breiman, 2001). La idea principal es reducir la varianza del modelo mediante el promedio (en regresión) o el voto mayoritario (en clasificación) de múltiples árboles entrenados de manera diversa.

La construcción de un Random Forest implica:

1. **Bootstrap Aggregating (Bagging):** Se generan múltiples subconjuntos de datos mediante muestreo con reemplazo del conjunto de entrenamiento original. Cada árbol se entrena en uno de estos subconjuntos.
2. **Selección Aleatoria de Características:** En cada nodo, en lugar de considerar todas las variables para la división, se selecciona aleatoriamente un subconjunto de m variables ($m < p$). Esto introduce diversidad adicional entre los árboles.

Matemáticamente, la predicción agregada del Random Forest para una nueva observación $\mathbf{x}$ es:

$$\hat{y}(\mathbf{x}) = \frac{1}{B} \sum b = 1^B \hat{y}^{(b)}(\mathbf{x}) \quad (32)$$

donde:

- B es el número total de árboles en el bosque.
- $\hat{y}^{(b)}(\mathbf{x})$ es la predicción del b -ésimo árbol.

El uso del bagging y la selección aleatoria de características reduce la correlación entre los árboles individuales, lo que disminuye la varianza total del modelo sin aumentar significativamente el sesgo. Esto resulta en un modelo más generalizable y robusto frente al sobreajuste.

La implementación del Random Forest requiere la elección de varios hiperparámetros clave:

- **Número de Árboles (B):** Un número mayor de árboles puede mejorar la performance pero aumenta el costo computacional.
- **Número de Variables a Considerar en Cada División (m):** Un valor común es $m = \sqrt{p}$ para problemas de clasificación y $m = p/3$ para problemas de regresión.
- **Profundidad Máxima de los Árboles:** Controla el equilibrio entre sesgo y varianza.

6 Metodología

6.1 Datos y división de la muestra

La base de datos utilizada en este estudio consiste en los retornos ajustados mensuales de los portafolios sectoriales de la economía elaborados por *Standard & Poor's* desde marzo del 2002 hasta diciembre 2021. Estos sectores representan diferentes ramas de actividad económica, tales como tecnología, salud, energía y finanzas, entre otros, proporcionando una visión segmentada del desempeño del mercado en cada área específica.

Los datos de los factores empleados en este análisis, incluyendo el factor de mercado, tamaño (SMB), valor (HML), rentabilidad (RMW), inversión (CMA) y momentum (MOM), fueron obtenidos de la biblioteca de datos disponible en la página web de Kenneth French (2023).

Para llevar a cabo el análisis, se realizó una división de la muestra en conjuntos de entrenamiento W_τ , validación V_τ y prueba T_τ , $\tau \in \{1, 2, \dots, 141\}$, mediante una ventana móvil mensual (*rolling window*). Los datos se ordenaron cronológicamente y se estructuraron de la siguiente manera para los métodos de *Machine Learning*: se utilizaron los primeros 60 meses para entrenar los modelos mediante el error cuadrático medio (*MSE*), seguidos de una ventana de validación de 24 meses y una ventana de prueba de 12 meses. Este enfoque permite capturar las tendencias históricas y patrones presentes en los datos, al tiempo que evalúa la capacidad predictiva de los modelos en períodos subsecuentes.

Para OLS, se utilizó una ventana de entrenamiento de 84 meses y una ventana de prueba de 12 meses.

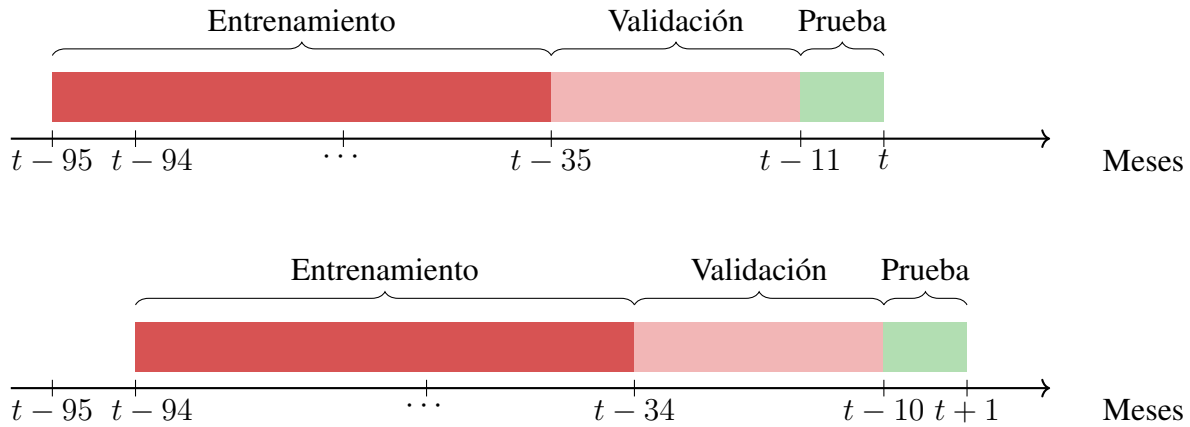


Figura 1: División de la muestra para dos ventanas consecutivas.

6.2 Hiperparametrización

En el caso de los modelos que incluyen hiperparámetros, se implementó un proceso iterativo de ajuste utilizando la ventana de validación. Inicialmente, el modelo se entrenó con un conjunto específico de hiperparámetros, luego se realizaron predicciones y se calculó el *MSE* de pronóstico en la ventana de validación. Posteriormente, el modelo se recalculó utilizando diferentes conjuntos de hiperparámetros. Este proceso se repitió sistemáticamente hasta identificar el conjunto de hiperparámetros que minimiza el *MSE* de pronóstico en la ventana de validación, asegurando así el mejor desempeño predictivo del modelo.

Matemáticamente, definimos los retornos de exceso pronosticados para un período futuro $t+k$, con k perteneciente al conjunto de pronósticos K , como los retornos pronosticados del activo menos la tasa libre de riesgo:

$$\hat{r}_{i,t+k} = \hat{R}_{i,t+k} - R_{f,t+k} \quad (33)$$

El pronóstico de los retornos de exceso se realiza utilizando parámetros (θ) e hiperparámetros (λ) estimados con información hasta el tiempo t , pero considera una matriz de regresoras del

mismo período de pronóstico X_{t+k} :

$$\hat{r}_{i,t+k} = E_{t+k}(r_{i,t+k}) = f(X_{t+k}, \theta, \lambda) \quad (34)$$

Donde:

- X_{t+k} es la matriz de regresoras correspondientes al período de pronóstico $t + k$.
- θ es el vector de parámetros estimados con información hasta el tiempo t .
- λ es el vector de hiperparámetros determinados hasta el tiempo t .
- $k \in K$ es el horizonte de pronóstico 1, 2, ..., 12.

En este contexto, es importante destacar que los parámetros θ y los hiperparámetros λ provienen de estimaciones basadas en datos hasta el tiempo t . Sin embargo, la matriz de regresoras X_{t+k} pertenece al período de pronóstico $t + k$, incorporando así la información más reciente.

A continuación se presenta un algoritmo detallado que explica el proceso de hiperparametrización para los métodos, a excepción de *OLS*:

Algorithm 1 Optimización de Hiperparámetros con Rolling Window

Definir espacio de hiperparámetros $\mathcal{H} = \{\lambda_1 \in \Lambda_1, \lambda_2 \in \Lambda_2, \dots, \lambda_n \in \Lambda_n\}$

for $\tau = 1$ hasta 141 **do**

Definir ventana de entrenamiento W_t , conjunto de validación V_t y conjunto de prueba T_t

for cada $\lambda \in \mathcal{H}$ **do**

Entrenar modelo $M_{\lambda,t}$ usando datos de W_t

Calcular MSE en el conjunto de validación V_t

if $\text{MSE}_{\lambda,t} < \text{MSE}^*$ **then**

$\lambda_t^* = \lambda$

$\text{MSE}_t^* = \text{MSE}_{\lambda,t}$

$M_t^* = M_{\lambda,t}$

end if

end for

end for

Tabla 1: Hiperparámetros para ElasticNet y Random Forest

Método	Parámetro	Rango/Valores
ElasticNet	α	$\in [10^{-4}, 1]$ (50 valores en escala logarítmica)
	ratio_l1	$\in [0, 1, 1, 0]$ (20 valores espaciados linealmente)
Random Forest	profundidad_máxima	$\in \{1, 2, 3, 4, 5, 6\}$
	n_estimadores	$\in \{300, 350, 400, 450, 500\}$
	características_máximas	$\in \{3, 4, 5, 6\}$

Nota: Las características máximas en Random Forest son el mínimo entre las características totales del modelo de valorización y el conjunto expuesto en la tabla.

6.3 Evaluación del rendimiento de los métodos/modelos

Para evaluar el poder predictivo extramuestral de los métodos/modelos y compararlos, agrupamos los errores asociados a los pronósticos de cada portafolio y calculamos el R^2 de la siguiente forma:

$$R^2 = 1 - \frac{\sum_{\tau=1}^{141} \sum_{t \in T_\tau} \sum_{i=1}^N (r_{i,t} - \hat{r}_{i,t})^2}{\sum_{\tau=1}^{141} \sum_{t \in T_\tau} \sum_{i=1}^N (r_{i,t} - \bar{r})^2} \quad (35)$$

donde T_τ indica que los retornos de exceso pertenecen al conjunto de prueba para cada ventana τ , y $\bar{r}$ es el promedio de los retornos de exceso en todos los conjuntos de prueba.

7 Resultados

En la Tabla 3 se presentan los valores de R^2 obtenidos para cada sector económico, método de estimación y modelo utilizado en el análisis de los retornos de exceso de los portafolios sectoriales. Esta tabla permite comparar el poder explicativo y predictivo de los distintos enfoques metodológicos aplicados.

La estructura de la tabla está organizada en columnas que representan los métodos estadísticos utilizados: Mínimos Cuadrados Ordinarios (*OLS*), *Elastic Net* y *Random Forest*. Dentro de cada método, se detallan los modelos empleados: el modelo de valoración de activos de capital (*CAPM*), el modelo de tres factores de Fama y French (*FF3*), el modelo de cinco factores de Fama y French (*FF5*) y el modelo de cinco factores más el factor de momentum (*FF5+M*). Las filas corresponden a los diferentes sectores económicos analizados.

Los valores de R^2 indican la proporción de la variabilidad total de los retornos de exceso que es explicada por el modelo. Un valor de R^2 más cercano a 1 implica un mejor ajuste del modelo a los datos, es decir, una mayor capacidad para explicar las variaciones observadas en los retornos. Los valores en negrita con un asterisco (*) representan el mejor desempeño en cada fila (sector), mientras que los valores en negrita sin asterisco indican el peor desempeño.

Al analizar los resultados, se observa que, en general, los métodos lineales (*OLS* y *Elastic Net*) presentan un mejor desempeño que el método no lineal (*Random Forest*). Por ejemplo, en el sector del Consumo Discrecional, el modelo *CAPM* estimado por *OLS* alcanza un R^2 de 0.839, el valor más alto para este sector, indicando un excelente ajuste. Por el contrario, el mismo sector presenta un R^2 de 0.660 bajo el método *Random Forest* con el modelo *FF5+M*, siendo este el valor

Tabla 2: Valores R^2 por Sector, Método y Modelo

Sector	OLS					Elastic Net			Random Forest		
	CAPM	FF3	FF5	FF5+M	FF3	FF5	FF5+M	FF3	FF5	FF5+M	FF3
Bienes Raíces	0.221*	0.029	0.033	0.024	0.106	0.107	0.097	0.121	0.121	0.136	0.121
Consumo Básico	0.459	0.533	0.571	0.560	0.551	0.584	0.587*	0.363	0.378	0.347	0.363
Consumo Discrecional	0.853*	0.825	0.834	0.833	0.842	0.844	0.841	0.722	0.714	0.706	0.722
Energía	0.514	0.570	0.504	0.536	0.592*	0.557	0.579	0.365	0.308	0.326	0.365
Finanzas	0.747	0.849	0.883*	0.867	0.861	0.883	0.862	0.695	0.675	0.671	0.695
Industriales	0.839	0.853	0.849	0.847	0.855*	0.847	0.849	0.680	0.663	0.660	0.680
Materiales	0.761	0.759	0.745	0.731	0.769*	0.769	0.762	0.637	0.631	0.634	0.637
Salud	0.549	0.560	0.570	0.571	0.593*	0.585	0.581	0.458	0.456	0.436	0.458
Serv. Públicos	0.022*	-0.072	-0.116	-0.114	-0.030	-0.033	-0.011	-0.007	0.014	-0.007	-0.007
Serv. de Comunicación	0.179	0.089	0.136	0.128	0.152	0.229*	0.215	0.027	0.052	0.022	0.027
Tecn. de la Información	0.745	0.768	0.737	0.725	0.796*	0.791	0.788	0.697	0.685	0.656	0.697
Promedio	0.535	0.524	0.522	0.519	0.553	0.560*	0.559	0.433	0.427	0.417	0.433

Nota: Los valores en negrita con un asterisco (*) son los mejores en la fila de la tabla. Los valores en negrita sin asterisco son los peores.

Tabla 3: Tiempos de ejecución en minutos por Sector, Método y Modelo

Sector	OLS				Elastic Net			Random Forest		
	CAPM	FF3	FF5	FF5+M	FF3	FF5	FF5+M	FF3	FF5	FF5+M
Bienes Raíces	0.001	0.002	0.003	0.002	9.325	9.398	9.425	191.212	192.977	194.163
Consumo Básico	0.002	0.002	0.003	0.003	9.220	9.387	9.459	191.776	193.651	194.218
Consumo Discrecional	0.001	0.001	0.004	0.000	9.256	9.389	9.364	191.306	193.199	193.733
Energía	0.001	0.002	0.003	0.004	9.273	9.445	9.448	191.349	193.781	193.288
Finanzas	0.001	0.002	0.002	0.003	9.233	9.315	9.498	190.797	192.865	194.123
Industriales	0.002	0.001	0.002	0.004	9.332	9.133	9.18	191.706	250.122	250.455
Materiales	0.001	0.002	0.002	0.003	8.968	9.089	9.139	248.193	249.781	250.582
Salud	0.001	0.002	0.002	0.003	9.067	9.064	69.804	249.564	249.773	189.510
Serv. Públicos	0.001	0.003	0.003	0.002	8.993	9.104	9.178	248.578	249.342	250.002
Serv. de Comunicación	0.001	0.002	0.003	0.004	8.953	9.137	9.097	248.312	249.787	250.058
Tecn. de la Información	0.001	0.002	0.004	0.004	9.036	9.098	4.632	248.414	249.317	241.972
Promedio	0.001	0.002	0.003	0.003	9.150	9.233	14.384	217.382	224.054	218.373

más bajo en esa fila.

En sectores como Consumo Básico y Salud, los modelos lineales también superan en desempeño al método no lineal. El sector de Consumo Básico muestra un R^2 máximo de 0.853 con *OLS* en el modelo *CAPM*, mientras que el mínimo es de 0.706 con *Random Forest* en el modelo *FF5+M*. De manera similar, el sector Salud alcanza un R^2 de 0.883 con *OLS* en el modelo *FF5*, destacando como el mejor ajuste para este sector.

El sector Energía presenta una mejora en el ajuste al incorporar más factores al modelo. El mejor desempeño se obtiene con *Elastic Net* en el modelo *FF5+M*, alcanzando un R^2 de 0.587, lo que sugiere que la inclusión de los factores adicionales de Fama y French más el factor de momentum contribuye a explicar mejor las variaciones en los retornos de este sector.

En el caso del sector de Tecnología de la Información, los valores de R^2 son notablemente más bajos en comparación con otros sectores, indicando una menor capacidad de los modelos para explicar la variabilidad de los retornos. Sin embargo, el modelo *FF5* estimado por *Elastic Net* logra el mejor desempeño dentro de este sector con un R^2 de 0.229, sugiriendo que la regularización implementada en *Elastic Net* ayuda a capturar mejor las relaciones presentes en los datos.

El sector de Materiales presenta valores de R^2 muy bajos e incluso negativos, lo cual indica que los modelos no están capturando adecuadamente la dinámica de los retornos en este sector. Un R^2 negativo sugiere que el modelo ajustado es peor que un modelo que simplemente predice el promedio de los datos. Este resultado apunta a la necesidad de explorar modelos alternativos o incluir variables adicionales que puedan explicar mejor el comportamiento de los retornos en el sector de Materiales.

Al comparar los métodos, *Elastic Net* suele obtener los valores de R^2 más altos en varios sectores, como el S&P 500, Finanzas e Industriales. Por ejemplo, en el sector Financiero, *Elastic*

Net con el modelo *FF3* alcanza un R^2 de 0.593, superando a los otros métodos y modelos. Esto indica que la combinación de regularización y selección de variables de *Elastic Net* mejora el ajuste del modelo en estos casos.

Por otro lado, el método *Random Forest* generalmente muestra un desempeño inferior en términos de R^2 en comparación con los métodos lineales. Esto puede deberse a que, aunque los métodos no lineales son capaces de capturar relaciones más complejas, pueden no ser tan efectivos cuando el número de observaciones es limitado o cuando las relaciones lineales predominan en los datos, sobre todo en el contexto de modelos con pocas características.

La Figura 4 presenta diagramas de caja y bigotes (*boxplots*) que ilustran la distribución del error cuadrático medio (MSE) de prueba para cada sector económico, método de estimación y modelo utilizado. Los gráficos están organizados en una cuadrícula, donde cada recuadro representa un modelo específico aplicado a diferentes sectores.

En el eje vertical de cada gráfico se muestra el valor del MSE, mientras que en el eje horizontal se representan los distintos sectores económicos, identificados por colores específicos según la leyenda incluida en la figura. Esta representación visual permite comparar el desempeño de los métodos y modelos en términos de precisión y consistencia en la predicción de los retornos sectoriales.

Al analizar los resultados, se observa que los métodos lineales, especialmente *Elastic Net*, presentan valores de MSE más bajos y una menor dispersión en comparación con el método no lineal *Random Forest*. Esto sugiere que *Elastic Net* logra un mejor ajuste en la mayoría de los sectores, posiblemente debido a su capacidad para manejar multicolinealidad y realizar selección de variables.

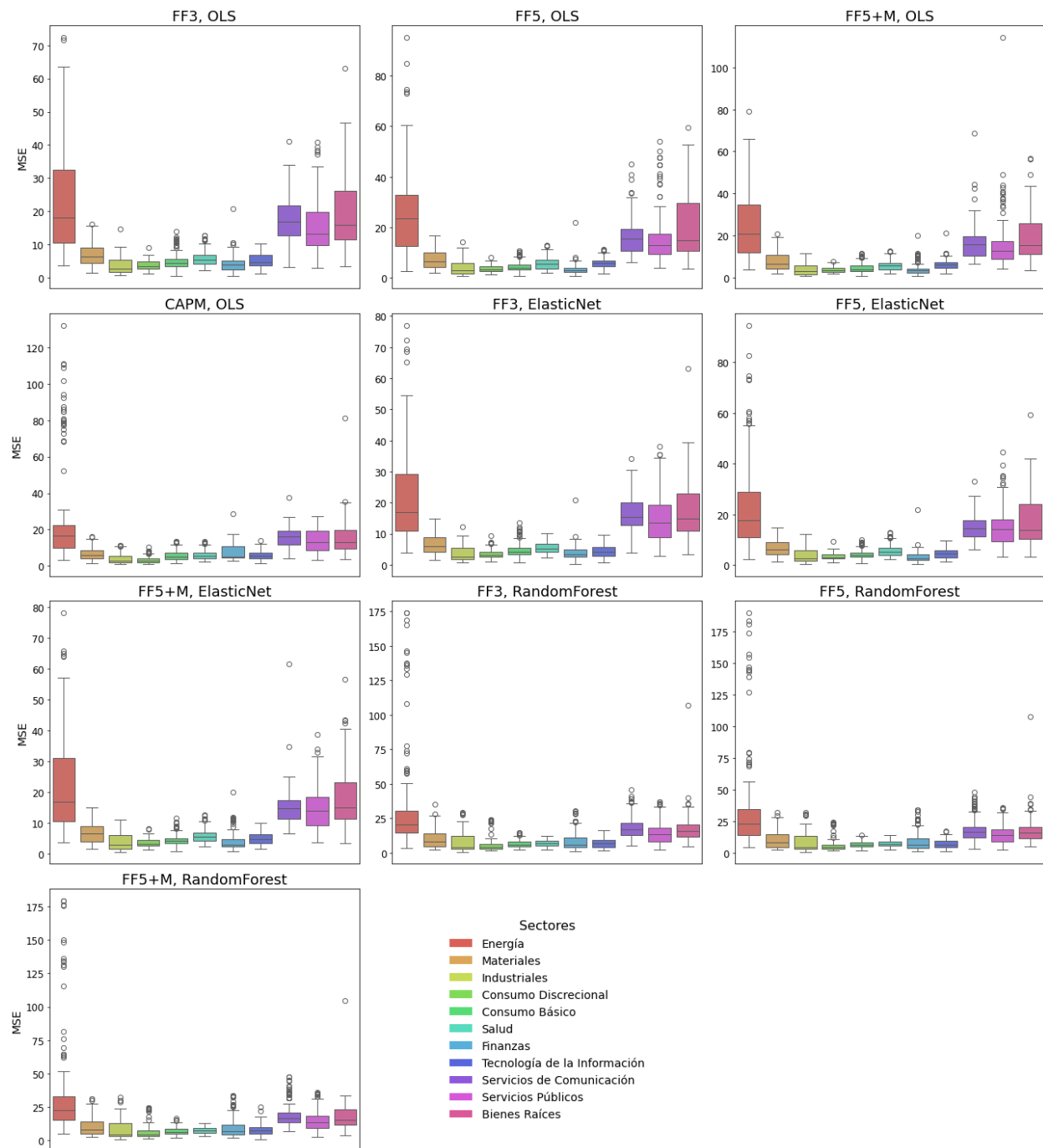


Figura 2: Boxplots de MSE de prueba agrupados por modelo, sector, para todas las ventanas

Por ejemplo, en los modelos FF3 y FF5 bajo *Elastic Net*, sectores como Finanzas, Industriales y Consumo Discrecional muestran MSE bajos y poca variabilidad, indicando predicciones más precisas y estables. En contraste, *Random Forest* exhibe MSE más altos y una mayor disper-

sión en estos mismos sectores, lo que implica una menor eficacia en la captura de las relaciones subyacentes en los datos.

El sector Energía destaca por presentar consistentemente los valores de MSE más altos y una mayor dispersión en todos los métodos y modelos. Esto indica que los retornos en este sector son más difíciles de predecir, posiblemente debido a factores externos como la volatilidad en los precios de los commodities o eventos geopolíticos que no están capturados por los modelos utilizados.

Además, se aprecia que la complejidad adicional al pasar de modelos más simples como CAPM a modelos más complejos como FF5+M no siempre se traduce en una mejora significativa en el MSE. En algunos casos, la incorporación de factores adicionales parece aumentar la variabilidad sin reducir sustancialmente el error de predicción. Esto sugiere que añadir más factores puede introducir ruido o redundancia, especialmente si los nuevos factores no aportan información relevante para ciertos sectores.

En el caso del método *OLS*, aunque muestra un desempeño aceptable en algunos sectores, generalmente presenta mayor variabilidad en los MSE en comparación con *Elastic Net*. Esto podría deberse a la falta de regularización en *OLS*, lo que lo hace más sensible a outliers y a la multicolinealidad entre variables explicativas.

Los resultados evidencian que la elección del método y del modelo es crucial y depende del sector analizado. Mientras que *Elastic Net* parece ser la opción más robusta para la mayoría de los sectores, es importante considerar las características específicas de cada sector al seleccionar el modelo. Sectores con alta volatilidad o influenciados por factores externos, como Energía, pueden requerir modelos más especializados o la inclusión de variables exógenas para mejorar las predicciones.

8 Conclusión

En el presente estudio se examina la efectividad de diversos modelos de valoración de activos en el contexto de distintos sectores económicos, comparando metodologías tradicionales, como el Modelo de Valoración de Activos de Capital (CAPM) y los modelos multifactoriales de Fama-French, con técnicas avanzadas de aprendizaje automático, específicamente Elastic Net y Random Forest. Mediante la implementación de una metodología de ventana móvil (rolling window), el trabajo se centra en evaluar el poder predictivo fuera de muestra de cada modelo, empleando métricas de rendimiento y validación para identificar la combinación más eficaz de modelo y método de estimación.

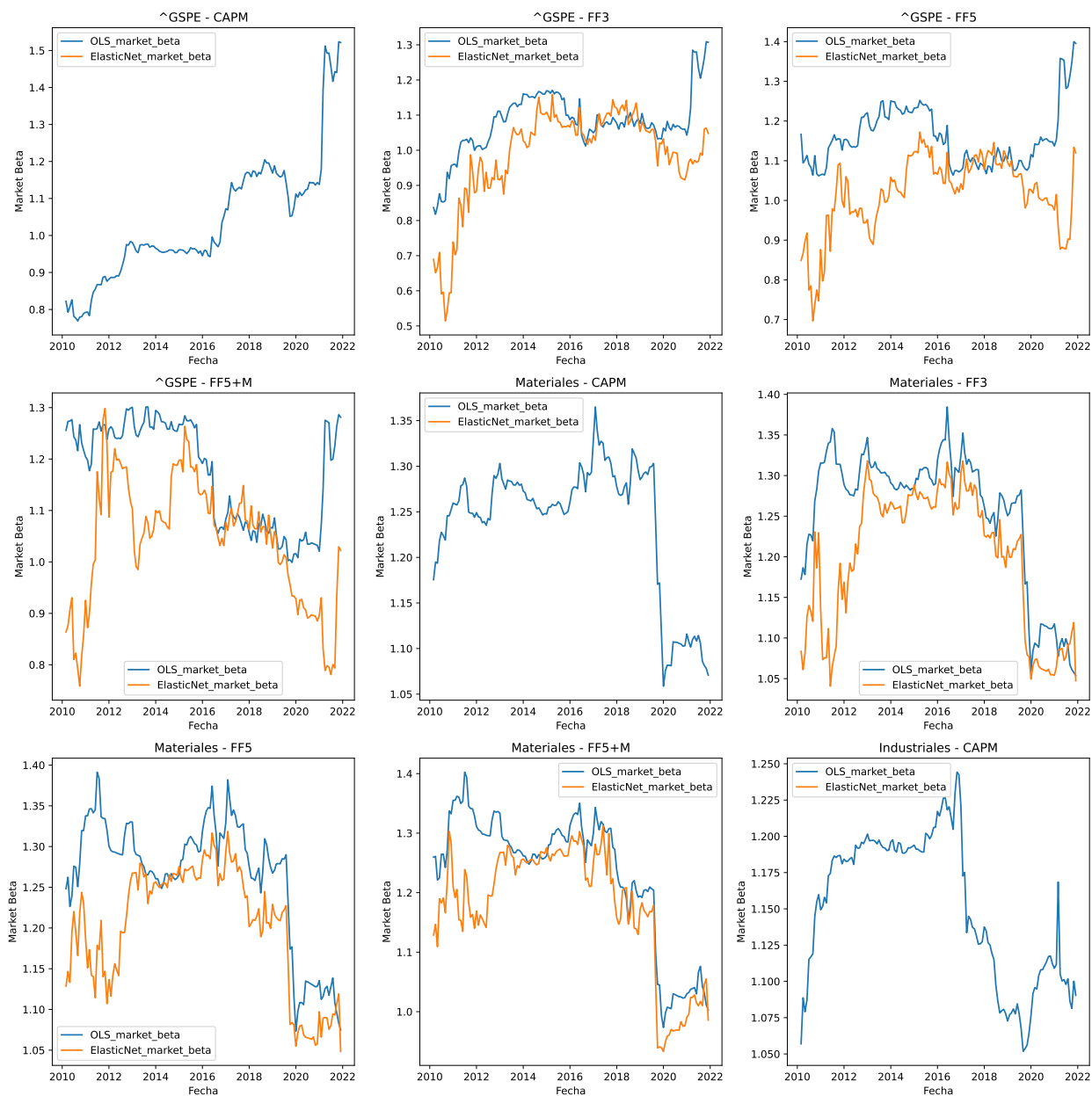
Los resultados obtenidos indican que, aunque el CAPM y los modelos de Fama-French (de tres y cinco factores) proporcionan una base sólida para la predicción de retornos en varios sectores, la incorporación de técnicas de aprendizaje automático permite una optimización adicional en términos de ajuste y precisión predictiva. En particular, Elastic Net demostró un desempeño superior en sectores con menor volatilidad, mientras que Random Forest, a pesar de su capacidad para capturar relaciones complejas, mostró limitaciones en sectores con alta colinealidad y predominio de relaciones lineales. Estos hallazgos sugieren que Elastic Net logra un equilibrio adecuado entre ajuste y capacidad de generalización, proporcionando predicciones más consistentes.

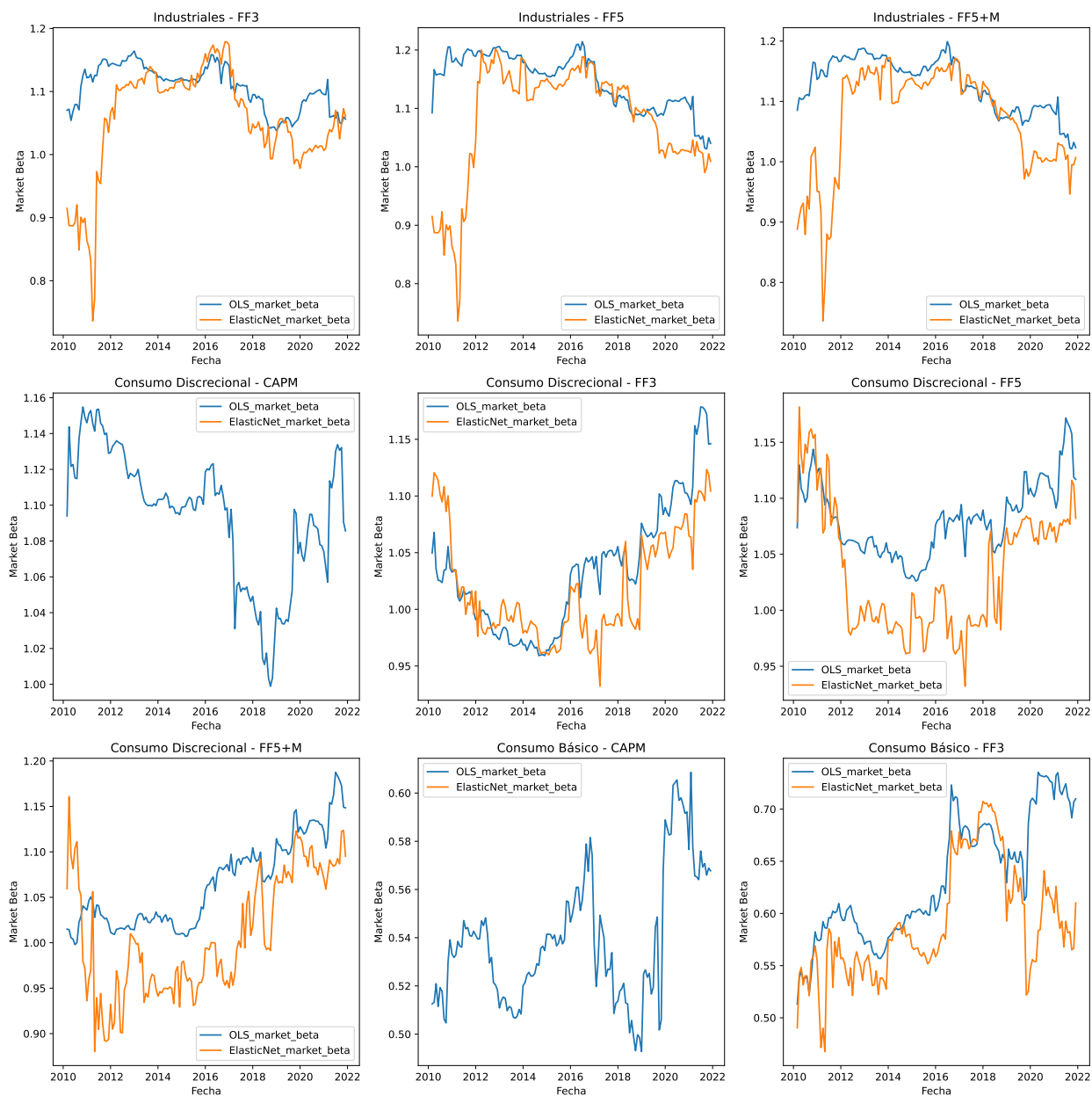
El análisis sectorial evidencia que la complejidad intrínseca de cada sector influye significativamente en la elección del modelo óptimo. Sectores como Finanzas e Industriales exhiben mayor estabilidad predictiva con modelos lineales, mientras que sectores como Energía presentan desafíos adicionales debido a la volatilidad externa y a las fluctuaciones abruptas en sus retornos. La inclusión de factores adicionales, como el momentum, contribuyó a mejorar el ajuste en ciertos

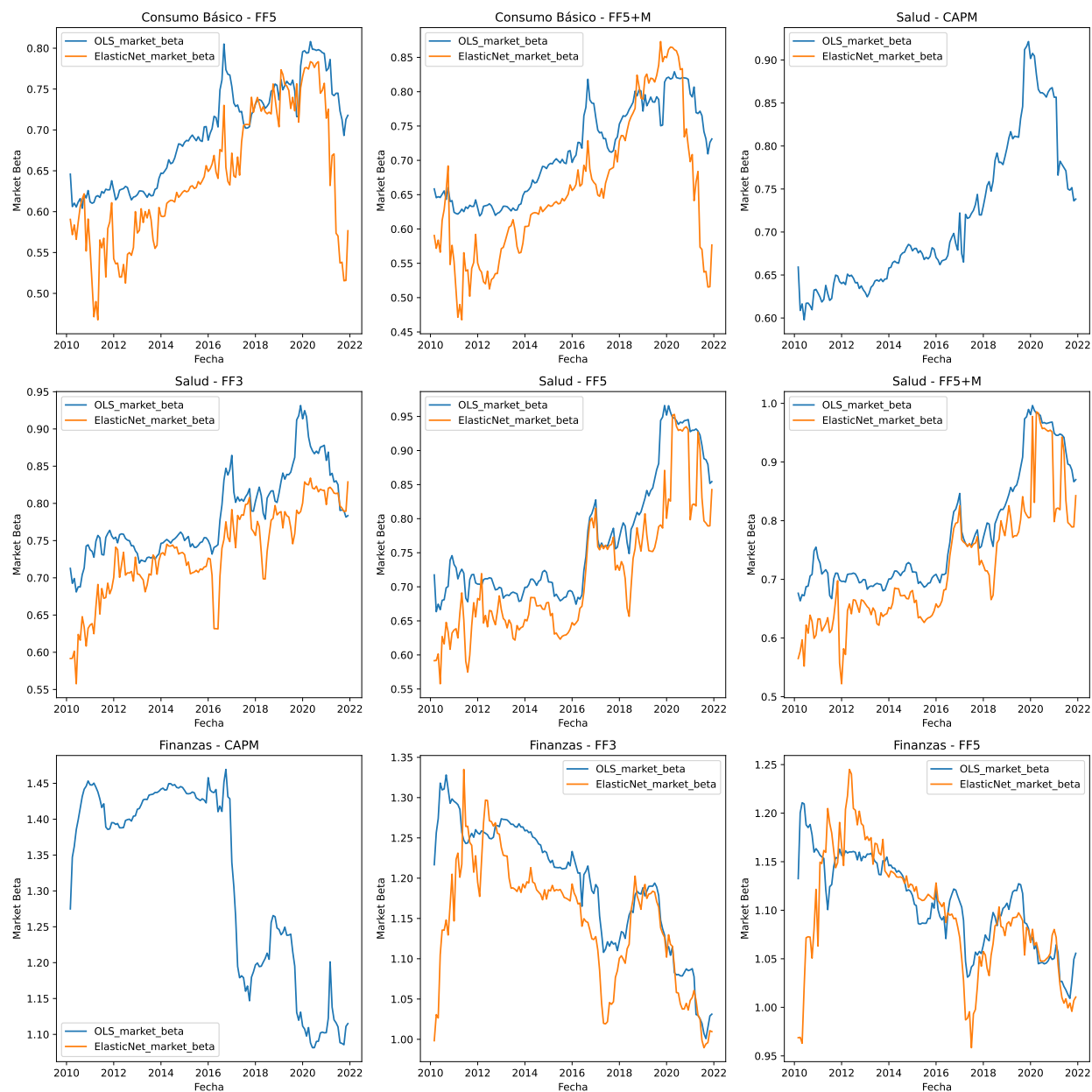
sectores, resaltando la importancia de adaptar los modelos a las características específicas de cada ámbito económico.

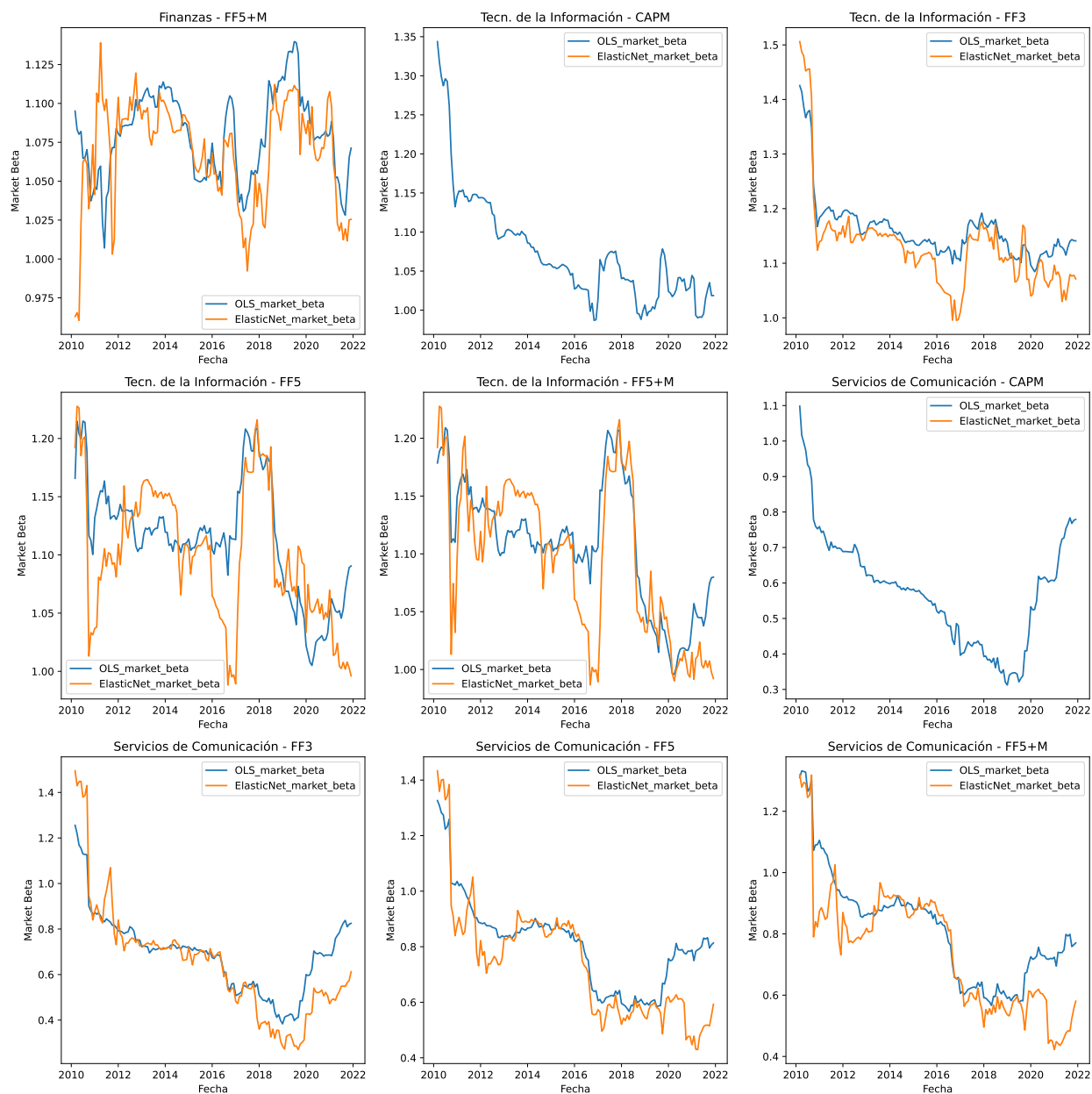
Este trabajo aporta al campo de la valoración de activos al demostrar que los modelos de aprendizaje automático pueden complementar y mejorar las metodologías tradicionales de valoración, especialmente en escenarios donde la multicolinealidad o la alta dimensionalidad de los datos pueden comprometer la precisión del modelo. La implementación de Elastic Net como método de regularización permite gestionar eficazmente la multicolinealidad y seleccionar variables clave sin sacrificar capacidad predictiva, lo cual es esencial en entornos de datos extensos y heterogéneos.

En conclusión, los hallazgos sugieren que la combinación de modelos tradicionales de valoración de activos con métodos de aprendizaje automático ofrece una estrategia robusta para la predicción de retornos en los mercados financieros. Esto representa una contribución significativa para la toma de decisiones de inversión y la gestión de carteras, al proporcionar una herramienta flexible y adaptable capaz de ajustarse a los cambios dinámicos del mercado. Futuros trabajos de investigación podrían centrarse en explorar técnicas más avanzadas de aprendizaje automático y en la integración de fuentes de datos alternativas para optimizar aún más los modelos de predicción, adaptándose a la evolución de los mercados y a la creciente disponibilidad de datos en el sector financiero.









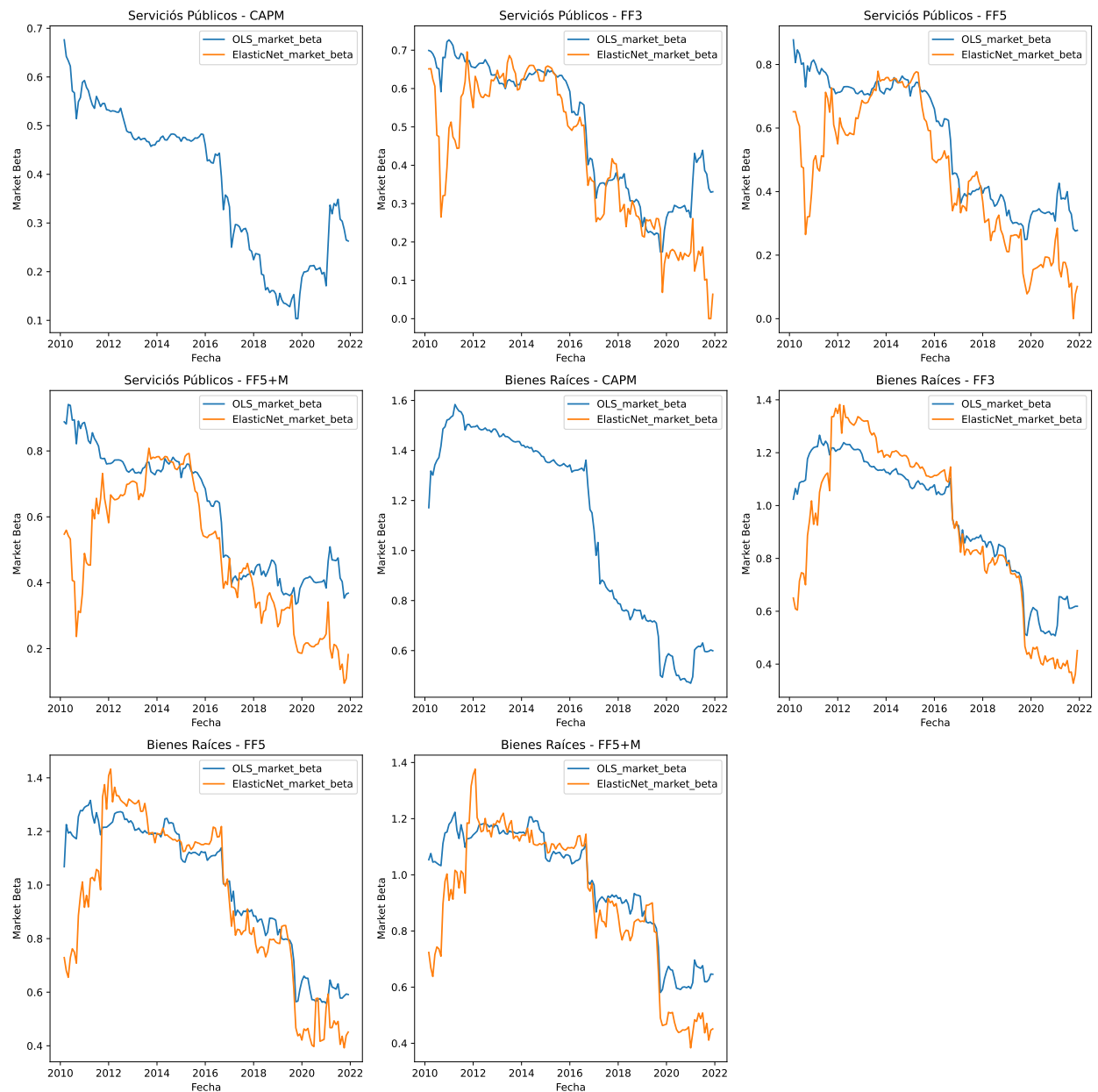
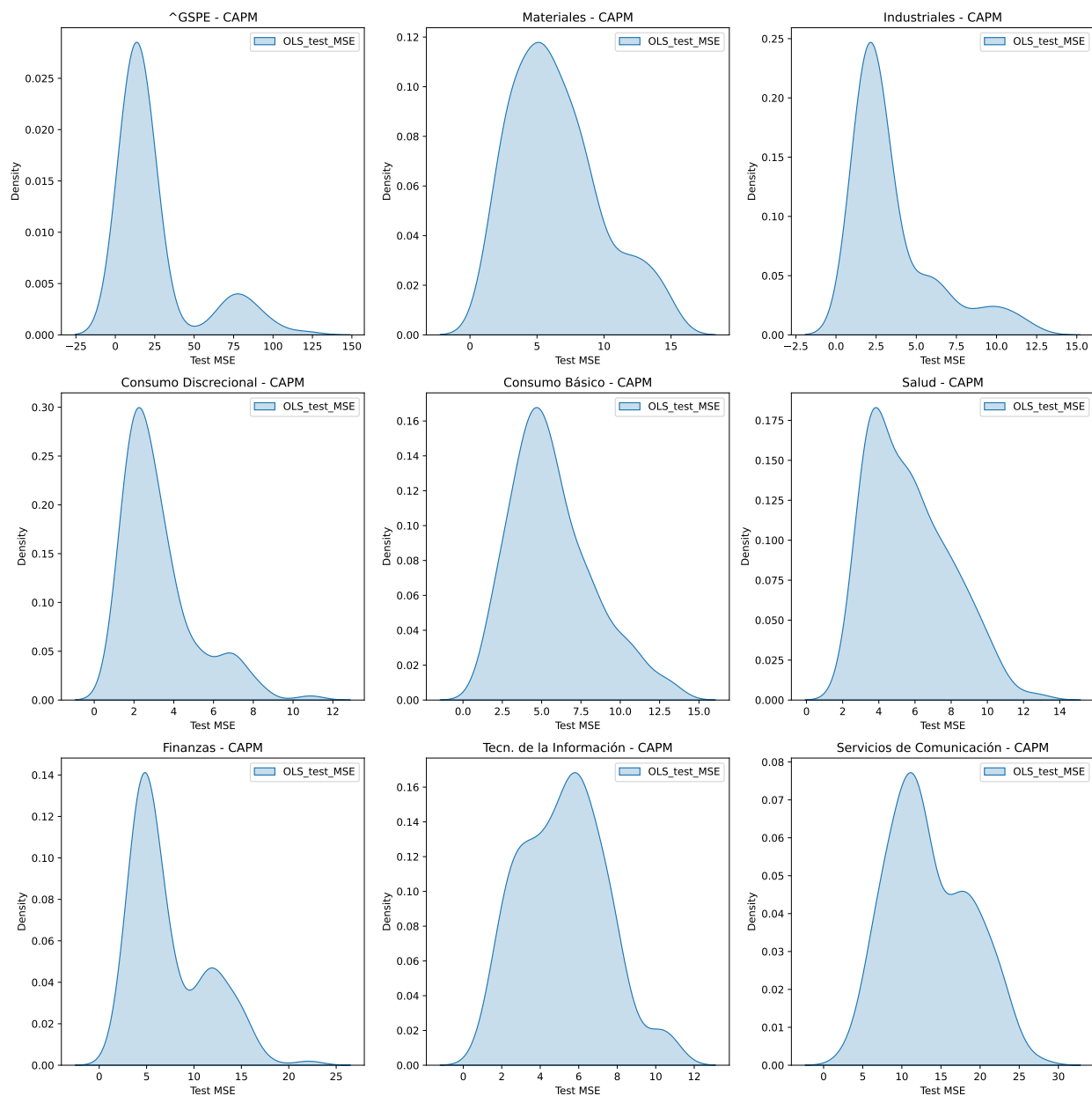
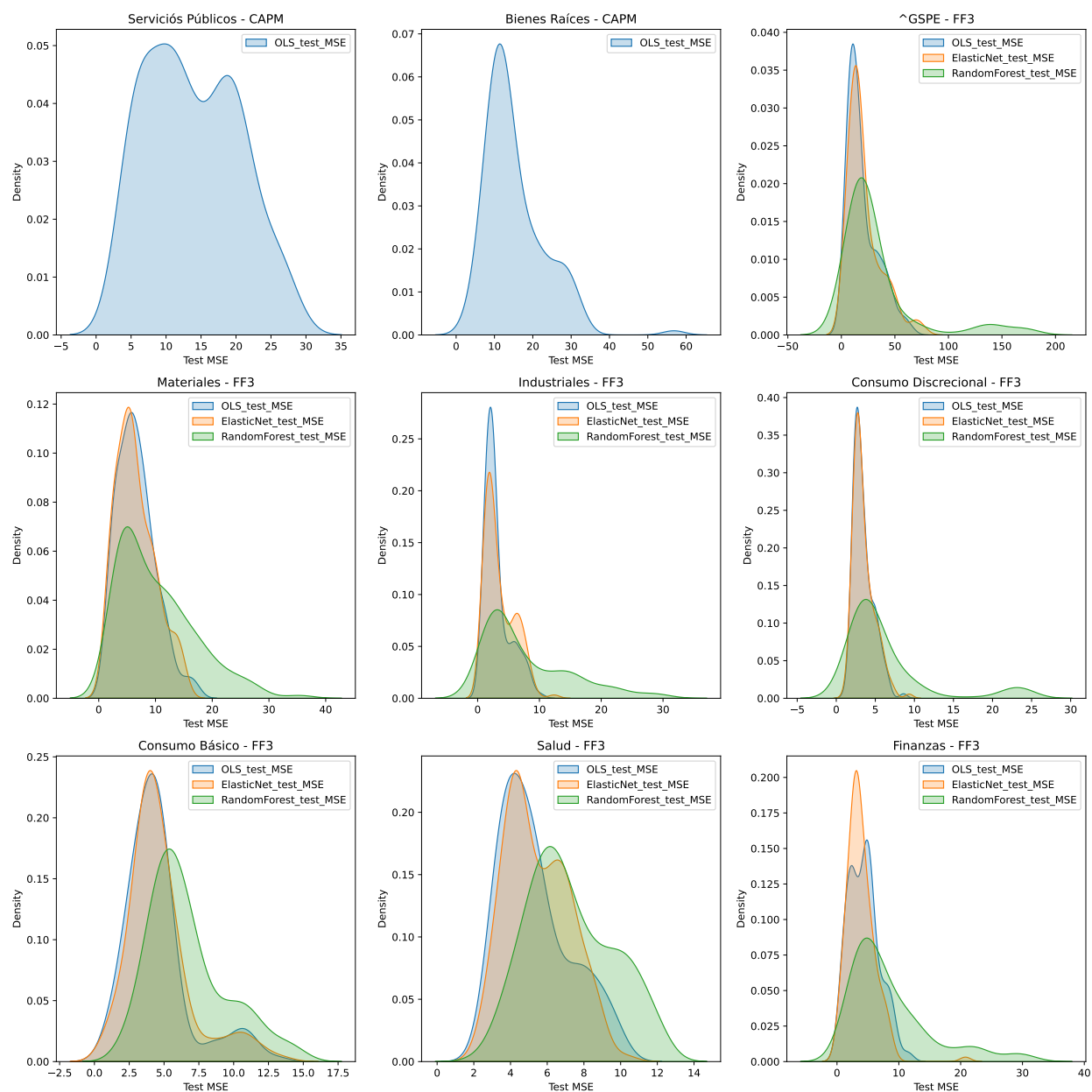
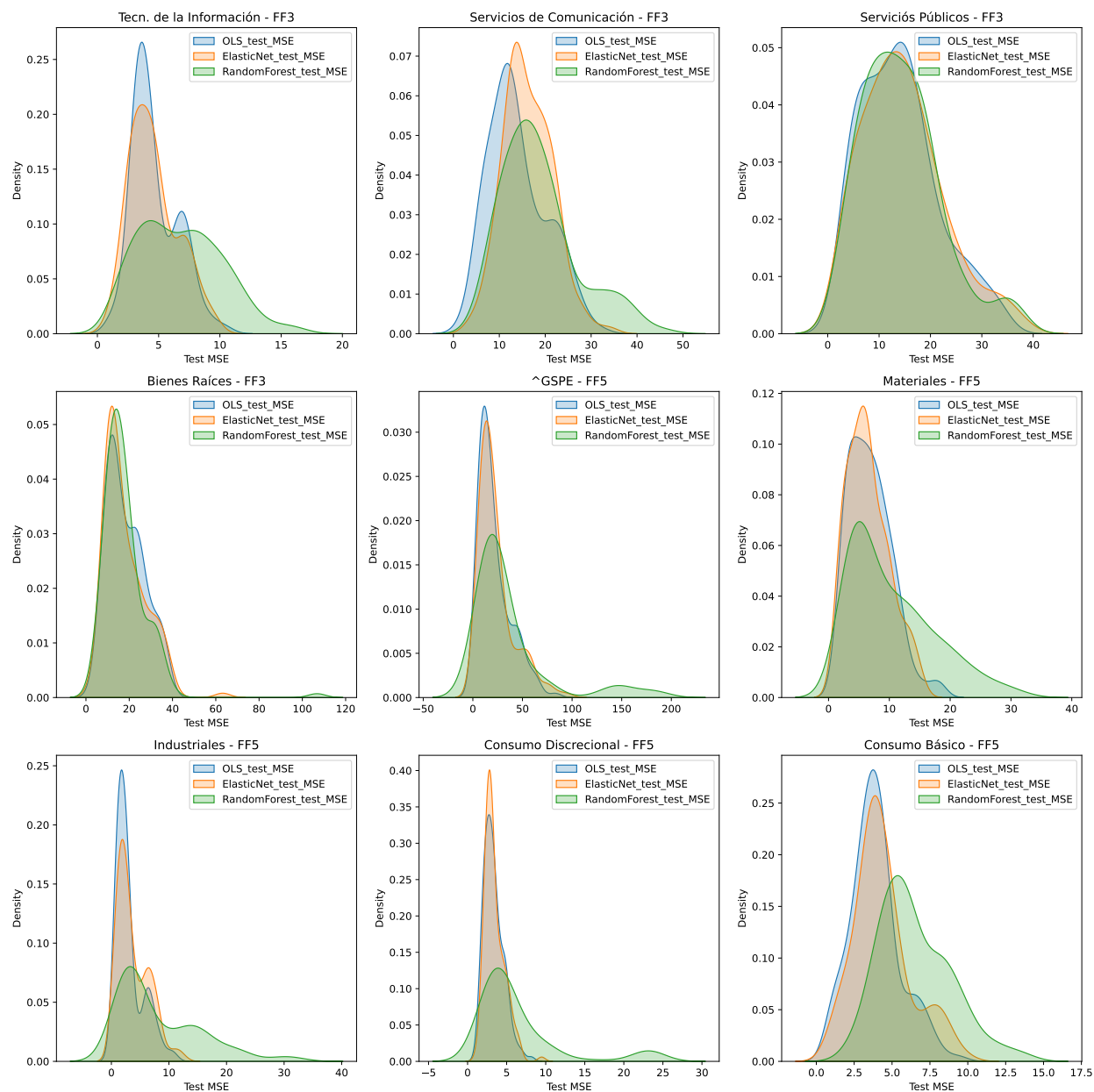
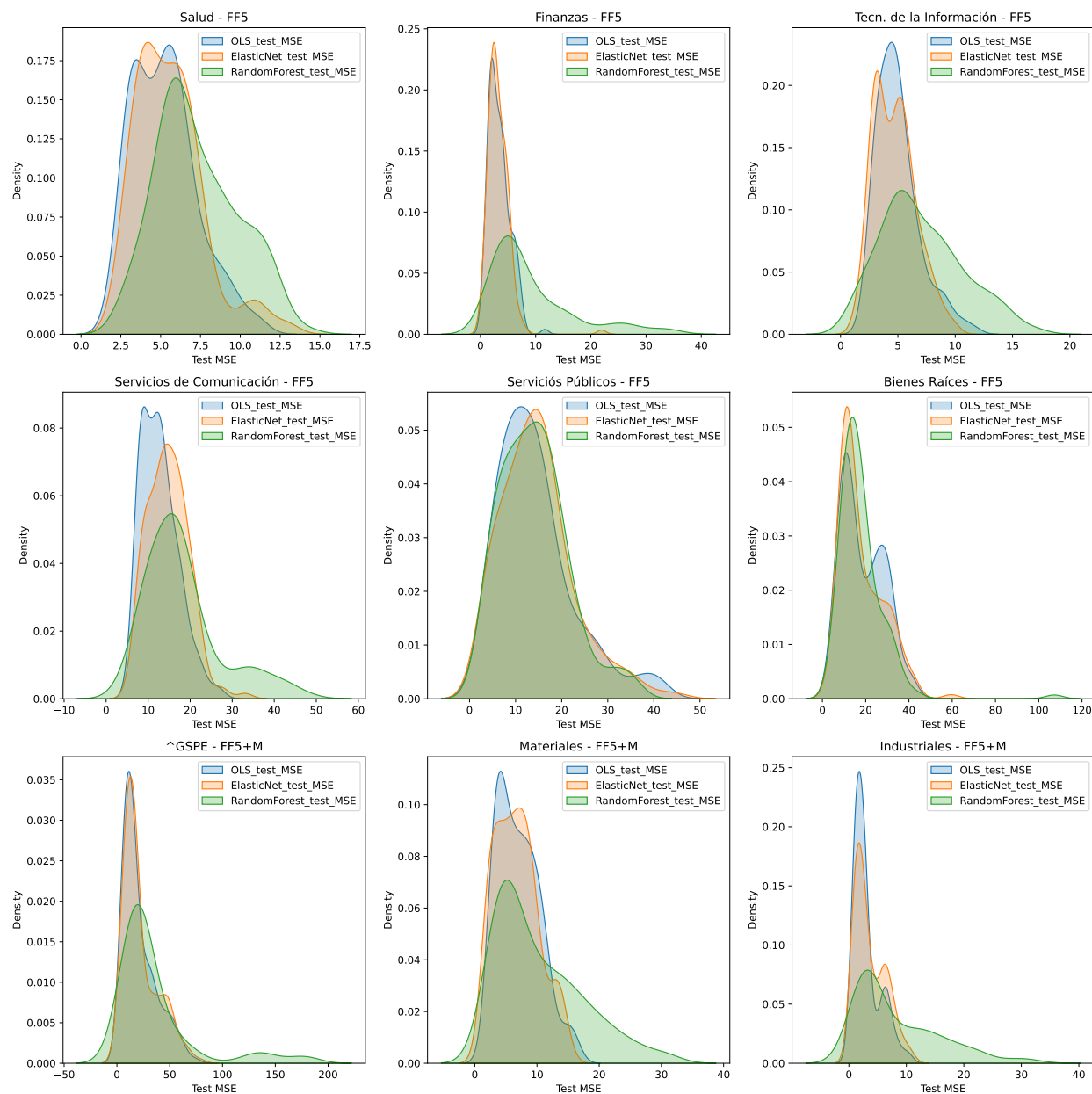


Figura 3: Market Betas a través del tiempo para las distintas combinaciones de método, modelo, sector.









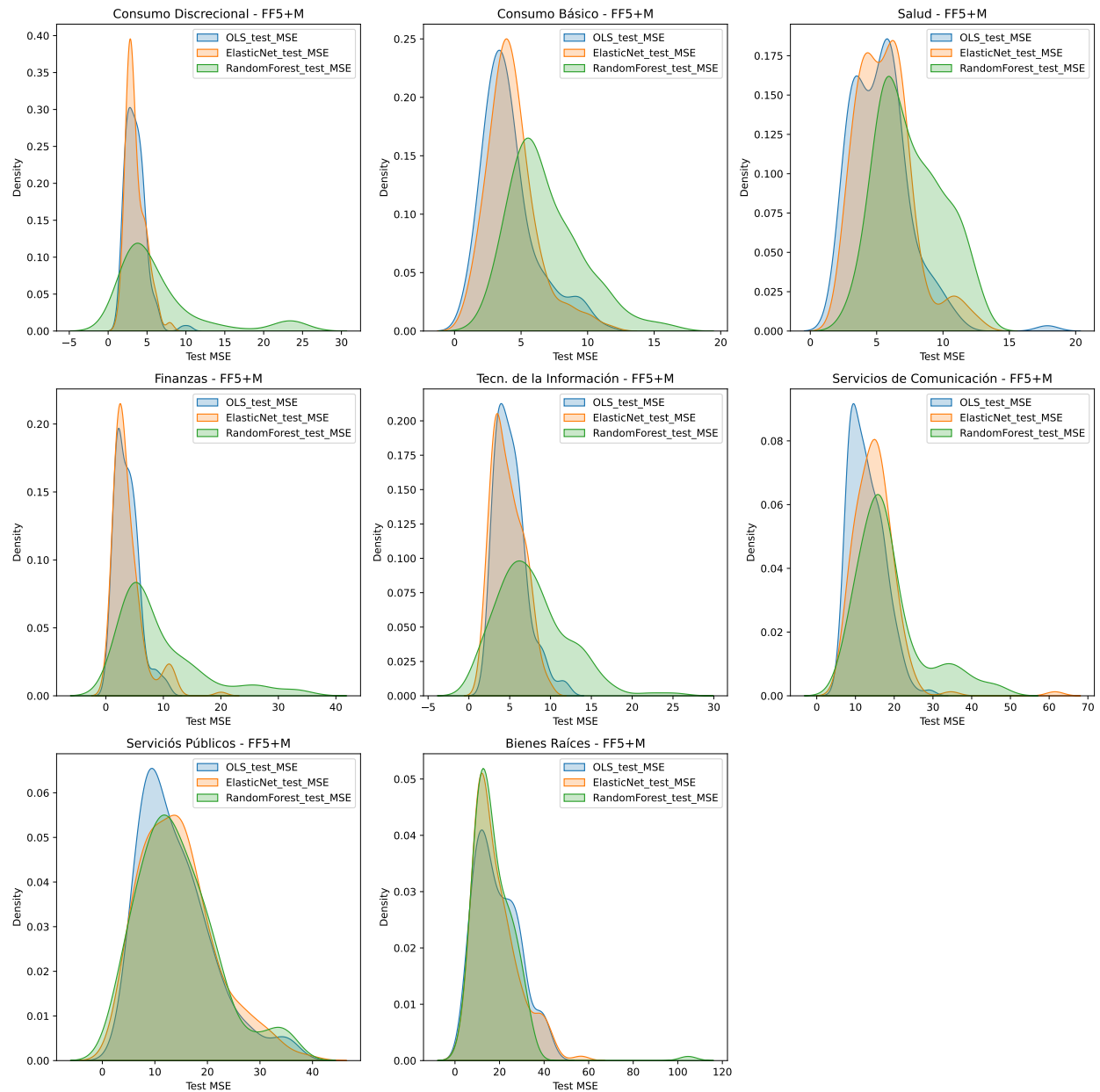


Figura 4: Gráficos de densidad del MSE de prueba para las distintas combinaciones de método, modelo, sector.

References

- Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1):3–18.
- Black, F., Jensen, M. C., y Scholes, M. (1972). The capital asset pricing model: Some empirical tests. *Studies in the theory of capital markets*, 81(3):79–121.
- Blume, M. E. (1975). Betas and their regression tendencies. *The Journal of Finance*, 30(3):785–795.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.
- Breiman, L., Friedman, J., Stone, C. J., y Olshen, R. A. (1984). *Classification and Regression Trees*. CRC press.
- Campbell, J. Y. y Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82.
- Cochrane, J. H. (2009). *Asset Pricing: (Revised Edition)*. Princeton University Press.
- Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of finance*, 66(4):1047–1108.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417.

- Fama, E. F. (1991). Efficient capital markets: Ii. *The journal of finance*, 46(5):1575–1617.
- Fama, E. F. y French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.
- Fama, E. F. y French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56.
- Fama, E. F. y French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1):55–84.
- Fama, E. F. y French, K. R. (2004). The capital asset pricing model: Theory and evidence. *Journal of economic perspectives*, 18(3):25–46.
- Fama, E. F. y French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.
- French, K. R. (2023). Data library. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
- Goyal, A. y Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508.
- Graham, J. R. y Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of financial economics*, 60(2-3):187–243.
- Gu, S., Kelly, B., y Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5):2223–2273.

- Harvey, C. R., Liu, Y., y Zhu, H. (2016). ...and the cross-section of expected returns. *The Review of Financial Studies*, 29(1):5–68.
- Hoerl, A. E. y Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67.
- Jegadeesh, N. y Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1):65–91.
- Kozak, S., Nagel, S., y Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, 135(2):271–292.
- Levhari, D. y Levy, H. (1977). The capital asset pricing model and the investment horizon. *The Review of Economics and Statistics*, pp. 92–104.
- Levy, R. A. (1971). On the short-term stationarity of beta coefficients. *Financial Analysts Journal*, 27(6):55–62.
- Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The journal of finance*, 20(4):587–615.
- Markowitz, H. (1952). Portfolio selection. *The journal of finance*, 7(1):77–91.
- McLean, R. D. y Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, 71(1):5–32.
- Miller, J. y Seidler, L. J. (1972). Sample size and the strength of evidence: A bayesian interpretation of binomial tests of the information content of qualified audit opinions. *Journal of Accounting Research*, pp. 42–55.

- Moskowitz, T. J., Ooi, Y. H., y Pedersen, L. H. (2012). Time series momentum. *Journal of financial economics*, 104(2):228–250.
- Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28.
- Roll, R. (1977). A critique of the asset pricing theory's tests part i: On past and potential testability of the theory. *Journal of financial economics*, 4(2):129–176.
- Rosenberg, B., Reid, K., y Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, 11(3):9–16.
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.
- Tobin, J. (1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies*, 25(2):65–86.
- Zou, H. y Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.