



Doble Grado E2 + Business Analytics

TRABAJO DE FIN DE GRADO

“Criterios de equity screening para poder filtrar de manera sistemática un universo ingente de valores como el índice S&P 500 para crear una cartera correctamente diversificada que sea capaz de batir a dicho benchmark en el largo plazo”

Autor: Benigno Méndez Mateo

Director: Ramon Bermejo Climent

Curso 2025 – 2026

RESUMEN

El presente Trabajo de Fin de Grado desarrolla un modelo cuantitativo de selección sistemática de acciones aplicado al universo del S&P 500, con el objetivo de construir una cartera diversificada capaz de generar rentabilidad superior al índice de referencia a lo largo de distintos ciclos económicos y monetarios. Para ello, se combina un conjunto de cinco criterios de screening fundamentados en la literatura académica de finanzas cuantitativas: el Piotroski F-Score (Piotroski, 2000), el Altman Z-Score (Altman, 1968), el momentum de precio (Jegadeesh y Titman, 1993), los indicadores de calidad empresarial (Novy-Marx, 2013) y la valoración relativa mediante el múltiplo EV/EBITDA (Greenblatt, 2006). Los umbrales de cada criterio se anclan en los valores establecidos por sus respectivos autores, evitando cualquier optimización retrospectiva sobre los datos del estudio.

El backtest cubre el período 2004–2024, un total de 21 años, con datos fundamentales y de precio extraídos de Capital IQ para una muestra de 902 empresas históricamente presentes en el S&P 500. La cartera se rebalancea anualmente de forma equiponderada y su rentabilidad se mide frente al S&P 500 Price Return como benchmark. Los resultados muestran que la estrategia genera una rentabilidad anual compuesta (CAGR) del 14,22% frente al 8,26% del índice, con un Sharpe Ratio de 0,790 frente a 0,597 y un máximo drawdown inferior al del benchmark (–27,84% vs. –38,49%). El alpha medio anual es de 6,41 puntos porcentuales, con una tasa de éxito del 57,14% de los años. El análisis por ciclos macroeconómicos revela que el alpha es especialmente pronunciado en períodos de recesión según la clasificación NBER (14,87% de media), así como en entornos de política monetaria restrictiva (7,78%).

El trabajo concluye que la combinación disciplinada de factores de calidad, solidez financiera, momentum y valoración permite identificar de forma sistemática empresas con rentabilidades superiores, independientemente del entorno de mercado, aunque con mayor eficacia en fases contractivas del ciclo. Se discuten asimismo las limitaciones del estudio, entre ellas la asimetría en el tratamiento de dividendos entre la cartera y el benchmark.

Palabras clave: selección sistemática de acciones, factor investing, Piotroski F-Score, Altman Z-Score, momentum, quality investing, EV/EBITDA, backtesting, ciclos macroeconómicos, S&P 500.

ABSTRACT

This dissertation develops a quantitative systematic equity screening model applied to the S&P 500 universe, with the objective of constructing a diversified portfolio capable of generating returns in excess of the benchmark across different economic and monetary policy cycles. The model combines five screening criteria grounded in academic finance literature: the Piotroski F-Score (Piotroski, 2000), the Altman Z-Score (Altman, 1968), price momentum (Jegadeesh and Titman, 1993), quality indicators (Novy-Marx, 2013), and relative valuation via the EV/EBITDA multiple (Greenblatt, 2006). All thresholds are anchored to values established in the original literature, avoiding any backward-looking optimisation on the study's own data.

The backtest covers the period 2004–2024, a total of 21 years, using fundamental and price data sourced from Capital IQ for a sample of 902 companies historically present in the S&P 500. The portfolio is rebalanced annually on an equal-weighted basis, and its performance is measured against the S&P 500 Price Return as the benchmark. Results show that the strategy achieves a compound annual growth rate (CAGR) of 14.22% versus 8.26% for the index, with a Sharpe Ratio of 0.790 against 0.597 and a lower maximum drawdown than the benchmark (−27.84% vs. −38.49%). Mean annual alpha stands at 6.41 percentage points, with a hit rate of 57.14% of years. Macroeconomic cycle analysis reveals that alpha is particularly pronounced during NBER-classified recession years (14.87% on average) and in tightening monetary policy environments (7.78%).

The study concludes that the disciplined combination of quality, financial strength, momentum and valuation factors allows for the systematic identification of outperforming equities across market environments, with particularly strong performance during contractionary phases. Limitations of the study are discussed, including the asymmetry in dividend treatment between the portfolio and the benchmark.

Keywords: systematic equity screening, factor investing, Piotroski F-Score, Altman Z-Score, momentum, quality investing, EV/EBITDA, backtesting, macroeconomic cycles, S&P 500.

INDICE

1. Introducción.....	6
1.1. Contexto y motivación.....	6
1.2. Justificación y relevancia.....	7
1.3. Pregunta de investigación y objetivos	7
1.4. Estructura del trabajo.....	8
2. Marco teórico.....	9
2.1. La hipótesis de los mercados eficientes y sus implicaciones para la selección de acciones	9
2.2. El Piotroski F-Score: solidez financiera como predictor de rentabilidad	10
2.3. El Altman Z-Score: riesgo de quiebra como filtro de exclusión	10
2.4. Momentum de precio: la persistencia de las tendencias.....	11
2.5. Calidad empresarial: rentabilidad como señal de ventaja competitiva.....	12
2.6. Valoración relativa: EV/EBITDA como criterio de precio.....	13
2.7. Ciclos macroeconómicos y performance de factores	14
3. Metodología.....	14
3.1. Diseño general del estudio.....	14
3.2. Fuentes de datos	15
3.3. Construcción del universo de inversión.....	16
3.4. El proceso de screening multifactorial	16
3.5. Construcción de la cartera y cálculo de rentabilidades.....	19
3.6. Métricas de evaluación de la performance	19
3.7. Clasificación de los ciclos macroeconómicos	20
4. Análisis de datos y resultados.....	20
4.1. Resultados agregados del backtest (2004–2024).....	20
4.2. Análisis del comportamiento anual	23
4.3. Análisis por ciclo macroeconómico: recesión frente a expansión.....	25
4.4. Análisis por ciclo de política monetaria	26
4.5. Distribución sectorial de la cartera	27
5. Conclusiones, limitaciones y líneas de investigación futura	28
5.1. Conclusiones principales	28
5.2. Limitaciones del estudio.....	30

5.3. Líneas de investigación futura.....	32
5.4. Reflexión final	33
6. Bibliografía.....	34
7. Anexos	36
7.1. Código Python del backtest	36

1. Introducción

1.1 Contexto y motivación

Los mercados de renta variable han sido históricamente el objeto de estudio de una de las preguntas más persistentes en finanzas: ¿es posible seleccionar acciones de forma sistemática y obtener rentabilidades superiores a las del mercado de manera consistente? La hipótesis de los mercados eficientes, formulada por Fama (1970), sostiene que los precios de los activos reflejan en todo momento toda la información disponible, haciendo imposible la generación de alpha de forma recurrente sin asumir un riesgo adicional. Sin embargo, décadas de investigación empírica han documentado la existencia de anomalías de mercado que contradicen esta hipótesis en su forma semifuerte: empresas con sólidos fundamentales, precios con momentum positivo o cotizaciones baratas en términos relativos han tendido a superar al mercado con una regularidad difícilmente atribuible al azar (Fama y French, 1992; Jegadeesh y Titman, 1993; Piotroski, 2000).

Este debate ha dado lugar a una rama de la gestión de inversiones conocida como factor investing o inversión sistemática por factores, que consiste en construir carteras a partir de criterios cuantitativos replicables, fundamentados en evidencia académica, en lugar de juicios discrecionales del gestor. A diferencia de la gestión activa tradicional, donde las decisiones de inversión dependen de la visión particular de un analista o equipo, el enfoque sistemático elimina el sesgo conductual, garantiza la consistencia metodológica y permite evaluar el proceso de forma transparente y objetiva. Esta filosofía ha ganado una presencia cada vez mayor en la industria de gestión de activos durante las últimas dos décadas, impulsada tanto por la creciente disponibilidad de datos fundamentales de calidad como por el aumento de la capacidad de procesamiento computacional.

La motivación personal de este trabajo nace de la confluencia entre mi formación en Business Analytics y el interés por aplicar técnicas cuantitativas a problemas reales de finanzas. El análisis de datos, la construcción de modelos y la interpretación rigurosa de resultados son herramientas que permiten abordar preguntas de inversión con un nivel de precisión y reproducibilidad que la gestión discrecional difícilmente puede igualar. Al mismo tiempo, la abundancia de literatura académica sobre factores individuales contrasta con la escasez de estudios que combinen múltiples criterios de screening de forma simultánea y los evalúen a través de distintos regímenes macroeconómicos, algo que este trabajo pretende abordar directamente.

1.2 Justificación y relevancia

El estudio presentado a continuación es relevante por tres razones fundamentales. En primer lugar, la inversión sistemática por factores constituye uno de los enfoques de mayor crecimiento dentro de la gestión de activos a nivel global. Según datos de MSCI, los activos bajo gestión en estrategias de factor investing superaban los tres billones de dólares a comienzos de la presente década, y su expansión ha sido particularmente notable en el segmento institucional. Comprender los fundamentos teóricos y empíricos de esta aproximación resulta, por tanto, de evidente relevancia tanto académica como profesional.

En segundo lugar, la mayor parte de la investigación existente sobre factores de inversión analiza cada criterio de forma aislada. Piotroski (2000) estudia el F-Score de manera independiente; Jegadeesh y Titman (1993) se centran exclusivamente en el momentum; Greenblatt (2006) combina únicamente rentabilidad y valoración en su denominada "fórmula mágica". Sin embargo, un inversor real no aplica un único criterio, sino que evalúa simultáneamente múltiples dimensiones de una empresa: su solidez financiera, su posición competitiva, su tendencia de precio y su valoración relativa. Esta integración multifactorial, que es precisamente la práctica habitual en la industria, ha recibido una atención académica comparativamente menor, lo que justifica la propuesta metodológica de este trabajo.

En tercer lugar, el análisis de la performance de una estrategia de inversión en función del ciclo macroeconómico aporta una dimensión analítica especialmente valiosa. Los mercados no se comportan de forma homogénea a lo largo del tiempo: las condiciones de expansión y recesión, así como los entornos de política monetaria restrictiva o acomodaticia, modifican sustancialmente la dinámica de los precios de los activos y la capacidad de los factores para generar alpha. Evaluar si una estrategia de screening es robusta en distintos regímenes de mercado, y no únicamente en condiciones favorables, es un requisito indispensable para valorar su utilidad práctica y su solidez como herramienta de inversión a largo plazo.

1.3 Pregunta de investigación y objetivos

El trabajo se articula en torno a la siguiente pregunta de investigación:

¿Es posible construir una estrategia sistemática de selección de acciones del S&P 500, basada exclusivamente en criterios fundamentados en la literatura académica, que genere rentabilidades superiores al índice de referencia de forma consistente a lo largo de distintos ciclos económicos y monetarios?

De esta pregunta general se derivan cuatro objetivos específicos. El primero es revisar la literatura académica relevante sobre los principales factores de inversión cuantitativa y establecer los criterios de screening del modelo con base en evidencia empírica contrastada. El segundo es construir un pipeline de backtesting sobre datos reales de 902 empresas del S&P 500 para el período 2004–2024, garantizando en todo momento la ausencia de look-ahead bias en la construcción de la cartera. El tercero es evaluar la performance de la estrategia frente al S&P 500 Price Return mediante métricas estándar de la industria: rentabilidad compuesta anual, ratio de Sharpe, máximo drawdown y alpha medio. El cuarto es analizar el comportamiento diferencial de la estrategia en función de la fase del ciclo económico, empleando las clasificaciones oficiales del NBER para las recesiones y la evolución del tipo de interés de los fondos federales para los regímenes de política monetaria.

1.4 Estructura del trabajo

El presente trabajo se organiza en cinco capítulos. Tras esta introducción, el capítulo 2 desarrolla el marco teórico del estudio, abordando la hipótesis de los mercados eficientes como punto de partida conceptual y revisando la evidencia académica sobre los cinco factores que componen el modelo: el Piotroski F-Score como medida de solidez financiera, el Altman Z-Score como indicador de riesgo de quiebra, el momentum de precio como factor de tendencia, los indicadores de calidad empresarial y el múltiplo EV/EBITDA como criterio de valoración relativa. El capítulo 3 describe en detalle la metodología empleada, incluyendo la construcción del universo de inversión, el proceso de screening secuencial, el tratamiento de los datos ausentes, la lógica del backtest y la clasificación de los ciclos macroeconómicos. El capítulo 4 presenta los resultados del backtest para el período 2004–2024, tanto a nivel agregado como desglosados por ciclo económico y monetario, apoyándose en los gráficos generados a partir del análisis cuantitativo. Por último, el capítulo 5 recoge las conclusiones del trabajo, discute las limitaciones del estudio —entre ellas la asimetría en el tratamiento de dividendos entre la cartera y el benchmark— y propone posibles líneas de investigación futura.

2. Marco teórico

2.1 La hipótesis de los mercados eficientes y sus implicaciones para la selección de acciones

El punto de partida conceptual de cualquier estudio sobre selección activa de acciones es necesariamente la hipótesis de los mercados eficientes (HME), formulada en su versión moderna por Fama (1970). En su artículo seminal, Fama define un mercado eficiente como aquel en el que los precios de los valores reflejan plenamente toda la información disponible, de modo que ningún inversor puede obtener de forma sistemática rentabilidades superiores a las del mercado sin asumir un riesgo adicional. Fama distingue tres formas de eficiencia: la forma débil, según la cual los precios incorporan toda la información histórica de precios; la forma semifuerte, según la cual los precios incorporan también toda la información pública disponible; y la forma fuerte, que extiende dicha incorporación incluso a la información privada o privilegiada.

La implicación directa de la forma semifuerte para el objetivo de este trabajo es clara: si los mercados fueran plenamente eficientes en este sentido, el análisis de los estados financieros públicos de las empresas —que es precisamente la base del Piotroski F-Score, el Altman Z-Score o los indicadores de calidad— no debería permitir la construcción de carteras con rentabilidades ajustadas al riesgo superiores a las del índice de referencia. Sin embargo, la evidencia empírica acumulada a lo largo de las décadas siguientes a la publicación de Fama (1970) ha puesto en cuestión esta conclusión de forma sistemática, documentando la existencia de anomalías de mercado persistentes que la HME en su forma estricta no es capaz de explicar.

El trabajo de Fama y French (1992) constituyó un punto de inflexión en este debate. En su análisis de la sección transversal de rentabilidades esperadas en el mercado estadounidense, los autores documentaron que dos variables observables —el tamaño de la empresa y el ratio book-to-market— tenían un poder explicativo sobre las rentabilidades futuras muy superior al del beta del mercado, la variable central del modelo CAPM de Sharpe (1964). Este hallazgo fue interpretado de dos formas radicalmente opuestas: como evidencia de que las anomalías documentadas representan primas de riesgo compensatorias no capturadas por el CAPM, y como evidencia de ineficiencia del mercado que permite la generación de alpha mediante estrategias sistemáticas. Este debate teórico, aún no resuelto en la literatura, constituye el trasfondo intelectual del presente trabajo, que adopta una postura empírica y no toma partido en la discusión

causal: lo que se evalúa es si la estrategia genera rentabilidades superiores de forma consistente, independientemente de si dichas rentabilidades representan una prima de riesgo o una ineficiencia explotable.

2.2 El Piotroski F-Score: solidez financiera como predictor de rentabilidad

El primer componente del modelo de screening desarrollado en este trabajo es el Piotroski F-Score, propuesto por Piotroski (2000) en su artículo publicado en el *Journal of Accounting Research*. El trabajo de Piotroski parte de la observación de que las empresas con alto ratio book-to-market —empresas de valor en la terminología de Fama y French— presentan una dispersión muy elevada en sus rentabilidades posteriores: algunas generan retornos excepcionales, pero muchas otras continúan deteriorándose. El objetivo del F-Score es precisamente separar a los ganadores de los perdedores dentro de este universo de empresas de valor, utilizando exclusivamente información contable pública disponible en el momento de la inversión.

El F-Score se construye como la suma de nueve señales binarias, cada una de las cuales toma el valor de uno si se cumple una condición favorable y cero en caso contrario. Estas nueve señales se agrupan en tres categorías. La primera, de rentabilidad, incluye cuatro indicadores: la rentabilidad sobre activos (ROA) positiva en el ejercicio corriente, el flujo de caja operativo (CFO) positivo, la mejora del ROA respecto al ejercicio anterior, y el hecho de que el flujo de caja operativo supere al beneficio neto —señal de calidad del beneficio contable—. La segunda categoría, de apalancamiento y liquidez, incorpora tres señales: la reducción de la deuda a largo plazo sobre activos, la mejora del ratio de liquidez corriente, y la ausencia de dilución de capital mediante nuevas emisiones de acciones. La tercera categoría, de eficiencia operativa, recoge dos indicadores: la mejora del margen bruto y la mejora del ratio de rotación de activos respecto al ejercicio anterior.

Piotroski (2000) demuestra que una estrategia que compra acciones con F-Score alto ($F \geq 7$) y vende en corto acciones con F-Score bajo ($F \leq 1$) dentro del universo de empresas de alto book-to-market genera un retorno anual del 23% en el período 1976–1996. En el modelo presentado en este trabajo se aplica únicamente el umbral de compra $F \geq 7$, en línea con la aplicación long-only del criterio, que es la más habitual en la práctica inversora institucional y la que resulta relevante para un inversor con una cartera diversificada convencional.

2.3 El Altman Z-Score: riesgo de quiebra como filtro de exclusión

El segundo criterio del modelo es el Altman Z-Score, desarrollado por Altman (1968) en su artículo publicado en el *Journal of Finance*. El contexto original del trabajo de Altman era la predicción del riesgo de quiebra empresarial mediante análisis discriminante multivariante, una técnica estadística que permite clasificar empresas en grupos — quebradas y no quebradas— a partir de un conjunto de ratios financieros. Mediante el análisis de una muestra de 66 empresas manufactureras, Altman identificó cinco ratios con mayor poder discriminante y los combinó en una función lineal que asigna a cada empresa una puntuación Z.

La fórmula del Z-Score original es la siguiente: $Z = 1,2 \cdot X_1 + 1,4 \cdot X_2 + 3,3 \cdot X_3 + 0,6 \cdot X_4 + 1,0 \cdot X_5$, donde X_1 es el capital circulante sobre activos totales, X_2 son los beneficios retenidos sobre activos totales, X_3 es el EBIT sobre activos totales, X_4 es el valor de mercado de los fondos propios sobre el total de deuda, y X_5 es la cifra de ventas sobre activos totales. Altman establece tres zonas de clasificación: empresas con $Z > 2,99$ se consideran financieramente sanas (zona segura); empresas con $Z < 1,81$ presentan un riesgo elevado de insolvencia (zona de peligro); y empresas con Z entre 1,81 y 2,99 se encuentran en una zona gris de incertidumbre.

En el modelo de este trabajo se utiliza el umbral $Z > 2,99$ como condición necesaria para que una empresa sea incluida en la cartera. La justificación es doble: por una parte, este umbral garantiza que el portfolio no contenga empresas en situación de dificultad financiera potencial, lo cual sería incompatible con una estrategia de inversión orientada a la calidad; por otra parte, el riesgo de quiebra suele materializarse precisamente en los entornos de mercado adversos que la estrategia pretende navegar con mayor eficacia, como las recesiones económicas. El Altman Z-Score actúa por tanto como un filtro de exclusión que complementa al Piotroski F-Score: mientras el F-Score selecciona empresas con fundamentales en mejora, el Z-Score asegura que dichas empresas se encuentran en una posición financiera suficientemente sólida como para no estar expuestas a riesgos de insolvencia.

2.4 Momentum de precio: la persistencia de las tendencias

El tercer criterio del modelo es el momentum de precio, cuyo fundamento académico descansa principalmente en el trabajo de Jegadeesh y Titman (1993), publicado en el *Journal of Finance*. Este artículo es uno de los más citados en la historia de las finanzas empíricas: los autores documentaron que estrategias que compren acciones con buen

comportamiento reciente y venden acciones con mal comportamiento reciente generan retornos positivos significativos en períodos de mantenimiento de entre tres y doce meses, con retornos anualizados que oscilan entre el 10% y el 15% para las combinaciones de formación y mantenimiento más efectivas.

El mecanismo subyacente al momentum ha sido objeto de un debate académico intenso. Las explicaciones basadas en el riesgo, en línea con la HME, sugieren que las acciones con mayor momentum son acciones más arriesgadas que requieren una prima de rentabilidad adicional. Sin embargo, la evidencia empírica apunta más hacia explicaciones de carácter conductual: la infrarreacción inicial de los inversores a las buenas noticias —que hace que el precio tarde en incorporar completamente la información positiva— y la posterior sobre-reacción por exceso de confianza generan la persistencia de tendencias que el momentum explota. Independientemente de la explicación, la robustez del efecto momentum ha sido replicada en múltiples mercados, períodos temporales y clases de activos, convirtiéndolo en uno de los factores más documentados y reconocidos en la literatura de finanzas cuantitativas.

En este trabajo se implementa el momentum en su variante más estándar, la denominada 12-1: se calcula el retorno de cada acción en los doce meses anteriores al momento de construcción de la cartera, excluyendo el mes más reciente para evitar el efecto de reversión a corto plazo documentado por Jegadeesh (1990). Dado que los datos disponibles son anuales, se aproxima esta medida como el retorno del año $t-1$ respecto al año $t-2$. Sólo se incluyen en la cartera las empresas que se sitúan en la mitad superior del universo por momentum en el año de screening, garantizando así la exposición al factor de forma relativa dentro del universo de inversión disponible.

2.5 Calidad empresarial: rentabilidad como señal de ventaja competitiva

El cuarto criterio es la calidad empresarial, cuya conceptualización moderna en el ámbito del factor investing se debe principalmente a Novy-Marx (2013), publicado en el *Journal of Financial Economics*. El argumento central de Novy-Marx es que la rentabilidad bruta sobre activos —medida como $(\text{ventas} - \text{coste de ventas}) / \text{activos totales}$ — tiene un poder predictivo sobre las rentabilidades futuras comparable al del ratio book-to-market, el factor de valor por excelencia. Las empresas más rentables generan rentabilidades significativamente superiores a las menos rentables, incluso controlando por otros factores de riesgo bien documentados.

Este hallazgo tiene una interpretación económica intuitiva: las empresas con márgenes elevados y alta rentabilidad son aquellas que poseen ventajas competitivas sostenibles —barreras de entrada, poder de fijación de precios, economías de escala— que les permiten generar valor para el accionista de forma persistente. En contraste con el factor de valor puro, que tiende a capturar empresas en dificultades que cotizan a descuento, el factor de calidad captura empresas con negocios de alta calidad que el mercado sistemáticamente infravalora en términos relativos.

En el modelo de este trabajo, la calidad se aproxima mediante dos indicadores complementarios: la rentabilidad sobre fondos propios (ROE) y el margen operativo (Operating Margin). Se exige que ambos superen el 10%, un umbral que refleja una rentabilidad empresarial genuinamente positiva y distingue a las empresas de calidad del conjunto del universo, al tiempo que resulta suficientemente holgado para no eliminar empresas de sectores con estructuras de capital distintas. La elección de estos dos indicadores responde a la lógica de capturar tanto la rentabilidad para el accionista —ROE— como la eficiencia operativa del negocio en sí mismo —margen operativo—, dos dimensiones complementarias de la calidad empresarial que Novy-Marx (2013) identifica como predictores robustos de rentabilidad futura.

2.6 Valoración relativa: EV/EBITDA como criterio de precio

El quinto y último criterio del modelo es la valoración relativa mediante el múltiplo EV/EBITDA, cuyo fundamento en el contexto del factor investing puede rastrearse hasta Greenblatt (2006), quien popularizó el uso combinado de rentabilidad y valoración en su denominada "fórmula mágica". El Enterprise Value to EBITDA es uno de los múltiplos de valoración más utilizados tanto en la práctica del análisis financiero como en la investigación académica, por su capacidad para comparar empresas con independencia de su estructura de capital y sus diferencias en tasas impositivas y políticas de amortización.

La lógica de incluir este criterio en el modelo es la de evitar pagar un precio excesivo por empresas que, aun siendo financieramente sanas, con buen momentum y de alta calidad, cotizan ya a valoraciones que descuentan expectativas de crecimiento muy optimistas. En términos prácticos, se exige que el EV/EBITDA de cada empresa candidata sea inferior a la mediana del universo de inversión en ese año concreto. Este umbral relativo —frente a un umbral absoluto— tiene la ventaja de adaptarse automáticamente a las condiciones

generales de valoración del mercado: en un entorno de mercado caro como el de finales de los años noventa o la era post-pandemia, la mediana del universo es más elevada, pero el filtro sigue seleccionando las empresas relativamente más baratas dentro de las condiciones disponibles. Ello garantiza que el criterio de valoración no se vuelve irrelevante en entornos de mercado alcista generalizado.

2.7 Ciclos macroeconómicos y performance de factores

Una dimensión analítica adicional de este trabajo es el estudio del comportamiento de la estrategia en función del ciclo macroeconómico. La literatura sobre la variación temporal del alpha de los factores de inversión ha documentado de forma consistente que distintos factores se comportan de forma diferente en función de la fase del ciclo económico y del régimen de política monetaria. Chordia y Shivakumar (2002) documentaron que el momentum presenta un comportamiento especialmente favorable en entornos de expansión económica y deteriora su performance en recesiones, mientras que los factores de valor tienden a recuperarse con fuerza en las fases iniciales de expansión tras períodos recesivos.

En este trabajo se emplea la clasificación oficial del NBER para identificar los años de recesión económica en Estados Unidos, y la evolución del tipo de interés objetivo de los fondos federales de la Reserva Federal para distinguir entre regímenes de política monetaria restrictiva —tightening, períodos de subida de tipos— y acomodaticia —easing, períodos de bajada o mantenimiento en niveles bajos—. Esta doble clasificación permite analizar si la estrategia multifactorial desarrollada en este trabajo es robusta a través de distintos entornos macroeconómicos, o si su capacidad de generar alpha se concentra en condiciones específicas de mercado, lo que tendría implicaciones importantes para su viabilidad como herramienta de inversión sistemática a largo plazo.

3. Metodología

3.1 Diseño general del estudio

El presente trabajo adopta un enfoque cuantitativo de backtesting, consistente en la simulación retrospectiva de una estrategia de inversión sobre datos históricos reales para evaluar su performance en condiciones lo más cercanas posible a las que habría enfrentado un inversor real. La metodología se articula en cinco fases secuenciales: la construcción del universo de inversión, la aplicación anual del proceso de screening multifactorial, la construcción de la cartera y el cálculo de rentabilidades, la evaluación

de la performance mediante métricas estándar, y el análisis diferencial por ciclo macroeconómico. A lo largo de todo el proceso se aplican medidas estrictas para garantizar la ausencia de look-ahead bias, es decir, la prohibición de utilizar en el proceso de selección cualquier información que no hubiera estado disponible para el inversor en el momento real de toma de decisiones.

El período de análisis cubre desde 2004 hasta 2024, lo que supone un total de 21 años de backtest. Esta ventana temporal ha sido elegida por tres razones. En primer lugar, es suficientemente larga para capturar múltiples ciclos económicos completos y obtener conclusiones estadísticamente robustas. En segundo lugar, incluye episodios de stress de mercado de naturaleza muy diversa —la crisis financiera global de 2008-2009, la corrección de 2020 asociada a la pandemia, y los mercados bajistas de 2022— que permiten evaluar la resiliencia de la estrategia en entornos adversos. En tercer lugar, la disponibilidad de datos fundamentales de calidad en Capital IQ para el universo considerado es alta y consistente a partir de los primeros años del siglo XXI, lo que garantiza la fiabilidad de los cálculos.

3.2 Fuentes de datos

Los datos utilizados en este trabajo proceden de dos fuentes diferenciadas. La primera es Capital IQ, plataforma de datos financieros de S&P Global, de la cual se han extraído los datos fundamentales y de precio para una muestra de 902 empresas históricamente presentes en el S&P 500, con cobertura desde el año 2000 hasta el año 2024. La base de datos contiene 19 variables financieras para cada empresa y año, organizadas en hojas individuales: rentabilidad sobre activos (ROA), flujo de caja operativo (CFO), activos totales, deuda a largo plazo, ratio de liquidez corriente, acciones en circulación, margen bruto, rotación de activos, capital circulante, beneficios retenidos, EBIT, capitalización bursátil, ingresos, pasivo total, rentabilidad sobre fondos propios (ROE), margen operativo, EBITDA, valor de empresa (TEV) y precio de cierre ajustado anual.

Un aspecto crítico del tratamiento de los datos es la coherencia de unidades entre variables. El TEV y la capitalización bursátil están expresados en millones de dólares, mientras que las variables de balance —activos totales, deuda, capital circulante, beneficios retenidos, EBIT, ingresos, pasivo total y EBITDA— están expresadas en miles de dólares. Esta diferencia de escala, documentada en los metadatos de Capital IQ, requiere las conversiones oportunas en el cálculo del EV/EBITDA y del Altman Z-Score:

en el primero, el EBITDA se divide entre 1.000 para expresarlo en millones antes de calcular el ratio; en el segundo, la capitalización bursátil se multiplica por 1.000 para expresarla en miles de dólares antes de dividirla entre el pasivo total.

La segunda fuente de datos es el S&P 500 Price Return mensual para el período enero 2000 – enero 2025, utilizado exclusivamente como benchmark de comparación. Es importante señalar que el benchmark empleado es el Price Return del índice, que no incluye los dividendos distribuidos por las empresas que lo componen. Esta elección introduce una asimetría metodológica que debe reconocerse explícitamente: mientras que los precios individuales de las acciones del portfolio capturan implícitamente el efecto de reinversión de dividendos —al basarse en precios ajustados—, el benchmark no lo hace. Ello puede generar una ligera sobreestimación del alpha real de la estrategia. Esta limitación se discute en detalle en el capítulo de conclusiones.

3.3 Construcción del universo de inversión

El universo de inversión está compuesto por las empresas presentes en la base de datos de Capital IQ que formaban parte del S&P 500 durante el año de screening. Para cada año t , se identifican las empresas elegibles aplicando dos condiciones: que la fecha de primera inclusión en el índice sea anterior al 31 de diciembre del año t , y que la fecha de última inclusión en el índice sea posterior al 1 de enero del año t —o bien que la empresa siga siendo miembro actual del índice—. Esta lógica de filtrado garantiza que el universo de cada año únicamente contiene empresas que efectivamente formaban parte del índice en ese período, evitando el denominado survivorship bias que se produciría si se utilizara únicamente la composición actual del S&P 500 como universo histórico.

La muestra contiene 902 empresas en total, de las cuales 519 son miembros actuales del índice y 383 son empresas históricas que en algún momento formaron parte del mismo pero que han sido eliminadas por fusiones, adquisiciones, quiebras o simplemente por no cumplir los criterios de inclusión en el índice. El universo válido por año oscila entre 434 empresas en 2003 y 510 empresas en 2024, con una media de 473 empresas por año a lo largo del período de backtest, lo que refleja tanto la expansión natural del universo cubierto por Capital IQ como la rotación normal en la composición del índice.

3.4 Proceso de screening multifactorial

El núcleo metodológico del trabajo es el proceso de screening secuencial que se aplica anualmente al universo de inversión. Para cada año t , se calcula el valor de cada uno de los cinco criterios para todas las empresas del universo y se van aplicando los filtros en el siguiente orden, eliminando en cada paso las empresas que no cumplen el criterio correspondiente.

El primer filtro es el Piotroski F-Score, calculado con base en los datos financieros del año t . Se requiere un F-Score mínimo de 7 sobre 9 puntos posibles. Para el cálculo de las nueve señales binarias se necesitan los datos de al menos dos ejercicios consecutivos — año t y año $t-1$ — para las señales comparativas. Se exige que estén disponibles como mínimo los valores de ROA, CFO y activos totales del año t para que la empresa sea elegible; en caso contrario, se asigna un valor nulo y la empresa es excluida. Para las señales que requieren comparación interanual, si el dato del año anterior no está disponible, la señal correspondiente se omite del cómputo total, de modo que el F-Score se calcula sobre el subconjunto de señales disponibles, requiriéndose un mínimo de seis señales válidas para que el resultado sea considerado fiable. Tras aplicar este primer filtro, el universo se reduce de una media de 473 a 148 empresas por año, lo que implica que aproximadamente el 31% de las empresas cumplen el requisito de solidez financiera básica.

El segundo filtro es el Altman Z-Score calculado con los datos del año t . Se requiere un valor superior a 2,99, umbral que Altman (1968) identifica como la frontera de la zona segura. El cálculo requiere siete variables del balance y la cuenta de resultados, todas ellas disponibles en la base de datos de Capital IQ. Los cálculos se realizan con la coherencia de unidades descrita anteriormente, convirtiendo la capitalización bursátil de millones a miles de dólares antes de dividirla entre el pasivo total. Tras aplicar este filtro, el universo se reduce a una media de 81 empresas por año, lo que refleja que aproximadamente el 55% de las empresas que superan el Piotroski también cumplen el requisito de solidez financiera frente al riesgo de insolvencia.

El tercer filtro es el momentum de precio. Dado que los datos de precio disponibles son anuales —el precio de cierre ajustado a 31 de diciembre de cada año—, el momentum se aproxima como el retorno del año $t-1$ respecto al año $t-2$. Esto es coherente con la convención estándar del momentum 12-1: se utiliza el retorno del año inmediatamente anterior al de screening como señal, excluyendo el año de screening en curso para evitar la reversión a corto plazo. El umbral aplicado es el percentil 50 del universo en cada año:

sólo las empresas con momentum igual o superior a la mediana del universo son seleccionadas, garantizando que la cartera siempre contiene empresas con impulso de precio relativo positivo independientemente de las condiciones absolutas de mercado. Tras este filtro, el universo se reduce a una media de 50 empresas por año.

El cuarto filtro aplica los criterios de calidad: se requiere simultáneamente una rentabilidad sobre fondos propios (ROE) superior al 10% y un margen operativo superior al 10%, ambos calculados con los datos del año t . La exigencia conjunta de ambas condiciones garantiza que las empresas seleccionadas son rentables tanto desde la perspectiva del accionista como desde la perspectiva operativa del negocio. Tras este filtro, el universo se reduce a una media de 37 empresas por año.

El quinto y último filtro es el de valoración relativa mediante EV/EBITDA. El ratio se calcula dividiendo el TEV —expresado en millones de dólares— entre el EBITDA convertido a millones de dólares. Se excluyen las empresas con EBITDA negativo o nulo, ya que el múltiplo pierde su significado económico en esos casos. El umbral es la mediana del ratio EV/EBITDA del universo completo en cada año t : sólo las empresas que cotizan por debajo de esta mediana son incluidas en la cartera final. Tras aplicar este último filtro, la cartera queda formada por una media de 15 empresas por año, con un mínimo de 8 en 2020 y un máximo de 25 en 2008.

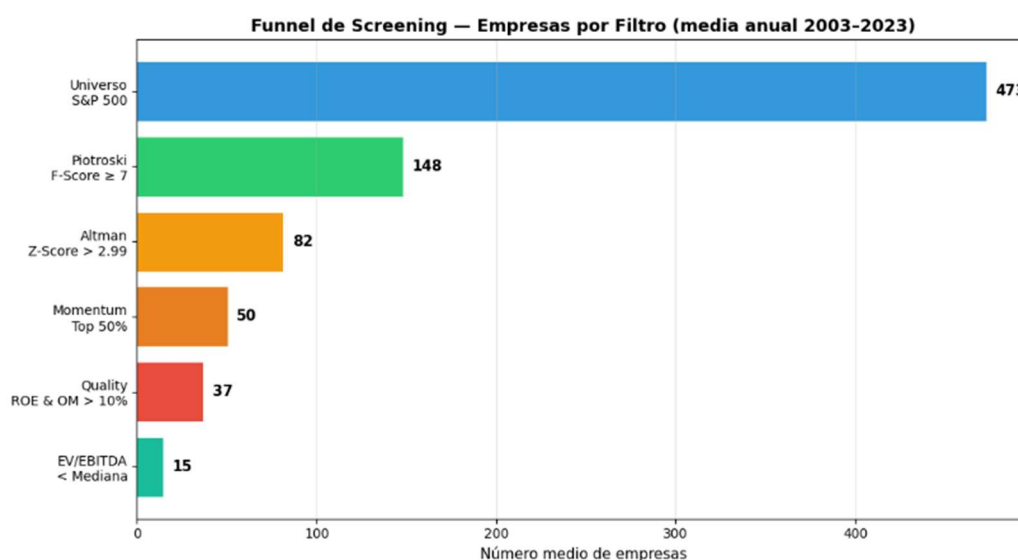


Figura 1. Número medio de empresas seleccionadas tras la aplicación secuencial de cada filtro de screening a lo largo del período 2003–2023. Fuente: elaboración propia a partir de datos de Capital IQ.

Una consideración metodológica importante es que todos los umbrales relativos —la mediana del EV/EBITDA y el percentil 50 del momentum— se calculan exclusivamente sobre el universo válido del año t , sin utilizar información de años futuros. De igual manera, todos los datos financieros utilizados en el screening corresponden al año t , cuyas cuentas anuales se publican típicamente en los primeros meses del año $t+1$. Esto implica que la cartera construida con datos del año t se implementa a comienzos del año $t+1$, y su rentabilidad se mide desde el precio de cierre de diciembre del año t hasta el precio de cierre de diciembre del año $t+1$. Este esquema de implementación es estándar en la literatura de backtesting de estrategias basadas en datos fundamentales y garantiza la completa ausencia de look-ahead bias.

3.5 construcción de la cartera y calculo de rentabilidades

La cartera se construye aplicando ponderación igual a todas las acciones seleccionadas en cada año, con rebalanceo anual. La elección de la ponderación igual responde a tres razones. En primer lugar, es la aproximación metodológicamente más limpia para evaluar la capacidad de selección de acciones del modelo, ya que no introduce un sesgo implícito hacia empresas de mayor capitalización —que podría confundir los resultados con una exposición encubierta a un factor de tamaño—. En segundo lugar, la ponderación igual maximiza la explotación del proceso de screening, otorgando el mismo peso a cada empresa seleccionada independientemente de su tamaño. En tercer lugar, es coherente con el espíritu del Piotroski F-Score y de los otros criterios utilizados, que son por naturaleza agnósticos respecto al tamaño de la empresa.

La rentabilidad anual de la cartera en el año $t+1$ se calcula como la media aritmética de los retornos individuales de cada acción, donde el retorno individual se define como $(P_{t+1} / P_t) - 1$, siendo P_t el precio de cierre ajustado de diciembre del año t y P_{t+1} el precio de cierre ajustado de diciembre del año $t+1$. El uso de precios ajustados garantiza que los retornos incorporan los ajustes por splits, distribuciones de acciones y otros eventos corporativos, aunque como se ha señalado anteriormente, pueden capturar parcialmente el efecto de los dividendos, lo que introduce una asimetría respecto al benchmark de Price Return.

La rentabilidad del benchmark S&P 500 para cada año $t+1$ se calcula de forma análoga como $(P_{S\&P,t+1} / P_{S\&P,t}) - 1$, utilizando los precios de cierre mensuales del índice correspondientes al mes de diciembre de cada año. El alpha anual de la estrategia

se define como la diferencia aritmética entre la rentabilidad de la cartera y la rentabilidad del benchmark en ese mismo año.

3.6 Métricas de evaluación de la performance

La evaluación de la estrategia se realiza mediante cinco métricas estándar en el análisis de carteras. La rentabilidad anual compuesta o CAGR se calcula como la media geométrica de los retornos anuales, expresada como $(\prod (1 + r_t))^{(1/n)} - 1$. La volatilidad anual se mide como la desviación típica de los retornos anuales. El ratio de Sharpe se calcula como el cociente entre la rentabilidad media anual y la volatilidad anual; dado que el objetivo del trabajo es comparar la cartera con el benchmark en términos relativos, se utiliza una versión simplificada sin deducción del tipo libre de riesgo, lo que es equivalente a asumir un tipo libre de riesgo de cero, convención habitual en los estudios de performance relativa de largo plazo. El máximo drawdown se calcula sobre el valor acumulado de la cartera como la mayor caída desde máximo hasta mínimo a lo largo del período completo. Finalmente, la tasa de éxito o win rate se define como la proporción de años en que la cartera supera al benchmark.

3.7 Clasificación de los ciclos macroeconómicos

Para el análisis diferencial por ciclo macroeconómico se emplean dos clasificaciones independientes. La primera utiliza las fechas oficiales de recesión del National Bureau of Economic Research (NBER), organismo de referencia en la datación de ciclos económicos en Estados Unidos. Se clasifican como años de recesión aquellos en que el NBER identifica al menos un trimestre de recesión: 2008, 2009 y 2020 dentro del período de backtest. El resto de años se clasifican como de expansión económica.

La segunda clasificación atiende al régimen de política monetaria de la Reserva Federal, distinguiendo entre períodos de endurecimiento monetario —tightening, caracterizados por una tendencia alcista en el tipo objetivo de los fondos federales— y períodos de relajación —easing, caracterizados por una tendencia bajista o por tipos en mínimos históricos—. Se clasifican como años de tightening los períodos 2004–2007, 2015–2018 y 2022–2023, correspondientes a los tres ciclos de subidas de tipos de la Reserva Federal documentados en el período analizado. El resto de años se clasifican como de easing. Esta clasificación es dicotómica a nivel anual y se asigna en función de la tendencia dominante del tipo de interés en cada año natural.

4. Resultados

4.1 Resultados agregados del backtest (2004-2024)

El backtest cubre un período de 21 años y genera resultados consistentes con la hipótesis de partida del trabajo: la combinación sistemática de criterios de solidez financiera, calidad empresarial, momentum y valoración relativa permite construir una cartera que supera de forma significativa al S&P 500 como benchmark a lo largo del período completo. La Tabla 1 recoge las métricas de performance agregadas para la cartera y el índice de referencia.

Tabla 1. Métricas de performance agregadas (2004-2024)

Métrica	Portfolio	S&P 500
CAGR (Rentabilidad Anual Compuesta)	14,22%	8,26%
Volatilidad Anual	20,38%	16,24%
Ratio de Sharpe	0,19	0,597
Máximo Drawdown	-27,84%	-38,49%
Alpha Medio Anual	6,41%	—
Win Rate (años > benchmark)	57,14%	—
NAV Final (\$1 invertido en 2003)	\$16,32	\$5,29

La cartera genera una rentabilidad anual compuesta del 14,22%, frente al 8,26% del S&P 500, lo que supone un diferencial de rentabilidad de 596 puntos básicos anuales en términos de CAGR. Expresado en términos de valor acumulado, un inversor que hubiera destinado un dólar a la estrategia a comienzos de 2004 habría obtenido 16,32 dólares al final de 2024, frente a los 5,29 dólares que habría generado la inversión pasiva en el índice. Esta diferencia de valor acumulado refleja el efecto del interés compuesto sobre una ventaja anual de rentabilidad mantenida durante dos décadas.

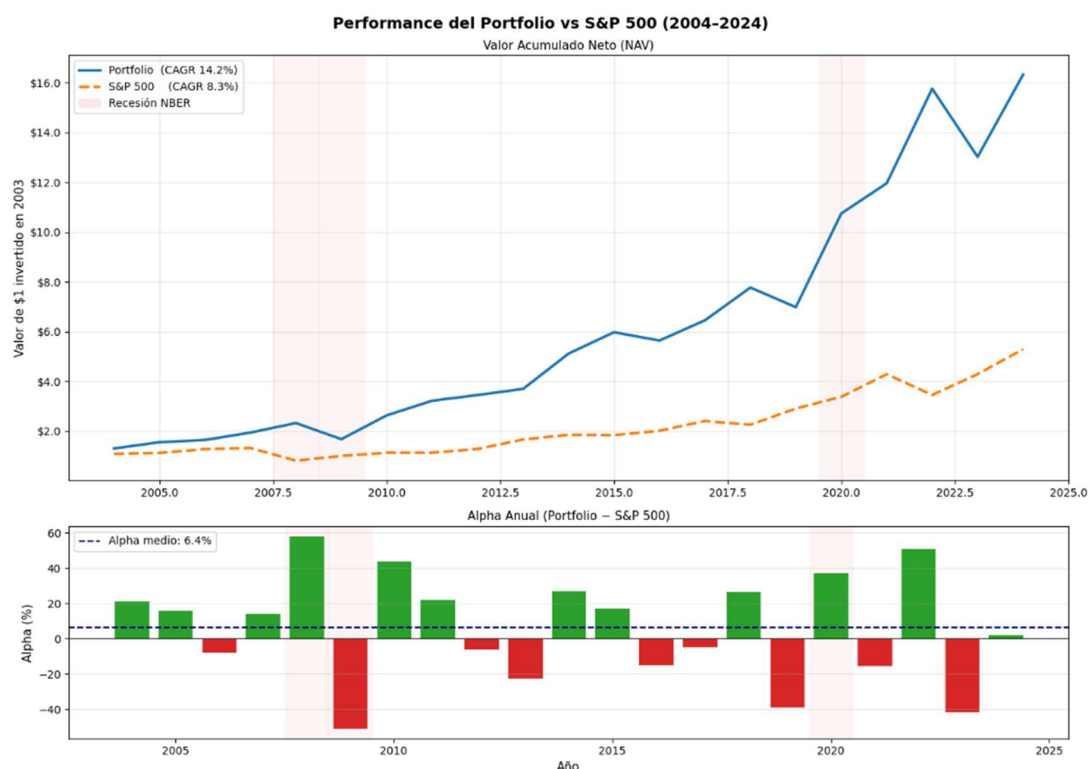


Figura 2. Valor acumulado neto (NAV) de \$1 invertido en 2003 y alpha anual generado por la estrategia frente al S&P 500 (2004–2024). Las zonas sombreadas en rojo corresponden a años de recesión según la clasificación NBER. Fuente: elaboración propia.

El ratio de Sharpe de la cartera, de 0,790, supera al del benchmark, de 0,597, lo que indica que la estrategia no sólo genera más rentabilidad sino que lo hace con una mejor relación rentabilidad-riesgo. Es importante señalar que la volatilidad anual de la cartera, del 20,38%, es superior a la del índice, del 16,24%. Esta mayor volatilidad es coherente con el reducido número de valores en cartera —una media de 15 acciones por año— y con el hecho de que la estrategia no está diseñada para minimizar el riesgo absoluto sino para maximizar la selección de empresas con fundamentales superiores. Sin embargo, el ratio de Sharpe superior indica que el exceso de rentabilidad generado compensa con creces el exceso de riesgo asumido.

El máximo drawdown de la cartera, del -27,84%, es notablemente inferior al del S&P 500, del -38,49%. Este resultado es especialmente relevante desde una perspectiva práctica: la estrategia no sólo genera más rentabilidad a largo plazo sino que además protege mejor el capital del inversor en los peores momentos de mercado, lo que tiene implicaciones directas sobre la experiencia psicológica y la capacidad de mantenimiento de la estrategia en condiciones adversas. La tasa de éxito anual, del 57,14%, implica que la cartera supera al benchmark en 12 de los 21 años del período analizado, lo que

evidencia que el alpha no se concentra en unos pocos años excepcionales sino que se distribuye de forma razonablemente amplia a lo largo del período.

4.2 Análisis del comportamiento anual

El análisis año por año de los resultados revela patrones interesantes que complementan y matizan la visión agregada. La Tabla 2 recoge los retornos anuales de la cartera y el benchmark, el alpha generado y el número de valores en cartera para cada año del período.

Tabla 2. Resultados anuales del backtest (2004-2024)

Año	Portfolio	S&P 500	Alpha	#valores	Recesión	Tightening
2004	30,24%	8,99%	21,25%	6		✓
2005	19,23%	3,00%	16,23%	11		✓
2006	5,84%	13,62%	-7,78%	8		✓
2007	18,03%	3,53%	14,50%	15		✓
2008	19,76%	-38,49%	58,24%	21	✓	
2009	-27,84%	23,45%	-51,30%	25	✓	
2010	57,05%	12,78%	44,27%	19		
2011	22,32%	0,00%	22,32%	13		
2012	7,13%	13,41%	-6,27%	11		
2013	7,12%	29,60%	-22,49%	15		
2014	38,40%	11,39%	27,01%	11		
2015	16,74%	-0,73%	17,47%	9		✓
2016	-5,57%	9,54%	-15,11%	16		✓
2017	14,41%	19,42%	-5,01%	17		✓
2018	20,40%	-6,24%	26,63%	17		✓
2019	-10,13%	28,88%	-39,01%	16		
2020	53,91%	16,26%	37,65%	24	✓	
2021	11,31%	26,89%	-15,58%	8		
2022	31,76%	-19,44%	51,20%	7		✓
2023	-17,38%	24,23%	-41,61%	18		✓
2024	25,38%	23,31%	2,07%	21		
MEDIA	16,10%	9,69%	6,41%	14,67		

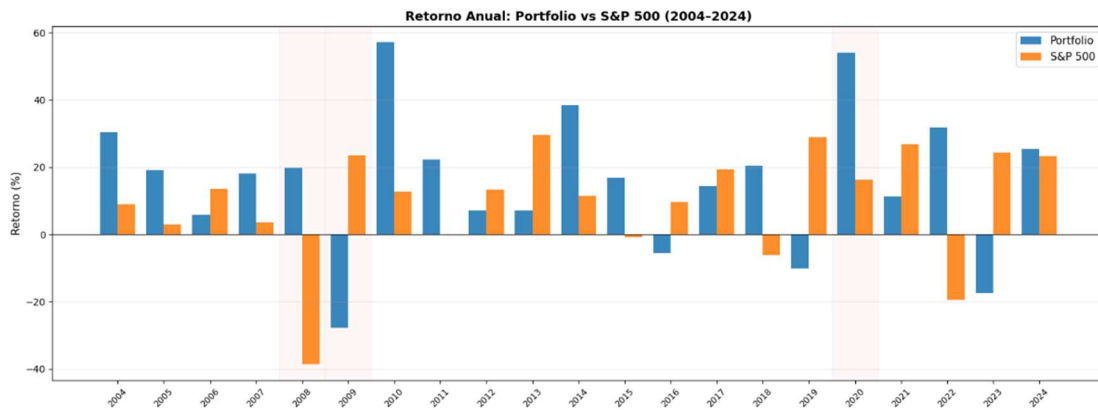


Figura 3. Retorno anual de la cartera sistemática y del S&P 500 para cada año del período 2004–2024. Fuente: elaboración propia a partir de datos de Capital IQ y S&P 500 Price Return.

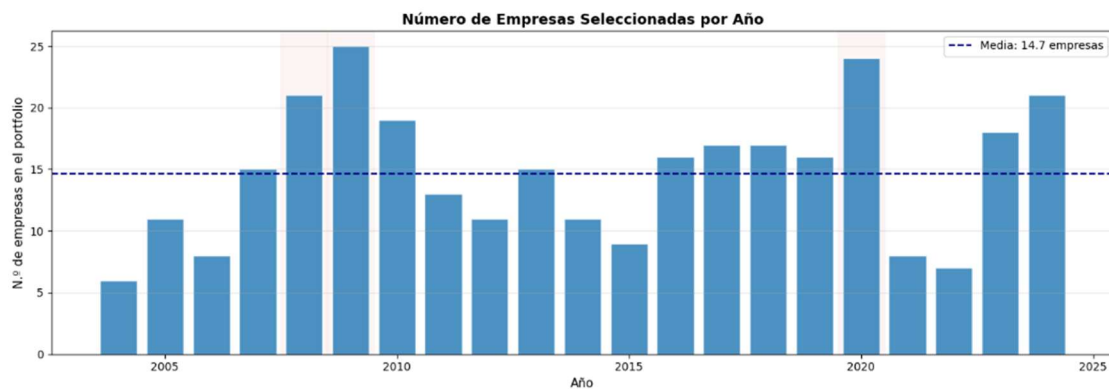


Figura 4. Número de valores incluidos en la cartera en cada año del período 2004–2024. La línea discontinua indica la media del período (15 valores). Las zonas sombreadas corresponden a años de recesión NBER. Fuente: elaboración propia.

El año de mayor alpha es 2008, con una diferencia de 58,24 puntos porcentuales: mientras el S&P 500 caía un 38,49% en el epicentro de la crisis financiera global, la cartera generaba un retorno positivo del 19,76%. Este resultado es especialmente significativo porque demuestra que los filtros de solidez financiera y ausencia de riesgo de quiebra — Piotroski F-Score y Altman Z-Score— fueron capaces de identificar empresas que resistieron el entorno de crisis sistémica de 2008 sin sufrir las caídas devastadoras que golpearon al conjunto del mercado. El año inmediatamente posterior, 2009, presenta el peor alpha del período, con −51,30 puntos porcentuales: el índice se recuperó con fuerza un 23,45% desde sus mínimos, mientras la cartera, que había evitado las empresas más golpeadas en 2008, no participó de la misma magnitud en el rebote. Este patrón de secuencia 2008-2009 es coherente con la naturaleza de la estrategia: el screening basado

en calidad y solidez tiende a excluir precisamente las empresas más cíclicas y endeudadas que lideran los rebotes tras crisis de crédito.

Otros años de alpha negativo significativo son 2013 (−22,49%), 2019 (−39,01%) y 2023 (−41,61%). En los tres casos se trata de años de expansión bursátil fuerte y generalizada en que el mercado fue liderado por un grupo reducido de empresas de elevada capitalización y crecimiento —fundamentalmente el sector tecnológico— cuyas valoraciones elevadas las habrían excluido del universo seleccionable por el filtro de EV/EBITDA y cuyas métricas de Piotroski no necesariamente se encontraban en los valores más altos del universo. Este comportamiento negativo en mercados alcistas concentrados en empresas de crecimiento de alta valoración es una característica intrínseca y esperada de cualquier estrategia orientada al valor y la calidad, y debe entenderse como el coste necesario para la protección que la estrategia ofrece en entornos adversos.

4.3 Análisis por ciclo macroeconómico: recesión frente a expansión

El análisis diferencial por fase del ciclo económico revela uno de los hallazgos más relevantes del trabajo. La Tabla 3 recoge los retornos medios anuales de la cartera y el benchmark, así como el alpha medio, desagregados por la clasificación NBER.

Tabla 3. Performance por ciclo NBER

Fase del Ciclo	N años	Portfolio (media)	S&P 500 (media)	Alpha Medio
Recesión	3	15,27%	0,41%	14,87%
Expansión	18	16,24%	11,23%	5,01%

El alpha medio en años de recesión, de 14,87 puntos porcentuales, es casi tres veces superior al alpha en años de expansión, de 5,01 puntos porcentuales. Este resultado confirma que la capacidad de la estrategia para generar rentabilidades superiores es especialmente pronunciada en entornos de contracción económica, precisamente los entornos en que la selección de empresas financieramente sólidas y sin riesgo de insolvencia tiene mayor valor. Cuando el ciclo económico se deteriora, la dispersión de rentabilidades entre empresas sanas y empresas frágiles se amplía considerablemente, y los filtros del Piotroski F-Score y el Altman Z-Score son especialmente efectivos para evitar las empresas más expuestas al riesgo de quiebra o deterioro severo de beneficios.

Cabe destacar que el número de años de recesión en la muestra es reducido —únicamente tres: 2008, 2009 y 2020—, lo que exige cierta cautela en la interpretación de los resultados para esta categoría. El elevado alpha medio en recesión está fuertemente influenciado por el año 2008, en que la diferencia de comportamiento entre la cartera y el índice fue extraordinariamente grande incluso para los estándares históricos de dispersión entre estrategias activas y el mercado.

4.4 Análisis por ciclo de política monetaria

La Tabla 4 desagrega los resultados en función del régimen de política monetaria de la Reserva Federal, distinguiendo entre los períodos de tightening y de easing definidos en la metodología.

Tabla 4. Performance por ciclo de política monetaria

Régimen Monetario	N años	Portfolio (media)	S&P 500 (media)	Alpha Medio
Tightening	10	13,37%	5,59%	7,78%
Easing	11	18,58%	13,41%	5,17%

La estrategia genera mayor alpha en entornos de política monetaria restrictiva, con 7,78 puntos porcentuales frente a los 5,17 puntos en entornos de easing. Este resultado es intuitivamente coherente: en períodos de subida de tipos de interés, el coste de la deuda aumenta y las empresas altamente apalancadas o con flujos de caja débiles se ven penalizadas de forma más severa por el mercado. La estrategia, al filtrar empresas con bajo riesgo de insolvencia y flujos de caja operativos positivos, tiende a construir carteras con menor exposición al riesgo de tipos de interés, lo que genera un alpha diferencial positivo en estos entornos. Este hallazgo es especialmente relevante en el contexto del ciclo de subidas de tipos de 2022-2023, uno de los más agresivos de las últimas décadas, en el que la cartera generó un retorno del 31,76% en 2022 mientras el índice caía un 19,44%.

En entornos de easing, el alpha de la estrategia, aunque inferior al de los períodos de tightening, se mantiene positivo en 5,17 puntos porcentuales, lo que indica que la estrategia no depende de un régimen monetario específico para generar retornos superiores al benchmark sino que muestra consistencia a través de distintas condiciones financieras.

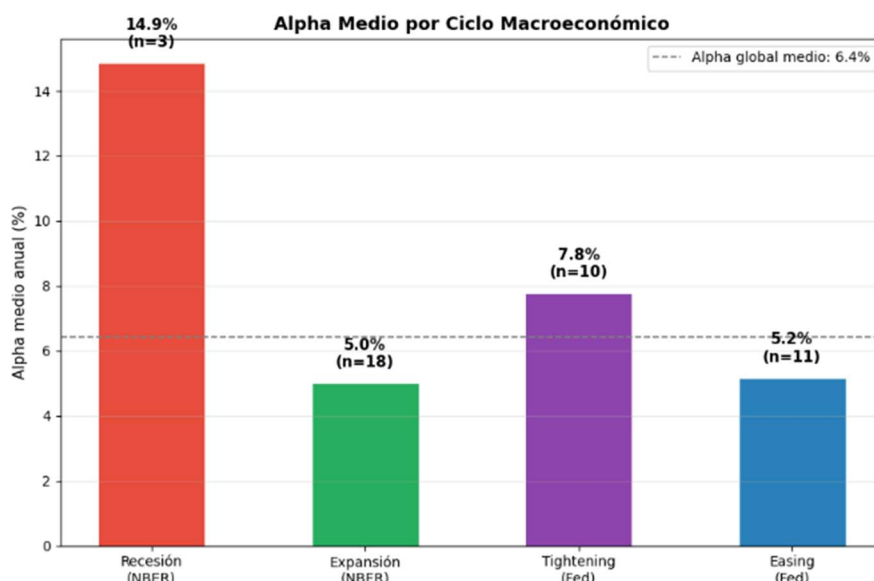


Figura 5. Alpha medio anual de la estrategia desagregado por fase del ciclo económico (NBER) y régimen de política monetaria (Fed) para el período 2004–2024. Fuente: elaboración propia.

4.5 Distribución sectorial de la cartera

A lo largo de los 21 años del backtest, el proceso de screening ha seleccionado un total de 309 acciones repartidas en 21 años, con una media de 15 valores por año como se ha señalado. El análisis de la distribución sectorial acumulada revela una concentración significativa en determinados sectores de la economía. El sector industrial representa el 23,0% de las selecciones acumuladas, seguido del sector tecnológico con el 20,1% y el consumo discrecional con el 16,2%. El sector energético representa el 11,3% de las selecciones y el de materiales el 9,7%. Los sectores financiero, de comunicaciones y de consumo básico presentan representaciones menores, del 1,9%, 1,3% y 4,5% respectivamente.

Esta distribución sectorial no es aleatoria sino que refleja las características inherentes de los criterios de screening utilizados. Los sectores industrial y de materiales tienden a presentar métricas de Piotroski favorables por la naturaleza cíclica de sus negocios, con fuertes mejoras de beneficios en fases expansivas y balances con activos tangibles que elevan el Altman Z-Score. El sector tecnológico aporta empresas con márgenes operativos elevados —favorables para el filtro de calidad— y frecuentemente cotizaciones dinámicas que favorecen el filtro de momentum. La baja representación del sector financiero es coherente con las limitaciones conocidas del Altman Z-Score original para empresas de este sector, cuyas ratios de balance tienen una interpretación diferente a la

de empresas industriales y para las que el modelo de Altman (1968) no fue diseñado originalmente.

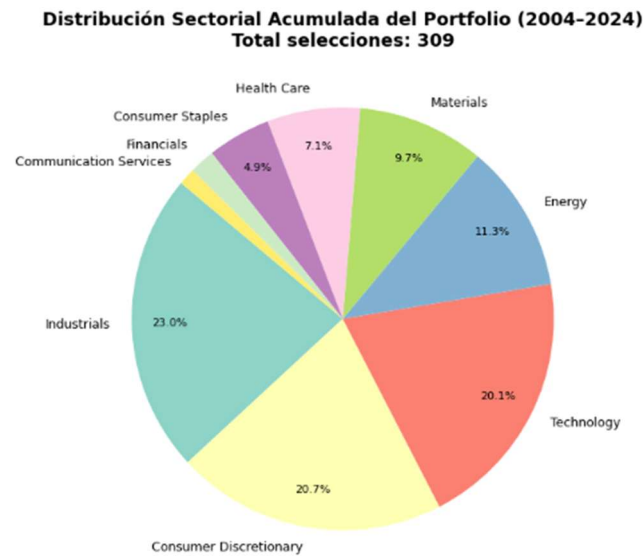


Figura 6. Distribución sectorial acumulada de las selecciones realizadas por la estrategia a lo largo del período 2004–2024 ($n = 309$ selecciones totales). Fuente: elaboración propia a partir de datos de Capital IQ.

5. Conclusiones, limitaciones y líneas de investigación futura

5.1 Conclusiones principales

El presente trabajo ha desarrollado y evaluado una estrategia de selección sistemática de acciones aplicada al universo del S&P 500 durante el período 2004–2024, combinando cinco criterios fundamentados en la literatura académica de finanzas cuantitativas: el Piotroski F-Score, el Altman Z-Score, el momentum de precio, los indicadores de calidad empresarial y el múltiplo EV/EBITDA. Los resultados del backtest permiten responder afirmativamente a la pregunta de investigación planteada en la introducción: la combinación disciplinada de estos criterios genera rentabilidades superiores al índice de referencia de forma consistente a lo largo del período analizado, con un alpha medio anual de 6,41 puntos porcentuales y un CAGR del 14,22% frente al 8,26% del S&P 500 Price Return.

La primera conclusión del trabajo es que la integración simultánea de múltiples factores de distinta naturaleza —solidez financiera, riesgo de insolvencia, momentum, calidad y valoración— produce resultados más robustos que los que cabría esperar de la aplicación aislada de cualquiera de ellos. Cada criterio aporta una dimensión de información

diferente sobre la empresa: el Piotroski F-Score evalúa la trayectoria reciente de sus fundamentales; el Altman Z-Score mide su distancia al riesgo de quiebra; el momentum captura la tendencia del precio como señal del consenso del mercado; la calidad identifica negocios con ventajas competitivas genuinas; y la valoración relativa garantiza que no se paga en exceso por ninguna de las empresas seleccionadas. La combinación secuencial de estos cinco filtros actúa como un proceso de depuración progresiva del universo de inversión que, al final, retiene únicamente las empresas que superan simultáneamente todos los criterios, reduciendo el universo inicial de 473 empresas de media a tan sólo 15 por año.

La segunda conclusión es que la estrategia presenta una robustez especialmente pronunciada en entornos de stress de mercado. El alpha medio en años de recesión según la clasificación NBER alcanza los 14,87 puntos porcentuales, casi tres veces superior al alpha generado en años de expansión. El ejemplo más ilustrativo es el año 2008, en el que la cartera generó un retorno positivo del 19,76% mientras el S&P 500 caía un 38,49%, una diferencia de 58 puntos porcentuales que refleja la capacidad de los filtros de solidez financiera —Piotroski y Altman— para evitar empresas en situación de fragilidad financiera precisamente cuando el ciclo económico se deteriora de forma severa. Este resultado confirma que la estrategia no es únicamente una herramienta de generación de rentabilidad en mercados favorables, sino también un mecanismo de protección del capital en entornos adversos, lo que la dota de un valor añadido significativo desde la perspectiva de la gestión del riesgo a largo plazo.

La tercera conclusión atañe al comportamiento diferencial de la estrategia en función del régimen de política monetaria. La estrategia genera un alpha superior en períodos de tightening, con 7,78 puntos porcentuales frente a los 5,17 de los períodos de easing. Esta asimetría es coherente con la lógica económica subyacente: cuando el coste del capital aumenta, el mercado penaliza con mayor severidad a las empresas con balances débiles, elevado apalancamiento o flujos de caja negativos, que son precisamente las empresas que los filtros de Piotroski y Altman excluyen del universo seleccionable. La estrategia, al concentrarse en empresas financieramente sólidas y generadoras de caja, resulta naturalmente menos expuesta al riesgo de tipos de interés que el conjunto del índice.

La cuarta conclusión, de carácter más matizado, es que la estrategia presenta episodios de underperformance significativa en años de expansión bursátil concentrada en valores de crecimiento de alta valoración. Los años 2009, 2013, 2019 y 2023 registran alphas

negativos sustanciales, en todos los casos asociados a mercados liderados por un reducido número de empresas tecnológicas o de crecimiento con valoraciones muy elevadas que el filtro de EV/EBITDA habría excluido sistemáticamente. Este comportamiento no es una anomalía del modelo sino una consecuencia directa y esperada de sus premisas de diseño: una estrategia orientada a la calidad y la valoración razonable sacrifica necesariamente parte de la participación en los mercados más especulativos a cambio de la protección que ofrece en los entornos adversos. El inversor que adopta esta estrategia debe ser consciente de este trade-off y disponer de la disciplina y el horizonte temporal necesarios para mantenerla en los períodos en que el mercado premia la euforia sobre los fundamentales.

En conjunto, los resultados son consistentes con la evidencia académica preexistente sobre la persistencia de las anomalías documentadas por Piotroski (2000), Altman (1968), Jegadeesh y Titman (1993), Novy-Marx (2013) y Greenblatt (2006), y contribuyen a la literatura al demostrar que la combinación de estos factores en un marco de screening secuencial produce resultados superiores a los que cabría esperar de su aplicación individual, especialmente en entornos de contracción económica y política monetaria restrictiva.

5.2 Limitaciones del estudio

El trabajo presenta varias limitaciones que deben ser reconocidas explícitamente y que condicionan la interpretación y la generalización de los resultados.

La primera y más relevante es la asimetría en el tratamiento de dividendos entre la cartera y el benchmark. El benchmark utilizado es el S&P 500 Price Return, que no incluye los dividendos distribuidos por las empresas que lo componen. Los precios ajustados de las acciones individuales de la cartera, procedentes de Capital IQ, incorporan ajustes por splits y distribuciones de acciones pero no necesariamente la reinversión de dividendos en efectivo de forma explícita. Esta asimetría introduce una posible sobreestimación del alpha real de la estrategia, cuya magnitud depende del yield de dividendo medio de las empresas seleccionadas en cada año. Aunque esta limitación es común en estudios de backtesting basados en datos fundamentales, su reconocimiento es necesario para una interpretación honesta de los resultados. Una comparación más rigurosa requeriría el uso del S&P 500 Total Return como benchmark, lo que reduciría el alpha estimado en una

magnitud aproximada al yield de dividendo anual del índice, históricamente en torno al 1,5%–2,0%.

La segunda limitación es la ausencia de costes de transacción en el modelo. El backtest calcula rentabilidades brutas sin deducir las comisiones de compraventa, los costes de impacto de mercado ni la horquilla bid-ask asociados al rebalanceo anual de la cartera. Aunque el rebalanceo es anual y el número de valores en cartera es reducido —entre 8 y 25 por año—, lo que limita el volumen de transacciones, los costes de transacción reales reducirían la rentabilidad neta de la estrategia. En un contexto de comisiones actuales cercanas a cero para la mayoría de los inversores minoristas, este efecto es probablemente menor que en décadas anteriores, pero no puede ser descartado completamente, especialmente para inversores institucionales que operan con posiciones de mayor tamaño y mayores costes de impacto.

La tercera limitación es el reducido número de observaciones en la categoría de años de recesión, con únicamente tres años —2008, 2009 y 2020— dentro del período de backtest. Este número es insuficiente para extraer conclusiones estadísticamente robustas sobre el comportamiento de la estrategia en recesión, y los resultados deben interpretarse como indicativos y no como evidencia estadísticamente significativa. El alpha medio en recesión está además fuertemente influenciado por el año 2008, cuyo comportamiento excepcional —una diferencia de 58 puntos porcentuales— no es extrapolable a la media histórica de episodios recesivos.

La cuarta limitación es el potencial survivorship bias residual en la base de datos. Aunque el proceso de construcción del universo incluye empresas históricas que en algún momento formaron parte del S&P 500, la cobertura de Capital IQ para empresas eliminadas del índice por quiebra puede no ser completa, especialmente en los primeros años del período analizado. En la medida en que algunas empresas quebradas no estén presentes en la base de datos, el universo disponible estaría sesgado hacia las supervivientes, lo que podría inflar ligeramente las métricas del Altman Z-Score al excluir los casos de insolvencia más extremos.

La quinta limitación es la naturaleza retroactiva del estudio. Los umbrales de los criterios de screening, aunque se han fijado a partir de los valores establecidos en la literatura original y no mediante optimización sobre los datos de la muestra, han sido seleccionados por el investigador con conocimiento de que dichos valores han funcionado bien en la

investigación previa. Ello introduce un grado de sesgo de selección de la literatura que es difícil de cuantificar pero que debe reconocerse como una limitación inherente a cualquier trabajo que aplique criterios preexistentes de la literatura a un conjunto de datos históricos.

5.3 Líneas de investigación futura

Los resultados y limitaciones del trabajo abren varias líneas de investigación que podrían ampliar y profundizar los hallazgos aquí presentados.

La primera y más inmediata sería la replicación del análisis utilizando el S&P 500 Total Return como benchmark, lo que permitiría cuantificar con precisión el efecto del tratamiento asimétrico de dividendos sobre el alpha estimado y ofrecer una comparación más rigurosa con la inversión pasiva en el índice.

La segunda línea consistiría en la incorporación de costes de transacción reales y el análisis de la implementabilidad práctica de la estrategia para distintos perfiles de inversor —minorista, institucional pequeño e institucional grande—, evaluando en qué medida el alpha bruto sobrevive a las fricciones de mercado reales.

Una tercera línea de investigación de gran interés sería el análisis de la robustez del modelo ante variaciones en los umbrales de cada criterio, mediante un análisis de sensibilidad sistemático que permitiera identificar qué filtros son determinantes para la generación de alpha y cuáles tienen un impacto marginal sobre los resultados. Esta línea sería especialmente valiosa para identificar versiones simplificadas del modelo con menor número de criterios pero performance similar.

Una cuarta extensión natural sería la aplicación de la misma metodología a universos de inversión distintos del S&P 500 —mercados europeos, emergentes o de pequeña capitalización— para evaluar si los factores académicos documentados en el mercado estadounidense mantienen su poder predictivo en otros contextos geográficos y de liquidez.

Finalmente, una quinta línea relevante sería la incorporación de técnicas de machine learning para la selección y ponderación dinámica de los factores en función del régimen macroeconómico vigente, reemplazando los umbrales fijos por modelos adaptativos capaces de identificar qué combinación de criterios resulta más eficaz en cada fase del ciclo. Esta aproximación, situada en la frontera entre el factor investing clásico y la

gestión cuantitativa avanzada, representaría una extensión metodológica coherente con la formación en Business Analytics que sustenta este trabajo.

5.4 Reflexión final

El presente trabajo demuestra que la disciplina cuantitativa aplicada a la selección de acciones, cuando se sustenta en criterios con fundamento teórico sólido y se implementa con rigor metodológico, es capaz de generar valor de forma consistente a lo largo de ciclos económicos y monetarios muy diversos. Los mercados no son perfectamente eficientes en el sentido estricto de la hipótesis de Fama (1970), y la información contable pública, correctamente analizada, sigue conteniendo señales con capacidad predictiva sobre las rentabilidades futuras de las acciones. Al mismo tiempo, los resultados confirman que no existe una estrategia que funcione bien en todos los entornos: el trade-off entre protección en mercados bajistas y participación en mercados alcistas especulativos es una característica intrínseca de cualquier enfoque orientado a la calidad y la valoración razonable, y gestionarlo con expectativas realistas es tan importante como el propio diseño del modelo.

6. Bibliografía

Altman, E.I. (1968) 'Financial ratios, discriminant analysis and the prediction of corporate bankruptcy', *The Journal of Finance*, 23(4), pp. 589–609. <https://doi.org/10.1111/j.1540-6261.1968.tb00843.x>

Board of Governors of the Federal Reserve System (US) (2024) *Federal Funds Effective Rate [FEDFUNDS]*. FRED, Federal Reserve Bank of St. Louis. Disponible en: <https://fred.stlouisfed.org/series/FEDFUNDS> (Consultado: junio 2025).

Chordia, T. y Shivakumar, L. (2002) 'Momentum, business cycle, and time-varying expected returns', *The Journal of Finance*, 57(2), pp. 985–1019. <https://doi.org/10.1111/1540-6261.00449>

Fama, E.F. (1970) 'Efficient capital markets: a review of theory and empirical work', *The Journal of Finance*, 25(2), pp. 383–417. <https://doi.org/10.1111/j.1540-6261.1970.tb00518.x>

Fama, E.F. y French, K.R. (1992) 'The cross-section of expected stock returns', *The Journal of Finance*, 47(2), pp. 427–465. <https://doi.org/10.1111/j.1540-6261.1992.tb04398.x>

Greenblatt, J. (2006) *The little book that beats the market*. Hoboken, NJ: John Wiley & Sons. ISBN: 978-0-471-73306-7.

Jegadeesh, N. (1990) 'Evidence of predictable behavior of security returns', *The Journal of Finance*, 45(3), pp. 881–898. <https://doi.org/10.1111/j.1540-6261.1990.tb05110.x>

Jegadeesh, N. y Titman, S. (1993) 'Returns to buying winners and selling losers: implications for stock market efficiency', *The Journal of Finance*, 48(1), pp. 65–91. <https://doi.org/10.1111/j.1540-6261.1993.tb04702.x>

National Bureau of Economic Research (NBER) (2023) *US business cycle expansions and contractions*. Cambridge, MA: NBER. Disponible en: <https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions>.

Novy-Marx, R. (2013) 'The other side of value: the gross profitability premium', *Journal of Financial Economics*, 108(1), pp. 1–28. <https://doi.org/10.1016/j.jfineco.2013.01.003>

Piotroski, J.D. (2000) 'Value investing: the use of historical financial statement information to separate winners from losers', *Journal of Accounting Research*, 38(Supplement), pp. 1–41. <https://doi.org/10.2307/2672906>

S&P Global Market Intelligence (2024) *Capital IQ Pro* [base de datos]. Nueva York: S&P Global. Disponible en: <https://www.spglobal.com/market-intelligence/en/solutions/products/sp-capital-iq-pro> (Consultado: junio 2025).

Sharpe, W.F. (1964) 'Capital asset prices: a theory of market equilibrium under conditions of risk', *The Journal of Finance*, 19(3), pp. 425–442. <https://doi.org/10.1111/j.1540-6261.1964.tb02865.x>

7. Anexos

7.1 Código

```
from google.colab import files
```

```
uploaded = files.upload()
```

```
import pandas as pd
```

```
db = pd.read_excel("Data base TFG Analytics.xlsx", sheet_name="UNIVERSE",  
header=0)
```

```
sp = pd.read_csv("S&P 500 closing price monthly 2000-2025.csv")
```

```
print("Base de datos OK:", db.shape)
```

```
print("S&P 500 OK:", sp.shape)
```

```
# _____
```

```
# MÓDULO 1 — CARGA Y LIMPIEZA DE DATOS
```

```
# _____
```

```
import pandas as pd
```

```
import numpy as np
```

```
import warnings
```

```
warnings.filterwarnings('ignore')
```

```

DB_PATH = "Data base TFG Analytics.xlsx"

SP5_PATH = "S&P 500 closing price monthly 2000-2025.csv"

# Hojas de métricas a cargar

SHEETS = [

    'ROA', 'CFO', 'Total_Assets', 'LT_Debt (total debt)',

    'Current_Ratio', 'Shares_Out', 'Gross_Margin', 'Asset_Turnover',

    'Working_Capital', 'Retained_Earn', 'EBIT', 'Market_Cap',

    'Revenue', 'Total_Liab', 'ROE', 'Oper_Margin', 'EBITDA', 'TEV', 'Price_Adj'

]

SHEET_ALIAS = {'LT_Debt (total debt)': 'LT_Debt'}

def load_metric(sheet_name):

    df = pd.read_excel(DB_PATH, sheet_name=sheet_name, header=1)

    df = df.set_index('CIQ_Ticker')

    df = df.drop(columns=['#', 'Company_Name'], errors='ignore')

    df.columns = [c.year if hasattr(c, 'year') else c for c in df.columns]

    df = df[[c for c in df.columns if isinstance(c, int)]]

    return df

print("Cargando base de datos...")

data = {}

for sheet in SHEETS:

```

```

alias = SHEET_ALIAS.get(sheet, sheet)

data[alias] = load_metric(sheet)

print(f" ✓ {alias}")


# Universo de empresas

universe = pd.read_excel(DB_PATH, sheet_name='UNIVERSE', header=0)

universe['First_In_SP500'] = pd.to_datetime(universe['First_In_SP500'], errors='coerce')
universe['Last_In_SP500'] = pd.to_datetime(universe['Last_In_SP500'], errors='coerce')

universe = universe.set_index('CIQ_Ticker')

print(f" ✓ UNIVERSE: {len(universe)} empresas")


# Benchmark S&P 500

def load_sp500():

    rows = []

    with open(SP5_PATH, 'r', encoding='utf-8') as f:

        for line in f:

            line = line.strip().strip('"')

            parts = line.split(',')

            if len(parts) == 2:

                date_str = parts[0]

                price_str = parts[1].strip('"').replace(',', '.')

                try:

                    price = float(price_str)

                    m, y = date_str.split('/')

```

```

        date = pd.Timestamp(year=int(y), month=int(m), day=1)

        rows.append({'Date': date, 'Price': price})

    except:

        pass

    return pd.DataFrame(rows).sort_values('Date').reset_index(drop=True)

sp500 = load_sp500()

sp500['Year'] = sp500['Date'].dt.year

sp500_annual = sp500[sp500['Date'].dt.month == 12].set_index('Year')['Price']

sp500_return = sp500_annual.pct_change().dropna()

sp500_return.name = 'SP500_Return'

print(f"    ✓ S&P 500: {len(sp500)} observaciones ({sp500_return.index.min()}-
{sp500_return.index.max()}")

# Función: empresas válidas por año (sin look-ahead bias)

def get_valid_universe(year):

    cutoff_start = pd.Timestamp(year=year, month=1, day=1)

    cutoff_end = pd.Timestamp(year=year, month=12, day=31)

    mask = (

        (universe['First_In_SP500'] <= cutoff_end) &

        (

            (universe['Last_In_SP500'] >= cutoff_start) |

            (universe['Status'] == 'Current')

        )

    )

```

```

return universe[mask].index.tolist()

print("\n✅ Módulo 1 completado.")

"""

=====
|| TFG — Business Analytics | ICADE / Comillas Pontificia ||
|| Systematic Equity Screening: S&P 500 Universe ||
|| Benchmark: S&P 500 Price Return ||
|| Período: 2004–2024 (21 años) ||
=====
"""

#

# 0. LIBRERÍAS

#

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import matplotlib.ticker as mticker

```

```
import matplotlib.patches as mpatches
```

```
import warnings
```

```
warnings.filterwarnings('ignore')
```

```
#
```

```
# 1. RUTAS
```

```
#
```

```
DB_PATH = "Data base TFG Analytics.xlsx"
```

```
SP5_PATH = "S&P 500 closing price monthly 2000-2025.csv"
```

```
#
```

```
# 2. CARGA DE DATOS
```

```
#
```

```
SHEETS = [
```

```
    'ROA', 'CFO', 'Total_Assets', 'LT_Debt (total debt)',
```

```
    'Current_Ratio', 'Shares_Out', 'Gross_Margin', 'Asset_Turnover',
```

```
    'Working_Capital', 'Retained_Earn', 'EBIT', 'Market_Cap',
```

```
    'Revenue', 'Total_Liab', 'ROE', 'Oper_Margin', 'EBITDA', 'TEV', 'Price_Adj'
```

```
]
```

```
SHEET_ALIAS = {'LT_Debt (total debt)': 'LT_Debt'}
```

```
def load_metric(sheet_name):
```

```
    """Carga una hoja de métricas. Índice = CIQ_Ticker, columnas = año (int)."""
```

```
    df = pd.read_excel(DB_PATH, sheet_name=sheet_name, header=1)
```

```
    df = df.set_index('CIQ_Ticker')
```

```
    df = df.drop(columns=['#', 'Company_Name'], errors='ignore')
```

```
    df.columns = [c.year if hasattr(c, 'year') else c for c in df.columns]
```

```
    df = df[[c for c in df.columns if isinstance(c, int)]]
```

```
    return df
```

```
print("Cargando base de datos...")
```

```
data = {}
```

```
for sheet in SHEETS:
```

```
    alias = SHEET_ALIAS.get(sheet, sheet)
```

```
    data[alias] = load_metric(sheet)
```

```
    print(f" ✓ {alias}")
```

```
# Universo de empresas con fechas de entrada/salida del índice
```

```
universe = pd.read_excel(DB_PATH, sheet_name='UNIVERSE', header=0)
```

```
universe['First_In_SP500'] = pd.to_datetime(universe['First_In_SP500'], errors='coerce')
```

```
universe['Last_In_SP500'] = pd.to_datetime(universe['Last_In_SP500'], errors='coerce')
```

```
universe = universe.set_index('CIQ_Ticker')
```

```
print(f" ✓ UNIVERSE: {len(universe)} empresas")
```

```

# S&P 500 Price Return mensual

def load_sp500():

    """Parsea el CSV con formato europeo (coma decimal, celda única)."""

    rows = []

    with open(SP5_PATH, 'r', encoding='utf-8') as f:

        for line in f:

            line = line.strip().strip("")

            parts = line.split(',')

            if len(parts) == 2:

                date_str = parts[0]

                price_str = parts[1].strip('').replace(',', '.')

                try:

                    price = float(price_str)

                    m, y = date_str.split('/')

                    date = pd.Timestamp(year=int(y), month=int(m), day=1)

                    rows.append({'Date': date, 'Price': price})

                except:

                    pass

    return pd.DataFrame(rows).sort_values('Date').reset_index(drop=True)

sp500 = load_sp500()

sp500['Year'] = sp500['Date'].dt.year

sp500_annual = sp500[sp500['Date'].dt.month == 12].set_index('Year')['Price']

```

```

sp500_return = sp500_annual.pct_change().dropna()

sp500_return.name = 'SP500_Return'

print(f' ✓ S&P 500: {len(sp500)} observaciones mensuales\n')

print("✅ Datos cargados correctamente.\n")

```

```

#

```

3. FUNCIÓN AUXILIAR: UNIVERSO VÁLIDO POR AÑO

```

#

```

```

def get_valid_universe(year):
    """
    Empresas que formaban parte del S&P 500 durante 'year'.

    Elimina look-ahead bias: sólo usa empresas ya presentes ese año.
    """
    cutoff_start = pd.Timestamp(year=year, month=1, day=1)
    cutoff_end = pd.Timestamp(year=year, month=12, day=31)
    mask = (
        (universe['First_In_SP500'] <= cutoff_end) &
        (
            (universe['Last_In_SP500'] >= cutoff_start) |
            (universe['Status'] == 'Current')
        )
    )

```

```

    )

    return universe[mask].index.tolist()

def get_val(metric, ticker, year):

    """Devuelve el valor de una métrica para un ticker/año. NaN si no existe."""

    try:

        v = data[metric].loc[ticker, year]

        return np.nan if pd.isna(v) else float(v)

    except:

        return np.nan

#


---




---



# 4. LOS 5 FILTROS DE SCREENING

#


---




---



# — FILTRO 1: Piotroski F-Score  $\geq 7$ 

# Fuente: Piotroski (2000). 9 señales binarias de salud financiera.

def piotroski_fscore(ticker, year):

    def g(m, y=year): return get_val(m, ticker, y)

    roa = g('ROA');      roa_p = g('ROA', year-1)

    cfo = g('CFO');      ta  = g('Total_Assets')

    ltd = g('LT_Debt');   ltd_p = g('LT_Debt', year-1)

```

```

cr = g('Current_Ratio'); cr_p = g('Current_Ratio', year-1)

sh = g('Shares_Out'); sh_p = g('Shares_Out', year-1)

gm = g('Gross_Margin'); gm_p = g('Gross_Margin', year-1)

at = g('Asset_Turnover'); at_p = g('Asset_Turnover', year-1)

if any(np.isnan(v) for v in [roa, cfo, ta]):

    return np.nan

sigs = []

# Rentabilidad

sigs.append(int(roa > 0))

sigs.append(int(cfo > 0))

sigs.append(int(roa > roa_p) if not np.isnan(roa_p) else np.nan)

sigs.append(int((cfo/ta) > (roa/100)) if ta != 0 else np.nan)

# Apalancamiento y liquidez

sigs.append(int(ltd <= ltd_p) if not np.isnan(ltd) and not np.isnan(ltd_p) else np.nan)

sigs.append(int(cr > cr_p) if not np.isnan(cr) and not np.isnan(cr_p) else np.nan)

sigs.append(int(sh <= sh_p) if not np.isnan(sh) and not np.isnan(sh_p) else np.nan)

# Eficiencia operativa

sigs.append(int(gm > gm_p) if not np.isnan(gm) and not np.isnan(gm_p) else
np.nan)

sigs.append(int(at > at_p) if not np.isnan(at) and not np.isnan(at_p) else np.nan)

valid = [s for s in sigs if not (isinstance(s, float) and np.isnan(s))]

return sum(valid) if len(valid) >= 6 else np.nan

```

```

# — FILTRO 2: Altman Z-Score > 2.99

# Fuente: Altman (1968). Zona segura:  $Z > 2.99$ .

# Unidades: balance sheet en $000, Market_Cap en $M → convertir a $000

def altman_zscore(ticker, year):

    def g(m): return get_val(m, ticker, year)

    wc = g('Working_Capital') # $000

    re = g('Retained_Earn') # $000

    eb = g('EBIT') # $000

    mc = g('Market_Cap') # $M → ×1000 para convertir a $000

    tl = g('Total_Liab') # $000

    rev = g('Revenue') # $000

    ta = g('Total_Assets') # $000

    if any(np.isnan(v) for v in [wc, re, eb, mc, tl, rev, ta]):

        return np.nan

    if ta == 0 or tl == 0:

        return np.nan

    mc_000 = mc * 1000

    return (1.2*(wc/ta) + 1.4*(re/ta) + 3.3*(eb/ta) +

            0.6*(mc_000/tl) + 1.0*(rev/ta))

```

— FILTRO 3: Momentum (12-1) — mitad superior del universo ese año

Fuente: Jegadeesh & Titman (1993).

Con datos anuales: retorno del año t-1 respecto al año t-2 (skip year t)

```
def momentum(ticker, year):
```

```
    p1 = get_val('Price_Adj', ticker, year-1)
```

```
    p2 = get_val('Price_Adj', ticker, year-2)
```

```
    if np.isnan(p1) or np.isnan(p2) or p2 == 0:
```

```
        return np.nan
```

```
    return (p1 / p2) - 1
```

— FILTRO 4: Quality — ROE > 10% Y Operating Margin > 10%

Fuente: Novy-Marx (2013).

```
def quality_pass(ticker, year):
```

```
    roe = get_val('ROE', ticker, year)
```

```
    om = get_val('Oper_Margin', ticker, year)
```

```
    if np.isnan(roe) or np.isnan(om):
```

```
        return np.nan
```

```
    return roe > 10 and om > 10
```

— FILTRO 5: EV/EBITDA < mediana del universo ese año

Fuente: Greenblatt (2006).

Unidades: TEV en \$M, EBITDA en \$000 → convertir EBITDA a \$M

```
def ev_ebitda(ticker, year):
```

```
    tev = get_val('TEV', ticker, year) # $M
```

```
    ebitda = get_val('EBITDA', ticker, year) # $000
```

```
    if np.isnan(tev) or np.isnan(ebitda):
```

```
        return np.nan
```

```
    ebitda_m = ebitda / 1000 # → $M
```

```
    if ebitda_m <= 0:
```

```
        return np.nan
```

```
    return tev / ebitda_m
```

```
#
```

```
# 5. SCREENING ANUAL
```

```
#
```

```
def screen_year(year):
```

```
    """
```

```
    Aplica los 5 filtros secuencialmente al universo válido del año.
```

```
    Umbrales relativos calculados sobre el universo de ese año.
```

```
    Sin look-ahead bias.
```

```
    """
```

```
    tickers = get_valid_universe(year)
```

```

ev_vals = {tk: ev_ebitda(tk, year) for tk in tickers}

ev_vals = {k: v for k, v in ev_vals.items() if not np.isnan(v)}

mom_vals = {tk: momentum(tk, year) for tk in tickers}

mom_vals = {k: v for k, v in mom_vals.items() if not np.isnan(v)}


if not ev_vals or not mom_vals:

    return [], {}


ev_median = np.median(list(ev_vals.values()))

mom_thresh = np.percentile(list(mom_vals.values()), 50) # top 50%


# Funnel: empresas tras cada filtro

after_f1, after_f2, after_f3, after_f4, after_f5 = [], [], [], [], []


for tk in tickers:

    f1 = piotroski_fscore(tk, year)

    if np.isnan(f1) or f1 < 7: continue

    after_f1.append(tk)


    f2 = altman_zscore(tk, year)

    if np.isnan(f2) or f2 <= 2.99: continue

    after_f2.append(tk)

```

```

f3 = mom_vals.get(tk, np.nan)

if np.isnan(f3) or f3 < mom_thresh: continue

after_f3.append(tk)


f4 = quality_pass(tk, year)

if f4 != True: continue

after_f4.append(tk)


f5 = ev_vals.get(tk, np.nan)

if np.isnan(f5) or f5 >= ev_median: continue

after_f5.append(tk)


funnel = {

    'Universe':    len(tickers),

    'After_Piotroski': len(after_f1),

    'After_Altman':  len(after_f2),

    'After_Momentum': len(after_f3),

    'After_Quality': len(after_f4),

    'Final':        len(after_f5),

}

return after_f5, funnel

```

```
#
```

```
# 6. MACRO CYCLE CLASSIFICATION
```

```
#
```

```
# Recesiones NBER (fechas oficiales)
```

```
RECESSION_YEARS = {2001, 2002, 2008, 2009, 2020}
```

```
# Años de política monetaria restrictiva (Fed Funds Rate en alza)
```

```
TIGHTENING_YEARS = {2004, 2005, 2006, 2007, 2015, 2016, 2017, 2018, 2022,  
2023}
```

```
#
```

```
# 7. BACKTEST PRINCIPAL
```

```
#
```

```
# Screening sobre financieros del año t → retorno medido año t → t+1
```

```
# Equal-weighted, rebalanceo anual.
```

```
print("Ejecutando backtest 2004–2024...")
```

```
results = []
```

```

funnel_log = []

for year in range(2003, 2024):

    portfolio, funnel = screen_year(year)

    funnel_log.append({'Year': year, **funnel})

    if not portfolio:

        continue

    returns = []

    for tk in portfolio:

        p_t = get_val('Price_Adj', tk, year)

        p_t1 = get_val('Price_Adj', tk, year + 1)

        if not np.isnan(p_t) and not np.isnan(p_t1) and p_t > 0:

            returns.append((p_t1 / p_t) - 1)

    if not returns:

        continue

    port_return = np.mean(returns)

    bench_return = sp500_return.get(year + 1, np.nan)

    results.append({

        'Year':      year + 1,

```

```

'Port_Return': port_return,

'Bench_Return': bench_return,

'Alpha':    port_return - bench_return if not np.isnan(bench_return) else np.nan,

'N_stocks': len(returns),

'Recession': (year + 1) in RECESSION_YEARS,

'Tightening': (year + 1) in TIGHTENING_YEARS,

})

```

```

df      = pd.DataFrame(results)

df_funnel = pd.DataFrame(funnel_log)

print(f"✅ Backtest completado: {len(df)} años válidos\n")

```

```

#

```

8. MÉTRICAS DE PERFORMANCE

```

#

```

```

port_rets = df['Port_Return'].values

bench_rets = df['Bench_Return'].dropna().values

n      = len(df)

port_cagr = (np.prod(1 + port_rets) ** (1/n)) - 1

```

```

bench_cagr = (np.prod(1 + bench_rets) ** (1/n)) - 1

port_vol = port_rets.std()

bench_vol = bench_rets.std()

port_sharpe = port_rets.mean() / port_rets.std()

bench_sharpe = bench_rets.mean() / bench_rets.std()


nav_p = np.cumprod(1 + port_rets)

peak_p = np.maximum.accumulate(nav_p)

max_dd_port = ((nav_p - peak_p) / peak_p).min()


nav_b = np.cumprod(1 + bench_rets)

peak_b = np.maximum.accumulate(nav_b)

max_dd_bench = ((nav_b - peak_b) / peak_b).min()


win_rate = (df['Alpha'] > 0).mean()


rec = df[df['Recession'] == True]

exp = df[df['Recession'] == False]

tight = df[df['Tightening'] == True]

ease = df[df['Tightening'] == False]


print("=" * 58)

print(f" RESULTADOS DEL BACKTEST (2004–2024, {n} años)")

print("=" * 58)

```

```

print(f'{'Métrica':<32} {'Portfolio':>10} {'S&P 500':>10}")

print("-" * 58)

print(f'{'CAGR':<32} {port_cagr:>9.2%} {bench_cagr:>9.2%}")

print(f'{'Volatilidad anual':<32} {port_vol:>9.2%} {bench_vol:>9.2%}")

print(f'{'Sharpe Ratio':<32} {port_sharpe:>9.3f} {bench_sharpe:>9.3f}")

print(f'{'Máximo Drawdown':<32} {max_dd_port:>9.2%} {max_dd_bench:>9.2%}")

print(f'{'Alpha medio anual':<32} {df['Alpha'].mean():>9.2%}")

print(f'{'Win Rate (años > benchmark)':<32} {win_rate:>9.2%}")

print(f'{'Años en muestra':<32} {n:>10}")

print(f"\n—— Por ciclo NBER
_____")

print(f" Recesión (n={len(rec)}): Port={rec['Port_Return'].mean():.2%}
Bench={rec['Bench_Return'].mean():.2%} Alpha={rec['Alpha'].mean():.2%}")

print(f" Expansión (n={len(exp)}): Port={exp['Port_Return'].mean():.2%}
Bench={exp['Bench_Return'].mean():.2%} Alpha={exp['Alpha'].mean():.2%}")

print(f"\n—— Por ciclo Fed
_____")

print(f" Tightening (n={len(tight)}): Port={tight['Port_Return'].mean():.2%}
Bench={tight['Bench_Return'].mean():.2%} Alpha={tight['Alpha'].mean():.2%}")

print(f" Easing (n={len(ease)}): Port={ease['Port_Return'].mean():.2%}
Bench={ease['Bench_Return'].mean():.2%} Alpha={ease['Alpha'].mean():.2%}")

print("\n—— Tabla anual completa ———")

print(f'{'Año':<6} {'Portfolio':>10} {'S&P
500':>10} {'Alpha':>9} {'N':>5} {'Recesión':>10} {'Tight.':>8}")

print("-" * 58)

```

```

for _, row in df.iterrows():

    rec_flag = '✓' if row['Recession'] else ''

    tight_flag = '✓' if row['Tightening'] else ''

    print(f'{int(row['Year']):<6} {row['Port_Return']:>9.2%} {row['Bench_Return']:>9.2%}'

    f'{row['Alpha']:>8.2%} {int(row['N_stocks']):>5} {rec_flag:>10} {tight_flag:>8}'))

```

```

#

```

9. GRÁFICOS

```

#

```

```

# Colores corporativos consistentes

```

```

C_PORT = '#1f77b4' # azul → portfolio

```

```

C_BENCH = '#ff7f0e' # naranja → S&P 500

```

```

C_POS = '#2ca02c' # verde → alpha positivo

```

```

C_NEG = '#d62728' # rojo → alpha negativo

```

```

C_REC = '#fadb8' # rosa → sombreado recesión

```

```

years = df['Year'].values

```

```

def shade_recessions(ax):

```

```

"""Sombrea los años de recesión NBER en un eje."""

for _, row in df[df['Recession']].iterrows():

    ax.axvspan(row['Year'] - 0.5, row['Year'] + 0.5,

               alpha=0.25, color=C_REC, zorder=0)

#
=====
=====

# FIGURA 1: NAV acumulado + Alpha anual

#
=====
=====

fig1, (ax1, ax2) = plt.subplots(2, 1, figsize=(14, 10),

                                gridspec_kw={'height_ratios': [2, 1]})

fig1.suptitle('Performance del Portfolio vs S&P 500 (2004–2024)',

              fontsize=14, fontweight='bold')

# Panel superior: NAV

ax1.plot(years, nav_p, color=C_PORT, linewidth=2.5,

         label=f'Portfolio (CAGR {port_cagr:.1%})')

ax1.plot(years, nav_b, color=C_BENCH, linewidth=2.5, linestyle='--',

         label=f'S&P 500 (CAGR {bench_cagr:.1%})')

shade_recessions(ax1)

ax1.set_ylabel('Valor de $1 invertido en 2003', fontsize=11)

ax1.yaxis.set_major_formatter(mticker.FuncFormatter(lambda x, _: f'${x:.1f}'))

ax1.legend(fontsize=10)

```

```

ax1.grid(alpha=0.3)

ax1.set_title('Valor Acumulado Neto (NAV)', fontsize=11)

rec_patch = mpatches.Patch(color=C_REC, alpha=0.6, label='Recesión NBER')

ax1.legend(fontsize=10, handles=ax1.get_legend_handles_labels()[0] + [rec_patch])


# Panel inferior: Alpha anual

colors_alpha = [C_POS if a > 0 else C_NEG for a in df['Alpha']]

ax2.bar(years, df['Alpha'] * 100, color=colors_alpha, edgecolor='white', linewidth=0.5)

ax2.axhline(0, color='black', linewidth=0.8)

ax2.axhline(df['Alpha'].mean() * 100, color='navy', linewidth=1.5,
            linestyle='--', label=f'Alpha medio: {df["Alpha"].mean():.1%}')

shade_recessions(ax2)

ax2.set_ylabel('Alpha (%)', fontsize=11)

ax2.set_xlabel('Año', fontsize=11)

ax2.legend(fontsize=10)

ax2.grid(alpha=0.3, axis='y')

ax2.set_title('Alpha Anual (Portfolio – S&P 500)', fontsize=11)


plt.tight_layout()

plt.savefig('fig1_nav_alpha.png', dpi=150, bbox_inches='tight')

plt.show()

print("✅ Guardado: fig1_nav_alpha.png")

```

```
#
```

```
# FIGURA 2: Retornos anuales comparados (barras agrupadas)
```

```
#
```

```
fig2, ax = plt.subplots(figsize=(16, 6))
```

```
x = np.arange(len(years))
```

```
w = 0.38
```

```
ax.bar(x - w/2, df['Port_Return'] * 100, w, label='Portfolio', color=C_PORT, alpha=0.88)
```

```
ax.bar(x + w/2, df['Bench_Return'] * 100, w, label='S&P 500', color=C_BENCH,
alpha=0.88)
```

```
ax.axhline(0, color='black', linewidth=0.8)
```

```
# Sombrear recesiones
```

```
for _, row in df[df['Recession']].iterrows():
```

```
    idx = list(years).index(row['Year'])
```

```
    ax.axvspan(idx - 0.5, idx + 0.5, alpha=0.2, color=C_REC, zorder=0)
```

```
ax.set_xticks(x)
```

```
ax.set_xticklabels(years, rotation=45, fontsize=9)
```

```
ax.set_title('Retorno Anual: Portfolio vs S&P 500 (2004–2024)',
```

```
            fontsize=13, fontweight='bold')
```

```
ax.set_ylabel('Retorno (%)', fontsize=11)
```

```
ax.legend(fontsize=11)
```

```
ax.grid(alpha=0.3, axis='y')
```

```
plt.tight_layout()
```

```
plt.savefig('fig2_retornos_anuales.png', dpi=150, bbox_inches='tight')
```

```
plt.show()
```

```
print("✅ Guardado: fig2_retornos_anuales.png")
```

```
#
```

```
# FIGURA 3: Alpha medio por ciclo macroeconómico
```

```
#
```

```
fig3, ax = plt.subplots(figsize=(9, 6))
```

```
cycle_labels = ['Recesión\n(NBER)', 'Expansión\n(NBER)',  
                'Tightening\n(Fed)', 'Easing\n(Fed)']
```

```
alpha_vals = [  
    rec['Alpha'].mean() * 100,  
    exp['Alpha'].mean() * 100,  
    tight['Alpha'].mean() * 100,  
    ease['Alpha'].mean() * 100,  
]
```

```
n_obs = [len(rec), len(exp), len(tight), len(ease)]
```

```
bar_colors = ['#e74c3c', '#27ae60', '#8e44ad', '#2980b9']
```

```

bars = ax.bar(cycle_labels, alpha_vals, color=bar_colors,
              edgecolor='white', linewidth=0.8, width=0.5)

ax.axhline(0, color='black', linewidth=0.9)

ax.axhline(df['Alpha'].mean() * 100, color='grey', linewidth=1.2,
          linestyle='--', label=f'Alpha global medio: {df["Alpha"].mean():.1%}')

for bar, val, n_o in zip(bars, alpha_vals, n_obs):
    ax.text(bar.get_x() + bar.get_width()/2,
            bar.get_height() + (0.4 if val >= 0 else -1.2),
            f'{val:.1f}%\n(n={n_o})',
            ha='center', va='bottom', fontsize=11, fontweight='bold')

ax.set_title('Alpha Medio por Ciclo Macroeconómico',
            fontsize=13, fontweight='bold')

ax.set_ylabel('Alpha medio anual (%)', fontsize=11)

ax.legend(fontsize=10)

ax.grid(alpha=0.3, axis='y')

plt.tight_layout()

plt.savefig('fig3_alpha_ciclo_macro.png', dpi=150, bbox_inches='tight')

plt.show()

print("✅ Guardado: fig3_alpha_ciclo_macro.png")

```

```
#
```

```
# FIGURA 4: Funnel de screening — empresas por filtro (media)
```

```
#
```

```
fig4, ax = plt.subplots(figsize=(11, 6))
```

```
funnel_means = df_funnel[['Universe','After_Piotroski','After_Altman',  
                          'After_Momentum','After_Quality','Final']].mean()
```

```
funnel_labels = ['Universo\nS&P 500',  
                 'Piotroski\nF-Score  $\geq 7$ ',  
                 'Altman\nZ-Score  $> 2.99$ ',  
                 'Momentum\nTop 50%',  
                 'Quality\nROE & OM  $> 10\%$ ',  
                 'EV/EBITDA\n $<$  Mediana']
```

```
funnel_colors = ['#3498db','#2ecc71','#f39c12','#e67e22','#e74c3c','#1abc9c']
```

```
bars = ax.barh(funnel_labels[:::-1], funnel_means.values[:::-1],  
               color=funnel_colors[:::-1], edgecolor='white', linewidth=0.8)
```

```
for bar, val in zip(bars, funnel_means.values[:::-1]):
```

```
    ax.text(bar.get_width() + 5, bar.get_y() + bar.get_height()/2,  
            f'{val:.0f}', va='center', fontsize=11, fontweight='bold')
```

```

ax.set_title('Funnel de Screening — Empresas por Filtro (media anual 2003–2023)',
             fontsize=13, fontweight='bold')

ax.set_xlabel('Número medio de empresas', fontsize=11)

ax.grid(alpha=0.3, axis='x')

plt.tight_layout()

plt.savefig('fig4_funnel_screening.png', dpi=150, bbox_inches='tight')

plt.show()

print("✅ Guardado: fig4_funnel_screening.png")

```

```
#
```

```
# FIGURA 5: Tamaño del portfolio por año
```

```
#
```

```
fig5, ax = plt.subplots(figsize=(14, 5))
```

```
ax.bar(years, df['N_stocks'], color=C_PORT, alpha=0.8, edgecolor='white')
```

```
ax.axhline(df['N_stocks'].mean(), color='navy', linewidth=1.5,
```

```
         linestyle='--', label=f'Media: {df["N_stocks"].mean():.1f} empresas')
```

```
shade_recessions(ax)
```

```
ax.set_title('Número de Empresas Seleccionadas por Año',
```

```
         fontsize=13, fontweight='bold')
```

```

ax.set_ylabel('N.º de empresas en el portfolio', fontsize=11)

ax.set_xlabel('Año', fontsize=11)

ax.legend(fontsize=10)

ax.grid(alpha=0.3, axis='y')

plt.tight_layout()

plt.savefig('fig5_portfolio_size.png', dpi=150, bbox_inches='tight')

plt.show()

print("✅ Guardado: fig5_portfolio_size.png")

```

```
#
```

```
# FIGURA 6: Distribución sectorial acumulada del portfolio
```

```
#
```

```
# Agrupar sectores en categorías más amplias para mejor visualización
```

```

sector_map = {

    'Semiconductors and Semiconductor Equipment': 'Technology',

    'Technology Hardware, Storage and Peripherals': 'Technology',

    'IT Services': 'Technology',

    'Software': 'Technology',

    'Electronic Equipment, Instruments and Components': 'Technology',

```

'Communications Equipment': 'Technology',
'Oil, Gas and Consumable Fuels': 'Energy',
'Energy Equipment and Services': 'Energy',
'Machinery': 'Industrials',
'Aerospace and Defense': 'Industrials',
'Professional Services': 'Industrials',
'Commercial Services and Supplies': 'Industrials',
'Air Freight and Logistics': 'Industrials',
'Ground Transportation': 'Industrials',
'Electrical Equipment': 'Industrials',
'Chemicals': 'Materials',
'Metals and Mining': 'Materials',
'Containers and Packaging': 'Materials',
'Paper and Forest Products': 'Materials',
'Construction Materials': 'Materials',
'Specialty Retail': 'Consumer Discretionary',
'Household Durables': 'Consumer Discretionary',
'Textiles, Apparel and Luxury Goods': 'Consumer Discretionary',
'Leisure Products': 'Consumer Discretionary',
'Hotels, Restaurants and Leisure': 'Consumer Discretionary',
'Broadline Retail': 'Consumer Discretionary',
'Automobile Components': 'Consumer Discretionary',
'Diversified Consumer Services': 'Consumer Discretionary',
'Passenger Airlines': 'Consumer Discretionary',

```

'Pharmaceuticals': 'Health Care',
'Health Care Providers and Services': 'Health Care',
'Health Care Equipment and Supplies': 'Health Care',
'Biotechnology': 'Health Care',
'Life Sciences Tools and Services': 'Health Care',
'Food Products': 'Consumer Staples',
'Beverages': 'Consumer Staples',
'Consumer Staples Distribution and Retail': 'Consumer Staples',
'Household Products': 'Consumer Staples',
'Tobacco': 'Consumer Staples',
'Personal Care Products': 'Consumer Staples',
'Capital Markets': 'Financials',
'Financial Services': 'Financials',
'Entertainment': 'Communication Services',
'Interactive Media and Services': 'Communication Services',
'Media': 'Communication Services',
'Trading Companies and Distributors': 'Industrials',
'Building Products': 'Industrials',
}

```

```

sector_counts = {}

for year in range(2003, 2024):

    portfolio, _ = screen_year(year)

    for tk in portfolio:

```

```

if tk in universe.index:

    raw_sector = universe.loc[tk, 'GICS_Sector']

    sector = sector_map.get(raw_sector, raw_sector)

    sector_counts[sector] = sector_counts.get(sector, 0) + 1


sector_df = pd.Series(sector_counts).sort_values(ascending=False)

total    = sector_df.sum()


fig6, ax = plt.subplots(figsize=(10, 6))

colors_sec = plt.cm.Set3(np.linspace(0, 1, len(sector_df)))

wedges, texts, autotexts = ax.pie(

    sector_df.values,

    labels=sector_df.index,

    autopct=lambda p: f'{p:.1f}%' if p > 3 else "",

    colors=colors_sec,

    startangle=140,

    pctdistance=0.82,

)

for text in texts:

    text.set_fontsize(9)

for autotext in autotexts:

    autotext.set_fontsize(8)


ax.set_title('Distribución Sectorial Acumulada del Portfolio (2004–2024)\n'

```

```

        f'Total selecciones: {total}',

        fontsize=13, fontweight='bold')

plt.tight_layout()

plt.savefig('fig6_sectores.png', dpi=150, bbox_inches='tight')

plt.show()

print("✅ Guardado: fig6_sectores.png")


print("\n" + "="*55)

print(" TODOS LOS GRÁFICOS GENERADOS CORRECTAMENTE")

print(" Archivos guardados en la sesión de Colab:")

for i in range(1, 7):

    names = ['nav_alpha', 'retornos_anuales', 'alpha_ciclo_macro',

            'funnel_screening', 'portfolio_size', 'sectores']

    print(f" fig{i}_{names[i-1]}.png")

print("="*55)

```