



Vigilada Mineducación

MAF CHALLENGE - Retornos de Equilibrio para Multiactivos.
MAF Challenge – Equilibrium Returns for Multi-Asset Portfolios

Daniel Aristizábal Arango
María Alejandra Giraldo García
María Alejandra Giraldo Restrepo
Juan Camilo Mejía Escobar

Tesis de Maestría en Administración Financiera

Asesor, docente
María Patricia Durango Gutiérrez
Carlos Alexander Grajales Correa

UNIVERSIDAD EAFIT

Escuela de Finanzas, Economía y Gobierno
Maestría en Administración Financiera.
Medellín.
2025

Resumen

Este estudio desarrolla un modelo cuantitativo para estimar los retornos de equilibrio en portafolios multiactivos mediante la integración de la Teoría de Precios por Arbitraje (APT) con el algoritmo de aprendizaje automático Extreme Gradient Boosting (XGBoost). El modelo incorpora factores macroeconómicos monetarios, reales, financieros y basados en expectativas, como inflación, empleo, política monetaria, liquidez, spreads de crédito y sentimiento del consumidor, clasificados según su naturaleza *leading*, *coincident* o *lagging*. La inclusión de rezagos permite capturar la transmisión gradual de los shocks macroeconómicos, mientras que XGBoost identifica relaciones no lineales y patrones complejos entre variables. A partir de esta arquitectura híbrida se estiman sensibilidades (betas) y primas de riesgo (λ), obteniendo retornos coherentes con un equilibrio multifactorial. La metodología se aplica a liquidez, renta fija, crédito High Yield, mercados emergentes, REITs y renta variable global. Los resultados confirman que tanto los fundamentos macroeconómicos como las expectativas de los inversionistas son determinantes centrales del retorno esperado.

Palabras clave: Asset pricing, Long-term expected returns, multi-asset portfolio, Capital markets, Local currency (COP)

Abstract

This study develops a quantitative model to estimate equilibrium returns in multi-asset portfolios by integrating the Arbitrage Pricing Theory (APT) with the machine learning algorithm Extreme Gradient Boosting (XGBoost). The model incorporates monetary, real, financial, and expectation-based macroeconomic factors—such as inflation, employment, monetary policy, liquidity, credit spreads, and consumer sentiment—classified as *leading*, *coincident*, or *lagging*. The inclusion of lag structures captures the gradual transmission of macroeconomic shocks, while XGBoost identifies nonlinear patterns and complex interactions among variables. This hybrid framework allows for the estimation of factor sensitivities (betas) and risk premia (λ), producing return estimates consistent with a multifactor equilibrium. The methodology is applied to liquidity instruments, sovereign bonds, high-yield credit, emerging market debt, REITs, and global equities. Results show that both macroeconomic fundamentals and investors' expectations play a central role in determining expected returns, highlighting the relevance of combining economic theory with modern machine learning techniques.

Keywords: Asset pricing, Long-term expected returns, multi-asset portfolio, Capital markets, Local currency (COP)

JEL Codes: C1, C5, E43, G11, G12, G15, G17

Tabla de contenidos

Capítulo 1: Introducción.....	4
Capítulo 2: Marco Teórico	6
2.1 Fundamentos teóricos del retorno de equilibrio	6
2.1 Modelos multifactoriales y Arbitrage Pricing Theory (APT)	8
Capítulo 3: Metodología	16
Metodología Propuesta (APT–XGBoost)	16
Estimación de modelo multifactor (APT) tradicional:	20
Integración APT–XGBoost:	22
Capítulo 4: Resultados.....	25
4.1 Principales factores utilizados por el modelo	25
4.2 Importancia de factores por activo	27
4.3 Interpretación de los resultados del modelo	28
4.4. Estimación de betas	31
4.5.1 Fundamentos Teóricos del Vector λ en el APT	42
5. Retornos de Equilibrio del Modelo APT–XGBoost	46
5.1 Conclusiones	48
5.2 Recomendaciones:	50
Capitulo 6. Referencias.....	51
Anexos	53

Tabla 1. Comparación de enfoques para estimar retornos esperados	15
Tabla 2. Parametrización	22
Tabla 3. Detalle de Factores Seleccionados.....	28
Tabla 4. Betas estimadas – Índice de Liquidez.....	32
Tabla 5. Betas estimadas – S&P 500	33
Tabla 6. Betas estimadas – ACWI.....	35
Tabla 7. Betas estimadas – Deuda USA	36
Tabla 8. Betas estimadas – Deuda HY.....	37
Tabla 9. Betas estimadas – Deuda EM	39
Tabla 10. Betas estimadas – REITs	40

Tabla 11. Primas de Riesgo Estimadas	43
Tabla 12. Retornos del modelo	46

Capítulo 1: Introducción

En la gestión de portafolios, la estimación precisa de los rendimientos esperados de los activos constituye un elemento central para la toma de decisiones estratégicas y la planeación financiera de largo plazo. Sin embargo, los modelos tradicionales de equilibrio presentan importantes limitaciones cuando se aplican en mercados complejos y dinámicos. En la práctica, muchos gestores aún dependen de promedios históricos de rendimiento o de supuestos lineales sobre la relación entre riesgo y retorno, lo que puede generar sesgos significativos ante cambios estructurales en los ciclos económicos o financieros (Ilmanen, 2011; Ma et al., 2024).

El Capital Asset Pricing Model (CAPM), desarrollado por Lintner (1965) y Sharpe, (1964) fue el punto de partida para formalizar la relación entre riesgo y rendimiento esperado. Este modelo plantea que el retorno de un activo debe ser igual al rendimiento libre de riesgo más una prima por riesgo proporcional a su sensibilidad al mercado (beta). Aunque el CAPM marcó un avance fundamental en la teoría financiera, su naturaleza unifactorial y sus supuestos de equilibrio perfecto limitan su capacidad para explicar las variaciones de retorno observadas en contextos reales, especialmente en mercados emergentes o durante regímenes macroeconómicos inestables (Michaud, 1989).

Ante estas limitaciones, Ross (1976) propuso la Arbitrage Pricing Theory (APT), que extiende el CAPM al incorporar múltiples factores de riesgo sistemático como determinantes de los rendimientos esperados. En lugar de depender exclusivamente del riesgo de mercado, la APT asume que los retornos de los activos se explican por su exposición a varios factores macroeconómicos —como tasas de interés, inflación, producción industrial o diferenciales de crédito— y las primas de riesgo asociadas a cada uno de ellos. Esta formulación multifactorial ofrece una representación más realista del equilibrio financiero, ya que reconoce que los movimientos en precios y rendimientos responden a un conjunto de riesgos interrelacionados. No obstante, los modelos lineales multifactoriales, como la APT o sus extensiones, siguen siendo sensibles a los supuestos de linealidad y estacionariedad, lo que limita su desempeño predictivo ante cambios estructurales o no lineales en las relaciones entre variables (Roll, 1977). En la última década, los avances en aprendizaje automático (machine learning) han permitido superar parte de estas restricciones mediante algoritmos capaces de capturar patrones complejos,

interacciones y efectos no lineales en grandes volúmenes de datos financieros y macroeconómicos (Gu et al., 2020).

Entre estas herramientas, el algoritmo Extreme Gradient Boosting (XGBoost), desarrollado por Chen y Guestrin (2016), ha demostrado una alta capacidad predictiva y eficiencia computacional en contextos financieros. A diferencia de los modelos de regresión tradicionales, XGBoost combina múltiples árboles de decisión de forma secuencial, ajustando cada uno para corregir los errores del anterior, lo que le permite modelar relaciones dinámicas y no lineales entre factores y rendimientos (James et al., 2023). En finanzas, este enfoque ha mostrado mejoras sustanciales en la predicción de retornos, la detección de regímenes de mercado y la estimación de riesgos (Abdi et al., 2025; Gu et al., 2020; Sürek & Lau, 2025).

En este contexto, la presente investigación propone la integración del modelo APT y, el algoritmo Extreme Gradient Boosting (XGBoost). La idea central consiste en utilizar el algoritmo de aprendizaje automático para identificar los factores macroeconómicos más relevantes y cuantificar su influencia sobre los retornos de los activos, para luego incorporar dichos factores en la estructura de equilibrio de la APT. Este enfoque busca estimar rendimientos esperados de largo plazo para portafolios multiactivos (renta fija, variable y alternativos), de manera robusta y coherente con los principios de equilibrio financiero.

El modelo APT–XGBoost pretende aportar una herramienta metodológica útil para la gestión de inversiones en un entorno caracterizado por alta volatilidad, interdependencia entre mercados y rápida transmisión de shocks macroeconómicos. Al integrar la teoría económica del equilibrio con la flexibilidad del aprendizaje automático, se espera mejorar la capacidad explicativa y predictiva de los modelos de retornos, contribuyendo así a los practicantes de asset-management en la implementación de esquemas de asignación de activos más estables, informados y adaptativos.

Pregunta de Investigación

¿Cómo puede integrarse el modelo Arbitrage Pricing Theory (APT) con el algoritmo de aprendizaje automático XGBoost para estimar los retornos de un portafolio multiactivo?

3. Objetivos general y específicos

Objetivo general

Desarrollar una metodología para estimar los retornos de equilibrio a largo plazo en diferentes activos del mercado de capitales, con el fin de que se pueda adaptar e implementar en la gestión de portafolios multiactivos.

Objetivos específicos

1. **Analizar los fundamentos teóricos** de los modelos de equilibrio financiero (CAPM, APT y extensiones multifactoriales), identificando sus limitaciones en la estimación de retornos esperados bajo diferentes condiciones macroeconómicas y de integración de mercados.
2. **Identificar y clasificar los factores de riesgo macroeconómicos y financieros** que determinan los rendimientos de los activos considerados (renta fija, variable y alternativos), diferenciando entre indicadores adelantados, coincidentes y rezagados, así como entre factores globales y locales.
3. **Diseñar, calibrar y entrenar el modelo APT–XGBoost** utilizando información histórica (2007–2025), incorporando los factores seleccionados y evaluando su capacidad para capturar relaciones no lineales entre variables macroeconómicas y retornos de activos.

Capítulo 2: Marco Teórico

2. 1 Fundamentos teóricos del retorno de equilibrio

La noción de retorno de equilibrio surge de la teoría financiera clásica, la cual postula que en mercados eficientes y competitivos los precios de los activos se ajustan de modo que ningún inversionista pueda obtener ganancias sin asumir riesgo, en concordancia con el principio de no arbitraje (Lintner, 1965; Mossin, 1966; Ross, 1976; Sharpe, 1964). Este concepto está estrechamente ligado a la teoría de portafolios de Markowitz (1952), que introdujo el marco media-varianza para la selección óptima de portafolios. Markowitz demostró que, dado un conjunto de activos con ciertos rendimientos esperados, desviaciones estándar y correlaciones, es posible calcular la *frontera eficiente* de portafolios que minimizan la varianza para cada nivel de retorno esperado (J. M. Chen, 2021). En dicha frontera, el cociente retorno/volatilidad es máximo; este óptimo se conoce como portafolio de tangencia cuando se incluye un activo libre de riesgo. Sobre esta base teórica, Tobin, (1958) estableció el principio de separación,

según el cual la decisión de inversión óptima se divide en elegir la combinación óptima de activos riesgosos (portafolio tangente) y luego combinarla con el activo libre de riesgo según la aversión al riesgo del inversor.

El **Capital Asset Pricing Model (CAPM)** surge a inicios de los años 1960 como una extensión de estos conceptos de media-varianza hacia un modelo de equilibrio general de precios de activos. Formulado independientemente por Lintner (1965), Mossin (1966) y Sharpe, (1964), el CAPM postula que en equilibrio todos los inversionistas poseen alguna combinación del portafolio de mercado (el portafolio eficiente tangente compuesto por todos los activos riesgosos, ponderados por su valor de mercado) y del activo libre de riesgo (J. M. Chen, 2021). En consecuencia, el único factor de riesgo relevante para explicar las diferencias de rendimiento esperado entre activos sería su *beta* respecto al portafolio de mercado, definido como una medida de sensibilidad o covarianza relativa del activo con los rendimientos del mercado. La fórmula clásica del CAPM es:

$$E(R_i) = R_f + \beta_i [E(R_m) - R_f], \quad \text{con } \beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)}$$

donde $E(R_i)$ es el rendimiento esperado del activo i , R_f es la tasa libre de riesgo, $E(R_m)$ es el rendimiento esperado del mercado, y $\beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)}$ representa el riesgo sistemático del activo i .

Intuitivamente, esta ecuación afirma que los inversores requieren un retorno extra - prima de riesgo - para invertir en activos riesgosos, y dicha prima debe ser proporcional al riesgo sistemático que cada activo añade al portafolio de mercado (J. M. Chen, 2021; Gökğöz, 2009). El coeficiente β_i actúa como multiplicador de la prima de riesgo de mercado $E(R_m) - R_f$, la cual es igual para todos los activos y refleja el exceso de rendimiento requerido por unidad de riesgo sistemático en la economía. En palabras de Sharpe, “el inversor recibe compensación solo por el riesgo que no puede eliminar mediante diversificación”, que es precisamente el medido por beta (J. M. Chen, 2021; Lintner, 1965; Sharpe, 1964).

Supuestos del CAPM: Para derivar la relación anterior, el modelo asume, entre otras cosas, que los inversionistas son racionales y **aversos al riesgo**, todos tienen horizontes de un solo período y expectativas homogéneas sobre los rendimientos esperados, varianzas y covarianzas de los activos (J. M. Chen, 2021). Asimismo, se supone que existen condiciones ideales de mercado: no hay costos de transacción, ni impuestos, la información es simétrica y está totalmente incorporada en los precios (hipótesis de eficiencia de mercado), y los inversionistas pueden pedir prestado o prestar al mismo tipo libre de riesgo de manera ilimitada. Bajo estos supuestos, todos los agentes optimizan sus portafolios según Markowitz y terminan eligiendo

la misma combinación de activos riesgosos (el portafolio de mercado), ajustando solo cuánto invierten en el activo libre de riesgo en función de su tolerancia al riesgo (J. M. Chen, 2021; Mossin, 1966; Sharpe, 1964). El resultado es que el portafolio de mercado está en equilibrio y la pendiente de la línea mercado de capitales (CML) –que une al activo libre de riesgo con el portafolio de mercado en el plano riesgo-retorno– corresponde a la prima de riesgo de mercado. Cada activo individual se ubica en la línea de mercado de valores (SML) según su beta.

Limitaciones y críticas

A pesar de su elegancia, pronto se hizo evidente que el CAPM adolecía de múltiples limitaciones empíricas. Fama & MacBeth (1973) desarrollaron pruebas en serie de tiempo y corte transversal que indicaron que beta por sí sola no capturaba completamente las variaciones en los rendimientos promedio de las acciones. Posteriormente, Fama & French (1992) documentaron de forma contundente que la relación empírica entre beta de mercado y rendimiento promedio era prácticamente plana en las décadas de 1960-80 una vez controlado por otros efectos, concluyendo que “la relación entre β y el rendimiento promedio es débil, quizás inexistente”. Este estudio y otros hallazgos de anomalías –como la prima de tamaño (empresas pequeñas tienden a rendir más) y la prima de valor (empresas con alto B/M rinden más que las de bajo B/M)– sugirieron la necesidad de modelos de equilibrio más flexibles. (Roll, 1977) también criticó el CAPM desde lo teórico, argumentando que el modelo no es falsable a menos que se observe el verdadero portafolio de mercado, el cual incluiría no solo acciones sino todo tipo de activos (capital humano, bienes raíces, etc.). Además, muchos supuestos del CAPM resultan poco realistas (p.ej., la posibilidad de endeudamiento ilimitado al mismo R_f , o la homogeneidad de expectativas) (J. M. Chen, 2021; Thalassinos et al., 2023).

No obstante, pese a sus deficiencias empíricas, el CAPM continúa utilizándose ampliamente como una referencia inicial por su simplicidad e intuición –es conocido en finanzas que “beta sigue siendo el primer aproximado razonable para los retornos de mercado en equilibrio” (J. M. Chen, 2021; Sürek & Lau, 2025) y porque proporciona una medida de costo de capital fácil de comunicar; misma que es usada, por ejemplo, para calcular el costo de capital accionario en evaluaciones corporativas. En resumen, el CAPM estableció el marco conceptual en el que el rendimiento esperado de un activo está ligado solo a riesgo sistemático, pero la evidencia llevó a ampliar ese marco (Illmer & Marty, 2007; Ilmanen, 2011).

2.1 Modelos multifactoriales y Arbitrage Pricing Theory (APT)

La **Teoría de Precios por Arbitraje (APT)**, introducida por Stephen A. Ross (1976), surge como una alternativa al CAPM que elimina muchos de sus supuestos restrictivos a cambio de renunciar a predicciones concretas sobre qué factores explican los retornos. El APT se basa

únicamente en el principio de *no arbitraje*: en equilibrio, ningún portafolio bien diversificado puede generar una ganancia segura superior a la libre de riesgo sin incurrir en riesgo (de lo contrario, los arbitrajistas explotarían esa oportunidad hasta eliminarla). Este principio implica que debe existir una relación lineal entre los retornos esperados de los activos y su exposición a cierto número de **factores de riesgo sistemático**. Si denotamos F_1, F_2, \dots, F_N como factores económicos (por ejemplo, inflación, crecimiento industrial, spreads de crédito, etc.), el modelo APT postula que para cada activo i :

$$E(R_i) = R_f + \beta_{i1}\lambda_1 + \beta_{i2}\lambda_2 + \dots + \beta_{iN}\lambda_N$$

donde β_{i1} es la sensibilidad del activo i al factor j (también llamado *loading* o beta de factor) y λ_1 es la prima de riesgo asociada a dicho factor (es decir, el exceso de retorno que un inversionista exige por estar expuesto unitariamente al factor j). De esta forma, cada activo tiene un retorno esperado que es una combinación lineal de las primas de los factores relevantes, ponderadas por sus respectivas exposiciones (Black & Litterman, 1992). Un resultado importante es que si un activo está libre de exposición a todos los factores (es decir, todas sus betas son cero), entonces su retorno esperado debe ser exactamente R_f ; de lo contrario, habría arbitraje.

A diferencia del CAPM, la APT no especifica ex ante cuáles ni cuántos son esos factores F_j . Identificarlos es tarea empírica: los economistas e inversores deben examinar qué variables macro o financieras parecen conducir los retornos de un conjunto de activos en un período dado (Burmeister & Wall, 1986; Roll & Ross, 1980). Ross sugirió que podrían ser cualquier fuente de riesgo económica no diversificable. Estudios posteriores ofrecieron pistas: por ejemplo, Chen et al., (1986) encontraron que shocks inesperados de inflación, cambios en la tasa de crecimiento industrial, variaciones en la pendiente de la curva de rendimientos, cambios en los diferenciales de bonos corporativos, entre otros, tenían poder explicativo significativo sobre los rendimientos del mercado accionario de EE.UU. En décadas subsiguientes, la idea de factores se expandió y entrelazó con la literatura de factores empíricos en acciones. Fama & French, (1993) incorporaron dos factores adicionales al factor de mercado: SMB (por sus siglas en inglés, refiere a Small Minus Big, rendimiento de acciones pequeñas menos grandes) y HML (High Minus Low, rendimiento de acciones de alto book-to-market menos bajo), capturando las primas de tamaño y valor que se observaban persistentemente. Más adelante, Carhart, (1997) añadió un cuarto factor de momentum (Up Minus Down, acciones ganadoras recientes menos perdedoras), reflejando la tendencia de precios a continuar subiendo o bajando en el corto plazo. Estos modelos multifactoriales (conocidos como de arbitraje o de factores en honor a la APT) han tenido gran éxito explicando los retornos históricos de acciones y otros activos, hasta el punto de que hoy existe un “zoológico de factores” propuestos que incluyen volatilidad, inversión, calidad, liquidez, sentimiento, etc (Cochrane, 2011; Feng et al., 2020).

Supuestos y diferencias con CAPM: El APT requiere supuestos mucho menos restrictivos. No asume un portafolio de mercado específico ni inversores idénticos; solo requiere que exista un número suficiente de activos de forma que los riesgos idiosincráticos puedan diversificarse, lo que garantiza que las relaciones de no arbitraje se apliquen a portafolios bien diversificados (Roll, 1977; Ross, 1976). En ese caso, cualquier mispricing respecto al modelo de factores sería explotado mediante una estrategia de arbitraje que consistiría en tomar posiciones largas y cortas en activos hasta neutralizar la exposición a los factores comunes y obtener un rendimiento libre de riesgo mayor a algo que en equilibrio no puede persistir. Asimismo, a diferencia del CAPM, el APT no garantiza que el portafolio de mercado esté en la frontera eficiente, ni que el mercado en su conjunto pueda describirse por un solo factor (Chambers et al., 2024; Ilmanen, 2011; Roll, 1977). En su lugar, admite múltiples fuentes de riesgo y por ende múltiples primas de riesgo.

Limitaciones de la APT: El principal desafío de la APT es precisamente determinar cuáles factores incluir en el modelo. Dado que la teoría no los especifica, el modelador enfrenta un riesgo de sobreajuste u omisión en la selección de factores. Empíricamente, los investigadores han utilizado tanto enfoques estadísticos (análisis factorial, componentes principales) como fundamentos económicos (variables macro, indicadores financieros) para extraer factores (Feng et al., 2020; Gladstone et al., 2021; Mainie, 2015; Shyu et al., 2006). Sin embargo, esto ha llevado a cierta proliferación de posibles factores el peligro de identificar correlaciones espurias propias del período analizado. Por ello, en la práctica se suele limitar el número de factores a unos pocos con justificación económica sólida.

Modelos de *Machine Learning* (XGBoost y otros) en retornos esperados.

La creciente disponibilidad de datos financieros y el avance del poder computacional han propiciado en años recientes la incorporación de técnicas de aprendizaje de máquina al pronóstico de rendimientos de activos. A diferencia de los modelos de equilibrio tradicionales (CAPM, APT), que suelen imponer relaciones lineales basadas en teoría económica, los modelos de machine learning adoptan un enfoque más flexible y orientado a predicción: permiten descubrir patrones en los datos sin especificar a priori una forma funcional estricta (Gu et al., 2020; James et al., 2023; Min et al., 2021). Esto resulta especialmente valioso para capturar posibles no linealidades, efectos de interacción entre variables y cambios de régimen temporales en los retornos –fenómenos difíciles de incorporar en modelos lineales paramétricos convencionales.

Dentro del amplio arsenal de métodos de machine (ML), las técnicas de ensemble basadas en árboles de decisión han mostrado particular eficacia en problemas de regresión y clasificación

financieros. En este contexto se destaca **XGBoost (Extreme Gradient Boosting)**, un algoritmo de boosting de árboles introducido por T. Chen & Guestrin (2016) que ha ganado popularidad por su alto rendimiento en competencias de datos y aplicaciones del mundo real. XGBoost construye de manera iterativa un conjunto de árboles de decisión de profundidad limitada, donde cada árbol nuevo se entrena para corregir los errores del conjunto de árboles previo (minimizando un objetivo de pérdida diferenciable). El resultado es un predictor ensamblado que logra gran capacidad de ajuste manteniendo controlado el sobreajuste mediante regularización explícitas y la naturaleza ensemble (James et al., 2023).

Aplicación a retornos de activos: Un modelo XGBoost puede usarse para predecir directamente el rendimiento futuro de un activo o la probabilidad de que supere cierto umbral, utilizando como features diversas variables relevantes. Estas pueden incluir factores macroeconómicos (ej: crecimiento PIB, inflación, tasas de interés), indicadores financieros (ej: valuaciones como P/E, yield de bonos, diferenciales de crédito), features técnicas (momentum, volatilidad reciente, etc.), e incluso medidas de sentimiento del mercado (por ejemplo, índices de confianza del consumidor, volumen de noticias positivas/negativas, sentiment en redes sociales) (Abdi et al., 2025; Gu et al., 2020; Malandri et al., 2018). Un punto a favor es que XGBoost puede manejar gran cantidad de variables de entrada, seleccionando internamente las más predictivas mediante la partición óptima en los árboles, similarmente a un proceso de selección de variables automático (James et al., 2023). Asimismo, captura interacciones: por ejemplo, podría aprender que una combinación alta de inflación y bajo crecimiento tiene cierto efecto no lineal adverso en rendimientos, algo que un modelo lineal con factores separados podría pasar por alto.

Intuición y supuestos: Los modelos de ML como XGBoost no requieren suposiciones distribucionales (no asumen normalidad de residuales ni linealidad). Sin embargo, **sí suponen** que el futuro tendrá patrones similares a los aprendidos en el pasado (estacionariedad en algún grado) y que se cuenta con suficientes datos para generalizar (James et al., 2023). Así, su desempeño depende críticamente de la calidad de datos y de la representatividad de las variables utilizadas. Además, a diferencia de los modelos de equilibrio, los modelos de ML son esencialmente cajas negras en términos de explicabilidad económica; aunque existen herramientas como SHAP values o importancias de variables que pueden ayudar a interpretar qué factores están siendo más utilizados por el modelo, los coeficientes no tienen una interpretación sencilla como betas de riesgo (Abdi et al., 2025). Por eso, es recomendable usarlos como complemento, manteniendo coherencia con la teoría financiera al seleccionar input features.

Evidencia empírica: La literatura reciente provee varios ejemplos exitosos de ML en la estimación de retornos o señales de inversión. Gu et al., (2020) que métodos como árboles, redes neuronales y bosques aleatorios (Random Forest) podían capturar de mejor manera la

relación entre características de acciones y retornos subsiguientes, generando predicciones más precisas que los modelos lineales tradicionales en la bolsa estadounidense. En horizontes más largos, Sürek & Lau, (2025) demostraron que un modelo XGBoost podía predecir correctamente la dirección (positiva/negativa) del rendimiento total a doce meses del S&P 500 en una proporción significativamente mayor que el azar, superando a la regresión logística benchmark. Su mejor modelo alcanzó 100% de acierto en los casos de retornos negativos en el período de prueba (con p-valor ~ 0.05), lo que sugiere que XGBoost fue capaz de identificar condiciones previas asociadas a caídas del mercado. En mercados emergentes, Abadi et al., (2025) compararon XGBoost contra regresión lineal para predecir tendencias de sobrereacción de precios en acciones iraníes, hallando que XGBoost entregaba un R^2 test 15% mayor y reducía el error a casi la mitad del de la regresión. Otros estudios híbridos, como Abdi et al., (2025), han integrado ML en el proceso de selección de portafolios: ellos usan Random Forest para filtrar variables, luego XGBoost para pronosticar rendimientos a corto plazo, y finalmente un esquema de selección dinámico de acciones. El enfoque logró incrementar el retorno del portafolio y a la vez reducir su volatilidad respecto a estrategias sin ML, evidenciando que estos métodos pueden mejorar tanto la rentabilidad como la gestión de riesgo al captar señales predictivas que escapan a métodos tradicionales.

Limitaciones y riesgos de ML: Un peligro latente al usar modelos de alta complejidad es el sobreajuste (*overfitting*). Los algoritmos de boosting pueden ajustarse demasiado a los datos pasados si no se regulan bien; por ejemplo, usando demasiados árboles o profundidad muy alta (*James et al., 2023*). Para evitarlo, se deben aplicar técnicas de validación cruzada, regularización (que XGBoost incluye) y reservar muestras de prueba para evaluar la verdadera capacidad predictiva. También es crucial actualizar periódicamente el modelo, ya que los patrones pueden cambiar (riesgo de *non-stationarity*) (Ma et al., 2024; McQuarrie, 2024). Otro punto es que los modelos ML no garantizan obedecer restricciones económicas; por ejemplo, podrían teóricamente producir pronósticos de retornos sistemáticamente inconsistentes con equilibrio (aunque en portafolios esto podría aprovecharse para arbitraje). Por ello, algunos investigadores integran *constrain* en los algoritmos para imponer cierta coherencia, o combinan ML con modelos estructurales. Finalmente, la interpretabilidad es menor. En tal sentido, en la práctica comunicar a un comité de inversión "el modelo predice X porque detectó una interacción no lineal entre 20 variables" es menos intuitivo que decir "el modelo asume que la prima de riesgo de mercado es 5%". En este sentido, se suele utilizar ML para apoyar la toma de decisiones, más que como sustituto total del juicio humano.

Datos alternativos y sentimiento: Un campo promisorio relacionado con ML es el uso de datos no tradicionales para predecir retornos, como **indicadores de sentimiento** derivados de noticias, redes sociales o búsquedas de internet. El estudio de Malandri et al., (2018) es ilustrativo: utilizaron datos de sentimiento público obtenidos de Twitter y otros medios para guiar un modelo automático de asignación de activos basado en redes LSTM (Long Short-

Term Memory). Encontraron que incluir el sentimiento financiero mejoraba las estrategias de portafolio, logrando hasta un +19% de rendimiento promedio por encima de un portafolio equiponderado, mientras que sin sentimiento la mejora era más modesta (~5%). En general, la incorporación de tales señales no tradicionales es factible gracias a ML, ya que estos algoritmos pueden procesar grandes cantidades de texto y datos no estructurados convirtiéndolos en indicadores numéricos útiles (James et al., 2023). Esto añade otra dimensión a la estimación de retornos de equilibrio, complementando los factores macro y financieros clásicos con el componente de psicología del mercado.

Comparación de enfoques

Habiendo descrito los principales modelos, es útil resumir hallazgos empíricos que comparan su desempeño y características. En la siguiente tabla se presentan de forma estructurada las diferencias entre cuatro enfoques relevantes: CAPM, APT (multifactorial), y modelos de Machine Learning (representados por XGBoost):

Criterio	CAPM (Sharpe 1964)	APT / Multifactor (Ross 1976, Fama-French 1993)	Machine Learning (XGBoost)
Naturaleza del modelo	Equilibrio unifactorial (mercado). Modelo analítico cerrado.	Equilibrio multifactorial (varios betas). Modelo analítico lineal.	Data-driven no paramétrico. Modelo predictivo (no impone estructura ex ante).
Fórmula básica	$E(R_i) = R_f + \beta_i [E(R_m) - R_f],$ Prima proporcional al beta de mercado.	$E(R_i) = R_f + \beta_{i1}\lambda_1 + \beta_{i2}\lambda_2 + \dots + \beta_{iN}\lambda_N$ Prima como combinación lineal de N factores.	No hay fórmula cerrada; el modelo aprende $f(x) \approx E(R)$ a partir de los datos. Por ejemplo, árboles de decisión en boosting minimizan $\sum(y - \hat{y})^2$.
Supuestos clave	Mercado eficiente, un periodo, homogeneidad, tasa libre de riesgo accesible y riesgo simétrico (varianza).	No arbitraje, muchos activos para	Distribución de datos aproximadamen

		diversificar riesgo idiosincrático y estructura lineal de factores.	te estacionaria. Relaciones históricas repetibles. Suficiente big data.
Inputs requeridos	β_i de cada activo vs mercado; prima de mercado: $E(R_m) - R_f$	Betas de cada activo vs cada factor; primas de cada factor λ_j	Gran conjunto de features potencialmente relevantes (macros, técnicos, sentimiento, etc.) con datos históricos de entrenamiento.
Racional económica	Un único riesgo sistemático (mercado) domina. Activos difieren solo en su sensibilidad al mercado.	Múltiples riesgos sistemáticos (crecimiento, inflación, etc.) influyen. Activos difieren en su exposición a cada factor.	Patrón complejo: el modelo detecta combinaciones de señales que históricamente anticiparon rendimientos. No siempre es evidente la causa económica.
Ventajas	Sencillez, pocas variables. Útil para estimar costo de capital y análisis exploratorio. Beta fácil de estimar.	Más poder explicativo cross-sectional. Captura anomalías (tamaño, valor, etc.). Flexible en elección de factores (macro, estilo, sector).	Muy flexible: capta no linealidades, interacciones y cambios de régimen. Puede manejar muchos datos y tipos de variables (incluso textuales). Alta precisión predictiva out-of-sample en varios estudios.
Desventajas	Empíricamente incompleto (no explica size, value, etc.) file:///file_000000001a3c71f5acdb9c9b3056f2e6/ . Difícil de testear y supone mercado proxy. Un solo periodo.	No indica cuáles factores usar; riesgo	Enfoque de “caja negra” y poca interpretabilidad

		de sobreajuste en selección. Puede requerir muchos datos para estimar betas con poca colinealidad	. Riesgo de sobreajuste si es mal regulado. Exige datos limpios y largos. No garantiza coherencia con teoría (puede violiar límites lógicos si no se controla).
Uso típico	Estimar costo de capital (CAPM sigue en finanzas corporativas). Medir performance ajustado a riesgo (alpha de Jensen).	Modelos de riesgo de portafolio (Barra, etc.). Asset pricing académico (pruebas de factores). Asignación táctica usando factor timing.	Trading algorítmico y <i>asset allocation</i> cuantitativo. Señales de mercado (ej: alertas de caídas). Soporte a gestores tradicionales con señales adicionales. Robo-advisors.

Tabla 1. Comparación de enfoques para estimar retornos esperados

En términos de rendimiento fuera de muestra, no hay una respuesta única sobre qué enfoque es superior; depende de la situación. Estudios sugieren que estrategias ingenuas como la diversificación equitativa son difíciles de superar consistentemente a menos que el modelo aproveche suficiente información predictiva. Sin embargo, combinaciones avanzadas sí logran mejoras: Black–Litterman, al estabilizar expectativas, superó el enfoque $1/N$ y media-varianza en la mayoría de los escenarios probados (Bessler et al., 2017) Los modelos multifactoriales ofrecen beneficios de diversificación y son fundamentos del *factor investing* actual, pero su capacidad predictiva a corto plazo suele ser limitada, siendo más útiles para explicar que para predecir (Thalassinos et al., 2023). Por otro lado, los modelos de machine learning han mostrado gran precisión en pronósticos de corto plazo y para identificar señales (e.g., prever crisis o correcciones), aunque su incorporación en portafolios de largo plazo aún es materia de investigación activa. Así pues, un metaanálisis reciente sugiere que combinar enfoques puede ofrecer lo mejor de ambos mundos. Por ejemplo, usar una estructura de

equilibrio para asegurar coherencia económica y aplicar ML para ajustar dinámicamente según datos recientes, surge como una estrategia prometedora.

Capítulo 3: Metodología

Metodología Propuesta (APT–XGBoost)

En esta sección se describe la metodología desarrollada para estimar retornos de equilibrio de portafolios multiactivos integrando un modelo multifactorial de equilibrio (APT) con técnicas de machine learning (XGBoost). El objetivo es estimar los rendimientos esperados de largo plazo para distintos activos (renta fija, renta variable y activos alternativos) incorporando tanto información macro-financiera fundamental como patrones no lineales identificados a partir de datos históricos. La metodología consta de los siguientes pasos generales:

1. Definición de asset classes: En primer lugar, se define el universo de activos sobre el cual se aplicará el modelo APT incorporando el algoritmo XGBoost. Dado que el objetivo de la presente investigación es estimar los retornos esperados de largo plazo para un portafolio multiactivo denominado en dólares estadounidenses, el universo estará compuesto por instrumentos financieros líquidos y representativos de los principales mercados internacionales, todos medidos o convertidos a USD. Específicamente, el portafolio incluye:

- **Renta variable global:** (ETF como son el ACWI o el S&P 500, que capturan la prima de riesgo de las acciones desarrolladas y emergentes).
- **Renta fija en dólares:** (bonos del Tesoro de EE.UU., deuda corporativa high yield y deuda de mercados emergentes en USD).
- **Activos alternativos listados:** REITs que reflejan la evolución y brindan exposición al mercado inmobiliario de Estados Unidos.
- **Activos de liquidez** (letras del Tesoro a corto plazo, que se asumen libres de riesgo).

Todos los activos se modelan en una moneda base única (USD), lo que permite integrar de forma coherente los factores macroeconómicos y financieros globales que inciden en los retornos, como tasas de interés de la Reserva Federal, spreads de crédito, confianza del consumidor, inflación o crecimiento del PIB de Estados Unidos, junto con variables relevantes de mercados emergentes (como el EMBI o el CDS soberano en dólares).

2. Selección de factores y variables explicativas: Con base en la literatura macro-financiera y en la lógica del modelo de Arbitrage Pricing Theory (APT), se seleccionaron diferentes factores económicos los cuales pueden influenciar los retornos de los activos en el

portafolio multiactivo denominado en dólares estadounidenses. Estos factores se estructuran en tres categorías según su relación temporal con el ciclo económico: adelantados, coincidentes y rezagados; además se dividen entre factores globales (EE. UU. y de mercados internacionales) y locales (Colombia).

El modelo APT–XGBoost desarrollado se compone principalmente en factores globales y de Estados Unidos, dada la denominación en dólares de los activos, integrando de forma complementaria algunas variables colombianas relevantes para capturar el componente emergente de riesgo.

Indicadores adelantados

Estos indicadores se caracterizan por anticipar los cambios en la actividad económica y en el apetito de riesgo de los inversionistas, reflejando variaciones esperadas en las condiciones microfinancieras (The Conference Board, 2001).

- **Índice de confianza del consumidor de Estados Unidos (UMCSENT):** mide la percepción y expectativas de los consumidores sobre la situación económica, anticipando posibles variaciones en el gasto de los hogares.
- **Pedidos manufactureros de bienes duraderos (AMTMNO):** refleja las nuevas órdenes de producción industrial, indicador sensible al ciclo económico y a la inversión empresarial.
- **Masa monetaria M2 (M2SL):** cuantifica la liquidez disponible en la economía, incluyendo efectivo, depósitos y cuasidinero; su crecimiento suele anticipar expansiones económicas.
- **Credit Default Swap soberano a 5 años de Estados Unidos (CDS 5Y USA):** representa el costo de asegurar la deuda soberana estadounidense, funcionando como termómetro de la percepción de riesgo global.
- **Rendimiento de bonos del Tesoro de Estados Unidos a cinco años (UST5Y):** refleja las expectativas de los mercados sobre tasas de interés e inflación a mediano plazo.
- **Spread ajustado por opciones de bonos corporativos (OAS):** mide la prima de riesgo exigida a los bonos corporativos sobre los títulos del Tesoro, indicador temprano de tensiones crediticias.
- **Inflación esperada a cinco años (T5YIFR):** derivada de instrumentos de mercado, refleja las expectativas inflacionarias a mediano plazo, anticipando decisiones de política monetaria.
- **Índice de volatilidad del mercado (VIXCLS):** conocido como “índice del miedo”, mide la volatilidad esperada del S&P 500 y sirve como indicador global de aversión o apetito por el riesgo.

- **Credit Default Swap soberano a 5 años de Colombia (CDS 5Y COL):** mide el riesgo de incumplimiento percibido para la deuda colombiana, reflejando expectativas sobre estabilidad macroeconómica.
- **Índice de confianza del consumidor de Colombia (ICC COL):** mide la percepción de los hogares colombianos sobre su situación financiera y la economía, anticipando el comportamiento del consumo interno.

Indicadores coincidentes

Esta clase de indicadores se mueven simultáneamente con el ciclo económico y reflejan la evolución actual de la actividad productiva y monetaria (The Conference Board, 2001).

- **Nóminas no agrícolas de Estados Unidos (PAYEMS):** mide el nivel de empleo en sectores no agrícolas; su crecimiento refleja expansión económica y fortaleza del mercado laboral.
- **Tasa de fondos federales de Estados Unidos (FEDFUNDS):** representa la tasa de interés de referencia de la Reserva Federal, ajustada según las condiciones macroeconómicas presentes.
- **Producción industrial de Estados Unidos (INDPRO):** mide el volumen de producción manufacturera, minera y de servicios públicos, indicador directo de la actividad económica.
- **Producto Interno Bruto real de Estados Unidos (GDPC1):** cuantifica el valor agregado real de bienes y servicios producidos, reflejando el nivel general de actividad económica.
- **Unidad de Valor Real (UVR):** indicador colombiano que refleja el valor real ajustado por inflación, utilizado en instrumentos financieros e hipotecarios.
- **Producto Interno Bruto de Colombia (PIB):** mide el nivel de producción nacional trimestral, indicador central de la evolución económica del país.

Indicadores rezagados

Los indicadores de naturaleza rezagada confirman las tendencias económicas después de que los cambios han ocurrido, consolidando información retrospectiva sobre inflación, empleo y riesgo (The Conference Board, 2001).

- **Índice de precios al consumidor de Estados Unidos (CPIAUCSL):** mide la variación promedio de los precios de bienes y servicios, confirmando presiones inflacionarias posteriores a cambios en la demanda.
- **Tasa de desempleo de Estados Unidos (UNRATE):** refleja la proporción de personas desempleadas dentro de la fuerza laboral, reaccionando con rezago a los movimientos del ciclo económico.
- **Índice de precios al consumidor de Colombia (IPC 12M):** mide la inflación anual colombiana, indicador rezagado de presiones de costos y demanda agregada.

- **Tasa de política monetaria – límite superior de la Reserva Federal (FED UPPER):** representa el techo del rango objetivo de la política monetaria estadounidense, ajustado tras los cambios en inflación y actividad.
- **Índice EMBI Global (EMBI GLOBAL):** mide el riesgo promedio de los mercados emergentes mediante el diferencial de rendimiento frente a bonos del Tesoro estadounidense, confirmando variaciones globales en riesgo soberano.
- **Tasa de política monetaria del Banco de la República (TASA BANREP):** indicador de la postura monetaria colombiana, que se ajusta en respuesta a variaciones previas en inflación y crecimiento.
- **Índice EMBI Colombia (EMBI COL):** mide la prima de riesgo de la deuda colombiana frente a los bonos del Tesoro estadounidense, confirmando la percepción de riesgo país posterior a eventos macroeconómicos relevantes.

En nuestro análisis de factores, incluimos indicadores tanto adelantados como rezagados, y tanto de flujo como de nivel, para obtener una visión integral. No obstante, es crucial ser coherente en la manera de utilizarlos. Los factores adelantados capturan expectativas del mercado y señales tempranas del ciclo, los coincidentes reflejan las condiciones actuales de la economía y los rezagados miden el impacto de decisiones pasadas de política monetaria o condiciones financieras (The Conference Board, 2001). En tal sentido, se combinó indicadores líderes y rezagados, así como variables de flujo (que miden cambios por período) y de nivel (que miden estados acumulados), para capturar las distintas dimensiones temporales de la economía. Esta clasificación busca reflejar de manera integral la transmisión de la información macroeconómica hacia los precios de los activos financieros.

Esta diversidad de indicadores en la selección de factores nos permite cubrir tanto la anticipación de escenarios (con los leading de flujo) como la verificación de la solidez de las tendencias (con los lagging de nivel), aportando robustez al modelo multifactor resultante (Comisión Europea, 2008). Cada indicador fue evaluado con la metodología XGBoost previamente descrita, asegurando que su inclusión en el modelo esté justificada por su importancia estadística y su relevancia teórica dentro del marco de la APT.

Desde una perspectiva metodológica, esta estructura permite que el modelo XGBoost aprenda e incorpore las relaciones no lineales y temporales entre los factores económicos y los retornos de los activos. La combinación de variables de distinto tipo y frecuencia (flujo y nivel) ofrece una representación robusta de la dinámica económica y su efecto sobre los precios de los activos (Stock & Watson, 1999), alineándose con la lógica multifactorial propuesta originalmente por Ross, 1976) en la teoría APT.

3. Preparación de datos históricos: Para la estimación del modelo APT–XGBoost se construyó un dataset histórico integrado que combina las series de retornos mensuales de los activos financieros con variables macroeconómicas y de riesgo sistémico. El periodo analizado abarca desde enero de 2007 hasta julio de 2025, permitiendo capturar diversas fases del ciclo económico y episodios de volatilidad financiera. Los precios y retornos fueron suministrados por la división de Asset Management de Bancolombia, mientras que las variables macroeconómicas provienen de la base FRED (Federal Reserve Bank of St. Louis). Todas las series se unificaron en frecuencia mensual utilizando el último dato disponible de cada mes. El conjunto completo de datos se encuentra documentado en el archivo `base_precios.xlsx` (véase Anexo A).

El procesamiento y depuración de los datos incluyó los siguientes pasos técnicos:

- **Control de valores faltantes:** las observaciones sin datos previos suficientes no fueron consideradas en el entrenamiento ni en las transformaciones derivadas (por ejemplo, diferencias o retornos logarítmicos), evitando así sesgos en la estimación.
- **Revisión de valores atípicos:** se realizó una inspección exploratoria de las series temporales para identificar posibles observaciones extremas. Sin embargo, no se aplicaron transformaciones ni criterios de eliminación automática, dado que dichas observaciones pueden reflejar episodios reales de volatilidad o shocks económicos relevantes para el comportamiento de los retornos financieros.
- **Normalización y transformaciones estadísticas:** se calcularon retornos logarítmicos y variaciones porcentuales, ajustando escalas y estabilizando varianzas para mejorar la comparabilidad entre series.
- **Alineación temporal y rezagos:** las variables macroeconómicas se sincronizaron con los retornos de los activos mediante la introducción de rezagos (*lags*) apropiados, reflejando el desfase con el que los factores económicos suelen impactar los precios financieros.

Estimación de modelo multifactor (APT) tradicional:

Como punto de comparación y componente base, se estima un modelo de regresión multifactorial para los retornos de cada activo. Es decir, para cada activo i se ajusta

$$R_{i,t} - R_f(t) = \alpha_i + \sum_{j=1}^N \beta_{ij} F_{j,t} + \varepsilon_{i,t}$$

donde $F_{j,t}$ son realizaciones de factores (ej: cambio inesperado de inflación, o simplemente nivel de la variable si se normaliza), y α_i es el intercepto (idealmente cercano a 0 si los factores explican bien los retornos). A partir de esta regresión obtenemos las betas $\beta_{i,j}$ de

cada activo a cada factor. Con un estimador de las primas esperadas de cada factor λ_j calculamos el **retorno de equilibrio APT** para cada activo:

$$E(R_i)_{APT} = R_f + \sum_j \hat{\beta}_{i,j} E(\lambda_j)$$

Esta sería la predicción según un modelo lineal de equilibrio. Por ejemplo, de manera simplificada, si un bono colombiano tiene $\beta_{Inflación} = 0.5$ a la inflación inesperada y $\beta_{CDS} = 1.2$ al shock de riesgo país, y esperamos a un año vista una prima de inflación de 2% y una prima de riesgo país de 3%, entonces $E(R_{Bono}) = R_f + 0.5(2\%) + 1.2(3\%) = R_f + 4.6\%$. Este enfoque produce ya un conjunto de retornos esperados para todos los activos, consistente con factores económicos.

5. Entrenamiento del modelo XGBoost: Para cada activo financiero se estimó un modelo independiente de regresión XGBoost (Extreme Gradient Boosting), en el cual la variable dependiente correspondió al retorno mensual del activo (target), y las variables explicativas fueron los factores macroeconómicos seleccionados junto con sus rezagos temporales. El entrenamiento se efectuó utilizando la totalidad de observaciones disponibles entre 2007 y julio del 2025, con el fin de aprovechar la máxima extensión temporal posible y capturar distintos regímenes económicos.

Se garantizó la consistencia temporal mediante el desplazamiento de las variables explicativas en un periodo $(t - 1)$, evitando así la utilización de información futura durante el proceso de ajuste. Este procedimiento asegura la causalidad temporal y preserva la secuencia cronológica de la información, en concordancia con las recomendaciones metodológicas de Hyndman & Athanasopoulos, (2021).

El modelo XGBoost se configuró con un conjunto de hiperparámetros homogéneos para todos los activos, definidos con base en criterios de parsimonia, estabilidad y control de sobreajuste. La parametrización final fue la siguiente:

Parametrización

<i>Parámetro</i>	<i>Valor</i>
<i>objective</i>	'reg:squarederror'
<i>n_estimators</i>	50,
<i>learning_rate</i>	0.05
<i>max_depth</i>	5
<i>subsample</i>	0.9
<i>Colsample_bytree</i>	0.9
<i>Random_state</i>	42

Esta configuración busca equilibrar la capacidad predictiva y la robustez del modelo, siguiendo las recomendaciones de (T. Chen & Guestrin, 2016) para la implementación de algoritmos de *boosting* en contextos de regresión financiera.

Previo al entrenamiento, las bases de datos fueron sometidas a un proceso de filtrado y control de calidad. Se eliminaron las variables con más del 30 % de valores faltantes o con menos de 36 observaciones válidas, garantizando una longitud mínima adecuada para la estimación.

Posteriormente, se descartaron las variables altamente correlacionadas (coeficiente de correlación superior a 0.95) con el fin de mitigar problemas de multicolinealidad. Las series restantes se alinearon temporalmente y, en caso necesario, se imputaron los valores ausentes utilizando la media de cada variable.

Cada modelo se estimó de forma independiente por activo, permitiendo identificar la estructura específica de sensibilidad entre los factores macroeconómicos y los rendimientos financieros. Los resultados del entrenamiento incluyeron medidas de importancia de variables, las cuales se emplearon posteriormente para derivar las betas y primas de riesgo dentro del marco teórico de la Teoría de Precios por Arbitraje (APT).

La evaluación de relevancia de las variables explicativas se realizó utilizando la métrica de importancia por ganancia interna (*gain importance*), que constituye la medida nativa del algoritmo XGBoost. Dicha métrica cuantifica la contribución promedio de cada variable a la reducción de la función de pérdida en los árboles de decisión, siendo un indicador directo de su relevancia estructural en el modelo (T. Chen & Guestrin, 2016). A partir de esta medida se identificaron los factores más influyentes en la predicción de los retornos de cada activo, agrupando los distintos rezagos de una misma variable bajo un único nombre base.

Con base en la importancia por ganancia, se seleccionaron los factores con mayor relevancia promedio para cada modelo, los cuales fueron empleados posteriormente como variables explicativas en la estimación de betas y primas de riesgo dentro del modelo APT. La elección de la métrica de ganancia como criterio principal se justifica por su estabilidad y su capacidad para reflejar la contribución directa de cada predictor al desempeño global del modelo (Gu et al., 2020).

Integración APT–XGBoost:

La metodología empleada combina las capacidades predictivas del modelo de aprendizaje automático XGBoost con la estructura teórica de la Teoría de Precios por Arbitraje (APT), con el propósito de estimar de manera empírica los rendimientos esperados de los activos financieros y analizar la sensibilidad de estos frente a los factores macroeconómicos subyacentes.

El proceso se llevó a cabo en dos etapas complementarias:

1. Etapa I: Estimación predictiva mediante XGBoost

En la primera fase se entrenaron modelos de regresión XGBoost independientes para cada activo financiero, utilizando como variable dependiente el retorno mensual y como variables explicativas los factores macroeconómicos seleccionados (globales y locales), junto con sus rezagos.

El modelo XGBoost permitió capturar relaciones no lineales y efectos de interacción entre los factores, superando las limitaciones de linealidad impuestas por los modelos multifactoriales tradicionales. Una vez entrenados, se extrajo la importancia por ganancia interna (gain importance) de cada predictor, la cual cuantifica su contribución marginal a la reducción del error cuadrático medio en el proceso de entrenamiento (Chen & Guestrin, 2016).

Dado que el modelo se entrenó individualmente por activo, la métrica de importancia permitió identificar qué variables macroeconómicas presentaban mayor poder explicativo sobre los retornos de cada instrumento financiero. Posteriormente, se agruparon los rezagos de un mismo factor bajo un único nombre base y se seleccionaron los cuatro factores con mayor ganancia promedio.

Estos resultados proporcionaron un mapa empírico de exposición de cada activo frente a los factores de riesgo, constituyendo la base para la construcción del modelo APT en la segunda etapa.

2. Etapa II: Estimación estructural mediante el modelo APT

En la segunda fase, los factores previamente identificados como relevantes por el modelo XGBoost se incorporaron dentro de un marco de estimación basado en la Teoría de Precios por Arbitraje (APT) propuesta por Ross (1976).

Para cada activo se especificó un modelo de regresión lineal múltiple del tipo:

$$R_{it} = a_i + \beta_{i1} F_{1t} + \beta_{i2} F_{2t} + \dots + \beta_{in} F_{nt} + \varepsilon_{it}$$

donde R_{it} representa el retorno del activo i en el periodo t , F_{jt} corresponde a los factores macroeconómicos seleccionados (identificados por XGBoost), β_{ij} son las sensibilidades o betas APT, y ε_{it} es el término de error idiosincrático.

El número de factores incluidos en cada modelo se limitó a los cuatro más relevantes (top-4) para cada activo, siguiendo el principio de parsimonia propuesto por Connor & Korajczyk, (1988) Las betas se estimaron mediante regresiones lineales ordinarias (OLS) aplicadas sobre los retornos y los factores seleccionados.

Una vez obtenidas las betas (β_{ij}), se procedió a estimar las primas de riesgo (λ_j) y los rendimientos esperados de equilibrio bajo el marco APT. Para ello, se resolvió el sistema:

$$E[R_i] - R_f = \sum \beta_{ij} \lambda_j$$

donde R_f representa la tasa libre de riesgo λ_j denota la prima de riesgo asociado al factor j . En la práctica se estimó λ_j mediante una regresión lineal sin intercepto Ross (1980) y replicado en estudios recientes de empirical asset pricing con machine learning (Gu et al., 2020). Finalmente, los rendimientos esperados $E[R_i]$ se calcularon como la suma del rendimiento libre de riesgo y la combinación ponderada de las primas de riesgo por las betas de cada activo.

En la segunda fase, los factores previamente identificados como relevantes por el modelo XGBoost se incorporaron dentro de un marco de estimación basado en la Teoría de Precios por Arbitraje (APT) propuesta por Ross (1976).

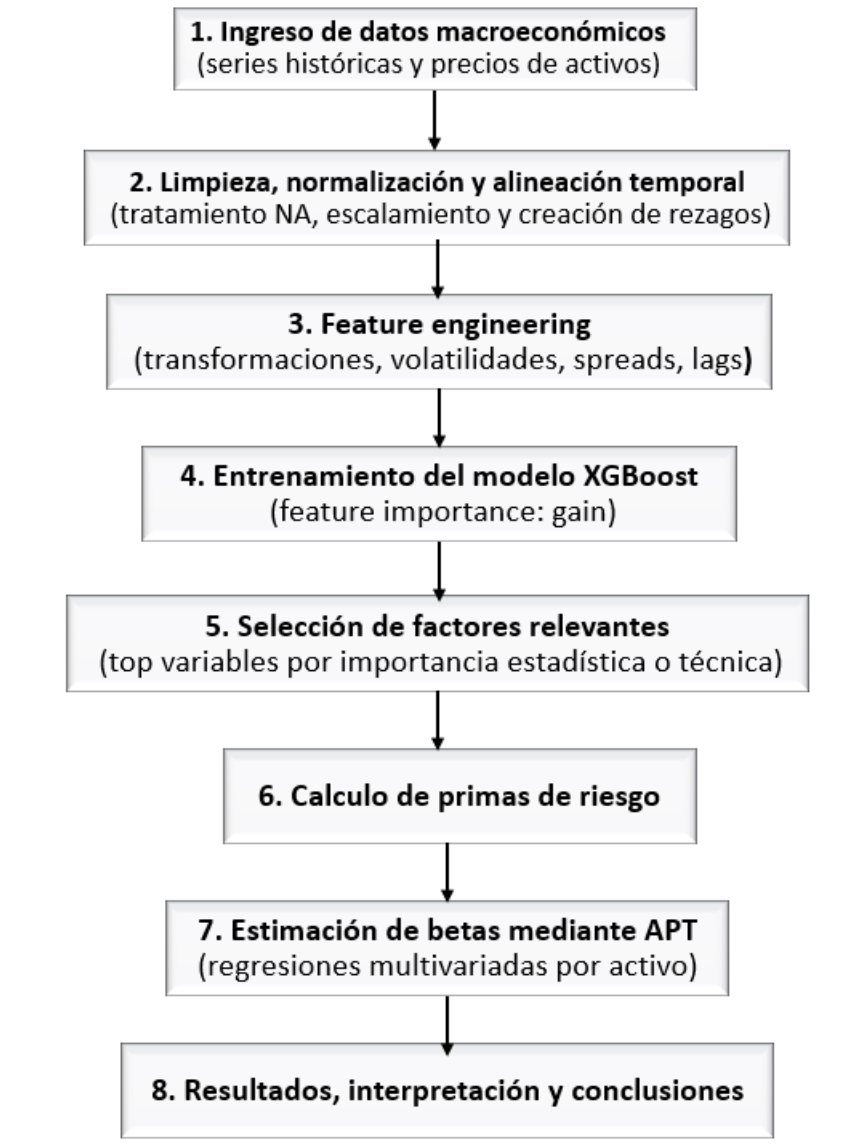
Para cada activo se especificó un modelo de regresión lineal múltiple del tipo:

$$R_{it} = \alpha_i + \beta_{i1} F_{1t} + \beta_{i2} F_{2t} + \dots + \beta_{in} F_{nt} + \varepsilon_{it}$$

donde R_{it} representa el retorno del activo i en el periodo t , F_{jt} corresponde a los factores macroeconómicos seleccionados (identificados por XGBoost), β_{ij} son las sensibilidades o betas APT, y ε_{it} es el término de error idiosincrático.

El número de factores incluidos en cada modelo se limitó a los cuatro más relevantes (top-4) para cada activo, siguiendo el principio de parsimonia propuesto por Connor & Korajczyk, (1988) Las betas se estimaron mediante regresiones lineales ordinarias (OLS) aplicadas sobre los retornos y los factores seleccionados.

Figura 1. Esquema descriptivo de la metodología implementada.



Capítulo 4: Resultados

4.1 Principales factores utilizados por el modelo

El modelo identifica para cada activo, los factores que mayor contribución relativa tienen en la explicación del retorno, con base en los valores de ganancia “gain” del algoritmo. Todos los factores corresponden a series oficiales provenientes de FRED, empleadas directamente en los

datasets del modelo. A continuación, se presenta una síntesis de los principales indicadores identificados:

- **CPIAUCSL (Consumer Price Index for All Urban Consumers):** índice de precios al consumidor para consumidores urbanos. Mide la variación del nivel general de precios en EE. UU. Su peso recurrente en varios activos refleja la sensibilidad de los mercados a la inflación observada.
- **PAYEMS (All Employees: Total Nonfarm):** nómina no agrícola total, que captura la evolución del empleo formal en EE. UU. Como indicador amplio del ciclo económico, aparece de forma sistemática en los factores seleccionados.
- **FEDFUNDS / FEDUPPER (Federal Funds Rate / Upper Target Limit):** tasa efectiva de fondos federales y límite superior del rango objetivo fijado por el FOMC. Representan la postura de política monetaria y afectan el costo de financiamiento y la liquidez.
- **US5Y (Market Yield on U.S. Treasury Securities at 5-Year Maturity):** rendimiento de los bonos del Tesoro a cinco años, que captura la parte media de la curva de rendimientos; su presencia en activos de renta fija y liquidez es consistente con su influencia directa en los precios de bonos.
- **T10Y3M (10-Year Treasury Constant Maturity Minus 3-Month Treasury Bill Rate):** diferencial entre el Treasury a 10 años y el de 3 meses, utilizado como aproximación de la pendiente de la curva de tipos. Se asocia con el estado del ciclo económico y con expectativas de actividad futura.
- **VIXCLS (Volatility Index):** índice de volatilidad implícita del mercado accionario estadounidense. Su aparición en los activos más expuestos al riesgo de mercado y liquidez (S&P 500 y REIT) indica sensibilidad a la incertidumbre financiera.
- **OAS (Option-Adjusted Spread, ICE BofA):** diferencial ajustado por opciones entre bonos corporativos y bonos del Tesoro. Refleja el riesgo crediticio y aparece con alta importancia en activos de crédito y mercados emergentes.
- **T5YIFR (5-Year Forward Inflation Expectation Rate):** expectativa implícita de inflación en un horizonte de cinco años. Su presencia en deuda de alto rendimiento sugiere sensibilidad a expectativas de inflación de mediano plazo.
- **INDPRO (Industrial Production Index):** índice de producción industrial que mide el output real de manufactura, minería y utilidades. Captura el estado del ciclo económico y aparece especialmente en bonos high yield.
- **GDPC1 (Real Gross Domestic Product):** PIB real de EE. UU ajustado por inflación, refleja el crecimiento económico. Su importancia en índices globales es coherente con el rol dominante del ciclo estadounidense.
- **UMCSENT (University of Michigan Consumer Sentiment Index):** indicador de percepción y expectativas económicas de los consumidores. Su presencia en acciones refleja la influencia del componente de confianza en la actividad económica.

- **UNRATE (Unemployment Rate):** tasa de desempleo total. Se relaciona con el mercado laboral y condiciones macroeconómicas generales; aparece en REITs como determinante de la actividad inmobiliaria.
- **M2SL (M2 Money Stock):** liquidez monetaria amplia (M1 + depósitos de ahorro y depósitos a plazo menores). Su relevancia en REITs sugiere relación entre condiciones de liquidez y apreciación de activos inmobiliarios financieros.

4.2 Importancia de factores por activo

En la siguiente tabla se resumen, para cada cartera, los cuatro factores más importantes y su contribución relativa respecto a la ganancia total del modelo. El parámetro Gain indica el porcentaje de la ganancia total que aporta cada factor, mientras que Total Top4 da cuenta de la suma del porcentaje total asociado a los cuatro factores seleccionados para cada activo que hace parte del portafolio. A su vez, Total Gain contiene la suma de la importancia de todas las variables del modelo para ese activo, es decir, tanto los factores que serán seleccionados como aquellos que serán descartados.

Detalle de Factores Seleccionados

Activo Factor	(1) IND LQ	Gain (%)	SP500	Gain (%)	ACWI	Gain (%)
1	US5Y_LEVEL_L1	43.7	CPIAUCSL_L2	3.8	CPIAUCSL_L2	4.7
2	PAYEMS_LEVEL_L1	18.8	VIXCLS_LEVEL_L1	3.0	GDPC1_L2	4.0
3	CPIAUCSL_LEVEL_L1	16.9	T10Y3M_L2	3.0	FEDFUNDS_LEVEL_L1	2.8
4	FEDFUNDS_LEVEL_L1	5.2	UMCSENT_LEVEL_L1	2.6	UMCSENT_LEVEL_L1	2.6
Total Top 4		84.6		12.3		14.16
Total Gain		93.7	29.6	29.6		31.3

Activo Factor	DEUDA USA	Gain (%)	DEUDA HY	Gain (%)	DEUDA EM	Gain (%)
1	CPIAUCSL_L2	5.7	PAYEMS_L2	7.1	PAYEMS_L1	7.6
2	US5Y_LEVEL_L1	3.3	INDPRO_L3	5.3	OAS_L1	7.3

3	FED_UPPER_L1	3.2	CPIAUCSL_L3	3.8	GDPC1_L1	6.1
4	FEDFUNDS_LEVEL_L1	2.9	T5YIFR_LEVEL_L1	3.7	T10Y3M_L2	5.1
Total Top 4		15.1		19.9		26.09
Total Gain		30.6		37.7		44.2

Factor	Activo REIT	Gain (%)
1	VIXCLS_L1	6.1
2	OAS_LEVEL_L1	3.9
3	UNRATE_LEVEL_L1	3.2
4	M2SL_LEVEL_L1	3.0
Total Top 4		16.1
Total Gain		33.8

Tabla 3. Detalle de Factores Seleccionados

4.3 Interpretación de los resultados del modelo

La incorporación de rezagos (*lags*) en los modelos econométricos y en los enfoques de aprendizaje automático obedece a la naturaleza dinámica de los procesos macroeconómicos y financieros. La literatura sobre modelos de rezagos distribuidos señala que los efectos de un choque macroeconómico no se manifiestan de manera inmediata, sino que se propagan gradualmente a través del tiempo, influyendo las decisiones de consumo, inversión y política monetaria, así como los mecanismos de formación de expectativas. Este comportamiento dinámico ha sido ampliamente documentado en modelos estructurales y de series de tiempo, donde la inclusión de rezagos permite capturar tanto la persistencia de las variables como la transmisión retardada de los shocks (Enders, 2014)

Desde esta perspectiva, la inclusión de variables con rezagos de uno, dos y tres periodos permite capturar:

- Efectos retardados de la política monetaria, coherente con la existencia de fricciones nominales y rigideces de ajuste.
- Reacciones diferidas del mercado laboral y la demanda agregada, que influyen en el ciclo económico con desfase.
- Dinámicas de transmisión financiera, donde la volatilidad, los spreads crediticios y las expectativas inflacionarias afectan rendimientos con distintos plazos de incidencia.
- Autocorrelaciones naturales de series macroeconómicas, que los modelos ML pueden explotar como predictores estructurados temporalmente.

En consecuencia, el uso de rezagos en el modelo no solo evita la utilización de información futura, sino que introduce una estructura explícita de transmisión temporal, coherente con la teoría macro-financiera y con la evidencia empírica sobre los mecanismos de ajuste económico.

- **IND LQ (Índice de liquidez):** En esta cartera los factores más relevantes presentan rezagos de un periodo, lo que evidencia una respuesta prácticamente inmediata a las condiciones monetarias y al ciclo económico. El rendimiento de los bonos del Tesoro a cinco años rezagado (US5Y_LEVEL_L1) concentra cerca de la mitad de la ganancia total del modelo, indicando que la estructura temporal de tasas internas incide de manera directa en los activos de muy baja duración. La relevancia del empleo no agrícola rezagado (PAYEMS_LEVEL_L1) y de la inflación del mes anterior (CPIAUCSL_LEVEL_L1) concuerda con la teoría del canal de expectativas y credibilidad del banco central: un mercado laboral robusto y presiones inflacionarias recientes tienden a elevar las tasas anticipadas, afectando instrumentos de liquidez de forma casi contemporánea.
- **S&P 500 (acciones estadounidenses):** En el caso del S&P 500, el modelo selecciona principalmente rezagos de dos periodos (CPIAUCSL_L2, T10Y3M_L2). Esto sugiere que los shocks inflacionarios y las variaciones en la pendiente de la curva afectan los retornos accionarios con un desfase que puede asociarse a retrasos en el ajuste de valuación, cambios en expectativas de beneficios corporativos y el proceso de transmisión de la política monetaria. La importancia del VIX con rezago de un periodo (VIXCLS_LEVEL_L1) revela que las condiciones de incertidumbre se trasladan rápidamente al mercado accionario, mientras que el sentimiento del consumidor rezagado (UMCSENT_LEVEL_L1) muestra la incidencia de expectativas agregadas sobre el consumo futuro en la valoración bursátil.
- **ACWI (acciones globales):** El ACWI presenta patrones similares al S&P 500, con la inflación y el PIB rezagados dos periodos (CPIAUCSL_L2 y GDPC1_L2) como los factores dominantes. Esto demuestra que los mercados globales internalizan el ciclo económico estadounidense con un retardo estructural. Los rezagos cortos en la tasa de fondos federales y en el sentimiento del consumidor (FEDFUNDS_LEVEL_L1, UMCSENT_LEVEL_L1) indican que la transmisión de información financiera de EE. UU. hacia mercados internacionales se produce rápidamente, pero la transmisión macroeconómica fue capturada con mayor rezago.
- **Deuda USA:** El principal determinante es la inflación de dos meses atrás (CPIAUCSL_L2), lo cual refleja que los precios de los bonos responden al proceso inflacionario con un desfase que coincide con el tiempo que tarda el mercado en

ajustar expectativas sobre el sendero futuro de la política monetaria.

Los factores FED_UPPER_L1 y FEDFUNDS_LEVEL_L1 evidencian una alta sensibilidad a los parámetros contemporáneos del rango objetivo de política monetaria. En otras palabras, los bonos internalizan tanto el rezago inflacionario como las señales recientes de la autoridad monetaria.

- **Deuda High Yield:** Esta clase de activo se caracteriza por rezagos más largos (L2 y L3), consistentes con la naturaleza del crédito corporativo de alto riesgo. El empleo rezagado dos periodos (PAYEMS_L2) y la producción industrial rezagada tres periodos (INDPRO_L3) sugieren que las condiciones de solvencia y riesgo idiosincrático de las empresas responden a fases atrasadas del ciclo económico. La inflación rezagada tres periodos (CPIAUCSL_L3) confirma que el *spread* requerido por los inversionistas incorpora presiones inflacionarias históricas, lo cual puede interpretarse como evidencia de rigidez en la actualización de primas de riesgo.
- **Deuda Emergente:** La importancia del empleo de EE. UU. con un rezago corto (PAYEMS_L1) refleja la inmediatez con la que la actividad económica estadounidense afecta el costo de financiamiento de economías emergentes. Sin embargo, variables como T10Y3M_L2 muestran que la transmisión vía condiciones financieras ocurre con desfase. El OAS_L1 confirma que el riesgo crediticio de los emergentes depende de la posición reciente del crédito corporativo en EE. UU., mientras que el crecimiento del PIB rezagado (GDPC1_L1) captura el componente real.
- **REIT:** Los REIT presentan una fuerte incidencia de factores financieros rezagados un periodo (VIXCLS_L1, OAS_LEVEL_L1). Este comportamiento es coherente con la naturaleza híbrida del activo: son instrumentos de renta variable (alta sensibilidad y volatilidad), pero con flujos de caja relativamente estables (sensibilidad al crédito y a la liquidez monetaria). La tasa de desempleo rezagada (UNRATE_LEVEL_L1) y la oferta monetaria (M2SL_LEVEL_L1) sugieren que los REIT reaccionan casi de inmediato al entorno macroeconómico, especialmente a condiciones de liquidez, lo cual coincide con su estructura de financiación intensiva en deuda.

Los hallazgos muestran que el modelo logra identificar estructuras de riesgo diferenciadas por clase de activo, coherentes con la teoría financiera y con estudios previos sobre determinación de retornos. La inflación, la política monetaria, el mercado laboral y los indicadores de riesgo (VIX, OAS) emergen consistentemente como variables clave, lo cual coincide con la evidencia empírica de modelos multifactoriales y con trabajos clásicos como Chen, Roll y Ross (1986). Un hallazgo relevante es la diferencia marcada entre clases de activos. En el

Índice de Liquidez, cuatro factores explican más del 80 % de la importancia total, lo que sugiere una estructura altamente concentrada dominada por tasas de interés y condiciones macroeconómicas básicas. En contraste, para índices accionarios como S&P 500 y ACWI, los factores se distribuyen de manera más difusa: la inflación, la volatilidad, el diferencial de tasas y el sentimiento del consumidor explican menos de la mitad de la importancia, reflejando un comportamiento más complejo y dependiente de un conjunto más amplio de riesgos.

El modelo también confirma que la política monetaria estadounidense influye de forma transversal en todos los activos, incluidos los de mercados emergentes, lo cual concuerda con estudios que resaltan el papel dominante de la liquidez global y del ciclo estadounidense en la formación de precios internacionales. Asimismo, la presencia del OAS como factor relevante en deuda corporativa y deuda emergente respalda la idea de que el riesgo crediticio agregado es un determinante central de la prima de riesgo exigida en estos mercados.

El uso del modelo como herramienta de selección de factores aporta evidencia adicional sobre la conveniencia de incorporar técnicas de machine learning en modelos de equilibrio. Aunque el APT clásico supone relaciones lineales y coeficientes constantes, el XGBoost facilita la identificación empírica de interacciones complejas y rezagos relevantes sin imponer una estructura funcional rígida. La estimación posterior de las betas mediante regresiones lineales mantiene, a su vez, la interpretabilidad del modelo, elemento esencial en un enfoque de equilibrio para decisiones de asignación estratégica.

En conjunto, los resultados muestran que la metodología propuesta captura adecuadamente las sensibilidades sistemáticas relevantes para cada activo y produce estimaciones de retorno de equilibrio coherentes con los patrones macrofinancieros observados en los mercados.

4.4. Estimación de betas

En esta sección se presentan los coeficientes beta estimados para cada clase de activo obtenidos tras la selección de factores realizada por el algoritmo XGBoost. Las betas cuantifican la sensibilidad marginal del retorno del activo ante variaciones en cada factor sistemático, manteniendo constantes los demás factores. Su interpretación es central para un modelo de equilibrio, pues determinan la prima exigida por exposición al riesgo sistemático.

Un aspecto clave de los resultados es la presencia simultánea de betas positivos y negativos, siendo estos últimos particularmente relevantes para comprender la dinámica de ajuste del retorno de equilibrio. A diferencia del CAPM, donde el signo suele estar condicionado por la covarianza con el mercado, en el APT una beta negativa no implica aversión al riesgo, sino que indica que: Un aumento en el factor reduce sistemáticamente el retorno esperado del activo, debido a la estructura económica que relaciona ese factor con los flujos de caja o con el costo del capital.

Índice de Liquidez – IND_LQ_USA_BILL1M_USD_001

Betas estimadas – Índice de Liquidez

<i>Factor</i>	<i>Beta</i>
<i>US5Y</i>	<i>0.0146</i>
<i>PAYEMS</i>	<i>-0.0010</i>
<i>CPIAUCSL</i>	<i>-0.0179</i>
<i>FEDFUNDS</i>	<i>0.1412</i>

Tabla 4. Betas estimadas – Índice de Liquidez

Interpretación

La estructura de sensibilidades del índice de liquidez es coherente con la naturaleza de un activo de duración casi nula, altamente expuesto a las condiciones monetarias de corto plazo. Las betas positivas frente a US5Y y FEDFUNDS reflejan que los instrumentos monetarios reaccionan de manera casi inmediata a los ajustes en las tasas de referencia: cuando la política monetaria se endurece, los rendimientos de los instrumentos líquidos aumentan rápidamente, elevando la rentabilidad del índice.

Las betas negativas requieren un análisis más matizado desde la perspectiva macroeconómica:

Beta negativa frente a CPIAUCSL (inflación)

La relación negativa con el nivel de precios es consistente con la teoría económica:

La inflación erosiona el retorno real de los activos de liquidez, cuyo rendimiento nominal es relativamente bajo.

En contextos de inflación alta, los inversionistas exigen mayor compensación real, pero los instrumentos de muy corto plazo solo ajustan su retorno gradualmente, generando un efecto negativo contemporáneo. Adicionalmente, la inflación persistente suele anticipar política monetaria más contractiva, lo que inicialmente presiona a la baja los rendimientos reales antes del ajuste pleno de las tasas cortas.

Beta negativa frente a PAYEMS (empleo total)

Aunque el empleo es un indicador coincidente del ciclo económico, su impacto sobre activos monetarios se produce a través de expectativas de política monetaria:

Un mercado laboral fuerte suele interpretarse como señal de mayor actividad económica y presión inflacionaria futura. Esto incrementa la probabilidad de endurecimiento monetario, lo que genera un costo de oportunidad temporal para los activos de liquidez, dado que los rendimientos tardan algunos periodos en ajustarse plenamente a las nuevas tasas.

En este sentido, la beta negativa no sugiere que “más empleo sea desfavorable”, sino que refleja la sensibilidad anticipada del índice de liquidez a cambios esperados en las condiciones financieras.

Renta Variable EE. UU. – S&P 500

Betas estimadas – S&P 500

<i>Factor</i>	<i>Beta</i>
<i>CPIAUCSL_L2</i>	<i>-1.2070</i>
<i>VIXCLS</i>	<i>-0.0094</i>
<i>T10Y3M_L2</i>	<i>0.0249</i>
<i>UMCSENT</i>	<i>-0.1176</i>

Tabla 5. Betas estimadas – S&P 500

La estructura de sensibilidades estimada para el S&P 500 refleja patrones consistentes con la teoría macro-financiera y la evidencia empírica del mercado accionario estadounidense. Cada factor captura un canal específico a través del cual las condiciones económicas afectan los retornos de la renta variable.

Beta negativa frente a CPIAUCSL_L2 (inflación rezagada)

La sensibilidad negativa y de magnitud alta indica que la inflación persistente deteriora el valor presente de los flujos de caja futuros, afectando directamente los múltiplos de valoración (P/E).

Esto ocurre por varios canales:

- Una inflación rezagada elevada implica mayor persistencia inflacionaria, lo que incrementa la probabilidad de políticas monetarias contractivas.
- El costo del capital (tasa de descuento) aumenta, reduciendo el valor presente de los ingresos corporativos futuros.
- La inflación tiende a comprimir márgenes cuando las firmas no pueden trasladar completamente sus costos a los precios.

Por esto, incluso con rezagos, la inflación tiene un efecto claramente contractivo sobre la renta variable, lo cual es coherente con la literatura de Chen, Roll y Ross (1986) que identifica a la inflación como un factor adverso para el equity premium.

Beta negativa frente a VIXCLS (volatilidad implícita)

El coeficiente negativo es totalmente consistente con la interpretación del VIX como un índice de aversión al riesgo y de incertidumbre financiera:

- Cuando la volatilidad esperada aumenta, los inversionistas demandan una prima de riesgo mayor, lo que presiona a la baja los precios accionarios.
- Los periodos de alta volatilidad suelen coincidir con shocks de riesgo sistémico, donde los mercados accionarios corrigen agresivamente.
- En el marco del APT, el VIX captura un factor de riesgo no diversificable relacionado con incertidumbre macroeconómica y tensiones financieras.

Beta negativa frente a UMCSSENT (sentimiento del consumidor)

La relación negativa implica que la caída en la confianza del consumidor se traduce en expectativas más débiles de crecimiento y beneficios corporativos.

Esto ocurre porque:

- El consumo privado representa un porcentaje importante del PIB en EE. UU. una reducción en las expectativas de gasto se transmite rápidamente a los ingresos empresariales proyectados.
- El deterioro en la confianza suele coincidir con fases tardías del ciclo económico.
- La confianza del consumidor es un indicador leading del ciclo: su caída anticipa desaceleraciones que afectan al mercado accionario.
- En conjunto, el UMCSSENT funciona como un factor de expectativas de demanda futura, alineado con estudios que muestran su capacidad predictiva sobre retornos accionarios.

Beta positiva frente a T10Y3M_L2 (pendiente de curva, rezagada)

El único coeficiente positivo indica que una curva más empinada es decir, mayores tasas largas respecto a las cortas se interpreta como señal de expansión económica futura. Este resultado es altamente coherente con la teoría del ciclo financiero:

- Una pendiente positiva implica expectativas de mayor crecimiento del PIB y de ventas corporativas.
- En los modelos macro-financieros, la curva de rendimientos es un predictor adelantado de la actividad económica (Harvey, 1988).
- Cuando la curva se empina, mejora la percepción de riesgo y aumenta el apetito por renta variable.

Acciones Globales – ACWI

Betas estimadas – ACWI

<i>Factor</i>	<i>Beta</i>
---------------	-------------

<i>FEDFUNDS</i>	4.2674
<i>CPIAUCSL_L2</i>	-2.8568
<i>UMCSENT</i>	-0.0815
<i>GDPC1_L2</i>	-0.3185

Tabla 6. Betas estimadas – ACWI

Interpretación

La estructura de betas estimada para ACWI muestra una marcada sensibilidad a las condiciones monetarias y macroeconómicas de Estados Unidos, lo cual es consistente con la dominancia del mercado estadounidense dentro del índice global.

La beta positiva frente a FEDFUNDS es especialmente relevante: sugiere que, en el periodo analizado, los episodios de normalización monetaria estuvieron asociados con recomposición de portafolios hacia activos globales de alta calidad, posiblemente debido al cierre de diferenciales de tasas internacionales o al rol del dólar como activo dominante en mercados globales. Esta dinámica implica que el endurecimiento monetario no necesariamente deprimió los mercados globales, sino que incentivó movimientos de capital hacia economías desarrolladas y activos más líquidos.

Las betas negativas aportan información clave sobre los canales de transmisión macroeconómica:

CPIAUCSL_L2 (inflación rezagada) Beta negativa y de gran magnitud

La inflación estadounidense, incluso con rezago, eleva la tasa de descuento global y deteriora los múltiplos de valoración en mercados desarrollados y emergentes.

Un repunte inflacionario persistente en EE. UU. tiende a: aumentar las expectativas de política monetaria contractiva, fortalecer al dólar, encarecer el financiamiento global, y reducir los flujos hacia renta variable mundial.

Por tanto, el signo negativo y su magnitud son coherentes con la centralidad de la política monetaria estadounidense en el sistema financiero internacional.

GDPC1_L2 (PIB real rezagado)

La beta negativa indica que un deterioro rezagado en la actividad económica de EE. UU. se transmite como un choque global, afectando tanto a empresas estadounidenses (que pesan más del 60% del ACWI) como a economías altamente dependientes del ciclo global.

El rezago (L2) sugiere que los efectos de desaceleraciones económicas tardan algunos trimestres en reflejarse plenamente en los retornos globales.

UMCSENT (sentimiento del consumidor)

La sensibilidad negativa frente al sentimiento del consumidor refleja que la demanda interna estadounidense actúa como ancla del crecimiento corporativo global.

Una caída en la confianza reduce las expectativas de consumo futuro, debilitando los ingresos proyectados de empresas multinacionales con exposición al mercado estadounidense.

Deuda del Tesoro – Deuda USA

Betas estimadas – Deuda USA

<i>Factor</i>	<i>Beta</i>
<i>US5Y</i>	<i>1.2449</i>
<i>FEDFUNDS</i>	<i>-1.1355</i>
<i>CPIAUCSL_L2</i>	<i>-1.0657</i>
<i>FED_UPPER_L1</i>	<i>-0.0584</i>

Tabla 7. Betas estimadas – Deuda USA

Interpretación

La estructura de betas estimada para la deuda del Tesoro estadounidense refleja con claridad los principios fundamentales de la valoración de renta fija y la fuerte dependencia de estos activos a la política monetaria. Las sensibilidades negativas frente a factores monetarios y de inflación son coherentes con la relación inversa entre tasas de interés y precios de los bonos.

FEDFUNDS y FED_UPPER_L1 – Betas negativas y consistentes con teoría de bonos

Los coeficientes negativos indican que aumentos en la tasa de política monetaria se traducen en reducciones en el precio de los bonos del Tesoro.

Esto responde a tres mecanismos bien documentados:

- **Aumento en la tasa de descuento:**
La política monetaria contractiva incrementa directamente la tasa libre de riesgo y, por ende, la tasa de descuento utilizada en la valoración de los bonos.
- **Ajuste inmediato en el tramo corto de la curva:**
La deuda del Tesoro, especialmente en horizontes cortos–medianos, incorpora rápidamente los movimientos del FOMC.
- **Reasignación de portafolios:**
Una subida de tasas aumenta la rentabilidad de instrumentos más cortos y reduce el atractivo relativo de los bonos ya emitidos, provocando presión bajista en sus precios. El diferencial entre FEDFUNDS (tasa efectiva) y FED_UPPER_L1 (límite superior del rango objetivo) confirma que los ajustes de política monetaria afectan el retorno de los bonos tanto contemporáneamente como con rezago, coherente con la transmisión gradual del FOMC.

CPIAUCSL_L2 – Beta negativa frente a la inflación rezagada

La beta negativa frente a la inflación tiene fundamentos sólidos:

- La inflación persistente erosiona el valor real de los pagos futuros de un bono nominal.
- Anticipa mayor presión para elevar la tasa de política monetaria, lo que incrementa el rendimiento exigido por los inversionistas.

- La inflación rezagada captura expectativas de continuidad inflacionaria, lo que presiona a la baja los precios de los bonos incluso antes de ajustes oficiales del FOMC.

Crédito High Yield – Deuda HY

Betas estimadas – Deuda HY

<i>Factor</i>	<i>Beta</i>
<i>PAYEMS_L2</i>	<i>-0.1989</i>
<i>INDPRO_L3</i>	<i>0.0249</i>
<i>CPIAUCSL_L3</i>	<i>-2.9407</i>
<i>T5YIFR</i>	<i>6.5873</i>

Tabla 8. Betas estimadas – Deuda HY

Interpretación

La estructura de betas estimada para la Deuda High Yield refleja los mecanismos centrales documentados en la literatura sobre riesgo crediticio, ciclos económicos y condiciones financieras. En particular, el comportamiento del HY combina características procíclicas, sensibilidad al riesgo de default y una respuesta diferenciada ante la inflación observada versus la inflación esperada.

PAYEMS_L2 – Beta negativa (empleo rezagado)

El coeficiente negativo frente a PAYEMS_L2 indica que una mejora rezagada en el empleo estadounidense se asocia con reducciones posteriores en los retornos del crédito HY. Este resultado es coherente con la evidencia empírica que muestra que los indicadores laborales suelen irse rezagando respecto al ciclo crediticio, mientras que los spreads High Yield comienzan a ampliarse antes de que el deterioro macroeconómico sea visible en variables reales (Altman, 1989, 2000).

Esto se explica porque, en etapas avanzadas del ciclo económico, los inversionistas suelen comenzar una rotación hacia activos de mayor calidad crediticia (“late-cycle rotation”), lo cual reduce la demanda por instrumentos HY aun cuando el empleo todavía luce fuerte. Asimismo, la literatura ha demostrado que el riesgo crediticio y los spreads HY reaccionan primero a las condiciones financieras y solo después a los indicadores reales, lo que justifica el uso de rezagos en este factor (Gilchrist & Zakrajšek, 2012).

CPIAUCSL_L3 – Beta negativa (inflación realizada rezagada)

La beta negativa y de magnitud significativa frente a CPIAUCSL_L3 indica que la inflación persistente afecta negativamente la solvencia de los emisores HY. Las empresas con baja calificación crediticia tienden a tener estructuras de deuda más frágiles, menores márgenes de

maniobra y una capacidad limitada para trasladar precios a los consumidores. La literatura muestra que en entornos de inflación elevada y sostenida se incrementan los costos operativos y financieros, aumentando la probabilidad de incumplimiento (Altman, 2012).

Adicionalmente, la inflación persistente suele anticipar ciclos de política monetaria contractiva, lo cual incrementa los costos de refinanciación para emisores HY, un mecanismo ampliamente discutido por Gilchrist y Zakrajšek (2012) y por Chen, Collin-Dufresne y Goldstein (2007). El uso de rezagos (L3) es coherente, dado que los efectos de la inflación observada tienden a transmitirse al riesgo crediticio con varios periodos de retraso.

INDPRO_L3 – Beta positiva pero de baja magnitud

La beta positiva frente a la producción industrial, aunque relativamente pequeña, es coherente con el carácter procíclico del crédito HY. Una mejora rezagada en la actividad industrial suele asociarse con mayores ingresos corporativos, menor riesgo de default y compresión moderada de spreads. Sin embargo, la magnitud contenida también es consistente con la evidencia empírica: los factores macroeconómicos reales explican una porción menor del spread HY en comparación con los factores financieros y de riesgo sistémico (Huang & Huang, 2012).

T5YIFR – Beta positiva y dominante (inflación esperada a 5 años)

El coeficiente positivo y dominante frente a la inflación esperada a 5 años refleja un fenómeno ampliamente documentado en los mercados de crédito: en periodos donde la inflación esperada aumenta junto con expectativas de crecimiento económico, los spreads HY tienden a comprimirse. Este comportamiento, conocido como “reflation trade”, ha sido analizado en reportes del Bank for International Settlements (2019), donde se observa que una mayor inflación esperada, cuando no se percibe como desanclada, coincide con mejores condiciones financieras para emisores de menor calidad crediticia.

Esto se debe a que:

- La inflación esperada suele acompañarse de crecimiento nominal más alto, lo que mejora ingresos corporativos.
- La deuda HY es mayoritariamente nominal, por lo que la inflación esperada reduce parte de la carga real de la deuda.
- Los inversionistas anticipan mejores flujos de caja en términos nominales, lo cual se traduce en reducción de spreads.

Este resultado es coherente con la evidencia empírica de Krishnamurthy y Vissing-Jorgensen (2012), quienes encuentran que las expectativas de inflación se relacionan positivamente con el apetito por riesgo en segmentos no-investment-grade.

Deuda de Mercados Emergentes – EM

Betas estimadas – Deuda EM

<i>Factor</i>	<i>Beta</i>
<i>T10Y3M_L2</i>	<i>-0.3978</i>
<i>PAYEMS_L1</i>	<i>-0.2804</i>
<i>OAS_L1</i>	<i>-0.1721</i>
<i>GDPC1_L1</i>	<i>0.1385</i>

Tabla 9. Betas estimadas – Deuda EM

Interpretación

La estructura de sensibilidades estimada evidencia que la Deuda de Mercados Emergentes (EM) está fuertemente condicionada por el ciclo económico y financiero de Estados Unidos. Tres factores presentan betas negativas, lo cual refleja la naturaleza subordinada del riesgo soberano emergente frente al desempeño de los mercados desarrollados.

T10Y3M_L2 – Beta negativa (pendiente de la curva de EE. UU.)

La pendiente positiva de la curva de rendimientos estadounidense (mayores tasas largas frente a las cortas) suele asociarse con expectativas de crecimiento económico en Estados Unidos, lo cual tiende a desplazar capital hacia activos domésticos de menor riesgo.

La beta negativa indica que, cuando la curva se empina, los inversionistas reducen su exposición al riesgo emergente en favor de activos estadounidenses, presionando al alza los rendimientos EM y reduciendo su precio.

El uso de rezagos (L2) sugiere que estos movimientos no son inmediatos: el ajuste de flujos hacia o desde mercados emergentes responde con desfase a las señales del ciclo estadounidense.

PAYEMS_L1 – Beta negativa (empleo estadounidense)

Una mejora en el empleo estadounidense incrementa la percepción de fortaleza económica interna y eleva la demanda por activos locales, especialmente instrumentos del Tesoro y crédito corporativo de alta calidad.

En ese contexto, los mercados emergentes pierden atractivo relativo, lo que se traduce en salidas de capital y ampliación de spreads soberanos. Por ello, un crecimiento rezagado del empleo en EE. UU. se asocia con menores retornos para la deuda EM.

OAS_L1 – Beta negativa (spreads corporativos de EE. UU.)

El coeficiente negativo frente a OAS_L1 indica que, cuando los spreads corporativos de EE. UU. se comprimen, es decir, cuando el riesgo crediticio de empresas estadounidenses disminuye, los inversionistas tienden a preferir crédito de alta calidad en EE. UU. frente a riesgo emergente.

Esto genera un movimiento de portafolios desde EM hacia activos corporativos estadounidenses, reduciendo el retorno de la deuda emergente.

GDPC1_L1 – Beta positiva (PIB real de EE. UU.)

El único coeficiente positivo indica que el crecimiento real de la economía estadounidense tiene un efecto favorable, aunque rezagado, sobre los retornos de mercados emergentes.

Una expansión del PIB estadounidense suele estar asociada con:

- mayor comercio internacional,
- mayor demanda por bienes y materias primas provenientes de economías emergentes,
- menores tensiones financieras globales.

Estos efectos mejoran la percepción de solvencia y flujo de capital hacia mercados emergentes, elevando los retornos de la deuda EM. El rezago (L1) es consistente con la transmisión gradual del ciclo estadounidense hacia economías emergentes.

REITs – Activos Alternativos

Betas estimadas – REITs

<i>Factor</i>	<i>Beta</i>
<i>VIXCLS_L1</i>	<i>-0.0017</i>
<i>OAS</i>	<i>1.2885</i>
<i>UNRATE</i>	<i>-0.0161</i>
<i>M2SL</i>	<i>-1.4567</i>

Tabla 10. Betas estimadas – REITs

Interpretación

La estimación de betas para los REITs revela una estructura de sensibilidades coherente con su naturaleza híbrida: combinan características de activos inmobiliarios con la dinámica financiera propia de instrumentos listados en mercados públicos. Los resultados muestran que los REITs responden principalmente a condiciones de riesgo, liquidez y ciclo laboral, mientras que presentan una relación particular con el riesgo corporativo.

VIXCLS_L1 – Beta negativa (volatilidad implícita)

El coeficiente negativo indica que un aumento rezagado en la volatilidad implícita del mercado reduce los retornos de los REITs.

Esto es consistente con su comportamiento como activo sensible al apetito de riesgo:

- Mayor incertidumbre implica mayores primas exigidas por los inversionistas.
- En periodos de aversión al riesgo, los activos inmobiliarios listados tienden a corregir debido a su naturaleza parcialmente apalancada.

La magnitud pequeña refleja que el impacto del VIX existe, pero es moderado comparado con otros sectores más procíclicos.

UNRATE – Beta negativa (tasa de desempleo)

La beta negativa muestra que incrementos en la tasa de desempleo deterioran el desempeño de los REITs.

Esto ocurre porque el mercado laboral está estrechamente relacionado con **la demanda de espacios inmobiliarios**:

- Peores condiciones laborales reducen la demanda por oficinas y comercio.
- Menor empleo implica menor ingreso disponible y presión sobre ocupación y capacidad de pago de arrendatarios.

M2SL – Beta negativa (liquidez monetaria agregada)

El coeficiente negativo frente a M2SL indica que una reducción en la liquidez monetaria afecta negativamente los retornos de los REITs.

Este resultado es consistente con que el sector inmobiliario:

- es intensivo en financiamiento,
- depende de condiciones crediticias holgadas,
- y es sensible al costo del capital.

Cuando la liquidez sistémica se reduce, el costo de financiamiento para empresas inmobiliarias aumenta y se restringen las condiciones crediticias, deteriorando retornos.

OAS – Beta positiva (spreads corporativos)

El coeficiente positivo frente al OAS es un resultado relevante porque plantea una dinámica distinta a los activos HY o IG.

Cuando los spreads corporativos se amplían —indicando mayor riesgo en crédito empresarial— los REITs pueden actuar como una **alternativa intermedia** entre renta variable y crédito corporativo:

- Los REITs mantienen ingresos relativamente estables basados en rentas.
- Su estructura de cash flows es menos volátil que la de emisores corporativos de menor calidad.

Durante episodios de ampliación moderada de spreads, ciertos inversionistas pueden rotar hacia REITs como opción de riesgo medio y con activos reales subyacentes. Por ello, la beta positiva sugiere que, en momentos donde el crédito corporativo se percibe como más riesgoso, los REITs pueden beneficiarse como un sustituto parcial dentro del portafolio.

4.5 Estimación e Interpretación de las Primas de Riesgo (λ) del Modelo

Esta sección presenta la estimación, estructura e interpretación del vector de primas de riesgo (λ) dentro del modelo. La estimación de λ es esencial en la teoría de valoración por arbitraje, pues determina el precio del riesgo sistemático asociado a cada factor macroeconómico y financiero. Su integración con las betas (β) estimadas permite obtener retornos de equilibrio coherentes con los fundamentos teóricos y con el comportamiento observado en los mercados financieros internacionales.

4.5.1 Fundamentos Teóricos del Vector λ en el APT

En el modelo de Arbitrage Pricing Theory (APT), el retorno esperado de un activo i viene dado por la ecuación:

$$E[R_i] = R_f + \sum \beta_{ik} \lambda_k$$

donde:

- β_{ik} : sensibilidad del activo i al factor k .
- λ_k : prima de riesgo asociada al factor k .
- R_f : tasa libre de riesgo.

En este marco, $\lambda_k > 0$ implica que un aumento en el factor incrementa el retorno esperado del activo, mientras que $\lambda_k < 0$ señala que dicho factor actúa como un descontador del retorno esperado, ya sea elevando la tasa de descuento o deteriorando expectativas macroeconómicas. La estimación desarrollada en esta tesis utiliza un enfoque híbrido: el modelo XGBoost identifica factores relevantes, mientras que el APT define la estructura lineal de precios.

4.5.2 Primas de Riesgo Estimadas

La Tabla 11 presenta las primas de riesgo estimadas (λ) del modelo. En el marco de la Teoría de Valoración por Arbitraje (APT), las primas de riesgo representan la remuneración marginal exigida por el mercado por unidad de exposición a cada factor sistemático (Ross, 1976). Es decir, cada λ captura cuánto aumenta (o disminuye) el retorno esperado de un activo cuando su sensibilidad a ese factor (su beta) se incrementa en una unidad, manteniendo constantes los demás factores.

Bajo este enfoque, el rendimiento esperado en exceso sobre la tasa libre de riesgo puede expresarse como:

$$E[R_i] - R_f = \sum \beta_{ik} \lambda_k$$

donde β_{ik} es la exposición del activo i al factor k , y λ_k es la prima de riesgo asociada a dicho factor (Cochrane, 2005).

En este sentido:

- Un λ positivo indica que el mercado recompensa la exposición a ese factor: los activos que se mueven más con ese factor tienden a exhibir mayores retornos esperados.
- Un λ negativo sugiere que el factor actúa como activo cobertura (hedge) o como “buen estado del mundo”: los activos que pagan más en esos escenarios ofrecen menor retorno esperado, porque proporcionan protección en estados relativamente favorables para el inversor.

Primas de Riesgo Estimadas

<i>Factor</i>	λ
<i>US5Y</i>	<i>-0.0123186800</i>
<i>PAYEMS</i>	<i>0.0045532160</i>
<i>CPIAUCSL</i>	<i>0.0853255800</i>
<i>FEDFUNDS</i>	<i>-0.0065794800</i>
<i>CPIAUCSL_L2</i>	<i>-0.0063145360</i>
<i>VIXCLS</i>	<i>0.0025219630</i>
<i>T10Y3M_L2</i>	<i>-0.0055936390</i>
<i>UMCSENT</i>	<i>0.0328811500</i>
<i>GDPC1_L2</i>	<i>-0.0536765400</i>
<i>FED_UPPER_L1</i>	<i>-0.0026803800</i>
<i>PAYEMS_L2</i>	<i>-0.0000071340</i>
<i>INDPRO_L3</i>	<i>0.0000008920</i>
<i>CPIAUCSL_L3</i>	<i>-0.0001054540</i>
<i>T5YIFR</i>	<i>0.0002362197</i>
<i>PAYEMS_L1</i>	<i>0.0030114890</i>
<i>OAS_L1</i>	<i>0.0018481890</i>
<i>GDPC1_L1</i>	<i>-0.0014878460</i>
<i>VIXCLS_L1</i>	<i>-0.0000008543</i>
<i>OAS</i>	<i>0.0006455936</i>
<i>UNRATE</i>	<i>-0.0000080868</i>
<i>M2SL</i>	<i>-0.0007310678</i>

Tabla 11. Primas de Riesgo Estimadas

4.5.3 Interpretación del Vector λ

A partir de las estimaciones, se observa que algunos factores presentan primas de riesgo positivas, mientras que otros muestran primas negativas, lo que permite caracterizar la forma en que el mercado valora distintos estados macroeconómicos.

Factores con λ positivo: riesgos que el mercado remunera, los factores con $\lambda > 0$ son aquellos para los cuales una mayor covarianza con el factor se asocia a mayores retornos esperados. En este grupo destacan:

CPIAUCSL (inflación, $\lambda > 0$)

La inflación actual es el factor mejor remunerado del sistema. Esto sugiere que los activos que logran mantener o incrementar sus retornos en entornos de inflación elevada obtienen una compensación adicional. Desde la perspectiva de equilibrio, los inversores exigen una prima por exposición a shocks inflacionarios, coherente con la literatura que identifica la inflación como un factor de riesgo central para los activos financieros.

UMCSENT (sentimiento del consumidor, $\lambda > 0$)

La prima positiva indica que los activos que se benefician en fases de mejora del sentimiento del consumidor ofrecen mayores retornos esperados. Dado que la confianza del consumidor es un indicador adelantado de la demanda agregada, una λ positiva sugiere que el mercado remunera la exposición a escenarios de expansión del gasto privado, típicamente asociados con mayores ventas y utilidades corporativas.

PAYEMS y PAYEMS_L1 (empleo, $\lambda > 0$)

Tanto el empleo contemporáneo como el rezagado presentan primas positivas, lo que indica que los activos que se mueven más intensamente con la fortaleza del mercado laboral ofrecen mayores retornos esperados. Esto es coherente con el rol del empleo como indicador del ciclo económico: una mayor sensibilidad al empleo implica mayor exposición a fases expansivas, que el mercado remunera con una prima de crecimiento.

OAS y OAS_L1 (spreads crediticios, $\lambda > 0$)

Los spreads de crédito reflejan el costo adicional por asumir riesgo corporativo. Las λ positivas indican que los activos que covarían positivamente con la ampliación de spreads ofrecen mayores primas de riesgo. En otras palabras, la exposición a escenarios de tensión crediticia (mayor OAS) es remunerada: los activos que sufren en estos episodios necesitan ofrecer un retorno esperado más alto para ser sostenidos en portafolio.

T5YIFR (inflación esperada a 5 años, $\lambda > 0$)

Aunque su coeficiente es pequeño, el signo positivo sugiere que la exposición a mayor inflación esperada de mediano plazo se asocia con una ligera prima adicional. Esto es consistente con episodios en los que la inflación esperada aumenta junto con expectativas de crecimiento nominal, generando una compensación marginal para activos expuestos a ese entorno.

Factores con λ negativo

Los factores con $\lambda < 0$ se comportan como variables asociadas a estados relativamente favorables o como proxies de “buenas noticias” macroeconómicas, de modo que los activos que pagan más en esos escenarios tienden a ofrecer retornos esperados menores. Entre ellos destacan:

US5Y (rendimiento del Tesoro a 5 años, $\lambda < 0$)

Un λ negativo indica que la exposición a aumentos en la tasa libre de riesgo de mediano plazo se asocia con menores retornos esperados en el margen. Este resultado es coherente con la interpretación de los Treasury como activos refugio: los activos que se comportan bien cuando suben las tasas libres tienden a ser vistos como coberturas, y por ello ofrecen una prima más baja en equilibrio.

FEDFUNDS y FED UPPER_L1 (política monetaria, $\lambda < 0$)

El endurecimiento monetario presenta λ negativas. En este caso, la exposición a estos factores funciona como cobertura frente a shocks de política monetaria, es decir, incrementos inesperados o más agresivos de lo anticipado en las tasas de interés. Los activos que resisten mejor estos episodios no necesitan ofrecer retornos esperados elevados, mientras que los que sufren en estos escenarios son los que capturan la prima positiva del sistema.

T10Y3M_L2 (pendiente de la curva, $\lambda < 0$)

Un λ negativo sobre el diferencial 10Y–3M implica que la exposición a una curva más empinada se asocia con menor prima de riesgo, mientras que los escenarios de aplanamiento o inversión de la curva históricamente vinculados a recesiones son estados “malos” que el mercado remunera. Es decir, los activos que tienden a comportarse peor cuando la curva se invierte necesitan ofrecer mayores retornos esperados ex ante, lo que se refleja en una λ negativa sobre el spread.

GDPC1_L1 y GDPC1_L2 (PIB real, $\lambda < 0$)

Las primas negativas asociadas al PIB real rezagado reflejan que la exposición a buenos tiempos económicos (mayor crecimiento) se comporta como un estado favorable para el inversor: los activos que pagan más en entornos de alto crecimiento ofrecen, en promedio, menor prima de riesgo, al actuar como cobertura frente a fluctuaciones en el bienestar agregado. Esta lógica es consistente con los modelos de valoración basados en consumo, donde los activos que rinden bien en “buenos estados” tienden a tener retornos esperados más bajos.

CPIAUCSL_L2 y CPIAUCSL_L3 (inflación rezagada, $\lambda < 0$)

A diferencia de la inflación contemporánea, la inflación acumulada rezagada presenta λ negativas de baja magnitud. Esto sugiere que una vez internalizada por el sistema, la inflación pasada no constituye un riesgo adicional sino un estado transitorio ya incorporado

M2SL (agregado monetario, $\lambda < 0$)

La prima negativa frente a M2SL indica que períodos de mayor liquidez monetaria se asocian a un entorno de menor prima de riesgo agregada: cuando la liquidez es abundante, las condiciones financieras son más favorables, los activos expuestos positivamente a la liquidez no necesitan ofrecer un retorno esperado elevado.

UNRATE, VIXCLS_L1, INDPRO_L3, PAYEMS_L2 ($\lambda < 0$)

Sus magnitudes reducidas indican que estos factores no constituyen riesgos sistemáticos remunerados una vez que el modelo controla por los factores dominantes (inflación, spreads y ciclo monetario).

4.5.4 Integración de λ y Betas para la Formación del Retorno de Equilibrio

El retorno de equilibrio de cada clase de activo surge de la interacción entre la exposición sistemática (β) y la prima de riesgo (λ). Factores cuyos signos coinciden entre β y λ incrementan el retorno esperado; aquellos con signos opuestos lo reducen. Asimismo, factores con β negativo y λ negativo funcionan como coberturas naturales.

Esta estructura explica las diferencias observadas entre clases de activos: la liquidez presenta retornos bajos, la renta variable retornos moderados condicionados por inflación y volatilidad, y activos como crédito HY y REITs presentan retornos elevados debido a su exposición a factores altamente remunerados.

5. Retornos de Equilibrio del Modelo APT–XGBoost

La presente sección presenta los retornos de equilibrio estimados para cada activo del portafolio, calculados a partir de la interacción entre las exposiciones sistemáticas (β) obtenidas mediante el modelo XGBoost y las primas de riesgo (λ) derivadas del e APT. La estimación del retorno esperado permite evaluar la contribución marginal de cada factor de riesgo y proporciona una medida coherente para la proyección de retornos de largo plazo dentro de un portafolio multiactivo.

5.1 Retornos estimados

Retornos del modelo

<i>Activo</i>	<i>Retorno nominal anual</i>
<i>Liquidez (IND_LQ_USA_BILLIM_USD_001)</i>	<i>0.83%</i>
<i>Renta Variable EE. UU. (S&P 500)</i>	<i>8.32%</i>
<i>Acciones Globales (ACWI)</i>	<i>7.20%</i>
<i>Bonos del Tesoro de EE. UU. (Deuda USA)</i>	<i>2.83%</i>
<i>Crédito High Yield (Deuda HY)</i>	<i>6.24%</i>
<i>Deuda de Mercados Emergentes (Deuda EM)</i>	<i>5.03%</i>
<i>REITs – Activos Alternativos</i>	<i>6.27%</i>

Tabla 12. Retornos del modelo

5.2 Interpretación General

Los resultados muestran una estructura de retornos que es consistente con los principios fundamentales de valoración de activos y con la evidencia empírica documentada en mercados desarrollados y emergentes. En particular, la jerarquía estimada sigue el patrón ampliamente reconocido en la literatura, donde los activos con menor riesgo sistemático presentan las menores tasas de retorno esperado, mientras que aquellos con mayor exposición a factores macroeconómicos remunerados alcanzan retornos superiores. Esta organización es coherente con la relación riesgo – retorno planteada desde los modelos pioneros de Sharpe (1964) y Lintner (1965), y posteriormente reafirmada en extensiones multifactoriales como el APT de Ross (1976).

En línea con esta teoría, la estructura observada ubica a los instrumentos de liquidez en la parte inferior de la jerarquía, seguidos por la deuda soberana de economías desarrolladas, los activos con mayor riesgo crediticio o exposición cíclica como mercados emergentes, High Yield y REITs, y finalmente la renta variable, que representa el activo con mayor sensibilidad a los factores macro-financieros que el modelo identifica como remunerados. Este comportamiento coincide con la evidencia empírica que muestra que activos más expuestos al ciclo económico, a la inflación y al riesgo financiero sistemático exhiben retornos esperados más altos (Fama & French, 1989; Cochrane, 2005)

5.3 Interpretación por Clase de Activo

Liquidez – IND_LQ_USA_BILL1M_USD_001 (0.83%)

El retorno esperado más bajo corresponde al activo de liquidez, lo cual es coherente con su mínima exposición a factores sistemáticamente remunerados. Su estructura de betas muestra sensibilidad negativa frente a inflación rezagada y señales de política monetaria contractiva, factores asociados a λ negativos en el modelo. En consecuencia, la liquidez actúa principalmente como activo refugio de corto plazo y no como generador de primas de riesgo.

Renta Variable EE. UU. – S&P 500 (8.32%)

La renta variable estadounidense se ubica como la clase de activo con mayor retorno esperado. Ello refleja su exposición a factores vinculados al ciclo económico —como sentimiento del consumidor, pendiente de la curva y expectativas futuras de crecimiento— que presentan primas de riesgo positivas. Este comportamiento es consistente con la evidencia que muestra que los activos accionarios capturan la mayor proporción del riesgo sistemático asociado al crecimiento económico y a las utilidades corporativas.

Acciones Globales – ACWI (7.20%)

El retorno esperado del ACWI es ligeramente inferior al del S&P 500. Aunque incorpora la exposición a factores globales y estadounidenses con primas remuneradas, la diversificación geográfica reduce su sensibilidad, particularmente a los factores domésticos de EE. UU. que presentan las λ más elevadas. Como resultado, su perfil de retorno refleja un equilibrio entre exposición al ciclo global y amortiguación del riesgo idiosincrático.

Bonos del Tesoro de EE. UU. – Deuda USA (2.83%)

Los bonos soberanos de EE. UU. presentan retornos moderados, coherentes con su menor

riesgo sistemático. La combinación de betas negativas frente a política monetaria y betas positivas frente a la tasa US5Y genera un equilibrio entre efectos contractivos y normalizaciones de tasas. Dado que los factores monetarios con λ negativo dominan la ecuación, el retorno esperado se mantiene en niveles bajos, acordes con su condición de activo seguro.

Crédito High Yield – Deuda HY (6.24%)

El crédito HY obtiene un retorno relativamente elevado debido a su marcada exposición a factores remunerados, en particular aquellos asociados al riesgo crediticio (OAS) y a expectativas inflacionarias de mediano plazo (T5YIFR). Al mismo tiempo, su sensibilidad negativa a inflación realizada y condiciones laborales rezagadas incrementa la prima exigida por el mercado, lo que se traduce en un mayor retorno esperado.

Deuda de Mercados Emergentes – EM (5.02%)

La deuda EM presenta retornos intermedios que reflejan tanto la compensación por riesgo soberano y riesgo crediticio como su vulnerabilidad a factores estadounidenses con λ negativo, tales como el ciclo económico y la pendiente de la curva. Su exposición simultánea a factores globales remunerados y no remunerados genera un retorno balanceado entre riesgo y diversificación geográfica.

REITs – Activos Alternativos (6.27%)

Los REITs exhiben un retorno esperado elevado y relativamente estable, consistente con su naturaleza híbrida entre renta variable y crédito. Su sensibilidad positiva frente a spreads corporativos (OAS), un factor con λ positivo, y su exposición moderada a condiciones de riesgo y liquidez explican este resultado. La combinación de flujos inmobiliarios estables con comportamiento financiero procíclico los ubica en un punto intermedio dentro del espectro de riesgo – retorno.

Capítulo 5: Conclusiones y recomendaciones

5.1 Conclusiones

Los resultados obtenidos en este trabajo permiten concluir que la integración entre la Teoría de Precios por Arbitraje (APT) y el algoritmo XGBoost constituye un enfoque metodológico sólido para la estimación de retornos de equilibrio en portafolios multiactivos. A lo largo del estudio se logró demostrar que un modelo de equilibrio multifactorial puede enriquecerse mediante técnicas de aprendizaje automático sin perder consistencia teórica, permitiendo identificar de manera empírica los factores macroeconómicos que determinan la estructura de riesgo de cada activo.

El análisis de los factores seleccionados por el modelo evidencia que las fuerzas macroeconómicas tradicionales continúan siendo determinantes centrales del retorno

esperado. Variables como la inflación, el empleo, la política monetaria, la liquidez y los spreads de crédito aparecen de manera consistente como factores dominantes, lo que confirma su papel estructural en la formación de precios a lo largo de distintos segmentos del mercado. La recurrencia de estos factores en carteras de renta fija, renta variable y activos alternativos demuestra que, aun en un entorno financiero caracterizado por mayor complejidad y relaciones no lineales, los fundamentos macroeconómicos siguen siendo puntos de referencia clave para explicar los movimientos del mercado.

No obstante, los resultados del modelo también muestran que la dinámica de los retornos no está determinada únicamente por los valores observados de las variables económicas, sino por las expectativas que los inversionistas forman respecto a su trayectoria futura. En varios casos, los factores asociados a anticipaciones como el sentimiento del consumidor, la pendiente de la curva de rendimientos o la inflación esperada presentan un peso explicativo igual o incluso superior al de las variables contemporáneas. Esto sugiere que los mercados incorporan de forma prioritaria la dirección esperada de la economía, y no únicamente su estado actual, reafirmando la idea de que el componente de expectativas es un motor esencial de los precios de los activos.

Asimismo, el ejercicio permitió identificar que la sensibilidad de los activos a los factores económicos no es homogénea. Mientras que las carteras asociadas a renta fija muestran una fuerte concentración de importancia en tasas de interés e inflación, los activos de renta variable y los REIT exhiben una estructura más diversificada, donde las expectativas, la percepción de riesgo y los indicadores de ciclo tienen un peso relativamente mayor. Estas diferencias confirman la premisa multifactorial de la APT: cada activo incorpora un conjunto específico de exposiciones sistemáticas, cuya estimación es esencial para la construcción de portafolios con criterios de equilibrio.

Desde la perspectiva metodológica, el modelo planteado probó ser una herramienta eficaz para capturar relaciones que podrían pasar desapercibidas en un enfoque puramente lineal. Su capacidad para procesar múltiples rezagos, identificar interacciones no triviales y priorizar variables relevantes permitió construir una matriz de betas más representativa que la que se obtendría aplicando únicamente un modelo clásico de regresión. Sin embargo, la posterior estimación lineal de las betas y primas de riesgo garantizó que los resultados mantuvieran interpretabilidad y se alinearan con los principios fundamentales de la APT, preservando la coherencia teórica requerida en un modelo de equilibrio.

En conjunto, el trabajo logró cumplir su objetivo de desarrollar una metodología replicable y conceptualmente consistente para estimar retornos esperados de largo plazo. La evidencia muestra que este marco metodológico puede convertirse en una herramienta útil para la asignación estratégica de activos, especialmente en contextos donde las instituciones necesitan estimaciones de equilibrio que incorporen información macroeconómica de manera sistemática y no dependan exclusivamente de promedios históricos o supuestos simplificados.

La estimación de retornos esperados mediante un esquema híbrido, que integra XGBoost para la identificación de factores relevantes y un modelo APT para la estimación estructural de betas, mostró resultados consistentes bajo una estrategia de validación del 70% entrenamiento y 30% prueba, evidenciando estabilidad en los coeficientes y un desempeño superior en capacidad explicativa frente a modelos univariados. La prueba de validación confirmó que el modelo mantiene una buena generalización fuera de muestra, lo que sugiere que la metodología es viable para apoyar procesos internos de estimación de retornos en contextos con múltiples fuentes de riesgo. No obstante, los hallazgos también destacaron los retos inherentes a la implementación del APT, especialmente la sensibilidad a la multicolinealidad entre factores, la complejidad del proceso de selección de variables y la dependencia del modelo respecto a la calidad de los datos. En conjunto, los resultados indican que este enfoque tiene potencial para complementar los modelos vigentes, siempre que se acompañe de prácticas robustas de validación continua.

Finalmente, el estudio abre la puerta a futuras aplicaciones más amplias del enfoque APT–XGBoost, tanto en la gestión de portafolios institucionales como en la modelación de escenarios de largo plazo. La combinación de teoría financiera y aprendizaje automático no solo mejora la precisión en la estimación de retornos, sino que ofrece un camino para integrar grandes volúmenes de información económica con criterios de equilibrio, preservando al mismo tiempo la transparencia y trazabilidad que exige la toma de decisiones en entornos profesionales.

5.2 Recomendaciones:

Realizar backtesting periódico: Implementar un esquema formal de backtesting que permita evaluar el comportamiento del modelo frente a periodos históricos no utilizados en el entrenamiento. Esto permitirá medir la estabilidad de las betas, la precisión predictiva y el ajuste del modelo en distintos ciclos económicos.

Comparar sistemáticamente contra el modelo CAPM corporativo: Se recomienda establecer una rutina de comparación entre los retornos estimados por el modelo propuesto y el modelo CAPM actualmente utilizado por el banco, con el fin de determinar si la incorporación de múltiples factores mejora la estimación del retorno de equilibrio y si aporta señales adicionales para la toma de decisiones en riesgo y portafolios.

Mitigar retos estructurales del APT (multicolinealidad y dimensionalidad)

Dado que el APT es sensible a la multicolinealidad, se sugiere incorporar procedimientos adicionales como VIF, PCA o regularización (Ridge/Lasso) para estabilizar las betas.

Asimismo, se recomienda monitorear la sobre parametrización del modelo y restringir el número de factores según criterios de significancia estadística y relevancia económica.

Entrenamiento iterativo para mejorar precisión: El modelo debe actualizarse de forma periódica con datos recientes. La naturaleza iterativa del XGBoost permite que, al reentrenar el modelo, este ajuste dinámicamente la relevancia de los factores y capture cambios en el comportamiento de mercado, mejorando la precisión del retorno estimado.

Ampliar el set de variables (macro, financieras y de sentimiento): Incorporar nuevas variables predictoras, por ejemplo, liquidez, riesgo país, indicadores sectoriales, sentimiento de mercado o estructuras de tasas, permitiría mejorar la robustez del modelo y su capacidad para reflejar fuentes de riesgo adicionales que no están incorporadas en el CAPM.

Establecer un procedimiento institucional para la selección de factores debido a la importancia crítica de escoger adecuadamente las variables del APT, se recomienda formalizar un protocolo que combine criterios estadísticos (importancia de variables, significancia, estabilidad temporal) con criterios económicos (validez conceptual y relevancia para el portafolio del banco).

Desarrollar una capa de gobernanza y explicabilidad del modelo para facilitar la adopción interna, se sugiere complementar el modelo con paneles de interpretación (SHAP values, análisis de sensibilidad y exposición a factores) que permitan explicar de forma clara cómo y por qué el modelo genera sus predicciones.

Capítulo 6. Referencias

- Abadi, M. S. N., Piri, H., & Sotudeh, R. (2025). Comparative Analysis of XGBoost Algorithm and Linear Regression in Predicting the Trend of Investor Overreaction. *Business, Marketing, and Finance Open*, 2(2), 125–137. <https://doi.org/10.61838/BMFOPEN.2.2.12>
- Abdi, F., Abolmakarem, S., & Yazdi, A. K. (2025). Predictive Stock Selection: A Hybrid RF-CNN XGBoost Model Integrated with Dynamic Adaptive Index and Stepwise Elimination Techniques. *Journal of OperationsIntelligence*, 3(1), 267–282. <https://jopi-journal.org/index.php/jopi/article/view/45/41>
- Bessler, W., Opfer, H., & Wolff, D. (2017). Multi-asset portfolio optimization and out-of-sample performance: an evaluation of Black–Litterman, mean-variance, and naïve diversification approaches. *European Journal of Finance*, 23(1), 1–30. <https://doi.org/10.1080/1351847X.2014.953699>
- Black, F., & Litterman, R. (1992). Global Portfolio Optimization. *Financial Analysts Journal*, 48(5), 28–43. <https://doi.org/10.2469/FAJ.V48.N5.28>
- Burmeister, E., & Wall, K. D. (1986). The Arbitrage Pricing Theory and Macroeconomic Factor Measures. *Financial Review*, 21(1), 1–20. <https://doi.org/https://doi.org/10.1111/j.1540-6288.1986.tb01103.x>
- Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57–82. <https://doi.org/10.1111/J.1540-6261.1997.TB03808.X>;REQUESTEDJOURNAL:JOURNAL:15406261;ISSUE:ISSUE:DOI

- Chambers, D., Dimson, E., Ilmanen, A., & Rintamäki, P. (2024). Long-Run Asset Returns. *The Annual Review of Financial Economics*, 16. <https://doi.org/10.1146/annurev-financial-082123>
- Chen, J. M. (2021). The Capital Asset Pricing Model. *Encyclopedia*, 1(3), 915–933. <https://doi.org/10.3390/encyclopedia1030070>
- Chen, N.-F., Roll, R., & Ross, S. (1986). Economic Forces and the Stock Market. *The Journal of Business*, 59(3), 383–403. <https://doi.org/10.1086/296344>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>;CSUBTYPE:STRING:CONFERENCE
- Cochrane, J. H. (2011). Presidential Address: Discount Rates. *Journal of Finance*, 66(4), 1047–1108. <https://doi.org/https://doi.org/10.1111/j.1540-6261.2011.01671.x>
- Comisión Europea. (2008). *Sistema de Cuentas Nacionales*. <https://unstats.un.org/unsd/nationalaccount/docs/sna2008spanish.pdf>
- Connor, G., & Korajczyk, R. A. (1988). Risk and Return in an Equilibrium Apt: Application of a New Test Methodology. *Journal of Financial Economics*, 21(2). <https://papers.ssrn.com/abstract=1917406>
- Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427–465. <https://doi.org/10.1111/J.1540-6261.1992.TB04398.X>;PAGE:STRING:ARTICLE/CHAPTER
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56. [https://doi.org/10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5)
- Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, 81(3), 607–636. <https://www.jstor.org/stable/1831028>
- Feng, G., Giglio, S., & Xiu, D. (2020). Taming the Factor Zoo: A Test of New Factors. *The Journal of Finance*, 75(3), 1327–1370. <https://doi.org/10.1111/JOFI.12883>
- Gladstone, S., Madhavan, A., Rana, A., & Ang, A. (2021). Macro Factor Model: Application to Liquid Private Portfolios. *The Journal of Portfolio Management*, 47(5), 72–90. <https://doi.org/10.3905/JPM.2021.1.231>
- Gökgöz, F. (2009). Mean variance optimization via factor models in the emerging markets: Evidence on the Istanbul Stock Exchange. *Investment Management and Financial Innovations*, 6(3), 43–53.
- Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5), 2223–2273. <https://doi.org/10.1093/RFS/HHAA009>
- Hyndman, R., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts. <https://research.monash.edu/en/publications/forecasting-principles-and-practice-3/>
- Illmer, S. J., & Marty, W. (2007). Return decomposition of absolute-performance multi-asset class portfolios. *Financial Markets and Portfolio Management*, 21(1), 121–134. <https://doi.org/10.1007/s11408-006-0028-0>
- Ilmanen, Antti. (2011). *Expected returns : an investor's guide to harvesting market rewards* (1st ed.). Wiley.
- James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, R. (2023). *An Introduction to Statistical Learning with Applications in Python* (1st ed.). Springer Cham. <https://doi.org/https://doi.org/10.1007/978-3-031-38747-0>
- Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1), 37. <https://doi.org/10.2307/1924119>
- Ma, R., Marshall, B. R., Nguyen, N. H., & Visaltanachoti, N. (2024). Estimating Long-Term Expected Returns. *Financial Analysts Journal*, 80(4), 134–154. <https://doi.org/10.1080/0015198X.2024.2358737>
- Mainie, S. (2015). *La historia de la inversión basada en factores*. <https://www.spglobal.com/spdji/es/documents/research/research-the-story-of-factor-based-investing-spanish.pdf#:~:text=en%20la%20Teor%C3%ADa%20de>
- Malandri, L., Xing, F. Z., Orsenigo, C., Vercellis, C., & Cambria, E. (2018). Public Mood-Driven Asset Allocation: the Importance of Financial Sentiment in Portfolio Management. *Cognitive Computation*, 10(6), 1167–1176. <https://doi.org/10.1007/S12559-018-9609-2>
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91. <https://doi.org/10.2307/2975974>
- McQuarrie, E. F. (2024). Stocks for the Long Run? Sometimes Yes, Sometimes No. *Financial Analysts Journal*, 80(1), 12–28. <https://doi.org/10.1080/0015198X.2023.2268556>
- Michaud, R. O. (1989). The Markowitz Optimization Enigma: Is “Optimized” Optimal? *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.2387669>
- Min, L., Dong, J., Liu, D., & Kong, X. (2021). A black-litterman portfolio selection model with investor opinions generating from machine learning algorithms. *Engineering Letters*, 29(2), 710–721.
- Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), 783. <https://doi.org/10.2307/1910098>

- Roll, R. (1977). A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. *Journal of Financial Economics*, 4(2), 129–176. [https://doi.org/10.1016/0304-405X\(77\)90009-5](https://doi.org/10.1016/0304-405X(77)90009-5)
- Roll, R., & Ross, S. A. (1980). An Empirical Investigation of the Arbitrage Pricing Theory. *The Journal of Finance*, 35(5), 1073–1103. <https://doi.org/https://doi.org/10.1111/j.1540-6288.1986.tb01103.x>
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3), 341–360. [https://doi.org/10.1016/0022-0531\(76\)90046-6](https://doi.org/10.1016/0022-0531(76)90046-6)
- Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3), 442. <https://doi.org/10.2307/2977928>
- Shyu, S., Jeng, Y., Ton, W. H., Lee, K., & Chuang, H. M. (2006). Taiwan multi-factor model construction: equity market neutral strategies application. *Managerial Finance*, 32(11), 915–947. <https://doi.org/10.1108/03074350610703858>
- Stock, J., & Watson, M. (1999). Business Cycle Fluctuations in US Macroeconomic Time Series. In J. B. Taylor & M. Woodford (Eds.), *Handbook of Macroeconomics* (Vol. 1, pp. 1–56). E. <https://scholar.harvard.edu/files/stock/files/businesscyclefluctuationsusmacroeconomic.pdf>
- Sürek, R. E., & Lau, W. Y. (2025). A refined methodological approach: Long-term stock market forecasting with XGBoost. *Journal of Intelligent Systems*, 34(1). <https://doi.org/10.1515/JISYS-2025-0027/XML>
- Thalassinos, E. I., Khan, N., Ahmed, S., Zada, H., & Ihsan, A. (2023). A Comparison of Competing Asset Pricing Models: Empirical Evidence from Pakistan. *Risks*, 11(4). <https://doi.org/10.3390/risks11040065>
- The Conference Board. (2001). *Business Cycle Indicators Handbook* (J. Norwood, A. S. Blinder, B. Bosworth, K. Eickhoff, R. Eisner, E. R. Fiedler, G. Fosler, P. Klein, S. J. Landefeld, R. H. McGuckin, G. H. Moore, M. P. Niemira, J. Popkin, C. A. Sims, J. Stock, M. Watson, & V. Zarnowitz, Eds.). The Conference Board. https://www.conference-board.org/pdf_free/economics/bci/BCI-Handbook.pdf
- Tobin, J. (1958). Liquidity preference as behavior towards risk. *Review of Economic Studies*, 25(2), 65–86. <https://doi.org/10.2307/2296205>
- Harvey, C. R. (1988). *The real term structure and consumption growth*. *Journal of Financial Economics*, 22(2), 305–333. [https://doi.org/10.1016/0304-405X\(88\)90070-0](https://doi.org/10.1016/0304-405X(88)90070-0)
- Altman, E. I. (1989). Measuring corporate bond mortality and performance. *The Journal of Finance*, 44(4), 909–922.
- Altman, E. I. (2000). Predicting financial distress of companies: Revisiting the Z-score and ZETA® models. New York University, Working Paper.
- Altman, E. I. (2012). Default recovery rates and LGD in credit risk modeling and practice. *Economic Notes*, 41(2), 111–129.
- Bank for International Settlements. (2019). Reflation trades and corporate bond spreads. BIS Markets Report.
- Chen, L., Collin-Dufresne, P., & Goldstein, R. S. (2007). On the relation between credit spread puzzles and fundamental factors. *Review of Financial Studies*, 22(9), 3367–3409.
- Gilchrist, S., & Zakrajšek, S. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, 102(4), 1692–1720.
- Huang, J., & Huang, M. (2012). How much of the corporate–treasury yield spread is due to credit risk? *Review of Asset Pricing Studies*, 2(2), 153–202.
- Krishnamurthy, A., & Vissing-Jorgensen, A. (2012). The aggregate demand for Treasury debt. *Journal of Political Economy*, 120(2), 233–267
- Enders, W. (2014). *Applied econometric time series* (4th ed.). Wiley

Anexos

ANEXO A. Base de Datos Histórica — base_precios.xlsx

El archivo base_precios.xlsx reúne la totalidad de la información utilizada para la estimación del modelo APT–XGBoost. Este conjunto de datos constituye la base empírica del estudio y consolida series históricas de precios, retornos e indicadores macroeconómicos para Estados Unidos y Colombia en diferentes frecuencias (diaria, mensual y trimestral), cubriendo el periodo 2007–2025.