



Universidad Nacional del Litoral

Facultad de Ciencias Económicas

Licenciatura en Economía

La Estructura Factorial del Mercado de Capitales Argentino

*Análisis empírico del modelo factorial de rendimientos financieros,
factores latentes y macroeconómicos en el periodo 2018–2025*

Autor:

Alexis Deppeler

Director y codirector:

Rogelio Villanueva y Gustavo Cabaña

Febrero 2026

Agradecimientos

Quiero expresar mi sincero agradecimiento a mi director, Rogelio Villanueva, y a mi codirector, Gustavo Cabaña, por su desinteresada ayuda, dedicación y compromiso con esta investigación. Sus orientaciones fueron fundamentales para garantizar el desarrollo y la solidez de esta tesis. Aun así, cualquier error u omisión que pudiera persistir es de mi exclusiva responsabilidad. Asimismo, agradezco a los miembros del jurado por su tiempo y disposición para evaluar este trabajo.

Deseo expresar también mi profundo agradecimiento a todas las personas que conocí y me acompañaron a lo largo de este trayecto académico. De manera especial, agradezco a mi pareja por su apoyo constante; a mis amigos, por las valiosas experiencias compartidas; y a mi familia, por su acompañamiento incondicional a lo largo de estos años de formación.

Alexis Deppeler

Índice general

1. Introducción	7
2. Marco teórico	12
2.1. El riesgo en el análisis financiero	12
2.2. El modelo factorial de rendimientos financieros	14
2.3. El análisis econométrico con rendimientos financieros	17
2.4. El análisis factorial	19
3. Revisión de antecedentes	29
3.1. Delimitación epistemológica del modelo factorial	29
3.2. La propuesta de Chen et al.	32
3.3. Antecedentes en la región	33
3.4. Evidencia empírica utilizando análisis factorial	35
4. Análisis econométrico	39
4.1. Factores latentes	43
4.1.1. Pruebas de adecuación	45
4.1.2. Determinación del número de factores a retener	46
4.2. Factores macroeconómicos	49
4.2.1. Construcción de factores macroeconómicos	52
4.2.2. Estimación de modelos de regresión	58
5. Reflexiones finales	69
Apéndices	71
A. Estadísticas descriptivas y resultados complementarios	71

B. Endogeneidad de las variables dummies	76
---	-----------

Bibliografía	78
---------------------	-----------

Índice de figuras

4.1. distribución del volumen monetario mensual real por activo a precios constantes de diciembre de 2025, período 01.2018–12.2025, en escala logarítmica base 10.	42
4.2. comparación entre autovalores observados y simulados	48
4.3. evolución temporal del rendimiento de la cartera, periodo 2018.01–2025.12 . . .	51
4.4. evolución temporal de los residuos de MCO del Modelo 2, periodo 2018.01–2025.12	64

Índice de cuadros

3.1. principales resultados de estudios que estiman factores macroeconómicos	37
3.2. principales resultados de estudios que estiman factores latentes	38
4.1. medida de adecuación muestral individual KMO	45
4.2. autovalores de la matriz de correlaciones	46
4.3. varianza explicada por factor en el modelo AFE de cuatro factores	49
4.4. fuentes de datos para variables independientes	50
4.5. Regresión del modelo MCO — especificación saturada	59
4.6. regresión del modelo MCO — especificación depurada	60
4.7. regresión del modelo MCO — especificación depurada alternativa	61
A.1. estadísticas descriptivas de los rendimientos	72
A.2. test ADF sobre rendimientos diarios	73
A.3. estadísticas descriptivas de los factores observables estimados	74

A.4. test ADF sobre factores macroeconómicos	75
B.1. regresión MCO de residuos sobre dummies	76

Abstract

This work focuses on the sources of systematic risk, a central concern in financial economics, as they shape the comovement of asset returns and mediate the transmission of macroeconomic shocks to capital markets. In Argentina—an economy characterized by high volatility, recurrent regime shifts, and persistent macroeconomic imbalances—identifying the underlying structure of risk factors is particularly relevant for both investors and policymakers.

This study examines the factorial structure of systematic risk in the Argentine equity market using a portfolio, based on the constituents of the S&P Merval index, between 2018 and 2025. The objective is to uncover the latent factors driving the common dynamics of stock returns and to evaluate the extent to which observable macroeconomic variables contribute to these movements.

Methodologically, the analysis proceeds in two stages. First, an Exploratory Factor Analysis is applied to daily returns to extract the statistical factors underlying market comovement. Second, multiple regression models relate monthly returns to unexpected macroeconomic changes in order to assess the influence of observable determinants.

The findings provide a clearer characterization of systematic risk in Argentina and offer a foundation for future research on the interaction between financial markets and macroeconomic fundamentals.

Keywords: systematic risk; factor models; exploratory factor analysis; emerging markets; asset pricing; macro-finance

JEL: G12, G15, E44, C38

Capítulo 1

Introducción

El denominado modelo factorial constituye una ampliamente utilizada herramienta en las finanzas modernas. El primero en emplearlo, aunque sin hacer mención explícita de dicho modelo, fue Sharpe (1964) al desarrollar el CAPM¹; en este caso, lo que utilizó, fue un modelo factorial de un sólo factor. Posteriormente, y ahora sí incorporándolo en forma explícita al proceso de formulación de su teoría, Ross (1976) lo emplea para proponer la Teoría de Precios de Arbitraje (APT, por sus siglas en inglés); en esta oportunidad, se utilizó el modelo factorial con múltiples factores.

Desde la publicación de los citados trabajos, la trascendente función del modelo factorial, ha sido la de explicar el riesgo sistemático, al permitir descomponerlo en función de sus determinantes, y así poder identificarlos. Por esta razón, en lo que podríamos denominar una segunda etapa de la vigencia de este modelo factorial, lo que ha desvelado a los investigadores, es la contrastación empírica del mismo. En este marco, es donde entra en escena el Análisis Factorial (AF), un modelo estadístico que aporta herramientas para realizar dichas contrastaciones empíricas, bajo la premisa de comprobar que detrás de la volatilidad del rendimiento de los activos, están presentes ciertos factores latentes.

A partir del trabajo de Roll y Ross (1980), que constituyó el primer test empírico que buscó mostrar la validez del empleo del modelo factorial (utilizando el AF), se han venido desarrollando investigaciones, tanto en países desarrollados como en economías emergentes, que buscan alcanzar el mismo objetivo; es decir, mostrar que es válido descomponer el rendimiento de los activos en función de sus determinantes, y así poder explicar las variables generadoras del llamado riesgo sistemático.

¹Modelo de valoración de activos de capital, por sus siglas en inglés

El presente trabajo se incluye en esta temática, dado que busca testear la validez del modelo factorial para descomponer el rendimiento de las acciones que cotizan en el mercado argentino en función de un conjunto de factores que permitan explicar el riesgo sistemático. No obstante, no es tan ambicioso como la mayoría de los trabajos analizados en los antecedentes. En dichos trabajos, mostrar la validez del modelo factorial es solo un paso para contrastar empíricamente la validez de un modelo de equilibrio, tal como podría ser la APT. La investigación que aquí se presenta, es algo más modesta que sus precedentes, dado que se concentra en testear la validez del modelo factorial. Además, los tests que se realizan, no tienen la generalidad que se habría deseado que tengan, dado que se limitan a emplear un reducido número de factores.

Estas limitaciones auto impuestas tienen sus razones. En primer lugar, el mercado financiero argentino debe calificarse como emergente; en consecuencia, múltiples factores afectan la eficiencia del mismo, generando problemas prácticos de difícil solución a la hora de realizar las estimaciones econométricas de los parámetros. En segundo lugar, también dificulta la estimación econométrica el hecho de trabajar con datos generados en el periodo analizado en la economía argentina. Las dificultades adicionales que esto provoca se relacionan con la elevada volatilidad en los precios de los activos, combinada con frecuentes cambios de régimen macroeconómico, que también contribuyen a generarla.

El trabajo que aquí se presenta, siendo específico, busca contrastar la validez de utilizar el modelo factorial para explicar los determinantes del riesgo sistemático en el mercado accionario argentino, considerando los activos que integraron el índice S&P Merval durante el tercer trimestre de 2025 y que a lo largo del período comprendido entre enero de 2018 y diciembre de 2025 han mantenido su liquidez. Además, se asume que dichos determinantes solo pueden ser factores macroeconómicos o financieros. Se considera que alcanzar este objetivo, además de realizar la tarea central de verificar si el modelo es adecuado para los datos que se disponen, lo cual se realiza en el Capítulo 4, también requiere dar respuesta a tres preguntas claves.

La primera de esas preguntas es: ¿En qué medida la variabilidad conjunta de los rendimientos de estas acciones sugiere la presencia de dimensiones subyacentes y cuántas serían necesarias para representarla mediante una estructura factorial común? La respuesta a esta pregunta nos indicara si estamos habilitados a utilizar el modelo factorial para intentar explicitar los

factores que determinan el riesgo sistemático de estas acciones. Para poder contestarla, lo que hacemos en la Sección 4.1 es emplear el AF para comprobar la existencia de factores latentes que determinan la variabilidad de dichos rendimientos. El resultado de esta primera parte del análisis nos habilita a utilizar el modelo factorial para tratar de indagar cuáles son esos factores.

La segunda de esas preguntas, se genera cuando ya estamos abocados al trabajo econométrico, y la formulamos así: ¿Cuáles son esos factores que definen el llamado riesgo sistemático? Esta pregunta es imprescindible contestar, porque el análisis factorial solo permite identificar la dimensión o cantidad de factores latentes, pero no identifica las variables de naturaleza económica que definen esos factores. Cabe aclarar, que la APT de Ross (1976) no identifica con nombres propios a dichos factores, tampoco supone que sean macroeconómicos específicos, y tampoco exige que sean directamente observables; para este autor, lo que es importante para formular su teoría es que dichos factores existen y son sistemáticos, es decir, afectan a todos los activos. Los factores comunes que se proponen para el análisis son los que se incluyen en la Sección 4.2. Como ya se dijo, no se responde esta pregunta con la más amplia generalidad que pueda desearse, dado que nos limitamos a considerar un reducido número de factores cuya selección siguen las recomendaciones de Chen et al. (1986).

La tercera de estas preguntas es la siguiente: ¿De qué manera los factores que determinan el rendimiento de los activos financieros deben construirse mediante procedimientos estadísticos?. De manera preliminar, se adelanta que dichos factores no son directamente observables, por lo que las variables utilizadas en la ecuación factorial deben ser construidas. Esto se debe a que el modelo factorial postula que son las innovaciones, shocks o sorpresas los determinantes que generan la variabilidad en el rendimiento de los activos financieros y, en consecuencia, el riesgo sistemático. El procedimiento para la construcción de estas innovaciones se desarrolla en la Sección 4.2.1, siguiendo el enfoque propuesto por Chen et al. (1986).

La pertinencia y relevancia de la investigación se apoya en tres dimensiones complementarias. En primer lugar, el aporte teórico-conceptual: El estudio contribuye a la comprensión del riesgo sistemático en mercados emergentes, al articular el papel de variables macroeconómicas observables con la identificación de factores latentes comunes, evitando imponer supuestos teóricos restrictivos. En segundo lugar, el enfoque metodológico: la utilización conjunta de

técnicas econométricas y el modelo factorial de rendimientos financieros ofrece una estrategia robusta para el análisis del riesgo sistemático. Esta complementariedad metodológica permite contrastar la incidencia de factores macroeconómicos con la existencia de estructuras latentes, reforzando la consistencia de los resultados y favoreciendo su replicabilidad en estudios sobre otros mercados emergentes. Finalmente, la relevancia práctica: los resultados de este trabajo son de utilidad para diversos actores del sistema financiero, ya que podrían facilitar a inversores y gestores de cartera a identificar los factores que explican el comportamiento conjunto de los activos y diseñar estrategias de diversificación más efectivas; ofrecen a los hacedores de política económica herramientas para analizar los vínculos entre variables macroeconómicas y dinámica bursátil; y contribuyen a la literatura aportando evidencia empírica sobre el comportamiento del riesgo sistemático en un mercado de capitales poco desarrollado, como el argentino.

El período analizado resulta particularmente interesante debido a la sucesión de episodios de inestabilidad económica y financiera que lo caracterizan, incluyendo procesos inflacionarios persistentes, cambios de régimen político-económico y volatilidad cambiaria. Este recorte temporal permite captar la sensibilidad del mercado ante distintos contextos macroeconómicos y constituye un escenario propicio para evaluar la interacción entre factores macroeconómicos y latentes.

No obstante debe señalarse, que en este trabajo el análisis se realiza sobre el período completo. Dado que la desagregación en subperíodos implicaría trabajar con muestras de tamaño reducido, lo que comprometería la potencia estadística de las inferencias econométricas. En consecuencia, se prioriza un enfoque de análisis agregado que permita identificar regularidades en la estructura de comovimiento y en la relación entre rendimientos y factores, dentro del horizonte temporal considerado.

La investigación se organiza en cinco capítulos. Primero se introduce la problemática, junto con los objetivos y las hipótesis que orientan el estudio. Luego se presenta el marco conceptual y la evidencia empírica, revisando los fundamentos del modelo factorial, la noción de riesgo sistemático y el rol de los factores macroeconómicos y latentes. Finalmente se desarrolla el enfoque empírico, detallando la recolección de datos, la extracción de factores latentes y la construcción de factores macroeconómicos, así como la evaluación del modelo factorial. Por

último, se exponen las reflexiones finales, donde se sintetizan los resultados, se discuten sus implicancias para el estudio del riesgo sistemático en mercados emergentes y se proponen líneas de investigación futura.

Capítulo 2

Marco teórico

El marco teórico de esta investigación se construye a partir de la noción de riesgo como elemento central en la determinación de los rendimientos financieros y de la idea de que los precios de los activos reflejan, de manera agregada, la exposición a fuentes comunes de riesgo sistemático. Bajo esta perspectiva, los modelos factoriales constituyen una herramienta fundamental para describir la estructura de covariación de los retornos, en este capítulo se desarrollan los fundamentos conceptuales necesarios para comprender dicha estructura.

2.1 El riesgo en el análisis financiero

Si nos guiamos por lo que podríamos llamar el enfoque clásico de riesgo de un activo, debemos definirlo como la variabilidad de sus rendimientos alrededor de su valor esperado. Por esta razón, si la variable aleatoria R_j representa el rendimiento del activo j , el riesgo total de dicho activo se mide, tal como lo propusiera Markowitz (1952), por la varianza: $\text{Var}(R_j) = \sigma_j^2$, o alternativamente por la desviación estándar: $\sqrt{\text{Var}(R_j)} = \sigma_j$, que también mide la volatilidad del activo.

La visión clásica del riesgo, que asocia el concepto con la variabilidad del rendimiento, trata de la misma forma tanto a las ganancias como a las pérdidas. No obstante, otro enfoque también muy utilizado, que se genera a partir de reconocer que los inversores son adversos al riesgo, vincula el concepto con la posibilidad de obtener pérdidas, dando lugar a medidas que se concentran en la cola inferior de la distribución de probabilidades que se le asigne a R_j ; es decir, dan más importancia a la posibilidad de obtener pérdida que ganancias. El primer autor en vincular el riesgo a la posibilidad de obtener pérdidas, fue Roy (1952), quien propone el criterio *safety first*, que consistió en una regla de decisión para construir carteras donde se evitan

las grandes pérdidas. Medidas modernas que siguen este enfoque, y son muy utilizadas en las finanzas actuales, son *Value at Risk* (VaR) y *Expected Shortfall* (ES).

En la teoría financiera clásica, el riesgo total que es medido por la varianza del rendimiento del activo, se puede descomponer en dos partes. Una de ellas, es el llamado riesgo sistemático o de mercado; sus dos características son: (i) afecta a todos los activos del sistema, dado que todos ellos están positivamente correlacionados; esta correlación existe porque los rendimientos de los activos dependen de ciertos factores comunes, tales como: La actividad industrial, tasa de interés, inflación, etc. Precisamente, esta investigación indaga sobre esos factores comunes que afectan los activos en Argentina; (ii) no es posible eliminarlo por la diversificación. La otra de las partes, es el llamado riesgo propio o idiosincrático, cuyas características son: (a) afecta sólo al activo cuya rentabilidad se está evaluando, dado que se asocia a decisiones de la empresa o eventos internos o propios de la misma; (b) puede eliminarse o reducirse por medio de la diversificación.

El modelo factorial que es el objeto de investigación en este trabajo, está mucho más asociado a la visión clásica de riesgo. Como vemos más adelante, en el modelo factorial, el riesgo total se cuantifica por medio de la varianza; además, cuando el mismo se descompone, el riesgo de mercado o sistemático se explica vía covarianzas.

Si hablamos de riesgo financiero, y queremos evaluarlo en una economía emergente como Argentina, es importante recordar que presenta características estructurales que lo distinguen de las economías desarrolladas, tal como lo señala Damodaran (2012). En primer lugar, el riesgo sistemático es más alto, dado que el riesgo macroeconómico es mayor que en países desarrollados, y también entra en escena el riesgo político; el riesgo macroeconómico es mayor, porque la inflación es más alta y volátil, existe riesgo cambiario, y además hay mayor inestabilidad fiscal y monetaria; el riesgo político es considerable, debido a frecuentes cambios regulatorios, incertidumbre institucional, inseguridad jurídica, y riesgo soberano; todos estos aspectos constituyen riesgo sistemático, dado que afectan a todos los activos y sus efectos en la rentabilidad no pueden neutralizarse con la diversificación de la cartera. Una medida aproximada de cuánto mayor es el riesgo sistemático en las economías emergentes viene dada por el riesgo país. En segundo lugar, como los mercados financieros son pequeños y poco profundos,

la posibilidad de diversificar para eliminar el riesgo propio es limitada. En tercer lugar, los mercados (en general) no suelen ser gaussianos; es decir, no es adecuado utilizar la distribución normal para modelar el rendimiento de los activos, dado que las distribuciones empíricas muestran: exceso de curtosis, lo que suele denominarse colas pesadas (distribuciones leptocúrticas), y genera la posibilidad que eventos extremos deban ser considerados; y asimetría, lo cual genera que la varianza no sea una medida completa de riesgo, sino que tiende a subestimarlos.

2.2 El modelo factorial de rendimientos financieros

El modelo factorial de rendimientos financieros argumenta, que los rendimientos de los activos financieros se ven afectados por un conjunto de factores comunes, es decir, por un conjunto de factores que afectan a todos los activos financieros. Por esta razón, el rendimiento de cualquier activo financiero j en el período t , denotado por $R_{j,t}$, se modela de la siguiente forma:

$$R_{j,t} = \alpha_j + \beta_{j,1} \cdot f_{1,t} + \beta_{j,2} \cdot f_{2,t} + \cdots + \beta_{j,K} \cdot f_{K,t} + \varepsilon_{j,t}, \quad (2.1)$$

donde las variables aleatorias $f_k, \forall k = 1, 2, \dots, K$, representan a los K factores comunes que afectan el rendimiento de todos los activos; el subíndice t identifica la observación, asumiendo que la misma es parte de una serie de tiempo. Estas variables se definen como innovaciones, es decir, como desviaciones de su valor esperado. De este modo, se supone que $E(f_k) = 0$, para todo $k = 1, 2, \dots, K$. Adicionalmente, también se supone que $E(f_k \cdot f_h) = 0, \forall h \neq k$, es decir, dichos factores no están correlacionados entre sí. Las $\beta_{j,k}$, que multiplican a los factores f_k , y deben estimarse para todo $k = 1, 2, \dots, K$, representan las sensibilidades que tiene el rendimiento del activo j a cada uno de dichos factores.

La Ecuación (2.1) también representa el modelo de regresión lineal múltiple con K regresores. Por esta razón, la variable $\varepsilon_{j,t}$ no es más que el término de perturbación estocástica, que recoge el efecto de la variabilidad no explicada por el modelo, pero, que en conjunto afectan el rendimiento de los activos. De este modo, también se supone que $E(\varepsilon_{j,t}) = 0$. Además, se asumen homoscedasticidad, es decir, $E(\varepsilon_{j,t}^2) = \sigma_j^2, \forall t$ y ausencia de autocorrelación, es decir,

$E(\varepsilon_{j,t} \cdot \varepsilon_{j,s}) = 0$, para todo $t \neq s$. Con estos supuestos, nos aseguramos que el residuo se comporte como ruido blanco, es decir, todos los factores que debían incluirse en el modelo para explicar la variación sistemática de la variable dependiente, ya están incluidos.

En el modelo de regresión lineal clásico, la matriz F de regresores, que contiene como elementos a los valores $f_{k,t} \forall k = 1, 2, \dots, K$ y para todo $t = 1, 2, \dots, T$, se supone no estocástica. No obstante, en el análisis financiero, los factores que determinan el rendimiento de los activos, son variables económicas que deben considerarse aleatorias; en consecuencia, no es aceptable suponer que la matriz F sea no estocástica. Para evitar que esto afecte las deseables propiedades del modelo de regresión, debemos agregar el siguiente supuesto: $E(f_{j,k} \cdot \varepsilon_{j,t}) = 0$ para $k = 1, 2, \dots, K$ y $t = 1, 2, \dots, T$; esto indica que los residuos no están correlacionados con los factores, permitiendo que dichos factores sean exógenos respecto del término de perturbación estocástica.

Por último, al modelo factorial que aquí se desarrolla, se le incorpora el siguiente supuesto: $E(\varepsilon_{j,t} \cdot \varepsilon_{i,t}) = 0, \forall j \neq i$, lo que implica que los términos de error de las regresiones de dos activos distintos no están correlacionados. Este supuesto se realiza para permitir que el riesgo total del activo sólo pueda separarse en dos componentes: el riesgo de mercado y el riesgo propio. Cabe aclarar que, si se permitiera que $\varepsilon_{j,t}$ y $\varepsilon_{i,t}$ estuvieran correlacionados, estaríamos aceptando que existen fuentes de riesgo comunes a ambos activos que no son capturadas por los factores f_k . Por ejemplo, si los activos j e i pertenecieran al mismo sector o industria, un factor exclusivamente sectorial podría hacer que los errores de ambos activos estén correlacionados. Como el modelo no desea lidiar con un supuesto riesgo sectorial que se suma a los riesgos de mercado y propio, se agrega este supuesto. Si aplicamos el operador E a la Ecuación (2.1), dados los supuestos antes mencionados, vemos que: $E(R_j) = \alpha_j$. Este resultado permite reescribir la ecuación del modelo de la siguiente forma:

$$R_{j,t} = E(R_j) + \beta_{j,1} \cdot f_{1,t} + \beta_{j,2} \cdot f_{2,t} + \dots + \beta_{j,K} \cdot f_{K,t} + \varepsilon_{j,t}. \quad (2.2)$$

Si aceptamos que la Ecuación (2.2) explica el rendimiento del activo j , podemos argumentar que éste se puede desagregar en dos componentes: el esperado, que viene dado por $E(R_j)$, y el no esperado, que se integra con todos los restantes términos que suman en el lado

derecho de la ecuación. El componente esperado define lo que el analista puede prever que ocurrirá en el futuro con el rendimiento del activo. El componente no esperado, tal como su nombre lo indica, determina la parte del rendimiento del activo que el analista no puede prever a la hora de realizar el análisis; en consecuencia, podemos interpretarlo como el generador de riesgo.

A su vez, ese componente generador de riesgo se puede subdividir en dos partes: una de ellas está integrada con todos los sumandos donde aparece un factor f_k ; la otra la integra el componente de error $\varepsilon_{j,t}$. Dado que se supone que esos factores f_k afectan a todos los activos, la parte del componente generador de riesgo que ellos determinan, pasa a definir el riesgo sistemático; es decir, el riesgo asociado al desempeño global de la economía. Por otra parte, en relación con el componente de error, como se explicó antes, se hacen dos supuestos claves: (i) $E(\varepsilon_{j,t} \cdot \varepsilon_{j,s}) = 0$, para todas las observaciones donde $t \neq s$, lo cual nos señala que los factores f_k incluidos capturan la variabilidad sistemática del modelo; y (ii) $E(\varepsilon_{j,t} \cdot \varepsilon_{i,t}) = 0, \forall j \neq i$, lo cual nos informa que los términos de error de dos activos distintos no están correlacionados y nos garantiza que no existen factores de tipo sectorial que puedan afectar el rendimiento de los activos. La presencia de estos dos supuestos nos permite interpretar al término $\varepsilon_{j,t}$, como el componente generador de riesgo propio; es decir, el riesgo asociado a cada activo de forma individual.

Dada la interpretación que podemos darle a (2.2), si consideramos que la varianza del rendimiento de un activo mide la cantidad total de riesgo asociada a ese activo, podemos descomponer dicha cantidad de riesgo total, en sus dos componentes: el riesgo de mercado, y el riesgo propio. Para lograr esta descomposición, partimos de la siguiente definición de varianza:

$$\sigma_{R_j}^2 = E[R_j - E(R_j)]^2. \quad (2.3)$$

Sustituyendo (2.2) en (2.3), luego de operar algebraicamente y dados los supuestos que hemos especificado para la ecuación factorial, se obtiene finalmente la siguiente expresión de la varianza del activo j :

$$\sigma_{R_j}^2 = \beta_{j1}^2 \cdot \sigma_{f_1}^2 + \beta_{j2}^2 \cdot \sigma_{f_2}^2 + \dots + \beta_{jK}^2 \cdot \sigma_{f_K}^2 + \sigma_{\varepsilon_j}^2. \quad (2.4)$$

La expresión (2.4) descompone el riesgo total del activo j en sus dos componentes: el riesgo de mercado y el propio. Los primeros K términos del lado derecho de (2.4), que se integran con las betas del activo respecto de cada factor, multiplicadas por las varianzas de dichos factores, miden la cantidad de riesgo sistemático asociada a este activo j , mientras que el último término, $\sigma_{\varepsilon_j}^2$, mide la cantidad de riesgo propio que se asocia a ese activo.

A partir de la Ecuación (2.2) y aceptando todos los supuestos que se han presentado, Ross (1976) deriva el modelo de valoración de activos financieros bajo ausencia de arbitraje, conocido por sus siglas APT. No obstante, el presente trabajo no se involucra en testear la validez de dicho modelo, sino sólo la ecuación factorial que le sirve de base. Por esta razón, no se desarrollan las explicaciones de dicha teoría.

2.3 El análisis econométrico con rendimientos financieros

Como ya se anticipó parcialmente en el punto anterior al explicitar los supuestos¹ de la Ecuación (2.1), el modelo que se utiliza para explicar el riesgo sistemático en función de sus factores determinantes es el modelo de regresión lineal, que se estima bajo el marco del modelo lineal clásico, utilizando Mínimos Cuadrados Ordinarios (MCO), con datos que constituyen series de tiempo.

Se indica que el modelo es de regresión lineal, dado que los parámetros β que se estiman se los supone estables, y la inferencia se realiza utilizando estadísticos tales como t y F ; no obstante, no es el modelo clásico tal como lo define Gujarati y Porter (2010), dado que: (i) la matriz de datos es estocástica, aunque se supone que los residuos no están correlacionados con los factores, tal como se explicó antes, para evitar que esto afecte las deseables propiedades del modelo de regresión lineal clásico; (ii) la matriz de datos es estimada, no observada; esto es así, dado que lo que afecta al rendimiento de los activos son las sorpresas, shocks, o innovaciones, que se producen en relación con los valores que se espera tomen las variables explicativas; el problema de la elección y estimación de dichas variables, se ve en detalle más adelante; (iii) los errores pueden ser heterocedásticos; para evitar que estos potenciales problemas afecten la

¹Los supuestos teóricos asociados a los factores se asumen válidos a nivel poblacional. No obstante, al trabajar con datos muestrales es posible que alguno de ellos no se cumpla exactamente, ya que se está analizando una muestra y no la población.

inferencia, porque los estadísticos t y F sean incorrectos y ello genere valores-p poco confiables, se utilizan errores estándar robustos que vuelven válidos a los test asintóticamente.

Una cuestión no menor, es que los datos que se utilizan son series de tiempo; es decir, constituyen una realización particular de un proceso estocástico. En consecuencia, si a partir de la serie observada queremos hacer inferencia sobre el proceso subyacente, el mismo debe ser estacionario. Cuando un proceso es no estacionario, las series que se obtienen de él también serán no estacionarias; en este caso, sólo podemos obtener conclusiones sobre su comportamiento, para el período al que corresponden los datos; es decir, cada conjunto de datos hace referencia a un caso particular, no pudiéndose generalizar las conclusiones.

En finanzas, los datos que se obtienen para un análisis como el propuesto consisten típicamente en series de precios diarios. Empíricamente se sabe que estas series de precios suelen ser no estacionarias, ya que, en general, siguen un proceso integrado de orden uno, $I(1)$. No obstante, su primera diferencia suele ser estacionaria. Esta es la razón por la cual, en la práctica financiera, se trabaja con rendimientos. Si bien los rendimientos no constituyen estrictamente una primera diferencia, representan cambios relativos en los precios.

En este trabajo se utilizaron rendimientos continuos o logarítmicos. Adicionalmente, para evitar problemas asociados a la no estacionariedad, se realizó el test de Dickey-Fuller aumentado (ADF), una prueba diseñada para contrastar la hipótesis nula de que una serie X_t es integrada de orden uno, es decir, $X_t \sim I(1)$. Este test se aplicó a todas las series utilizadas.

Otro problema que potencialmente se enfrenta, cuando se utilizan series de tiempo en las regresiones que se proponen realizar, es que aparezca cambio estructural entre la variable regresada y las regresoras. Con cambio estructural se quiere indicar, que los valores de los parámetros no permanezcan constantes a lo largo de todo el período que comprenden los datos. Para testear este potencial problema de validez del análisis econométrico, en la Sección [4.2.2](#) se realiza el conocido test de Bai-Perron de quiebres estructurales, afortunadamente, el mismo no estaría presente en los datos.

2.4 El análisis factorial

El Análisis Factorial (AF) es una técnica estadística multivariante, que busca identificar si la variabilidad conjunta o covariación de un conjunto de variables observadas puede ser explicada a partir de un pequeño número de variables latentes o no observadas, las cuales se denominan factores. Lo que intenta es dar respuesta a la pregunta: ¿puedo explicar la estructura de covarianzas de ese conjunto de variables observadas con unos pocos factores latentes? Para el análisis que nos proponemos es pertinente, ya que si ese conjunto de variables observadas son rentabilidades de acciones que cotizan en bolsa, su estructura de covarianzas contiene el mencionado riesgo sistemático; consecuentemente, el AF es la técnica estadística ideal, para verificar si estamos habilitados a utilizar el modelo factorial para explicitar las variables que determinan el riesgo sistemático de dichas acciones.

Tal como lo indica Peña (2002), el modelo de AF establece que, si se observa una muestra \mathbf{x} de elementos de una población, de dimensiones $p \times 1$, este vector de datos se genera por la siguiente relación:

$$\mathbf{x} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f} + \mathbf{u}, \quad (2.5)$$

donde \mathbf{f} es el vector de factores latentes o no observados de tamaño $m \times 1$; $\boldsymbol{\Lambda}$ es una matriz de $p \times m$, con $m < p$, llamada matriz de cargas, que se integra con constantes desconocidas que, en el AF, se denominan saturaciones factoriales, valores de saturación o cargas factoriales, y representan los coeficientes que describen cómo los factores \mathbf{f} afectan a las variables observadas; $\boldsymbol{\mu}$ es un vector de $p \times 1$ de constantes; y \mathbf{u} es el vector de $p \times 1$ de perturbaciones no observadas, que recoge los efectos de todas las variables distintas a los factores \mathbf{f} que influyen sobre \mathbf{x} . Los supuestos que usualmente se utilizan en relación con el AF, son de dos tipos: por un lado están los estructurales o esenciales, que son necesarios para su formulación básica; y por otro, los distribucionales, útiles para ciertos métodos de estimación o inferencia.

Los supuestos esenciales que se utilizan en el AF explicitado en la Ecuación (2.5) son los siguientes:

1. $E(\mathbf{f}) = \mathbf{0}$. Esto indica que todos los factores comunes tienen media cero.

2. $E(\mathbf{u}) = \mathbf{0}$. Esto señala que todas las perturbaciones no observadas tienen media cero.
3. $\text{Cov}(\mathbf{f}, \mathbf{u}) = E(\mathbf{f}\mathbf{u}^\top) = \mathbf{0}$. Es decir, las covarianzas entre los componentes del vector \mathbf{f} y del vector \mathbf{u} son cero.
4. Si $\Psi = \text{Cov}(\mathbf{u}) = E(\mathbf{u}\mathbf{u}^\top)$, entonces Ψ es una matriz diagonal de dimensiones $p \times p$. Esto significa que las perturbaciones u_i y u_j asociadas a x_i y x_j , con $i \neq j$, no están correlacionadas.
5. Si $\Phi = \text{Cov}(\mathbf{f}) = E(\mathbf{f}\mathbf{f}^\top)$, entonces Φ es una matriz diagonal de dimensiones $m \times m$. Así, los factores comunes no están correlacionados entre sí, de manera análoga al modelo factorial de la Ecuación (2.1).

De los anteriores supuestos, se tiene que $\boldsymbol{\mu} = E(\mathbf{x})$, el vector de medias. Si operamos algebraicamente en la Ecuación (2.5), pasando al lado izquierdo el vector de valores esperados de \mathbf{x} , multiplicando por la derecha ambos miembros de la ecuación por el vector de factores latentes transpuesto, \mathbf{f}^\top , y aplicando el operador E en ambos miembros, llegamos a:

$$E[(\mathbf{x} - E(\mathbf{x}))\mathbf{f}^\top] = E(\Lambda\mathbf{f}\mathbf{f}^\top + \mathbf{u}\mathbf{f}^\top). \quad (2.6)$$

Usando la linealidad del operador E en la Ecuación (2.6), dados los supuestos que hemos especificado, se tiene:

$$E[(\mathbf{x} - E(\mathbf{x}))\mathbf{f}^\top] = \Lambda E(\mathbf{f}\mathbf{f}^\top) = \Lambda\Phi. \quad (2.7)$$

Multiplicando ambos términos de la Ecuación (2.7) por Φ^{-1} (la inversa de la matriz de varianzas de los factores latentes), arribamos a la siguiente propiedad del modelo factorial con factores ortogonales que se está desarrollando:

$$E[(\mathbf{x} - E(\mathbf{x}))\mathbf{f}^\top] \Phi^{-1} = \Lambda. \quad (2.8)$$

La expresión (2.8) nos está señalando que, como asumimos que los factores latentes no están correlacionados, los elementos de la matriz de cargas Λ son coeficientes beta que se podrían obtener con una hipotética regresión, utilizando el modelo de regresión lineal múltiple.

En efecto, si nos concentramos en el lado izquierdo de la Ecuación (2.8),

$$E[(\mathbf{x} - E(\mathbf{x})) \mathbf{f}^T] = \text{Cov}(\mathbf{x}, \mathbf{f}),$$

además, como

$$\Phi = \text{diag}(\sigma_{f_1}^2, \sigma_{f_2}^2, \dots, \sigma_{f_m}^2),$$

su inversa será

$$\Phi^{-1} = \text{diag}\left(\frac{1}{\sigma_{f_1}^2}, \frac{1}{\sigma_{f_2}^2}, \dots, \frac{1}{\sigma_{f_m}^2}\right).$$

Consecuentemente, el producto matricial que dicho lado izquierdo expresa nos señala que cada elemento de la matriz Λ , que se denota λ_{il} , para todo $i = 1, 2, \dots, p$ y para todo $l = 1, 2, \dots, m$, será:

$$\lambda_{il} = \frac{\text{Cov}(x_i, f_l)}{\sigma_{f_l}^2} = \beta_{il}.$$

Este resultado es muy importante para el análisis que aquí se propone, porque permite conectar el AF que aquí se desarrolla, con la ecuación factorial (2.2) con rendimientos financieros, dado que en aquel análisis también se supone factores explicativos no correlacionados

$$E(f_k \cdot f_h) = 0, \quad \forall h \neq k.$$

Este resultado nos permite decir que los coeficientes beta que se obtienen cuando se resuelve la ecuación factorial utilizando el modelo de regresión lineal MCO bien podrían ser estimaciones de los elementos λ_{il} de la matriz Λ . No obstante, esta conclusión hay que interpretarla con cautela, dado que la Ecuación (2.8) sólo muestra una equivalencia algebraica que permite asociar los parámetros del modelo factorial, que en principio están indeterminados, con los coeficientes de una hipotética regresión que utiliza como variables independientes factores latentes no observables.

En el AF, ni la matriz de cargas Λ ni el vector de factores \mathbf{f} son observables, lo cual plantea un problema de indeterminación. En consecuencia, no puede concluirse de la Ecuación (2.8) que dicha matriz Λ sea única y que sus elementos sean idénticos a los coeficientes beta de una regresión lineal. La interpretación correcta de la Ecuación (2.8) podría enunciarse así: si,

gracias a una restricción adicional impuesta al modelo factorial, es posible elegir una realización concreta de los factores \mathbf{f} que permita que éstos puedan ser tratados como variables observables, asumiendo además que son ortogonales, entonces, si usamos esos factores para realizar una regresión lineal, los coeficientes beta que se obtengan coincidirán con los elementos que integran la matriz de cargas Λ .

Si volvemos a operar algebraicamente en la Ecuación (2.5), pasando al lado izquierdo el vector de valores esperados de \mathbf{x} , multiplicando por la derecha ambos miembros de la ecuación por el vector $[\mathbf{x} - E(\mathbf{x})]^\top$, y aplicando el operador E en ambos miembros, luego de considerar los supuestos antes establecidos para el modelo factorial, llegamos a:

$$E[(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^\top] = \Lambda E(\mathbf{f}\mathbf{f}^\top)\Lambda^\top + E(\mathbf{u}\mathbf{u}^\top). \quad (2.9)$$

El lado izquierdo de la Ecuación (2.9) es la matriz de varianzas y covarianzas de la muestra de datos observados, $V(\mathbf{x})$, y en el lado derecho tenemos que $E(\mathbf{f}\mathbf{f}^\top) = \Phi$ y $E(\mathbf{u}\mathbf{u}^\top) = \Psi$, ambas matrices diagonales. Entonces, al sustituir se tiene:

$$\mathbf{V} = \Lambda\Phi\Lambda^\top + \Psi. \quad (2.10)$$

La Ecuación (2.10) muestra la propiedad fundamental del modelo factorial que estamos considerando, al señalar que la matriz de varianzas y covarianzas de las variables observadas puede descomponerse como la suma de dos partes. La primera, dada por $\Lambda\Phi\Lambda^\top$, muestra la matriz de covarianzas comunes; es decir, la parte común de la variabilidad del conjunto de las variables incluidas en el vector \mathbf{x} . La otra parte, dada por la matriz diagonal Ψ , contiene la parte específica de la variabilidad de cada variable incluida en el vector \mathbf{x} .

Si, aprovechando la descomposición que muestra la Ecuación (2.10), definimos la matriz de comunalidades como

$$\mathbf{H} = \Lambda\Phi\Lambda^\top,$$

entonces las varianzas de las variables observadas pueden expresarse de la siguiente forma:

$$\sigma_{x_i}^2 = \sum_{l=1}^m \lambda_{il}^2 \sigma_{f_l}^2 + \sigma_{u_i}^2, \quad \forall i = 1, 2, \dots, p. \quad (2.11)$$

La Ecuación (2.11) nos indica que las varianzas de las x_i que integran el vector \mathbf{x} de variables observadas pueden descomponerse en dos partes. La primera de ellas, h_i^2 , denominada comunalidad, y formada por la sumatoria del primer sumando del lado derecho:

$$h_i^2 = \sum_{l=1}^m \lambda_{il}^2 \sigma_{f_l}^2,$$

representa la parte de la varianza de x_i que es explicada por los factores comunes. La segunda parte, ψ_i , denominada especificidad, y formada por el segundo sumando del lado derecho:

$$\psi_i = \sigma_{u_i}^2,$$

representa la parte de la varianza de x_i que es específica de dicha variable. Esta descomposición de las varianzas de las variables incluidas en \mathbf{x} también muestra la estrecha relación entre el modelo y la ecuación factorial (2.2) con rendimientos financieros. En efecto, si por lo dicho en párrafos precedentes aceptamos que las betas que se obtienen con el modelo de regresión lineal son estimaciones de los elementos λ_{il} de la matriz Λ , la Ecuación (2.11) es análoga a la Ecuación (2.4), que descompone el riesgo total del activo j , medido por la varianza de su rendimiento, en sus dos componentes: el riesgo de mercado y el propio. En efecto, este es el motivo por el cual se ha decidido utilizar AF en esta investigación en lugar del conocido Análisis de Componentes Principales (ACP), dado que a diferencia de este último, el AF descompone la varianza de cada variable siguiendo la Ecuación (2.11) mientras que el ACP se limita a descomponer la varianza total del conjunto de variables observadas.

Como explica Peña (2002), el AF fue impulsado por Karl Pearson y Charles Spearman en los años 30 del siglo pasado, quienes tenían interés en conocer las dimensiones de la inteligencia humana; por esta razón, muchos de sus avances y usos están en el área de la psicometría. Consecuentemente, en el AF clásico, la matriz de datos \mathbf{X} se forma con n vectores \mathbf{x} de dimensión p , $\mathbf{x} \in \mathbb{R}^p$, donde p es el número de variables observadas, y cada vector corresponde a un sujeto distinto. Ahora, si el AF se aplica para individualizar los factores latentes en la rentabilidad de las acciones, el vector \mathbf{x} estará formado por la rentabilidad de p

empresas distintas en un momento dado del tiempo, $\mathbf{x} \in \mathbb{R}^p$, y la matriz de datos \mathbf{X} se integrará con T vectores \mathbf{x} de dimensión p , uno para cada momento $t \in T$.

A partir de la matriz de datos $\mathbf{X} \in \mathbb{R}^{T \times p}$, cuyas filas son los vectores \mathbf{x}_t de variables observadas en cada momento $t = 1, 2, \dots, T$, se calcula la matriz de varianzas y covarianzas de la muestra, \mathbf{S} , de la siguiente forma:

$$\mathbf{S} = \frac{1}{T-1} \sum_{t=1}^T (\mathbf{x}_t - \bar{\mathbf{x}})(\mathbf{x}_t - \bar{\mathbf{x}})^\top. \quad (2.12)$$

La matriz \mathbf{S} que se obtenga será de dimensiones $p \times p$ e integrada con los elementos s_{ij} , donde: si $i = j$, tendremos $s_{ii} = \sigma_{x_i}^2$, y si $i \neq j$, tendremos $s_{ij} = \text{cov}(x_i, x_j)$. Esta matriz constituye el insumo fundamental del AF, ya que todo el análisis que se realiza gira en torno a su descomposición de la forma que se explicitó en la Ecuación (2.10).

El AF que nos proponemos realizar es el llamado exploratorio, ya que se realiza sin imponer hipótesis a priori sobre el número de factores o la forma en que se relacionan los datos sobre las variables con esos factores (estructura de cargas); es decir, se deja que los datos hablen. Este tipo de análisis es pertinente para nuestro objetivo, dado que el uso del AF en esta investigación sólo está destinado a estimar la dimensión del riesgo común o sistemático.

Los pasos requeridos para realizar el Análisis Factorial Exploratorio (AFE) que se propone son los siguientes: 1) análisis de la matriz de correlación; 2) extracción de factores; y 3) determinación del número de factores. En un AF más completo, correspondería incorporar un cuarto paso relativo a la rotación de factores; sin embargo, en esta investigación sólo se incluyen los tres pasos citados. La rotación de factores se realiza dado que la solución del modelo factorial no es única; lo que con ella se intenta es reorientar los factores estimados para obtener una estructura más simple de interpretar (rotación del espacio vectorial); se utiliza cuando el análisis tiene por finalidad avanzar más allá de la simple identificación del número de factores latentes que explican la variabilidad conjunta o covariación del conjunto de rentabilidades observadas. Como nuestro objetivo es usarlo como herramienta estadística para individualizar el número de factores, la rotación no es necesaria y el análisis puede considerarse completo sin ella.

El primer paso, como se dijo, es analizar la matriz de correlación. Para poder cumplirlo, antes se debe calcular dicha matriz, que simbolizamos \mathbf{R} . Ésta se obtiene estandarizando la matriz \mathbf{S} por medio de las desviaciones estándar que correspondan. Para esta tarea, se define la matriz

$$\mathbf{D} = \text{diag}(\sqrt{s_{11}}, \sqrt{s_{22}}, \dots, \sqrt{s_{pp}}),$$

y con ella se realiza el siguiente producto matricial:

$$\mathbf{R} = \mathbf{D}^{-1}\mathbf{S}\mathbf{D}^{-1} \quad (2.13)$$

Una vez obtenida \mathbf{R} , el siguiente paso es analizarla. Un requisito clave que debe cumplirse para que el AF tenga sentido es que las variables \mathbf{x} deben estar altamente intercorrelacionadas; la pregunta que debemos contestar es: ¿son lo suficientemente altas las correlaciones como para que las variables \mathbf{x} puedan ser explicadas por factores comunes? Una forma muy conocida de responder a esta pregunta es examinando \mathbf{R} mediante el test de esfericidad de Bartlett; sin embargo, este test no se utiliza aquí dado que demanda que al vector \mathbf{x} pueda asignársele una distribución normal multivariante, cosa que no se hace en este trabajo. Las razones que hacen que no se incorporen supuestos distribucionales son tanto teóricas como empíricas; desde el punto de vista teórico, el trabajar con la ecuación factorial (2.2) de rendimientos financieros no demanda este tipo de supuestos; desde el punto de vista práctico, el hecho de trabajar con datos generados por la economía argentina en los últimos años hace cuestionar que se pueda asumir una distribución normal multivariante al vector \mathbf{x} , dado los shocks generados por los frecuentes cambios de régimen macroeconómico; más adelante se le realiza a los datos el test de normalidad de Mardia, comprobándose la hipótesis. En su lugar, utilizamos el test de Kaiser-Meyer-Olkin (KMO), el cual constituye una alternativa más robusta, dado que no requiere normalidad del vector \mathbf{x} .

El KMO se basa en comparar las correlaciones simples, que son los elementos r_{ij} de la matriz de correlación \mathbf{R} , con las correlaciones parciales, que se obtienen a partir de los

elementos de \mathbf{R}^{-1} , p_{ij} , con la siguiente fórmula:

$$p_{ij} = -\frac{r_{ij}}{\sqrt{r_{ii}} \sqrt{r_{jj}}},$$

es decir, son los elementos normalizados de \mathbf{R}^{-1} . Las correlaciones parciales entre x_i y x_j miden la asociación de estas variables una vez eliminada la influencia lineal de todas las demás variables; consecuentemente, si las correlaciones parciales son pequeñas, las correlaciones observadas se deben a factores comunes. El índice KMO se construye de la siguiente forma:

$$\text{KMO} = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} p_{ij}^2}. \quad (2.14)$$

El resultado obtenido con la Ecuación (2.14) se evalúa con la siguiente regla práctica. Si $\text{KMO} < 0,50$, se considera que los datos no son adecuados para el AF; para que se consideren buenos, $\text{KMO} > 0,70$.

El índice KMO puede evaluarse tanto a nivel global como a nivel individual de cada activo; en ambos casos se usa la Ecuación (2.14), difiriendo únicamente en las correlaciones que se suman. El KMO global mide la adecuación global de los datos para realizar AF. El KMO individual se utiliza para detectar activos problemáticos y justificar su exclusión del análisis.

Comprobado que los datos son adecuados para el AF, pasamos a la extracción de factores. La misma consiste en aplicar un procedimiento que permita una identificación preliminar de la estructura de factores latentes que explica la covariación entre las variables. En esta investigación se utiliza el método de ejes principales (PAF) (*Principal Axis Factoring*, por sus siglas en inglés); el objetivo de este método es aislar una estimación de la parte común de la matriz de varianzas y covarianzas de \mathbf{x} , $\mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^T$, a partir de comunalidades estimadas. Los pasos para realizar este trabajo son los siguientes:

- Como no se conoce $\mathbf{\Lambda}$, se parte haciendo una estimación de las comunalidades h_i^2 , suponiendo que ésta es igual al coeficiente de determinación de una regresión múltiple de la variable x_i , R_i^2 , respecto de todas las otras variables que integran el vector \mathbf{x} .

- Se construye la matriz reducida \mathbf{S}^* , reemplazando los valores de la diagonal principal de \mathbf{S} por dichas comunalidades estimadas $h_i^2 = R_i^2$. Lo que aquí se hace conceptualmente podría aproximarse diciendo que $\mathbf{S}^* \approx \mathbf{S} - \mathbf{\Psi}$.
- Se diagonaliza la matriz reducida haciendo:

$$\mathbf{Q}^\top \mathbf{S}^* \mathbf{Q} = \mathbf{\Lambda}_Q,$$

donde la matriz \mathbf{Q} surge de resolver el problema de autovalores de \mathbf{S}^* ; en consecuencia,

$$\mathbf{\Lambda}_Q = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p),$$

donde λ_i son los autovalores y las columnas de \mathbf{Q} sus autovectores asociados. Como \mathbf{S}^* es simétrica, todos los autovalores y autovectores son reales y no existen problemas de cálculo.

- Se calculan las cargas factoriales iniciales, haciendo:

$$\hat{\mathbf{\Lambda}} = \mathbf{Q}_m \mathbf{\Lambda}_m^{1/2},$$

donde $\mathbf{\Lambda}_m$ es la matriz diagonal que sólo toma los autovalores que se deciden relevantes, y la matriz \mathbf{Q}_m se forma con los autovectores asociados a éstos. Estas cargas factoriales se califican como iniciales, dado que el procedimiento itera con un recálculo de comunalidades y repetición del procedimiento, hasta que se observa que dichas comunalidades convergen. En esta investigación, el procedimiento se detiene con este primer cálculo, dado que con el mismo ya se tiene el valor m de factores latentes relevantes.

El último paso de nuestro AFE consiste en la determinación del número de factores. Esto implica fijar la dimensión m del vector \mathbf{f} , lo cual determina el número de columnas de la matriz de cargas $\mathbf{\Lambda}$. El número de factores relevantes, simbolizado en nuestro modelo por m , suele determinarse a partir de diversos criterios.

En esta investigación se consideraron dos criterios habituales para evaluar el número de factores: la Regla de Kaiser y el análisis paralelo. La Regla de Kaiser propone retener aquellos factores cuyos autovalores sean mayores que uno ($\lambda > 1$). Sin embargo, este criterio presenta limitaciones conocidas, ya que puede sobreestimar el número de factores cuando el tamaño de la muestra o el número de variables es elevado, al retener componentes asociados al error, y subestimarlos en muestras pequeñas, descartando factores potencialmente relevantes.

Por ello, la decisión sobre el número de factores a extraer se basó en el análisis paralelo, el cual contrasta los autovalores observados con los obtenidos a partir de matrices aleatorias de igual dimensión generadas mediante simulaciones de Monte Carlo y procedimientos de remuestreo. Bajo este enfoque, sólo se conservan aquellos factores cuyos autovalores empíricos superan a los esperados por azar, garantizando que la varianza explicada por cada factor sea mayor que la atribuible al ruido aleatorio.

De este modo, la Regla de Kaiser se utilizó como un criterio inicial, el análisis paralelo constituyó el criterio decisorio para la determinación del número de factores retenidos. Es importante aclarar que todos los procedimientos descritos fueron implementados mediante las funciones del paquete `psych` del software R, sin realizar estos cálculos manualmente, garantizando así la reproducibilidad y consistencia de los resultados.

Capítulo 3

Revisión de antecedentes

La literatura empírica es amplia y heterogénea, tanto en términos de enfoques teóricos como de estrategias metodológicas. En este capítulo se presenta una revisión selectiva de los principales antecedentes relevantes para el objetivo de la investigación, con especial énfasis en aquellos trabajos que analizan la organización del riesgo sistemático a partir de modelos factoriales. Ya sea mediante factores macroeconómicos, contruidos a partir de variables observables, o factores latentes extraídos estadísticamente. La revisión no persigue una exposición exhaustiva ni cronológica sino que se estructura destacando: contribuciones que permiten delimitar el marco epistemológico del enfoque adoptado, discutir los problemas de identificación de factores y contextualizar la evidencia empírica disponible para mercados emergentes.

3.1 Delimitación epistemológica del modelo factorial

Aunque gran parte de la literatura financiera ha enfocado sus esfuerzos en contrastar empíricamente la APT de Ross (1976), resulta fundamental destacar que esta investigación se orienta hacia un objetivo algo menos ambicioso: testear la validez del modelo factorial en el mercado argentino, el cual constituye un insumo básico para el desarrollo y aplicación del APT, pero sin pretender someter este modelo específico a prueba directa. En otras palabras, el propósito central es comprender cómo se organiza el riesgo en este contexto particular, no determinar el precio de equilibrio financiero de los activos según los postulados del APT.

La cuestión en todo esto, naturalmente, no es si la APT es verdadera o falsa. Como todas las teorías que no son vacías, es falsa en algún grado de precisión en la contrastación: si se prueba durante suficiente tiempo, todas las teorías interesantes son

rechazadas. Más bien, la pregunta es qué aprenderemos de estas pruebas respecto de qué tan bien se desempeña la teoría en competencia con alternativas específicas. Lo que está en juego es la intuición básica de la APT: que solo la variabilidad sistemática afecta los rendimientos esperados, y este es el tema central de la teoría moderna de valoración de activos. (Roll y Ross, 1980, p. 1100) (traducción propia).

En este sentido, aún cuando el objetivo no sea contrastar las relaciones de equilibrio del modelo, los resultados de estas investigaciones ofrecen evidencia útil acerca de la presencia y relevancia de factores comunes en la explicación de los rendimientos. En consecuencia, tanto los trabajos centrados en el análisis de la estructura factorial como aquellos que abordan la APT de manera explícita constituyen antecedentes valiosos para el presente estudio.

En términos prácticos Roll y Ross (1980), en uno de los primeros trabajos empíricos que emplean el APT, realizan una adaptación de la metodología propuesta por Fama y MacBeth (1973) para estimar las primas de riesgo asociadas a los factores sistemáticos. En este enfoque, se lleva a cabo una regresión en dos etapas: en la primera etapa, se estiman las sensibilidades (betas) respecto a los factores sistemáticos mediante regresiones de los rendimientos históricos; en la segunda etapa, se realiza una regresión cruzada de los rendimientos promedio de los activos sobre las betas estimadas para obtener las primas de riesgo asociadas a cada factor. Este enfoque fue ampliamente replicado, incluso en estudios empíricos para mercados emergentes tales como Cortés y Porras (2013) en México.

No obstante, para que las inferencias realizadas en esta regresión de segunda etapa sean válidas es necesario que se satisfagan los supuestos del APT caracterizados en el Capítulo 8 del libro de Marín y Rubio (2011), entre los cuales se destacan: (i) la ausencia de oportunidades de arbitraje, (ii) la existencia de mercados completos y (iii) la posibilidad de formar carteras diversificadas que eliminen el riesgo idiosincrático. El cumplimiento de éstos es esencial para garantizar que las primas de riesgo estimadas reflejen adecuadamente la compensación por el riesgo sistemático asociado a los factores considerados. Su verificación empírica constituye un desafío interesante para futuras investigaciones, pero excede a los objetivos de esta tesis. Por ello, solo la primera etapa del proceso propuesto por Roll y Ross (1980) es relevante a nuestro propósito.

Un elemento central para comprender el vínculo entre el modelo factorial y el APT es el concepto mismo de arbitraje en el sentido planteado por Ross (1976): una oportunidad de arbitraje surge cuando es posible construir una cartera que, sin requerir inversión neta y sin asumir riesgo alguno, genera un beneficio cierto. La existencia de estas oportunidades implica una inconsistencia fundamental en los precios, ya que permitiría obtener ganancias sin exposición a innovaciones sistemáticas ni idiosincráticas. El APT se construye precisamente sobre la premisa opuesta: si los mercados permiten un grado adecuado de diversificación, entonces las oportunidades de arbitraje desaparecen o se vuelven insignificantes, lo que impone una estructura de rendimientos esperados. Bajo esta condición, cualquier activo debe ofrecer un rendimiento esperado que sea la suma de los premios por riesgo asociados a sus sensibilidades a los factores sistemáticos, pues solo estos generan riesgos no diversificables que el mercado debe remunerar (tal como vimos en la Ecuación (2.2) del capítulo anterior). En consecuencia, el principio de ausencia de arbitraje opera como el puente conceptual que convierte al modelo factorial en un modelo de valoración: no es la dinámica estadística de los retornos la que impone los precios, sino la necesidad de evitar inconsistencias que permitan beneficios libres de riesgo. De este modo, el arbitraje constituye el fundamento lógico que justifica la relevancia económica de la estructura factorial.

El principal problema gira en torno a que no se conoce explícitamente cuáles deben ser los factores que integren el modelo factorial, generando un desafío importante en su identificación empírica, como explican Chamberlain y Rothschild (1982) y Shanken (1982). Por ello técnicas estadísticas como el AF se han consolidado como métodos útiles para extraer factores latentes que expliquen la variación común de los retornos, tal como en los trabajos de Cagnetti (2001), Groenewold y Fraser (1997) y Lehmann y Modest (1985), entre otros. No obstante, es pertinente explorar los hallazgos empíricos de la literatura que aporta el enfoque macroeconómico del modelo factorial.

3.2 La propuesta de Chen et al.

Desde el punto de vista empírico estudios como Chen et al. (1986) demostraron, por medio del método de Fama y MacBeth (1973), que cambios en la producción industrial, la inflación no esperada y cambios en la estructura temporal de la tasa de interés y la insolvencia financiera, explican con robustez la dinámica conjunta de los precios de los activos para el mercado estadounidense en el período comprendido entre los años 1958 y 1984. La mayor contribución de este trabajo radica en su propuesta metodológica para la representación de las innovaciones macroeconómicas y construcción de variables. En relación a este desafío, plantea Swoboda (2002, p. 6) que “Las variables a emplear deben reunir algunas características tales como: ser fáciles de interpretar, ser relevantes a lo largo del período de análisis y a su vez explicar una proporción elevada de las variaciones de los precios de los activos financieros elegidos para el estudio.” en la Sección 4.2 lo desarrollaremos con mayor detalle.

Quienes reconocen la centralidad bibliográfica de este trabajo son Burmeister y Wall (1986, p. 3) afirmando que solo hay un artículo bien conocido que no utiliza Análisis Factorial, y es justamente el de Chen et al. Numerosos trabajos empíricos, además del de Burmeister y Wall han replicado esta construcción de factores, por ejemplo: Cagnetti (2001), McElroy y Burmeister (1988) o el de Chan (1985), este último concluye que, para el mercado americano en el período 1953–1977, los rendimientos de firmas de menor tamaño generan primas de riesgo mayores a las firmas grandes (justificado en una mayor exposición al riesgo). Y en particular el trabajo de Rayón (2016), quien lo replica para el caso de fondos de cobertura para el mercado americano en el período 2002.01–2006.06, concluyendo que solo la primera etapa de la estimación por el método de Fama y MacBeth (1973) presenta un ajuste favorable, i.e. aceptando el modelo factorial y rechazando la APT.

En mercados emergentes, López Herrera (2006) adecúa la metodología de Chen et al. (1986) al mercado mexicano entre 1991 y 2003, encontrando que los factores macroeconómicos propuestos, más una serie de variables dummies que capturan sorpresas políticas de relevancia, permiten una adecuada caracterización de la estructura de rendimientos de los activos, aunque con menor poder explicativo que en mercados desarrollados.

Karan (2010) realiza una síntesis del modelo de Chen et al. (1986) y de los factores propuestos por Fama y French (1993), los denominados *firm specific factors*. Estos últimos se basan en características propias de las empresas, como el tamaño y el valor contable sobre valor de mercado, que han demostrado tener poder explicativo adicional sobre los retornos de las acciones. La incorporación de estos factores específicos queda fuera del alcance de nuestra investigación como ya se explicó en el capítulo anterior.

3.3 Antecedentes en la región

Swoboda (2002) examina el funcionamiento del mercado de capitales argentino durante la década de 1990, destacando su crecimiento y modernización tecnológica, aunque también su limitada profundidad y baja eficiencia informativa. En este contexto el autor subraya la escasez de estudios empíricos que apliquen modelos teóricos de formación de precios en el país y la necesidad de avanzar en esquemas multifactoriales, como el propuesto por la APT que permitan capturar múltiples fuentes de riesgo sistemático. El trabajo resalta además la relevancia de estos modelos no solo para estimar rendimientos esperados, sino también para analizar la eficiencia del mercado y las inconsistencias entre precios observados y los fundamentos macroeconómicos.

En la extensión multifactorial del análisis, Swoboda (2002) emplea el APT con factores macroeconómico construidos como innovaciones no anticipadas, similares a los propuestos en nuestro trabajo. En particular retiene principalmente tres factores, el diferencial de tasas de interés en dólares de largo plazo respecto de la tasa libre de riesgo ($tycdol$), la variación del índice bursátil brasileño Bovespa ($boves$) como proxy de condiciones externas, la tasa de variación interanual de la actividad industrial ($tafiel$). Además, incluye un factor residual destinado a captar componentes no explicados por las variables observadas. Mediante regresiones MCO encuentra que los factores financieros y externos, en particular el spread de tasas en dólares y el Bovespa, junto con el factor residual presentan coeficientes estadísticamente significativos en la mayoría de los activos, mientras que el factor de actividad real muestra resultados menos robustos. Sin embargo, en la segunda etapa del procedimiento APT, los precios de riesgo asociados a estos factores no resultan significativamente distintos de cero.

En esta línea, Guerra (2012) analiza la relación entre la tasa de inflación y los rendimientos accionarios en Argentina, evaluando la capacidad de los activos financieros para funcionar como cobertura frente a la pérdida de poder adquisitivo. La autora interpreta la inflación como un factor sistemático con impacto directo sobre los precios de los activos. Los resultados de la investigación sugieren una relación negativa de corto plazo entre el rendimientos de los activos y la inflación, aunque “El estudio de largo plazo no permitió confirmar la existencia de una relación positiva entre el rendimiento de las acciones y la tasa de inflación, por lo cual resulta posible que el crecimiento conjunto del precio de las acciones y la tasa de inflación sea provocado por el componente tendencial de ambas series y no por la existencia de relación entre las variables.” (p. 22).

Fuentes et al. (2017) examinan la influencia de factores macroeconómicos sobre los rendimientos accionarios chilenos en el período 1990–2003. El estudio incorpora cuatro factores macroeconómicos —entre ellos la tasa de crecimiento de la actividad económica, los precios del cobre y del petróleo— siguiendo la metodología de Roll y Ross (1980) y McElroy y Burmeister (1988). Los resultados sugieren evidencia muestral favorable al APT. En particular, las sorpresas en la actividad económica y en los precios de los commodities muestran efectos significativos sobre los retornos, mientras que la inflación carece de impacto relevante.

Más recientemente, Pesce y Pedroni (2021) retoman el estudio del mercado argentino con un enfoque teórico-empírico que integra el rol de la inflación en un país caracterizado por alta intervención cambiara y escasa profundidad financiera. Utilizando datos de panel para 19 empresas entre 2005 y 2014, los autores estiman un modelo que incluye como variables explicativas el rendimiento libre de riesgo, el rendimiento de mercado, efectos temporales y sectoriales, además de la inflación y el tipo de cambio. Los resultados muestran efectos robustos de las primeras variables, pero no evidencia significativa para la inflación ni para el tipo de cambio.

3.4 Evidencia empírica utilizando análisis factorial

Como se mencionó en secciones anteriores, el AF como técnica de extracción de factores ha sido ampliamente empleado por la literatura especializada. En particular, el trabajo elaborado por Roll y Ross (1980) es uno de los primeros en contrastar empíricamente el APT con el método de Fama y MacBeth (1973) y lo hace justamente utilizando factores latentes obtenidos por AF. Los autores extraen factores comunes a partir de los rendimientos diarios de acciones estadounidenses entre 1962 y 1976. Los resultados indican que un modelo con cuatro factores latentes explica una proporción considerable de la varianza común de los retornos (74.3 %), y que las primas de riesgo asociadas a estos factores son estadísticamente significativas en la regresión cruzada de Fama y MacBeth (1973).

Del mismo modo, Chen (1983) utilizando frecuencias diarias para el periodo 1963–1978, demuestra que los rendimientos esperados pueden explicarse adecuadamente por las cargas factoriales asociadas a factores sistemáticos mientras que características individuales de los activos como el tamaño de la firma o la varianza propia no aportan poder explicativo adicional. El autor argumenta, en relación al problema de identificación de factores que:

“Hay dos formas, algo equivalentes, de resolver ese problema: podemos hacer supuestos y producir una teoría que especifique qué variables deberían entrar en la ecuación de precios y luego probarla; o podemos examinar el rendimiento realizado de los activos y determinar empíricamente a qué variables macroeconómicas (sugeridas por las teorías) corresponden. El APT está más en el espíritu del segundo enfoque. El cálculo de cargas factoriales en este artículo permitiría la construcción de una cartera correspondiente a cada uno de los factores comunes.”

(Chen, 1983, p. 1409, traducción propia).

En este sentido, la estrategia metodológica adoptada en esta investigación se encuentra alineada con las consideraciones de Chen (1983); por un lado, mediante teorías formuladas que especifiquen variables macroeconómicas relevantes, cómo los antecedentes discutidos en la Sección 4.2; Y por otro, utilizando técnicas de análisis factorial para extraer factores latentes directamente de los datos de retornos del mercado argentino. Esta estrategia permite capturar la

estructura subyacente del riesgo sistemático sin imponer restricciones teóricas previas sobre la naturaleza de los factores, facilitando así una comprensión más profunda de los determinantes empíricos de los rendimientos en este contexto específico.

Groenewold y Fraser (1997) aplican ambas herramientas al mercado australiano entre 1980 y 1994, comparando el desempeño de modelos con factores macroeconómicos y artificiales obtenidos mediante AF. En el caso del modelo con factores macroeconómicos, el autor los construye siguiendo a Chen et al. (1986) y obteniendo un $R^2_{ajustado} = 0.665$. El análisis solo ofrece evidencia favorable a la inflación como factor valorado por el mercado. Por su parte, el AF muestra un mejor ajuste con tres factores que explican el 74.5 % de la variabilidad. Además, evidencia que el desempeño del modelo mejora al utilizar períodos de tiempo más extensos en comparación con los subperíodos estimados.

Por su parte, Cagnetti (2001) extiende el análisis al mercado italiano durante el período 1990–2001. Encuentra, por medio del AF, 5 factores con autovalores asociados mayores a 1, que explican un 72 % de la varianza común. Además, los interpreta como carteras con sensibilidades a distintos factores macroeconómicos por medio de regresiones lineales entre las cargas factoriales y variables observables, sus resultados son ambiguos en su interpretación, dado que cada factor se asocia a múltiples variables macroeconómicas.

El precedente quizá más relevante para nuestro caso sea el ya citado trabajo de Swoboda (2002), quien por medio de esta metodología, concluye que la varianza de los retornos en el mercado de capitales argentino durante la década de 1990, puede sintetizarse adecuadamente con dos factores latentes, concluyendo que esta técnica logra un ajuste ligeramente superior al obtenido por medio de regresión contra variables macroeconómicas.

A modo de resumen los siguientes cuadros presentan los principales resultados de los trabajos empíricos revisados en este capítulo. Naturalmente esta descripción no agota la literatura existente, no obstante, la selección de estudios presentada permite fundamentar la utilización de factores latentes y macroeconómicos en el análisis de la estructura del riesgo sistemático.

Cuadro 3.1: principales resultados de estudios que estiman factores macroeconómicos

Autor	Período	País	Nº activos	Frec.	Método	Factores macroeconómicos	$R^2_{ajustado}$ ^a
Chen et al. (1986)	1953–1983	Estados Unidos	20	Mensual	MCO	Producción industrial; Inflación no esperada; Cambio en expectativas de inflación; Riesgo de default; Estructura temporal	N/I
Swoboda (2002) ^b	1995–2000	Argentina	19	Mensual	MCO	Estructura temporal; Riesgo de default; Producción industrial; Factor residual	0.28
Burmeister & Wall (1986)	1971–1981	Estados Unidos	Índice S&P500	Mensual	MCO-SUR	Riesgo de default; Estructura temporal; Inflación no esperada; Actividad económica	0.31
Rayón (2016)	1996–2006	Estados Unidos	Fondos de cobertura	Mensual	MCO	Inflación no esperada; Riesgo de default; Estructura temporal; Producción industrial; Factor residual	0.25–0.73 ^c
Guerra (2012) ^d	1993–2013	Argentina	Índice Merval	Mensual	MCO-ARCH	Inflación; Dummies	0.24–0.26
Fuentes et al. (2017)	1990–2003	Chile	60	Mensual	MCO	Inflación; Actividad económica; Precio del cobre; Precio del petróleo; Factor residual	N/I
Pesce & Pedroni (2021)	2005–2014	Argentina	19	Mensual	MCO-PLM	Inflación; Tipo de cambio; Riesgo de default; Rendimiento del mercado; Dummies	0.32
Groenewold & Fraser (1997) ^b	1980–1994	Australia	19 (carteras)	Mensual	MCO	Inflación; Producción industrial; Tipo de cambio	N/I

^a Algunos trabajos no informan el ajuste de sus regresiones de primera etapa al concentrarse en las primas de riesgo del APT.

^b También estima factores latentes mediante PAF-FA.

^c Según el fondo de cobertura considerado.

^d No busca inferir estructura factorial sino estimar el efecto de la inflación.

N/I: No informado.

SUR: Modelo de regresiones aparentemente no relacionadas

ARCH: Modelo autorregresivo de heterocedasticidad condicional

Fuente: elaboración propia.

Cuadro 3.2: principales resultados de estudios que estiman factores latentes

Autor	Período	País	Nº activos	Frec.	Método	Nº factores	Var. explicada (%)
Roll & Ross (1980)	1962–1972	Estados Unidos	30	Diaría	ML-FA	5	56
Cortés & Porras (2013)	2000–2006	México	22	Semanal	ML-FA	5	84
Cagnetti (2001)	1990–2001	Italia	30	Diaría	ML-FA	5	40
Groenewold & Fraser (1997) ^a	1980–1994	Australia	19 (carteras)	Mensual	PAF-FA	3	72
Lehmann & Modest (1985)	1963–1982	Estados Unidos	750	Diaría	PAF-FA	5–10	40–60
Swoboda (2002) ^a	1995–2000	Argentina	19	Mensual	PAF-FA	2	33
Chen (1983)	1963–1978	Estados Unidos	180	Diaría	ML-FA	5	N/I

^a También estima un modelo macroeconómico por MCO.

ML: Metodo de máxima verosimilitud

N/I: No informado.

Fuente: elaboración propia.

Capítulo 4

Análisis econométrico

En este capítulo se aborda la implementación empírica de la investigación, cuyo propósito es examinar la existencia de una estructura factorial de riesgo sistemático en el mercado accionario argentino, en base a una cartera representativa de activos construida para el período comprendido entre enero de 2018 y diciembre de 2025. En consonancia con los objetivos planteados, se busca analizar tanto el rol de determinantes macroeconómicos como el de factores latentes, con el fin de explicar el comovimiento de los rendimientos accionarios y caracterizar la forma que adopta el riesgo sistemático en este mercado.

La estrategia metodológica se organiza en dos etapas complementarias. En primer lugar, se aplican técnicas de AF, con el objetivo de explorar la existencia de factores latentes que subyacen al comovimiento de los activos. Esta aproximación posibilita identificar la estructura subyacente de riesgo sistemático desde una perspectiva puramente estadística. En segundo lugar, se estiman modelos econométricos de regresión múltiple en línea con la literatura revisada, que buscan vincular los rendimientos de los activos con un conjunto de factores macroeconómicos construidos en base a variables macroeconómicas observables. Este paso permite evaluar empíricamente el poder explicativo de factores económicos concretos y su incidencia en la dinámica de los precios de los activos.

De esta manera, el capítulo avanza desde la detección de factores latentes hasta la identificación de factores macroeconómicos, integrando ambos enfoques en una visión que busca aportar evidencia empírica robusta sobre las fuentes del riesgo sistemático en un mercado emergente.

El procesamiento y análisis empírico se llevaron a cabo con el software *R* (versión 4.4.1)¹, utilizando un conjunto de librerías especializadas: La descarga y organización de precios bursátiles desde Yahoo Finance se realizó mediante *quantmod*, a partir de lo cual se construyeron las series de rendimientos utilizadas en el estudio; la estimación de factores latentes se llevó a cabo con *psych*. A su vez para la estimación, validación y diagnóstico de los modelos se emplearon *tseries*, *lmtest*, *sandwich*, *strucchange*, *car* y *dynlm*.

Al momento de seleccionar los datos, hay dos aristas que es necesario distinguir: (i) la que incluye las acciones que se utilizarán en el análisis, que constituyen las variables a explicar por el modelo factorial. (ii) la que incluye la información requerida para construir los shocks o innovaciones que se emplean para explicar la variabilidad de los rendimientos. Los criterios empleados para el abordaje de este último punto se desarrolla a detalle en la Sección 4.2.1.

Para la selección del panel de activos se parte de los veintitrés títulos que integraron durante el tercer trimestre de 2025 el índice S&P Merval, manteniendo un universo fijo a lo largo de todo el período de estudio. En particular los activos considerados son: YPF S.A. (YPFD.BA), Grupo Financiero Galicia S.A. (GGAL.BA), Banco Macro S.A. (BMA.BA), Telecom Argentina S.A. (TECO2.BA), Pampa Energía S.A. (PAMP.BA), Banco BBVA Argentina S.A. (BBAR.BA), Transportadora de Gas del Sur S.A. (TGSU2.BA), Ternium Argentina S.A. (TXAR.BA), Central Puerto S.A. (CEPU.BA), Aluar Aluminio Argentino S.A.I.C. (ALUA.BA), Edenor S.A. (EDN.BA), Bolsas y Mercados Argentinos S.A. (BYMA.BA), Loma Negra C.I.A.S.A. (LOMA.BA), Grupo Supervielle S.A. (SUPV.BA), Cablevisión Holding S.A. (CVH.BA), Transener S.A. (TRAN.BA), Cresud S.A.C.I.F. y A. (CRES.BA), Holcim Argentina S.A. (HARG.BA), Sociedad Comercial del Plata S.A. (COME.BA), Mirgor S.A.C.I.F.I.A. (MIRG.BA), Transportadora de Gas del Norte S.A. (TGNO4.BA), Banco de Valores S.A. (VALO.BA) y Adecoagro S.A. (AGRO.BA).

Esta elección permite preservar la coherencia transversal del panel y evitar distorsiones asociadas a cambios en la composición del índice. El período de análisis se restringe al ya mencionado, con el objetivo de garantizar homogeneidad institucional. A partir de mediados de 2017 se implementa un cambio estructural en la metodología de cálculo del índice Merval,

¹Códigos empleados disponibles a solicitud.

asociada a la creación de Bolsas y Mercados Argentinos (BYMA) y a la adopción del índice S&P Merval, que introduce un esquema de ponderación basado en la capitalización bursátil ajustada por flotación, en reemplazo del criterio previo basado en liquidez operada. Por tal motivo, no se consideró adecuado trabajar con observaciones previas a 2018. Los datos de precios y volúmenes utilizados fueron obtenidos a partir de Yahoo Finance.

Con el objetivo de garantizar que los precios de los activos incluidos reflejen de manera adecuada la información relevante y los fundamentos macroeconómicos que este trabajo busca analizar, la depuración del universo de acciones se realiza atendiendo a su nivel de liquidez. Ya que si esta es persistentemente baja puede dar lugar a precios poco representativos, afectados por escasa negociación y discrecionalidad en las transacciones o episodios de ausencia de cotización. Lo que introduce ruido adicional en el análisis empírico y debilita la relación entre los rendimientos observados y los factores macroeconómicos subyacentes.

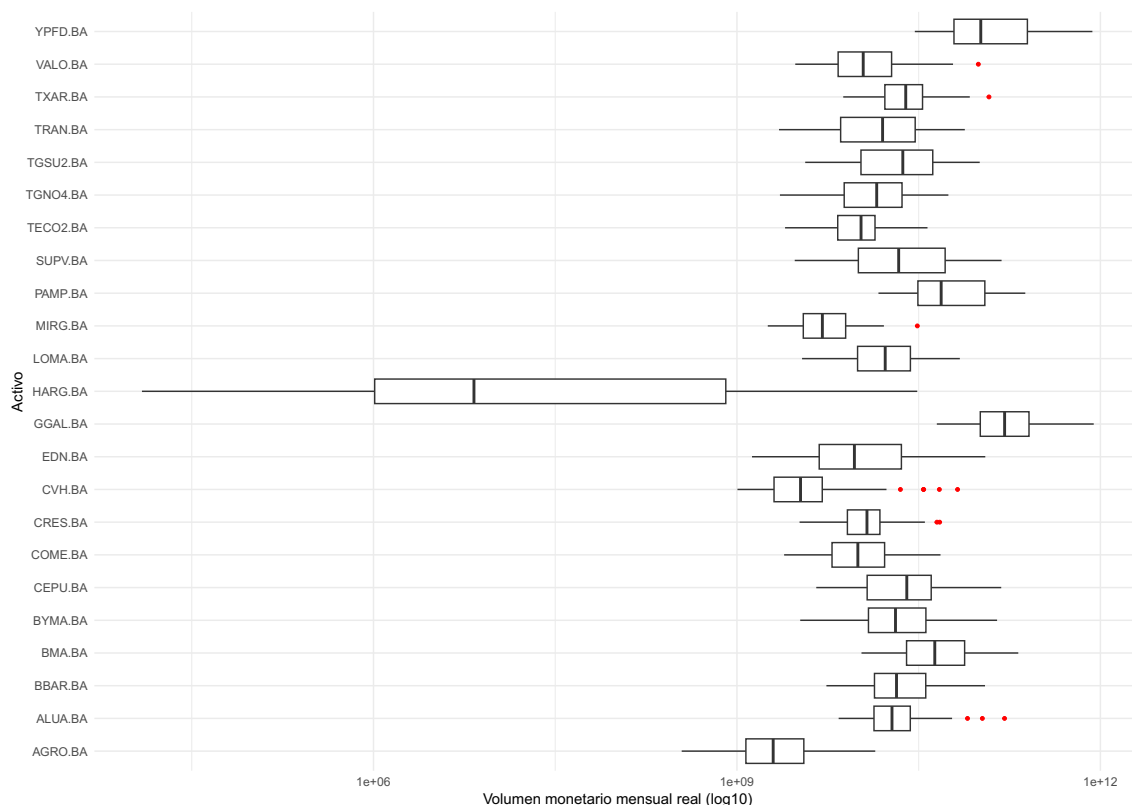
Por ello se analiza el volumen monetario operado mensual² para cada activo del panel a lo largo del período considerado. Dicho volumen se construye como el producto entre precio y cantidad negociada y se expresa en términos reales, deflactándolo por el Índice de Precios al Consumidor (IPC) y llevándolo a precios constantes de diciembre de 2025. Esta corrección resulta necesaria dado el contexto de alta inflación del período analizado, y permite comparar los niveles de negociación entre distintos momentos del tiempo sin que las variaciones nominales de precios distorsionen la medición de la liquidez efectiva.

A partir de esta medida de volumen monetario en términos reales, se analiza la distribución histórica de negociación de cada activo con el fin de evaluar el grado de homogeneidad en la liquidez del panel a lo largo del período muestral. Dado que el volumen negociado presenta una distribución fuertemente sesgada a la derecha y se caracteriza por la presencia de valores extremos, particularmente en los activos de mayor capitalización, el análisis se realiza mediante diagramas de cajas en escala logarítmica. Esta transformación permite comparar diferencias relativas en los niveles de liquidez entre activos, evitando que episodios de negociación excepcionalmente elevada distorsionen la interpretación de aquellos con menor volumen.

²Se analiza en frecuencia mensual y no diaria ya que no se busca capturar episodios puntuales de liquidez o de iliquidez extrema propios de la dinámica bursátil diaria.

La Figura 4.1 muestra la distribución del volumen monetario mensual real por activo. En general, las acciones del panel presentan distribuciones relativamente estables y concentradas en niveles adecuados de negociación. Sin embargo, se identifica algún caso que evidencia episodios de iliquidez.

Figura 4.1: distribución del volumen monetario mensual real por activo a precios constantes de diciembre de 2025, período 01.2018–12.2025, en escala logarítmica base 10.



Fuente: elaboración propia en base a datos de Yahoo Finance.

En particular, Holcim Argentina S.A. exhibe una distribución desplazada hacia valores sustancialmente inferiores en comparación con el resto del panel, reflejando una menor liquidez relativa y una alta frecuencia de períodos con escasa negociación. Si bien el activo alcanza en 2025 niveles de liquidez compatibles con su inclusión en el índice S&P Merval, esta condición no se verifica de manera consistente a lo largo del horizonte temporal de la muestra. Por este motivo, se decide excluirlo del panel definitivo utilizado en las estimaciones econométricas.

Finalmente, el panel de activos considerado en el análisis empírico queda conformado por los veintitrés títulos que integran el índice S&P Merval al tercer trimestre de 2025, excluyendo Holcim Argentina S.A., abarcando el período enero de 2018–diciembre de 2025.

Para calcular los rendimientos de las acciones consideradas, se utilizan los precios formados en el mercado local (BYMA), expresados en pesos argentinos. Se emplean precios diarios de cierre ajustados por dividendos y *splits*, lo que permite calcular los rendimientos de forma directa sin necesidad de realizar correcciones adicionales sobre las series.

Los rendimientos del activo j se expresan como $R_{j,t} = \ln(P_{j,t}) - \ln(P_{j,t-1})$ donde $P_{j,t}$ representa el precio ajustado de j al cierre en el período t . Se opta por rendimientos continuos (logarítmicos) en lugar de discretos, dado que presentan propiedades analíticas convenientes: son aditivos en el tiempo y simétricos ante variaciones positivas y negativas de igual magnitud relativa.

Posteriormente se realizó el test de raíz unitaria de Dickey-Fuller aumentado (ADF) para verificar la estacionariedad de los rendimientos de los activos considerados. Los resultados, disputables en el apéndice, indican que todos los rendimientos son estacionarios al nivel de significancia del 1 %.

4.1 Factores latentes

Para identificar la estructura subyacente en los rendimientos de los activos se realiza AFE mediante el método de PAF. La elección de esta estrategia no es arbitraria: el método de ejes principales constituye una alternativa adecuada para la estimación de la matriz de cargas ya que evita la resolución de las ecuaciones de máxima verosimilitud, mucho más restrictivas (Peña, 2002, p. 362).

En conjunto se adopta un esquema de frecuencias mixtas para los distintos procedimientos empíricos. En particular, la estimación de factores latentes mediante AFE se realiza sobre rendimientos diarios lo que arroja un total de 1949 observaciones por activo, mientras que los modelos que incorporan factores macroeconómicos se estiman a frecuencia mensual con 96 observaciones. Esta decisión puede considerarse en principio discutible, dado que el uso de rendimientos diarios introduce mayor variabilidad asociada a fluctuaciones de corto plazo y a componentes de ruido de mercado.

No obstante, la utilización de datos diarios en la etapa de estimación factorial responde a razones principalmente econométricas. La identificación estable de la estructura factorial requiere un número elevado de observaciones en relación con la cantidad de activos incluidos en el análisis, esto se discute en trabajos como Shukla y Trzcinka (1990). Trabajar con rendimientos mensuales reduciría significativamente el tamaño muestral disponible, generando matrices de datos menos adecuadas para el AF. En este sentido, el uso de información diaria permite estimar de manera más robusta la estructura de covariación transversal entre activos, que es precisamente el objeto del AFE.

Por otra parte, en la etapa de estimación de modelos con factores macroeconómicos —siguiendo el enfoque de Chen et al. (1986)— se trabaja con rendimientos y variables explicativas a frecuencia mensual. Desde el punto de vista económico, la variación relevante de los rendimientos esperados y de las primas de riesgo ocurre a baja frecuencia, ya que los rendimientos diarios contienen en gran medida componentes de ruido impredecible (Cochrane (2005)).

Adicionalmente, la disponibilidad de series macroeconómicas a frecuencia diaria es limitada o inexistente para la mayoría de las variables relevantes, lo que imposibilita la construcción consistente de factores macroeconómicos en dicha frecuencia. En consecuencia, el enfoque adoptado combina estimación factorial en alta frecuencia —para asegurar robustez estadística en la identificación de factores latentes— con estimación de modelos explicativos en baja frecuencia —para asegurar coherencia económica y compatibilidad de datos—, constituyendo la estrategia metodológica del estudio.

Ambos ejercicios no son directamente comparables ni necesariamente convergentes, ya que no es posible asegurar que la estructura factorial sea la misma en ambas frecuencias. El AFE con datos diarios cumple una función diagnóstica, confirmar que existe covariación común no explicable por azar, mientras que la regresión mensual evalúa si ciertas variables macroeconómicas observables pueden aproximar empíricamente dicha covariación. La coherencia del estudio no reside en la correspondencia numérica entre ambos procedimientos, sino en la consistencia de sus resultados con la hipótesis de existencia de riesgo sistemático multifactorial en el mercado argentino.

4.1.1 Pruebas de adecuación

Antes de llevar a cabo el AFE es necesario evaluar si los datos cumplen con los supuestos básicos que justifican su aplicación³. Para ello se utiliza, como se explicó en la Sección 2.4, la medida de adecuación muestral KMO empleada por trabajos como Cagnetti (2001) y Swoboda (2002), entre otros.

La medida KMO evalúa la proporción de la varianza entre las variables que podría ser explicada por factores comunes. El índice global arrojó un valor de 0.97, lo que se interpreta como una excelente adecuación muestral. Asimismo, el KMO se calculó individualmente para cada activo, observándose que todos los valores superan el umbral mínimo aceptable (0.50), lo que respalda la pertinencia del análisis. El Cuadro 4.1 resume los valores de KMO para cada activo analizado.

Cuadro 4.1: medida de adecuación muestral individual KMO

Activo	KMO	Activo	KMO
ALUA	0.91	LOMA	0.99
BBAR	0.97	MIRG	0.99
BMA	0.96	PAMP	0.98
BYMA	0.99	SUPV	0.98
CEPU	0.99	TECO2	0.98
COME	0.99	TGNO4	0.97
CRES	0.98	TGSU2	0.98
CVH	0.99	TRAN	0.97
EDN	0.99	TXAR	0.92
GGAL	0.96	YPFD	0.98
VALO	0.98	AGRO	0.99

Fuente: elaboración propia.

La normalidad multivariante de los rendimientos se evaluó mediante el test de Mardia. Los resultados muestran valores de asimetría y curtosis que permiten rechazar la hipótesis nula de normalidad multivariante al 1 %, indicando desviaciones significativas respecto a la distribución normal. Esto confirma la necesidad de optar por el método de estimación de ejes principales (PAF), que no depende de dicha condición.

³El AF no exige independencia temporal de las observaciones ni ausencia de autocorrelación, ya que modela la estructura de covariación transversal entre variables y no su dinámica en el tiempo. Asimismo, los valores atípicos en rendimientos financieros suelen reflejar shocks reales de mercado, por lo que no se aplican filtros que puedan distorsionar la matriz de correlaciones.

4.1.2 Determinación del número de factores a retener

La selección del número de factores a extraer se realizó combinando dos criterios complementarios ampliamente utilizados en la literatura: el criterio de Kaiser y el análisis paralelo. El objetivo fue evaluar la robustez de la estructura factorial y balancear capacidad explicativa con parsimonia del modelo. Aquí buscamos identificar un número parsimonioso de factores que explique una proporción adecuada de la estructura de comunalidades.

Criterio de Kaiser

El criterio de Kaiser establece que deben retenerse aquellos factores cuyos autovalores sean mayores que la unidad. El Cuadro 4.2 presenta los primeros autovalores ordenados de mayor a menor, de acuerdo con este criterio deben retenerse tres factores.

Cuadro 4.2: autovalores de la matriz de correlaciones

Factor	1	2	3	4	5	6	7	8	(...)
Autovalor	12.365	1.176	1.041	0.830	0.741	0.660	0.593	0.511	(...)

Fuente: elaboración propia en base a datos de Yahoo Finance.

El marcado predominio del primer autovalor respecto de los restantes sugiere la presencia de una dimensión dominante en la matriz de correlaciones de los rendimientos. Este resultado indica que una fracción importante de la variabilidad conjunta de los activos se encuentra asociada a un movimiento agregado común, reflejando un elevado grado de comovimiento entre los títulos analizados. Este patrón es consistente con la evidencia empírica habitual en mercados financieros, donde las correlaciones positivas entre activos generan que el primer componente concentre una proporción significativa de la variabilidad total del sistema.

Es importante señalar que los autovalores presentados en el Cuadro 4.2 corresponden a la descomposición espectral de la matriz de correlaciones y se utilizan exclusivamente como herramienta para evaluar la dimensionalidad de los datos y aplicar los criterios de selección del número de factores. En esta etapa, los componentes se construyen de manera tal que maximizan sucesivamente la varianza explicada de la matriz original y son ortogonales entre sí, por lo que

cada componente adicional captura la mayor proporción posible de variabilidad remanente no explicada por los anteriores.

No obstante, el objetivo de este trabajo no es realizar un ACP, sino estimar un modelo de AFE mediante el método de PAF. En consecuencia, una vez determinado el número de factores a retener, la estimación del modelo factorial se realiza a partir de la matriz reducida de comunalidades \mathbf{S}^* , definida en el marco teórico como

$$\mathbf{S}^* \approx \mathbf{S} - \mathbf{\Psi},$$

la cual aproxima la parte común de la matriz de varianzas y covarianzas de \mathbf{x} , a diferencia de la matriz de correlaciones \mathbf{D} , eliminando la contribución de la varianza específica.

Como consecuencia de esta transformación, los factores estimados explican únicamente la varianza común entre los activos y no la varianza total observada en la matriz de correlaciones original. Por esta razón, la proporción de varianza explicada reportada posteriormente en el Cuadro 4.3 se calcula sobre la base de la matriz reducida \mathbf{S}^* y no es directamente comparable con la magnitud de los autovalores presentados en el Cuadro 4.2.

Análisis paralelo

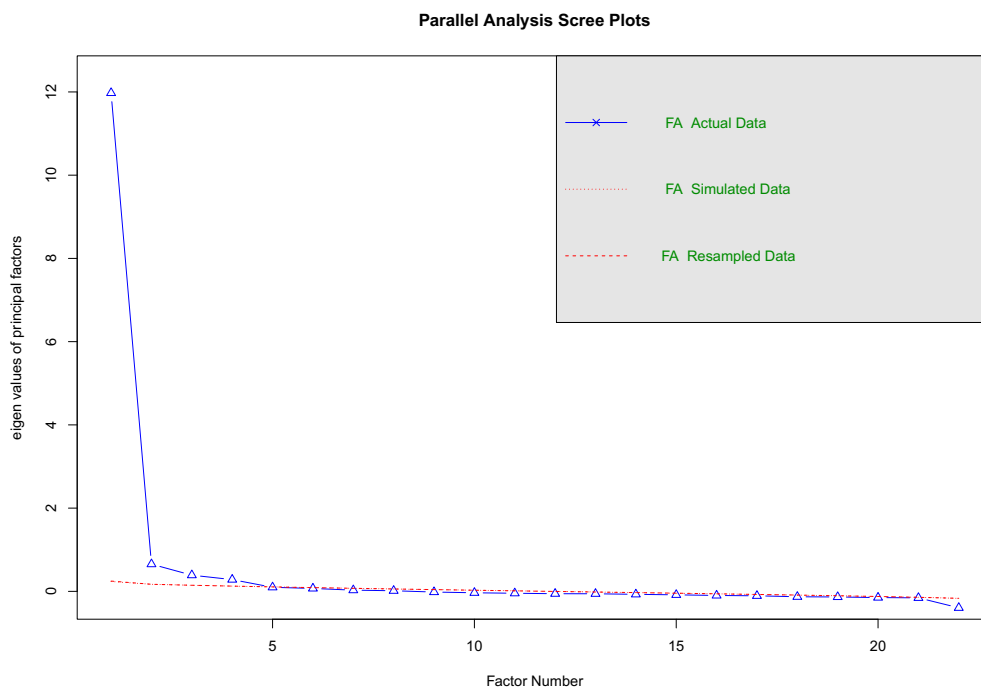
El análisis paralelo se incorporó como criterio superador al de Kaiser debido a las limitaciones ya mencionadas en la Sección 2.4. En contraste, el análisis paralelo proporciona una regla de decisión basada en comparar los autovalores observados con aquellos que se obtendrían en matrices de datos generadas bajo ausencia de estructura factorial. En su implementación tradicional, dichos valores de referencia se obtienen mediante simulaciones Monte Carlo (en este caso, 500 iteraciones) generando matrices aleatorias con igual dimensión muestral, bajo el supuesto de normalidad multivariada e independencia entre variables.

Sin embargo, dado que los rendimientos financieros suelen exhibir distribuciones leptocúrticas, asimetrías y desviaciones respecto de la normalidad, se complementó este procedimiento con un procedimiento de remuestreo. Este enfoque genera matrices a a partir de permutaciones o remuestreos de los datos originales, preservando sus propiedades distributivas pero eliminando la estructura de correlación. De este modo, se obtiene un umbral de compa-

ración más robusto frente a no normalidad y colas pesadas, características típicas de las series financieras.

Este procedimiento permite retener únicamente aquellos factores cuyos autovalores observados superan los autovalores promedio obtenidos bajo ausencia de estructura factorial. La Figura 4.2 presenta la comparación entre los autovalores empíricos y los generados por el análisis paralelo —tanto mediante simulaciones Monte Carlo bajo normalidad, como a través de remuestreo—.

Figura 4.2: comparación entre autovalores observados y simulados



Nota: el scree plot o gráfico de sedimentación es una representación de los autovalores ordenados de mayor a menor, utilizada para identificar el número de factores o componentes relevantes.

Fuente: elaboración propia en base a datos de Yahoo Finance.

En la Figura 4.2 se observa que los cuatro primeros autovalores empíricos superan los valores de referencia obtenidos bajo la hipótesis de ausencia de correlación entre las variables. En consecuencia, el análisis paralelo sugiere la retención de cuatro factores, ya que solo hasta el cuarto componente la varianza explicada excede la atribuible al azar. Este resultado es consistente con las críticas al criterio de Kaiser señaladas en la Sección 2.4, motivo por el cual en este trabajo se opta por la extracción de cuatro factores.

Cuadro 4.3: varianza explicada por factor en el modelo AFE de cuatro factores

Factor	Proporción de varianza	Varianza acumulada
Factor 1	0.2154	0.2154
Factor 2	0.1565	0.3718
Factor 3	0.1528	0.5247
Factor 4	0.1051	0.6298

Fuente: elaboración propia en base a datos de Yahoo Finance.

En suma, los resultados del AFE muestran que los rendimientos de los activos pueden describirse mediante un conjunto reducido de factores comunes, confirmando la existencia de una estructura factorial en el mercado. Los elevados valores del índice KMO indican que las correlaciones entre activos no son aleatorias, sino que responden a fuentes comunes de variación. La extracción de cuatro factores que explican el 62.98 % de la varianza total sugiere que gran parte del comportamiento de los rendimientos puede ser representado por un número acotado de factores latentes. Siguiendo la interpretación clásica de Connor y Korajczyk (1988), estos factores pueden entenderse como carteras factoriales: combinaciones lineales de activos que representan shocks agregados no observados y que resumen las distintas dimensiones del riesgo sistemático. Así, la evidencia obtenida sugiere que el mercado argentino presenta una estructura de riesgo.

Futuras investigaciones podrían profundizar en la interpretación económica de los factores extraídos a partir del análisis de las cargas factoriales y su vinculación con la estructura sectorial de los activos. No obstante, como se explicó previamente, el objetivo de este trabajo es utilizar el AFE simplemente como una herramienta para testear formalmente la existencia de una estructura común subyacente, complementando este enfoque con la identificación de factores de naturaleza macroeconómica. En consecuencia, el análisis de las cargas factoriales y su interpretación económica excede el alcance del presente estudio.

4.2 Factores macroeconómicos

En esta sección, aunque se toma como referencia la propuesta de Chen et al. (1986) para la reconstrucción de factores, el objetivo no es replicar su trabajo ni estimar precios del riesgo de manera estricta. En cambio, se adopta su metodología para proponer factores macroeconómicos

plausibles —actividad real, inflación, riesgo de default, condiciones financieras y shocks discretos relevantes—, adaptando su construcción a la información disponible y complementando con variables adicionales pertinentes para adecuar la especificación al contexto local.

El Cuadro 4.4 presenta las fuentes utilizadas para la obtención de las variables macroeconómicas⁴.

Cuadro 4.4: fuentes de datos para variables independientes

Variable	Fuente
Índice de Producción Industrial Manufacturero (IPIM)	Instituto Nacional de Estadística y Censos (INDEC)
Ventas de supermercados	INDEC
Expectativas de inflación (REM)	Banco Central de la República Argentina (BCRA)
Expectativas de tipo de cambio nominal (REM)	BCRA
Tipo de cambio nominal de referencia Comunicación A3500 (ARS/USD)	BCRA
Índice de precios internacionales de commodities energéticos	Banco Mundial
Índice de precios internacionales de commodities no energéticos	Banco Mundial
Índice de precios internacionales de commodities agrícolas	Banco Mundial
Riesgo país de Argentina (Emerging Markets Bond Index, EMBI)	Yahoo Finance

Nota: todas las series se expresan en frecuencia mensual y corresponden período al analizado.

Fuente: elaboración propia.

Por otro lado, la variable dependiente utilizada es el rendimiento de una cartera equiponderada compuesta por los activos previamente seleccionados. Los rendimientos mensuales se obtuvieron como la suma de los rendimientos diarios dentro de cada mes calendario. Luego, el rendimiento de la cartera se computó como:

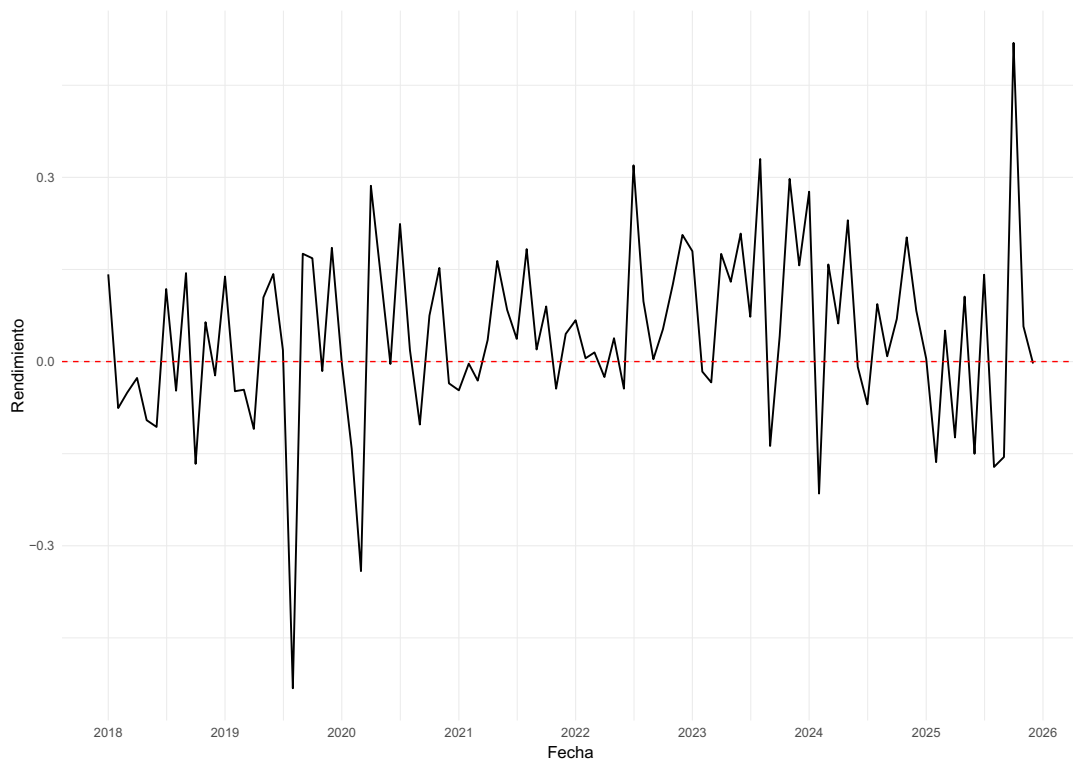
$$R_t^C = \frac{1}{N} \sum_{i=1}^N R_{i,t}, \quad (4.1)$$

⁴Las series de ventas de supermercados y el índice de producción industrial manufacturero se utilizan en su versión desestacionalizada publicada por INDEC. El procedimiento de desestacionalización se basa en un esquema multiplicativo (X-13ARIMA-SEATS). Dado que estos factores se reestiman a medida que se incorporan nuevos datos, algunos valores de la muestra —en particular hacia fines de 2025, cercanos al momento de elaboración de este trabajo— pueden presentar revisiones posteriores. Esto constituye una posible limitación para la replicabilidad exacta de los resultados aquí presentados de no considerarse.

donde $N = 22$ es el número de activos.

La Figura 4.3 muestra la evolución del rendimiento de la cartera a lo largo del período muestral. Se observan fluctuaciones de magnitud considerable, con episodios de mayor volatilidad, característica propia de mercados en contextos macroeconómicos inestables (estos episodios motivan la necesidad de incorporar variables de control mas adelante presentadas). La serie no exhibe una tendencia determinística marcada, sino que oscila alrededor del cero, comportamiento consistente con un proceso estacionario, en línea con los resultados de las pruebas formales de raíz unitaria reportadas mas adelante.

Figura 4.3: evolución temporal del rendimiento de la cartera, periodo 2018.01–2025.12



Fuente: elaboración propia.

Esta especificación permite diversificar el componente de riesgo propio, como demuestran Marín y Rubio (2011), a medida que se incrementa el número de activos incluidos en la cartera, el componente idiosincrático del rendimiento converge a cero. De modo que la variabilidad remanente en R_t^C representa una aproximación adecuada del riesgo sistemático, el cual por medio de los factores propuestos el modelo pretende capturar.

Una extensión natural de este trabajo podría dirigirse hacia un modelo de datos de panel, tal como propone Pesce et al. (2015). Este enfoque permitiría controlar efectos fijos individuales, capturando así la sensibilidad de las firmas a los factores. Alternativamente, también podría adoptarse un enfoque de regresiones aparentemente no relacionadas *Seemingly Unrelated Regressions* (SUR), como emplean McElroy y Burmeister (1988), que permite estimar conjuntamente un sistema de ecuaciones por activo sujetos a restricciones consistentes con el APT. No obstante, estas extensiones exceden el alcance del presente trabajo, que se limita a evaluar la capacidad explicativa de los factores macroeconómicos sobre una cartera diversificada representativa del mercado argentino.

4.2.1 Construcción de factores macroeconómicos

Siguiendo la notación de Chen et al. (1986), a continuación se expone la metodología empleada para la construcción de variables orientadas a capturar las innovaciones o sorpresas no anticipadas consideradas como posibles factores relevantes para el mercado argentino.

Inflación no anticipada (UI) y el cambio en la inflación esperada (DEI)

Siguiendo la especificación de los autores, la inflación observada se descompone en un componente anticipado y otro no anticipado. La inflación mensual se aproxima con el Índice de Precios al Consumidor (IPC) elaborado por el INDEC, mientras que las expectativas inflacionarias del Relevamiento de Expectativas de Mercado (REM) del BCRA.

Sea I_t la inflación observada en el mes t y $\mathbb{E}_{t-1}[I_t]$ la inflación esperada en $t - 1$ para el mes t . La inflación no anticipada se construye como una sorpresa inflacionaria relativa, definida como el desvío relativo de la inflación observada respecto de la esperada:

$$UI_t = \frac{I_t - \mathbb{E}_{t-1}[I_t]}{\mathbb{E}_{t-1}[I_t]}. \quad (4.2)$$

Se incluye además una segunda medida asociada a la dinámica de las expectativas, que captura shocks derivados de revisiones en la inflación esperada (DEI):

$$DEI_t = \frac{\mathbb{E}_t[I_{t+1}] - \mathbb{E}_{t-1}[I_t]}{\mathbb{E}_{t-1}[I_t]}. \quad (4.3)$$

Aunque en su trabajo los autores deducen la inflación esperada a partir de la conocida ecuación de Fisher, aquí se emplean directamente las expectativas relevadas por el REM, lo que resulta consistente con el espíritu de la formulación teórica y permite una medición empírica directa de los componentes anticipados y no anticipados. Por otro lado, Chen et al. (1986) definen la inflación no anticipada como la diferencia simple entre el valor observado y el esperado, en este trabajo se utiliza el desvío relativo, con el fin de evitar que episodios de alta volatilidad inflacionaria distorsionen la magnitud del efecto sorpresa y sesguen la estimación de su impacto sobre los rendimientos.

Variación de la actividad real: producción industrial (MP y YP)

Los autores utilizan el crecimiento industrial como proxy de actividad real. Para el caso argentino se empleó el Índice de Producción Industrial Manufacturero (IPIM) del INDEC e igualmente se consideraron dos medidas.

MP: Variación mensual desestacionalizada de la producción industrial.

$$MP_t = \ln(IPIM_t) - \ln(IPIM_{t-1}). \quad (4.4)$$

YP: Variación interanual de la producción industrial.

$$YP_t = \ln(IPIM_t) - \ln(IPIM_{t-12}). \quad (4.5)$$

Ambas medidas capturan distintos horizontes temporales: mensual e interanual respectivamente. Podría argumentarse que, un índice de actividad económica que contemple la dinámica de diversos sectores como podría ser el EMAE, Estimador Mensual de Actividad Económica elaborado por el INDEC, proporcionaría una representación mas adecuada de la evolución de

la actividad en la economía argentina, no obstante se decide emplear el IPIM a los fines de mantener coherencia, en la medida de lo posible, con el trabajo de Chen et al. (1986).

Prima de riesgo inesperada (UPR)

La prima por riesgo de default se aproxima mediante la variación del índice EMBI+ Argentina elaborado por J.P. Morgan. Esta variable captura shocks financieros globales que afectan la tasa de descuento de los flujos futuros de las empresas. En particular, se utiliza la diferencia logarítmica del índice.

$$UPR_t = \Delta \ln(EMBI_t) = \ln(EMBI_t) - \ln(EMBI_{t-1}). \quad (4.6)$$

Estructura temporal de tasas de interés (UTS)

En el modelo de Chen et al. (1986) se incorpora un factor asociado a la estructura temporal de tasas de interés (UTS, *Unexpected Term Structure*), destinado a capturar variaciones en la pendiente de la curva de rendimientos, y con ello cambios en las condiciones financieras intertemporales y en las expectativas macroeconómicas. Este factor es de uso extendido en la literatura académica como vimos en el Cuadro 3.1.

Siguiendo dicha especificación, el factor se define como el diferencial entre una tasa de interés de bonos soberanos de largo plazo y una de corto plazo, de modo de aproximar la prima por plazo implícita en la estructura temporal:

$$UTS_t = TL_t - TS_t. \quad (4.7)$$

No obstante, para el caso argentino no fue posible reconstruir este factor de manera consistente para todo el período analizado, dado que se basa en la Tasa Interna de Retorno (TIR) de bonos soberanos de largo plazo, cuya disponibilidad y continuidad resultan muy limitadas especialmente en los primeros años de la muestra.

Este problema es abordado por Swoboda (2002), quien en su lugar aproxima este factor calculando la diferencia entre la TIR de un bono del Tesoro de los EEUU de largo plazo y la tasa de interés local para depósitos en dólares a treinta días. Este criterio, si bien válido, no se utiliza

aquí dado que en esencia, el UTS busca capturar variaciones en la curva de rendimientos local, las cuales reflejan condiciones específicas del contexto. La utilización de tasas internacionales podría introducir ruido y sesgos en la medición del factor, al no reflejar las condiciones del mercado doméstico.

Por ello, si bien se presenta su definición por tratarse de un factor ampliamente utilizado, el UTS no se prueba formalmente en esta investigación. Esta exclusión se reconoce como una limitación derivada de restricciones de información, frecuentes en aplicaciones sobre mercados emergentes.

Precios de commodities y consumo doméstico

Los autores sugieren que, además de los factores previamente desarrollados, el cambio en el logaritmo del precio del petróleo podría representar un factor sistemático, aunque no encuentran evidencia estadística suficiente a favor de esta hipótesis. Para el caso argentino, en este sentido, se analiza la sensibilidad a cambios en las diferencias logarítmicas de los precios de commodities. Utilizando los índices de precios de commodities del Banco Mundial, incluyendo la variación del índice de energía (V.ENERGY), el índice de productos no energéticos (V.NON.ENERGY) y el índice agrícola (V.AGRICULTURE).

También se propone testear las variaciones en el consumo doméstico. Para el caso argentino, este efecto se analiza utilizando la serie de ventas de supermercados, variación mensual desestacionalizada, (V.SUP) del INDEC. Se reconoce que una posible limitación de esta variable para medir consumo es que tiene la desventaja de no capturar cambios en los patrones de consumo. Sin embargo, dada la disponibilidad de datos, se considera una proxy adecuada para este análisis.

Variables dummies (shock.POS; shock.NEG)

En línea con López Herrera (2006), quien incorpora para el mercado mexicano variables binarias para controlar episodios extraordinarios como el “tequila”, “dragón”, “vodka” o la crisis brasileña de 1999, se incorporan variables dummies con el objetivo de controlar el efecto de

eventos políticos discretos con efectos sobre las expectativas potencialmente disruptivos sobre el mercado argentino.

Para la identificación de estos episodios se empleó un criterio estrictamente institucional y cronológico, definido ex ante a la estimación del modelo y con independencia de la evolución observada en los rendimientos de los activos analizados. Los eventos fueron identificados a partir de hitos políticos y macroeconómicos de relevancia pública. Asimismo, su clasificación como eventos de impacto positivo o negativo se basó en la interpretación dominante al momento del hecho en el debate económico y político, y no en la magnitud ni en el signo de los retornos observados en la muestra utilizada para la estimación⁵.

■ **Eventos con impacto negativo:**

- Crisis cambiaria de octubre de 2018.
- Elecciones primarias (PASO) de agosto de 2019.
- Anuncio del confinamiento por COVID-19 en marzo de 2020.
- Elecciones provinciales en la Provincia de Buenos Aires de septiembre de 2025.

■ **Eventos con impacto positivo:**

- Nombramiento de Sergio Massa como ministro de Economía en julio de 2022.
- Elecciones primarias (PASO) de agosto de 2023.
- Elección presidencial en segunda vuelta (ballotage) de noviembre de 2023.
- Elecciones legislativas nacionales de octubre de 2025.

Las variables se definieron como:

$$shock.POS_t = \begin{cases} 1 & \text{si en } t \text{ ocurrió un evento con impacto positivo,} \\ 0 & \text{en caso contrario,} \end{cases}$$

⁵Una crítica a la inclusión de estas variables dummies es la posible endogeneidad, dada su clasificación entre impactos positivos y negativos. Este problema surge de la dificultad de construir factores que capturen sorpresas de índole política de manera objetiva. En el Apéndice B se presenta un chequeo ex post que analiza la relación de estas dummies con los residuos del modelo, mostrando que no existe correlación significativa.

$$shock.NEG_t = \begin{cases} 1 & \text{si en } t \text{ ocurrió un evento con impacto negativo,} \\ 0 & \text{en caso contrario.} \end{cases}$$

Cabe señalar que la incorporación de estas variables busca aislar perturbaciones extraordinarias de carácter discreto asociadas a eventos de naturaleza institucional, política o macroeconómica identificables. Su ocurrencia responde a hechos externos al modelo, por lo que se consideran exógenos en sentido econométrico. Estas variables permiten capturar movimientos abruptos que, de no considerarse, sesgarían la estimación de los coeficientes. Por este motivo se incorporan a los factores propuestos por Chen et al. (1986).

Riesgo cambiario (FX)

Siguiendo un enfoque análogo al utilizado para la inflación, el tipo de cambio nominal se descompone en un componente anticipado y otro no anticipado. Para ello se emplea como medida observada el tipo de cambio de referencia Comunicación A-3500 del BCRA, mientras que las expectativas del mercado se obtienen del REM (para este mismo tipo de cambio de referencia).

Esta descomposición permite identificar tanto la sorpresa cambiaria —asociada a nueva información no incorporada previamente por los agentes— como las revisiones en las expectativas de depreciación. Ambos componentes resultan relevantes para la valuación de activos financieros locales, en la medida en que afectan las decisiones de cobertura, los flujos de portafolio y las primas de riesgo.

Sea TC_t el tipo de cambio observado en el mes t y $\mathbb{E}_{t-1}[TC_t]$ el tipo de cambio esperado en $t - 1$ para el mes t . El componente no anticipado del tipo de cambio se define como una sorpresa cambiaria relativa, medida como el desvío relativo del valor observado respecto del esperado:

$$UFX_t = \frac{TC_t - \mathbb{E}_{t-1}[TC_t]}{\mathbb{E}_{t-1}[TC_t]}. \quad (4.8)$$

Adicionalmente, se incorpora una medida asociada a la dinámica de las expectativas cambiarias, que captura revisiones en el tipo de cambio esperado entre períodos consecutivos. Esta variable refleja shocks provenientes de cambios en las proyecciones del mercado:

$$DEFX_t = \frac{\mathbb{E}_t[TC_{t+1}] - \mathbb{E}_{t-1}[TC_t]}{\mathbb{E}_{t-1}[TC_t]}. \quad (4.9)$$

4.2.2 Estimación de modelos de regresión

A continuación se expone el procedimiento econométrico empleado para estimar y validar la especificación factorial macroeconómica adecuada al mercado argentino.

Dado que la inferencia en modelos de series temporales requiere trabajar con variables estacionarias, se aplicó la prueba ADF a cada una de las series consideradas. Los resultados, disponibles en el apéndice, indican que el rendimiento de la cartera (R), así como las variables MP, DEI, UPR, V.AGRICULTURE, V.NON.ENERGY, V.ENERGY, UFX y V.SUP, resultan estacionarias en niveles. En contraste, las variables YP, UI y DEFX no permiten rechazar la presencia de una raíz unitaria en niveles, consecuentemente, se procedió a calcular la primera diferencia de dichas series (dif.YP, dif.UI y dif.DEFX), verificándose posteriormente que estas transformaciones resultan estacionarias según el test ADF al nivel de significancia del 5 %, por lo que se las clasifica como procesos $I(0)$.

Se parte de una especificación saturada estimada por MCO, que incluye el conjunto completo de factores macroeconómicos. Esta regresión general funciona como punto de partida para una estrategia de depuración progresiva, basada en criterios de significatividad estadística individual, significatividad conjunta y parsimonia. A partir de ella se construyen especificaciones depuradas alternativas, conservando únicamente los factores con mayor aporte empírico y coherencia económica.

Modelo 1 — modelo saturado

El modelo presenta elevada significatividad global y un coeficiente de determinación ajustado de 0.534, lo que indica una buena capacidad explicativa conjunta. A nivel individual, resultan estadísticamente significativos la prima de riesgo (UPR) y los shocks positivos y

negativos. La variación de la producción industrial (MP) presenta significatividad marginal. El resto de los regresores no muestra contribución estadísticamente relevante.

Este patrón es consistente con la presencia de sobreparametrización, por lo que se procede a construir especificaciones más parsimoniosas mediante un proceso de depuración.

Cuadro 4.5: Regresión del modelo MCO — especificación saturada

Variable	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.03699	0.01129	3.28	0.0015**
MP	0.65695	0.36466	1.80	0.0753.
dif.YP	0.00920	0.24212	0.04	0.9698
dif.UI	-0.05895	0.03827	-1.54	0.1273
DEI	0.00645	0.05490	0.12	0.9068
UPR	-0.22458	0.09304	-2.41	0.0180*
SHOCK.POS	0.37310	0.05799	6.43	7.9e-09***
SHOCK.NEG	-0.19086	0.06516	-2.93	0.0044**
V.ENERGY	-0.15341	0.14256	-1.08	0.2850
V.NON.ENERGY	1.00555	0.96931	1.04	0.3026
V.AGRICULTURE	-1.21124	1.04263	-1.16	0.2487
V.SUP	-0.00237	0.00394	-0.60	0.5483
UFX	-0.34629	0.34628	-1.00	0.3202
dif.DEFX	-0.04838	0.07974	-0.61	0.5457
Residual SE: 0.101 (82 gl)				
R ² = 0.588 Ajustado = 0.523				
F = 9.01 (13, 82 gl), p = 3.66e-11				

Fuente: elaboración propia.

Si bien en la especificación saturada los coeficientes asociados a V.AGRICULTURE, dif.UI y UFX no resultan estadísticamente significativos en forma individual —con valores-p en torno a 0.25—, ello no implica necesariamente la ausencia de relevancia económica. En modelos con múltiples regresores, la sobreparametrización puede inflar las varianzas estimadas y reducir la precisión de los coeficientes individuales, dificultando la detección de efectos marginales.

En este contexto, se consideró pertinente evaluar dichas variables en especificaciones más parsimoniosas, donde su contribución pudiera estimarse con mayor precisión y menor varianza. Si bien no se encontró evidencia estadística que respalde a V.AGRICULTURE como factor de riesgo, a continuación se presentan modelos depurados que conservan dif.UI y UFX, a fin de analizar su desempeño en un entorno más restringido.

Modelo 2 — modelo depurado

Como primera especificación depurada se conserva el conjunto de factores con mayor respaldo estadístico, incorporando la prima de riesgo, la actividad, la inflación inesperada y las dummies de shocks. En esta especificación todos los coeficientes resultan estadísticamente significativos al 5 %. El modelo presenta significatividad conjunta y un coeficiente de determinación ajustado de 0.536.

La inflación inesperada (dif.UI) presenta signo negativo y significativo, consistente con un efecto contractivo de sorpresas inflacionarias sobre los retornos reales de la cartera. Este resultado sugiere que aumentos no anticipados en el nivel de precios afectan negativamente la rentabilidad de los activos domésticos, posiblemente a través de mayores tasas de descuento o deterioro de expectativas macroeconómicas.

La prima de riesgo (UPR) mantiene signo negativo y significativo, indicando que incrementos en el riesgo sistemático se asocian con menores retornos contemporáneos. La actividad industrial (MP) conserva signo positivo y significativo, en línea con la intuición general. Finalmente, las dummies de shocks políticos mantienen magnitud y signos estables.

Cuadro 4.6: regresión del modelo MCO — especificación depurada

Variable	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.0368	0.0108	3.42	0.001***
MP	0.6514	0.2513	2.59	0.011*
UPR	-0.2511	0.0785	-3.20	0.002**
dif.UI	-0.0754	0.0309	-2.44	0.017*
shock.POS	0.3787	0.0534	7.09	2.89e-10***
shock.NEG	-0.1765	0.0618	-2.86	0.005**

Residual SE: 0.1001 (90 gl)
 $R^2 = 0.5604$ Ajustado = 0.5360
 $F = 22.95$ (5, 90 gl), $p = 8.80e-15$

Fuente: elaboración propia.

Modelo 3 — modelo depurado alternativo

Como especificación alternativa se estima un modelo depurado que incorpora el componente no anticipado del tipo de cambio (UFX) como medida de shock nominal, junto con la prima de riesgo, la actividad y las variables dummies. En esta versión todos los coeficientes resultan significativos al 5 % y el modelo presenta elevada significatividad conjunta.

El coeficiente del componente no anticipado del tipo de cambio (UFX) es negativo y estadísticamente significativo, consistente con un efecto contractivo de sorpresas de depreciación sobre los retornos de la cartera. Mientras que el resto de los regresores mantienen signos y niveles de significancia similares a los observados en la especificación previa.

El coeficiente de determinación ajustado (0.528) resulta levemente inferior al obtenido en la especificación anterior, lo que sugiere una capacidad explicativa marginalmente menor.

Cuadro 4.7: regresión del modelo MCO — especificación depurada alternativa

Variable	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.0392	0.0108	3.62	0.0005***
MP	0.7199	0.2515	2.86	0.005**
UPR	-0.1979	0.0832	-2.38	0.019*
UFX	-0.6273	0.3023	-2.08	0.041*
shock.POS	0.3520	0.0521	6.76	1.37e-09***
shock.NEG	-0.1986	0.0619	-3.21	0.0019**
Residual SE: 0.1010 (90 gl)				
R ² = 0.5528 Ajustado = 0.5279				
F = 22.25 (5, 90 gl), p = 1.88e-14				

Fuente: elaboración propia.

En conjunto, los resultados indican que el componente no anticipado del tipo de cambio representa un canal nominal potencial para explicar los retornos de la cartera, aunque su poder explicativo agregado resulta ligeramente inferior al de la especificación alternativa considerada previamente.

Dado que los estadísticos VIF⁶ se mantienen en niveles moderados —todos por debajo de los umbrales usuales de aceptación—, la variación observada en la significatividad de *dif.UI* entre especificaciones no puede atribuirse a problemas severos de multicolinealidad. En particular, en el modelo saturado el componente no anticipado del tipo de cambio (UFX) presenta un VIF de 1.47, mientras que la inflación inesperada (*dif.UI*) registra un VIF de 1.74. Estos valores se encuentran muy por debajo de los niveles comúnmente considerados problemáticos (usualmente VIF >10, o de forma más conservadora VIF >5).

No obstante, ambas variables exhiben una correlación positiva moderada ($\rho = 0.36$), lo que sugiere que capturan información macroeconómica parcialmente relacionada. En este contexto, la menor significatividad individual observada cuando se incluyen conjuntamente

⁶Factor de inflación de varianza, este estadístico mide cuánto se incrementa la varianza del estimador de un coeficiente debido a la correlación entre las variables explicativas del modelo.

puede interpretarse como resultado de un solapamiento parcial en la información contenida en ambas proxies, más que como consecuencia de problemas severos de multicolinealidad.

Asimismo, como se muestra más adelante, no se encuentra evidencia muestral de heterocedasticidad ni de autocorrelación serial en los residuos. En consecuencia, no se rechazan los supuestos clásicos del modelo lineal, por lo que la inferencia basada en los valores- p puede interpretarse bajo dicho marco.

En línea con la interpretación de Guerra (2012), quien advierte que la comovilidad entre inflación y precios de los activos puede estar asociada a componentes macroeconómicos comunes más que a una relación causal directa, resulta relevante considerar la forma en que se construye la variable dependiente en este estudio. Los rendimientos se definen como variaciones logarítmicas de precios y se agregan mensualmente sobre una cartera equiponderada de activos, lo que reduce el componente idiosincrático y enfatiza la variación asociada a factores sistemáticos.

En este contexto, puede argumentarse que $dif.UI$ y UFX capturan distintas manifestaciones de shocks nominales que afectan el entorno macroeconómico, tales como sorpresas inflacionarias o variaciones inesperadas en el tipo de cambio. Aunque estas variables no son estadísticamente redundantes, es plausible que reflejen parcialmente una dimensión nominal común vinculada a la evolución del poder adquisitivo del dinero y a las condiciones monetarias que inciden sobre la valuación de los activos.

Esto no representa una limitación de la metodología propuesta, dado que estimar los rendimientos sobre precios deflactados implicaría sustraer de la serie información relevante para analizar.

En consecuencia, y considerando el conjunto de resultados obtenidos, se adopta como especificación final el Modelo 2, que incorpora la actividad industrial (MP), la inflación inesperada ($dif.UI$), la prima de riesgo (UPR) y las variables dummy de shocks positivos y negativos.

La elección de esta especificación responde, en primer lugar, a su adecuado desempeño econométrico: todos los coeficientes resultan estadísticamente significativos, presentan signos consistentes con la interpretación teórica desarrollada y el modelo exhibe elevada significati-

vidad conjunta, aunque la diferencia en el ajuste respecto de la especificación alternativa es cuantitativamente reducida.

En segundo lugar, la utilización de la inflación no anticipada como medida de shock nominal se encuentra ampliamente documentada en la literatura empírica revisada, donde esta variable constituye una aproximación estándar para capturar sorpresas macroeconómicas, tal como se observa en el Cuadro 3.1.

De este modo, la especificación seleccionada mantiene un equilibrio adecuado entre parsimonia y capacidad explicativa, al tiempo que preserva la comparabilidad con los antecedentes empíricos y la coherencia con el enfoque factorial macroeconómico adoptado para el análisis del mercado argentino.

Diagnóstico post-estimación

Con el objetivo de validar la inferencia del modelo MCO seleccionado (Modelo 2), se aplicó una batería de pruebas diagnósticas sobre los residuos y los regresores, verificando el cumplimiento de los supuestos clásicos del modelo lineal.

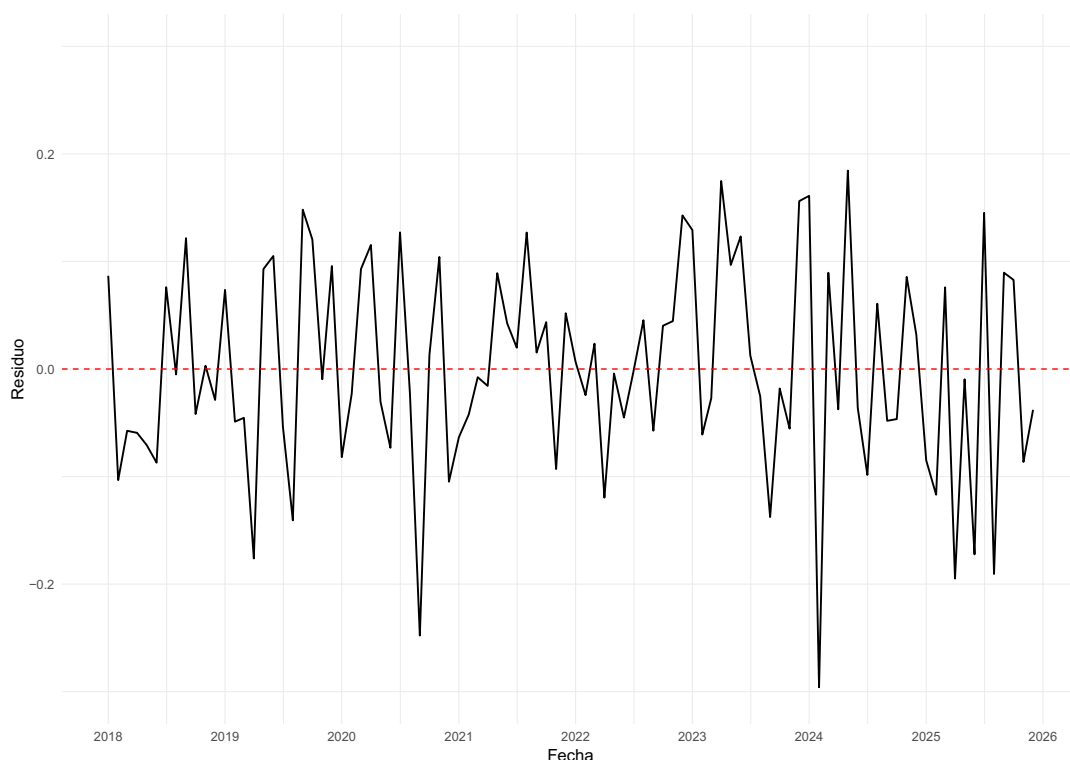
En primer lugar, se evaluó la posible presencia de autocorrelación en los residuos. El test de Durbin-Watson no permite rechazar la hipótesis nula de ausencia de autocorrelación de primer orden ($DW = 2.276$, $p = 0.906$). Este resultado es consistente con el test de Breusch-Godfrey para autocorrelación hasta 12 rezagos, el cual tampoco rechaza la hipótesis nula ($LM = 11.845$, $p = 0.458$). En conjunto, ambos contrastes no proporcionan evidencia estadísticamente significativa de dependencia temporal en los errores.

A continuación, se examinó la homocedasticidad mediante el test de Breusch-Pagan. Los resultados no permiten rechazar la hipótesis nula de varianza constante ($BP = 7.062$, $p = 0.216$), por lo que no se detectan problemas de heterocedasticidad y no resulta necesario utilizar errores robustos.

La normalidad de los residuos se evaluó mediante el test de Shapiro-Wilk, que no rechaza la hipótesis nula de normalidad al 5 % ($W = 0.978$, $p = 0.104$). Esto respalda la validez de la inferencia basada en distribuciones t y F en muestra finita.

Como complemento visual, la Figura 4.4 presenta la serie temporal de los residuos. Se observa que fluctúan alrededor de cero, sin patrones persistentes y con dispersión aproximadamente estable, comportamiento compatible con un proceso de ruido blanco. Se identifican algunos valores atípicos aislados que no alteran la dinámica general de la serie. En conjunto, estos resultados sugieren que el modelo captura adecuadamente la dinámica de los rendimientos sin evidencia de heterocedasticidad ni autocorrelación significativa.

Figura 4.4: evolución temporal de los residuos de MCO del Modelo 2, periodo 2018.01–2025.12



Fuente: elaboración propia.

En cuanto a la posible multicolinealidad entre los regresores, se calcularon los factores de inflación de la varianza (VIF). Todos los valores obtenidos son bajos —MP (1.15), UPR (1.36), dif.UI (1.13), shock.POS (1.09) y shock.NEG (1.46)— y se sitúan ampliamente por debajo de los niveles a partir de los cuales la multicolinealidad se considera problemática, lo que indica ausencia de multicolinealidad relevante en el modelo.

La correcta especificación funcional del modelo se evaluó mediante el test RESET de Ramsey. El contraste no resulta significativo ($F = 0.165$, $p = 0.848$), por lo que no se encuentra evidencia de formas funcionales inadecuadas ni de omisión sistemática de variables relevantes.

Finalmente, se analizó la estabilidad estructural del modelo y la posible presencia de observaciones atípicas. El procedimiento de Bai-Perron no identifica quiebres estructurales en los parámetros durante el período analizado⁷. Asimismo, el test de outliers con corrección de Bonferroni no detecta observaciones atípicas estadísticamente significativas al 5 %.

Bajo la formulación del teorema de Gauss-Markov, los estimadores MCO son BLUE (*Best Linear Unbiased Estimators*) cuando se verifica linealidad en parámetros, ausencia de multicolinealidad perfecta, exogeneidad condicional de los regresores, homocedasticidad e independencia de los errores. Los contrastes diagnósticos aplicados no encuentran evidencia muestral suficiente para rechazar las hipótesis de homocedasticidad, ausencia de autocorrelación ni colinealidad, y la exogeneidad resulta plausible dada la construcción de los factores. En consecuencia, no se encuentra evidencia para rechazar la hipótesis de que los estimadores obtenidos pueden considerarse BLUE dentro de la clase de estimadores lineales e insesgados.

Interpretación económica de los coeficientes

Los coeficientes estimados del modelo MCO permiten cuantificar cómo cada factor incide, *ceteris paribus*, sobre el rendimiento mensual de la cartera. Dado que las variables explicativas se encuentran expresadas en tasas de crecimiento, shocks discretos e innovaciones no anticipadas, cada parámetro mide el efecto marginal de nuevas realizaciones de información macroeconómica y financiera sobre R_t , manteniendo constantes los demás factores.

- **Intercepto:** el término constante resulta positivo y estadísticamente significativo al 1 % ($\hat{\beta}_0 \approx 0.037$, $p = 0.001$). Esto sugiere que, en ausencia de shocks en los factores incluidos, la cartera exhibe un rendimiento medio mensual cercano a 0.037 puntos porcentuales (pp).
- **MP:** el coeficiente asociado al crecimiento mensual de la producción industrial es positivo y estadísticamente significativo ($\hat{\beta} \approx 0.651$, $p = 0.011$). Dado que la variable se define como diferencia logarítmica, un aumento de 1 pp en la tasa de crecimiento mensual del

⁷Si bien a lo largo del período de estudio se registran reiterados cambios de régimen macroeconómico, ello no implica necesariamente una alteración en los coeficientes estimados. En particular, dado que las betas capturan la sensibilidad de los rendimientos ante factores expresados como componentes no anticipados, un cambio de régimen podría modificar el nivel o la varianza de dichas sorpresas, sin que ello implique, per se, un cambio en la elasticidad o respuesta marginal de los rendimientos frente a las mismas.

índice de producción se asocia, en promedio, con un incremento de aproximadamente 0.65 pp en el rendimiento de la cartera.

- **UPR:** el coeficiente estimado es negativo y significativo ($\hat{\beta} \approx -0.251$, $p = 0.002$). Como esta variable aproxima la tasa de crecimiento del riesgo soberano, un aumento de 1 punto porcentual en dicha tasa se asocia con una disminución cercana a 0.25 pp en el rendimiento de la cartera.
- **dif.UI:** la inflación no anticipada presenta un coeficiente negativo y estadísticamente significativo ($\hat{\beta} \approx -0.075$, $p = 0.017$). Esto implica que un aumento de 1 pp en la inflación inesperada se asocia con una caída aproximada de 0.08 pp en el rendimiento de la cartera.
- **shock.POS y shock.NEG:** ambos coeficientes presentan los signos esperados y son estadísticamente significativos.
 - Los episodios con impacto positivo incrementan los rendimientos en aproximadamente 0.38 pp ($\hat{\beta} \approx 0.379$, $p < 0.001$).
 - Los shocks negativos reducen los rendimientos en torno a 0.18 pp ($\hat{\beta} \approx -0.177$, $p = 0.005$).

Por todo lo anterior, la Ecuación (4.10) no solo constituye una especificación empíricamente parsimoniosa, sino que puede interpretarse como una identificación adecuada de los factores de riesgo sistemático de la Ecuación (2.1).

$$\begin{aligned}
 R_t^C &= 0.037 && \text{(rendimiento esperado)} \\
 &+ 0.65 MP_t && \text{(producción)} \\
 &- 0.25 UPR_t && \text{(riesgo país)} \\
 &- 0.08 dif.UI_t && \text{(inflación inesperada)} \\
 &+ 0.38 shock.POS_t && \text{(sorpresas político-institucionales positivas)} \\
 &- 0.18 shock.NEG_t && \text{(sorpresas político-institucionales negativas)} \\
 &+ \varepsilon_t && \text{(término de error)}
 \end{aligned} \tag{4.10}$$

En primer lugar, los resultados del análisis sugieren que la producción constituye un componente relevante en la explicación de los rendimientos agregados. Las variaciones en el nivel de actividad parecen asociarse con cambios en las expectativas de beneficios futuros y, en consecuencia, en la valuación de los activos. Este hallazgo es compatible con la idea de que el ciclo real forma parte del riesgo sistemático en economías emergentes.

En segundo término, la evidencia indica que el riesgo país se encuentra vinculado de manera significativa con la dinámica de los retornos. Esto sugiere que las variaciones en la percepción de solvencia soberana y en las condiciones de financiamiento externo podrían trasladarse al conjunto del mercado accionario, reflejando una relación entre la credibilidad y el desempeño de los activos domésticos.

Asimismo, los resultados muestran que la inflación no anticipada presenta una relación negativa con los rendimientos, lo cual es consistente con la hipótesis de que las sorpresas nominales afectan la valuación de los flujos de caja reales y elevan la incertidumbre intertemporal. Este resultado adquiere particular relevancia en una economía como la argentina, caracterizada históricamente por episodios recurrentes de alta inflación y volatilidad nominal. En ese contexto, las desviaciones respecto de la inflación esperada podrían amplificar la incertidumbre macroeconómica y trasladarse con mayor intensidad a los precios de los activos financieros.

Por último, la significancia estadística de los shocks político-institucionales sugiere que los eventos de naturaleza institucional pueden actuar como verdaderos factores sistemáticos que inciden sobre las expectativas agregadas.

En conjunto, los resultados del modelo sugieren que los rendimientos del mercado argentino se encuentran asociados a cuatro dimensiones principales⁸: el ciclo real, el riesgo soberano, la pérdida del poder adquisitivo del dinero y las sorpresas político-institucionales. Estas dimensiones parecen interactuar en la configuración del riesgo sistemático, en línea con las características macroeconómicas e históricas del país.

En este marco, una conclusión relevante que se desprende del análisis es que no serían los resultados observables de la política económica los que determinan por sí solos la dinámica

⁸Si bien el número de dimensiones identificadas coincide con la cantidad de factores latentes extraídos mediante el AFE, debe señalarse que los resultados no son directamente comparables debido a las distintas frecuencias temporales empleadas en cada procedimiento. En consecuencia, esta coincidencia podría obedecer a una regularidad subyacente, pero también podría ser meramente circunstancial.

común de los rendimientos, sino la interacción entre ellos y las expectativas formadas por los agentes. En particular, los desvíos respecto de lo anticipado —esto es, las innovaciones o shocks no previstos— parecen desempeñar un papel central en la generando desviaciones de los rendimientos bursátiles respecto de sus valores esperados.

De este modo, el riesgo sistemático no surgiría exclusivamente de los fundamentos macroeconómicos en sí mismos, sino de la medida en que estos difieren de lo esperado por el mercado. A su vez, este último aspecto constituye quizá la principal dificultad metodológica de la presente investigación: la necesidad de aproximar y reconstruir estadísticamente aquellas innovaciones que, por definición, no son directamente observables. Naturalmente, futuros trabajos podrían perfeccionar las estrategias empíricas aquí empleadas para dicha tarea.

Capítulo 5

Reflexiones finales

El objetivo de esta investigación fue contrastar la validez del modelo factorial para caracterizar los determinantes del riesgo sistemático en un conjunto de acciones argentinas con adecuada liquidez durante el período 2018–2025, bajo el supuesto de que dichos determinantes pueden vincularse con factores macroeconómicos y financieros. Para ello se emplearon dos aproximaciones complementarias: la identificación de factores latentes mediante AFE y la construcción de factores macroeconómicos a partir de innovaciones. Los resultados obtenidos permiten dar respuesta a las preguntas planteadas al inicio del trabajo.

En relación con la primera pregunta —En qué medida la variabilidad conjunta de los rendimientos de estas acciones sugiere la presencia de dimensiones subyacentes y cuántas serían necesarias para representarla mediante una estructura factorial común— las pruebas de adecuación muestral y el AFE mostraron que existe una estructura de comovimiento estadísticamente significativa. En particular, la presencia de cuatro factores latentes capaces de explicar aproximadamente el 63 % de la varianza conjunta de los rendimientos indica que una parte sustancial de la variabilidad responde a fuentes comunes de riesgo, lo que valida el uso del enfoque factorial para este conjunto de activos.

Respecto de la segunda pregunta —cuáles son esos factores comunes que definen el riesgo sistemático— el modelo de regresión especificado muestran que, en particular, las sorpresas en la actividad económica, en el riesgo soberano, los shocks político-institucionales y la pérdida del poder adquisitivo del dinero, resultaron estadísticamente relevantes, explicando en conjunto aproximadamente el 53.6 % de la variación observada. No obstante, debemos recordar que los factores capturados por el AFE no son directamente comparables con los factores macroeconómicos identificados por MCO, dadas las diferencias en las frecuencias de los datos.

En cuanto a la tercera pregunta —de que manera deben construirse los factores que determinan los rendimientos— los resultados apoyan la necesidad de trabajar con innovaciones o shocks construidos econométricamente. La estimación de factores macroeconómicos mediante regresión múltiple permitió identificar un conjunto reducido de innovaciones macroeconómicas con poder explicativo significativo sobre los rendimientos de la cartera. Esto confirma que no son los niveles de las variables macroeconómicas los que inciden directamente, sino sus componentes no anticipados.

Consideradas en conjunto, ambas aproximaciones ofrecen evidencia convergente a favor de una caracterización factorial del riesgo sistemático en el mercado accionario argentino. Por un lado, la existencia de una estructura latente robusta nos habilita a usar el modelo factorial. Por otro, la identificación de innovaciones macroeconómicas con capacidad explicativa aporta contenido económico a esa estructura común, vinculándola con dimensiones reales, financieras e institucionales.

Como líneas futuras de investigación se sugieren las siguientes: (i) profundizar el análisis abordando el problema de la generalidad de los resultados, dado que el estudio se basó en un conjunto acotado de factores macroeconómicos y en una especificación particular de los mismos. En este sentido, la ampliación del universo de factores considerados, así como la evaluación de especificaciones y combinaciones alternativas, podría aportar evidencia adicional sobre la robustez de la estructura factorial identificada. (ii) incorporar enfoques alternativos de formación de expectativas y modelos que permitan contrastar distintos mecanismos de transmisión entre las innovaciones macroeconómicas y los rendimientos de los activos financieros, enriqueciendo la interpretación económica de los resultados obtenidos. (iii) avanzar en la contrastación empírica de los supuestos fundamentales de la APT. La evaluación de estos supuestos permitiría, eventualmente, extender la estructura factorial estimada hacia un modelo de equilibrio financiero, aspecto que excede el alcance del presente trabajo.

Apéndice A

Estadísticas descriptivas y resultados complementarios

En este Apéndice se presentan, en primer lugar, las estadísticas descriptivas de los rendimientos diarios de los activos analizados, junto con los resultados del test ADF. A continuación, se reportan también las estadísticas descriptivas correspondientes a las variables macroeconómicas y financieras consideradas en el estudio, en frecuencia mensual.

Para cada conjunto de variables se incluyen medidas de tendencia central y dispersión (media, mediana y desviación estándar), así como indicadores de forma de la distribución (asimetría y curtosis), con el objetivo de caracterizar sus principales propiedades estadísticas.

Cuadro A.1: estadísticas descriptivas de los rendimientos

	Observaciones	Media (%)	Desv.Est (%)	Mínimo (%)	Máximo (%)	Mediana (%)	Asimetría	Curtosis
YPDF.BA	1949	0.25	3.59	-30.09	32.65	0.20	-0.10	9.07
GGAL.BA	1949	0.23	3.96	-63.61	24.49	0.09	-2.11	37.13
BMA.BA	1949	0.23	4.15	-62.58	25.18	0.15	-1.64	28.81
TECO2.BA	1949	0.18	3.41	-23.52	35.37	0.00	0.60	9.39
PAMP.BA	1949	0.24	3.55	-59.57	19.57	0.19	-2.24	42.50
BBAR.BA	1949	0.23	4.19	-63.57	26.98	0.17	-1.65	29.27
TGSU2.BA	1949	0.25	3.67	-44.86	27.19	0.20	-0.74	14.78
TXAR.BA	1949	0.21	3.37	-28.31	16.03	0.00	-0.47	6.35
CEPU.BA	1949	0.23	3.95	-63.90	25.06	0.13	-2.05	37.56
ALUA.BA	1949	0.23	3.33	-22.03	17.37	0.16	-0.39	4.18
EDN.BA	1949	0.20	4.04	-67.21	28.21	0.12	-1.94	42.38
BYMA.BA	1949	0.23	3.31	-57.40	19.39	0.00	-2.20	48.96
LOMA.BA	1949	0.20	3.61	-64.13	18.53	0.19	-3.09	53.47
SUPV.BA	1949	0.19	4.56	-65.02	31.43	0.00	-1.45	25.81
CVH.BA	1949	0.14	3.40	-43.17	28.14	0.00	-0.74	17.15
TRAN.BA	1949	0.24	4.02	-68.06	33.59	0.13	-2.09	46.60
CRES.BA	1949	0.22	3.50	-28.06	29.99	0.11	0.19	7.72
VALO.BA	1949	0.24	2.88	-32.45	16.79	0.00	-0.50	11.23
COME.BA	1949	0.22	3.36	-32.88	21.99	0.00	-0.20	8.01
MIRG.BA	1949	0.21	3.08	-34.46	17.89	0.07	-0.21	13.36
TGNO4.BA	1949	0.23	3.91	-43.74	33.19	0.14	-0.26	13.37

Nota: Los rendimientos fueron convertidos a porcentaje para mayor legibilidad.

Fuente: elaboración propia en base a datos de Yahoo Finance.

Cuadro A.2: test ADF sobre rendimientos diarios

Activo	ADF_stat	p_value	Estacionaria
YPFD.BA	-11.73	0.01	Si
GGAL.BA	-11.13	0.01	Si
BMA.BA	-11.72	0.01	Si
TECO2.BA	-12.14	0.01	Si
PAMP.BA	-12.71	0.01	Si
BBAR.BA	-11.89	0.01	Si
TGSU2.BA	-12.77	0.01	Si
TXAR.BA	-12.79	0.01	Si
CEPU.BA	-12.31	0.01	Si
ALUA.BA	-12.44	0.01	Si
EDN.BA	-11.63	0.01	Si
BYMA.BA	-12.35	0.01	Si
LOMA.BA	-12.11	0.01	Si
SUPV.BA	-11.36	0.01	Si
CVH.BA	-11.57	0.01	Si
TRAN.BA	-11.63	0.01	Si
CRES.BA	-13.21	0.01	Si
VALO.BA	-10.58	0.01	Si
COME.BA	-10.46	0.01	Si
MIRG.BA	-11.39	0.01	Si
TGNO4.BA	-11.68	0.01	Si

Fuente: elaboración propia en base a datos de Yahoo Finance.

Cuadro A.3: estadísticas descriptivas de los factores observables estimados

	N	Media	Desv.Est	Mínimo	Máximo	Mediana	Asimetría	Curtosis
R	96	0.04	0.15	-0.53	0.52	0.04	-0.32	2.20
MP	96	-0.00	0.04	-0.21	0.19	-0.00	-1.09	11.62
YP	96	-0.01	0.11	-0.41	0.45	-0.01	0.35	3.39
dif.YP	96	-0.00	0.06	-0.20	0.23	-0.00	0.12	3.95
UI	96	0.08	0.33	-0.61	1.28	0.03	0.55	0.60
DEI	96	0.02	0.24	-0.32	1.57	-0.01	3.08	15.78
dif.UI	96	-0.00	0.35	-1.14	0.93	-0.01	-0.11	1.10
UPR	96	0.01	0.15	-0.43	0.66	0.01	0.86	3.91
UFX	96	0.00	0.04	-0.09	0.17	-0.00	2.04	7.66
DEFX	96	0.04	0.14	-1.00	0.55	0.03	-3.34	32.53
dif.DEFX	96	-0.01	0.16	-1.01	0.45	-0.00	-2.36	17.75
V.SUP	96	-0.16	2.96	-12.50	12.80	-0.10	-0.20	6.97
V.ENERGY	96	0.00	0.09	-0.44	0.22	0.00	-1.10	5.02
V.NON.ENERGY	96	0.00	0.03	-0.09	0.07	0.00	-0.29	1.35
V.AGRICULTURE	96	0.00	0.02	-0.08	0.07	0.00	-0.08	0.60

Nota: se presentan estadísticas descriptivas de las variables continuas utilizadas para la identificación de factores macroeconómicos. Las variables binarias fueron excluidas.

Fuente: elaboración propia.

Cuadro A.4: test ADF sobre factores macroeconómicos

Variable	ADF_stat	p_value	Estacionaria
R	-4.3746	0.0100	Sí
MP	-4.7431	0.0100	Sí
YP	-2.5658	0.3426	No
dif.YP	-4.6155	0.0100	Sí
UI	-3.0543	0.1408	No
dif.UI	-6.5468	0.0100	Sí
DEI	-4.7277	0.0100	Sí
UPR	-4.7135	0.0100	Sí
UFX	-5.4130	0.0100	Sí
DEFX	-2.0445	0.5581	No
dif.DEFX	-3.9568	0.0145	Sí
V.SUP	-5.2398	0.0100	Sí
V.ENERGY	-3.7353	0.0252	Sí
V.NON.ENERGY	-3.4702	0.0487	Sí
V.AGRICULTURE	-4.1592	0.0100	Sí

Fuente: elaboración propia.

Apéndice B

Endogeneidad de las variables dummies

Una crítica válida al modelo de regresión con factores macroeconómicos presentado (modelo 2) es la posible presencia de endogeneidad, dado que, si bien la elección de variables dummies se realiza ex ante, su clasificación entre eventos con impacto positivo (shock.POS) y eventos con impacto negativo (shock.NEG) podría generar problemas de endogeneidad si, de manera implícita, estas variables incorporaran información de los rendimientos observados. Esto refleja, en parte, la dificultad de construir factores que capturen de manera objetiva sorpresas de índole política, ya que los efectos de ciertos eventos pueden ser anticipados o interpretados de distinta forma por el mercado.

Para evaluar esta preocupación se realizó un chequeo ex post consistente en extraer los residuos del modelo y analizar su relación con las dummies de eventos. En particular, se estimaron regresiones auxiliares de los residuos sobre las variables shock.POS y shock.NEG para verificar significancia estadística. Los resultados obtenidos se presentan en el Cuadro B.1:

Cuadro B.1: regresión MCO de residuos sobre dummies

Variable	Estimate	Std. Error	t value	Pr(> t)
<i>Regresión: residuos ~ shock.POS</i>				
(Intercept)	1.34e-17	1.021e-02	0.00	1.000
shock.POS	-4.96e-17	5.002e-02	0.00	1.000
Residual SE: 0.09794 (94 gl)	$R^2 = 8.144e-32$	Ajustado = -0.01064	F = 7.655e-30	p = 1.000
<i>Regresión: residuos ~ shock.NEG</i>				
(Intercept)	1.25e-17	1.021e-02	0.00	1.000
shock.NEG	-2.84e-17	5.002e-02	0.00	1.000
Residual SE: 0.09794 (94 gl)	$R^2 = 4.313e-31$	Ajustado = -0.01064	F = 4.054e-29	p = 1.000

Fuente: elaboración propia.

Como puede observarse, los valores-p de las regresiones auxiliares son iguales a 1, lo que indica que no hay evidencia empírica de que estas variables estén correlacionadas con los residuos. Esto respalda su inclusión como controles discretos para capturar perturbaciones

extraordinarias, sin que su presencia genere sesgo en la estimación de los coeficientes de los factores macroeconómicos. Si bien este análisis no elimina la discusión teórica sobre la definición de las dummies, permite afirmar que, dentro del modelo, su utilización no compromete la validez de los resultados.

Bibliografía

- Burmeister, E., & Wall, K. D. (1986). The Arbitrage Pricing Theory and Macroeconomic Factor Measures. *Financial Review*, 21(1), 1-20.
- Cagnetti, A. (2001). *Capital Asset Pricing Model and Arbitrage Pricing Theory in the Italian Stock Market: An Empirical Study* [Tesis doctoral, University of Edinburgh].
- Chamberlain, G., & Rothschild, M. (1982). *Arbitrage, factor structure, and mean-variance analysis on large asset markets* (inf. téc. N.º 996). National Bureau of Economic Research.
- Chan, K. (1985). An exploratory investigation of the firm size effect. *Journal of Financial Economics*, 14(3), 451-471.
- Chen, N. (1983). Some Empirical Tests of the Theory of Arbitrage Pricing. *The Journal of Finance*, 38(5), 1393-1414.
- Chen, N., Roll, R., & Ross, S. A. (1986). Economic Forces and the Stock Market: Testing the APT and Alternative Asset Pricing Theories. *Journal of Business*, 59(3), 383-403.
- Cochrane, J. H. (2005). *Asset Pricing* (Revised). Princeton University Press.
- Connor, G., & Korajczyk, R. A. (1988). *Risk and return in an equilibrium APT: Application of a new test methodology* (inf. téc. N.º 9). University of California, Berkeley; Northwestern University.
- Cortés, R. L., & Porras, S. T. (2013). Estimation of the underlying structure of systematic risk with the use of principal component analysis and factor analysis. *Contaduría y Administración*, 59(3), 197-234.
- Damodaran, A. (2012). *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset* (3.ª ed.). John Wiley & Sons.
- 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.

- Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, 81(3), 607-636.
- Fuentes, R., S. M., Gregoire C., J., & Zurita L., S. (2017). Factores macroeconómicos en rendimientos accionarios chilenos. *El Trimestre Económico*, 73(289), 125-138.
- Groenewold, N., & Fraser, P. (1997). Share Prices and Macroeconomic Factors. *Journal of Business Finance & Accounting*, 24(9-10), 1367-1383.
- Guerra, A. L. (2012). *Rendimiento de las acciones en contextos inflacionarios: Análisis empírico del mercado accionario argentino* [Programa de Formación 2012, Bolsa de Comercio de Rosario].
- Gujarati, D. N., & Porter, D. C. (2010). *Econometría* (5.^a ed.). McGraw-Hill/Interamericana Editores.
- Karan. (2010, abril). Autoregressive Multifactor APT Model for U.S. Equity Markets [Posted 22 June 2010 08:35 UTC].
- Lehmann, B., & Modest, D. (1985). *The Empirical Foundations of the Arbitrage Pricing Theory I: The Empirical Tests* (inf. téc. N.º w1725). National Bureau of Economic Research.
- López Herrera, F. (2006). Riesgo sistemático en el mercado mexicano de capitales: Un caso de segmentación parcial. *Contaduría y Administración*, (219), 109-140.
- Marín, J. M., & Rubio, G. (2011). El modelo de valoración de activos financieros bajo ausencia de arbitraje: el APT. En *Economía financiera*. Antoni Bosch Editor.
- Markowitz, H. M. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77-91.
- McElroy, M. B., & Burmeister, E. (1988). Arbitrage Pricing Theory as a Restricted Nonlinear Multivariate Regression Model Iterated Nonlinear Seemingly Unrelated Regression Estimates. *Journal of Business & Economic Statistics*, 6(1), 29-42.
- Peña, D. (2002). *Análisis de Datos Multivariantes*. McGraw-Hill.
- Pesce, G., Pedroni, F., Esandi, J. I., & Milanese, G. Teoría de fijación de precios por arbitraje: Evidencias del mercado de capitales argentino. En: *En XXXV Jornadas Nacionales de Administración Financiera*. Universidad Nacional del Sur. 2015.

- Pesce, G., & Pedroni, F. V. (2021). Inflación y rendimientos en mercados emergentes: El caso de Argentina. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 32, 341-375.
- Rayón, E. L. (2016). Modelo multifactorial APT para el análisis de los factores de riesgo macroeconómico a los que se exponen los hedge funds [Trabajo de universidad, Universidad de Las Américas Puebla]. *EconoQuantum*, 14(1).
- Roll, R., & Ross, S. A. (1980). An Empirical Investigation of the Arbitrage Pricing Theory. *The Journal of Finance*, 35(5), 1073-1103.
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3), 341-360.
- Roy, A. D. (1952). Safety First and the Holding of Assets. *Econometrica*, 20(3), 431-449.
- Shanken, J. (1982). The Arbitrage Pricing Theory: Is it Testable? *The Journal of Finance*, 37(5), 1129-1140.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425-442.
- Shukla, R., & Trzcinka, C. (1990). Sequential Tests of the Arbitrage Pricing Theory: A Comparison of Principal Components and Maximum Likelihood Factors. *The Journal of Finance*, 45(5), 1541-1564.
- Swoboda, C. (2002). *Teoría del arbitraje de precios: Una investigación empírica para la Argentina* [Versión al 23/11/02. Trabajo seleccionado para ser expuesto durante el III Encuentro Internacional de Finanzas organizado por la Facultad de Administración y Economía de la Universidad de Santiago de Chile].