



UNIVERSIDAD  
DEL CEMA

**UCEMA**

Tesis de Maestría en Finanzas

**Estrategia secreta para ganarle al mercado de capitales de Estados Unidos basada en un modelo probit instrumental y estimaciones por máxima verosimilitud**

Tutor:  
Prof. Dr. José Pablo Dapena

Candidato:  
David Quiroz Sillo

Año Académico 2021 / 2022



Declaro que el abajo firmante, así como el autor del documento es responsable de su contenido, y de las partes tomadas de otras obras, estas se declaran expresamente citando las fuentes.

## **Dedicatoria**

*Dedicado a mi esposa **Sara** y mis bebes por sacrificar su tiempo conmigo y permitir que su tiempo lo emplee en esta maestría ¡Muchas gracias!*

## Resumen

Es evidente que, en periodos recesivos de la economía, tanto la incertidumbre como el riesgo se incrementan en el mercado de capitales. De hecho, en estos periodos en la mayoría de los casos se registran pérdidas elevadas, por tanto, no realizar inversiones en estos periodos mejoraría los retornos y minimizaría el riesgo para los inversionistas. El presente trabajo de investigación plantea una estrategia de inversión basado en un modelo de pronóstico de recesiones para lograr un mayor retorno y menor riesgo en comparación al retorno del mercado de capitales de Estados Unidos. El método de pronóstico utilizado es un modelo probit instrumental que marca la estrategia y se probaron los resultados de la estrategia con un modelo de valoración de activos financieros con estimaciones de máxima verosimilitud. Después de verificar la hipótesis, el empleo de la estrategia basada en un modelo de pronóstico no lineales de las recesiones de Estados Unidos, se concluye que esta estrategia alcanza un mayor retorno y menor riesgo en el mismo periodo de tiempo comparada con el retorno y riesgo observado en el mercado de capitales.



# Índice

<b>1. Introducción</b>	<b>6</b>
<b>2. Tema de tesis</b>	<b>8</b>
<b>3. Planteamiento del problema</b>	<b>8</b>
3.1. Formulación del problema . . . . .	11
<b>4. Objetivos</b>	<b>12</b>
4.1. Objetivo general . . . . .	12
4.2. Objetivos específicos . . . . .	12
<b>5. Marco teórico</b>	<b>12</b>
5.1. Modelo de valoración de activos financieros . . . . .	12
5.2. Estimación del CAPM . . . . .	13
<b>6. Marco referencial</b>	<b>14</b>
6.1. El modelo de pronóstico probit . . . . .	14
6.2. Estrategias de inversión . . . . .	16
<b>7. Planteamiento de la hipótesis</b>	<b>17</b>
7.1. Hipótesis . . . . .	18
<b>8. Marco metodológico</b>	<b>19</b>
8.1. Modelo de pronóstico probit . . . . .	19
8.1.1. Especificación . . . . .	19
8.2. Modelo de pronóstico probit . . . . .	19
8.2.1. Estimación . . . . .	19
8.3. Modelo de regresión lineal . . . . .	23
<b>9. Métodos y técnicas de recolección de datos</b>	<b>26</b>
9.1. Datos para el modelo econométrico de pronóstico de las recesiones económicas . . . . .	26
9.2. Datos para el modelo econométrico de estimación del modelo CAPM . . . . .	28
<b>10. Resultados</b>	<b>28</b>
10.1. Resultados del modelo de pronóstico probit . . . . .	28
10.1.1. Evaluación fuera de muestra . . . . .	30
10.2. Resultados del modelo de regresión lineal simple . . . . .	31
10.2.1. Contrastación de la hipótesis . . . . .	32
10.2.2. Retornos anuales . . . . .	33
<b>11. Conclusiones</b>	<b>33</b>
<b>12. Apéndice 1: Cálculo de la anticipación de la propuesta de (Harvey, 2011)</b>	<b>37</b>
<b>13. Apéndice 2: Resultado de estimación del modelo probit</b>	<b>38</b>

14. Apéndice 3: Curva ROC y Bondad de Ajuste	39
15. Apéndice 4: Resultado de estimación del modelo CAPM	40
16. Apéndice 5: Matriz de correlaciones	42
17. Autorización para publicar los trabajos finales	43

## Índice de cuadros

1. Predicción y realización . . . . .	21
2. Tabla de contingencia . . . . .	23
3. Comparación de resultados: febrero de 1985 a julio de 2022 . . . . .	33

## Índice de figuras

1. Retornos históricos S&P 500 . . . . .	10
2. Recesiones y probabilidades dentro de la muestra . . . . .	30
3. Recesiones y probabilidades fuera de muestra . . . . .	31
4. Fechas antes que se observe una recesión económica . . . . .	37
5. Días antes que se observe una recesión económica en Estados Unidos . . . . .	37



## 1. Introducción

Una de las variables macroeconómicas más angustiantes en el comportamiento de la economía es el crecimiento de la misma (se mide por la variación de un periodo a otro de la producción total en un país a través del Producto Interno Bruto (PIB)). Particularmente, la economía de Estados Unidos ha registrado en el segundo trimestre de 2022 una variación interanual del PIB de 1,8 % inferior a la observada en el primer trimestre (3,7 %) <sup>1</sup>. Asimismo, el spread de los bonos del tesoro de Estados Unidos a 10 años con vencimiento constante menos los bonos del tesoro a 2 años con vencimiento constante se encuentra en valores negativos desde julio de 2022 <sup>2</sup>. Este spread muestra que los bonos a menor plazo tienen tasas de interés más elevadas que los bonos de mayor plazo, evidenciando que las curvas de rendimiento se han invertido. (Harvey, 2011) evidenció que la curva de rendimientos se invierte durante al menos tres meses antes de que se observe una contracción económica (variación negativa del PIB). Por tanto, se podrá esperar una recesión económica en Estados Unidos <sup>3</sup> en el tercer trimestre de 2022.

No me cabe la menor duda de que una recesión económica tiene un impacto en el comportamiento del mercado de capitales debido a que provoca un incremento de incertidumbre y riesgos a los que afrontan no solo las empresas que cotizan en la bolsa de valores, sino también en el conjunto de empresas en su totalidad y nuevos proyectos de inversión. Por ejemplo, si consideramos los retornos promedios mensuales del S&P500 como referente del mercado de capitales en periodos de recesión versus periodos de crecimiento de la economía de Estados Unidos, se observa que en el periodo de 1985 hasta 2022 para los meses en los que la economía norteamérica se encontraba en recesión el retorno promedio mensual fue -2,1 % y el desvío estándar muestral fue de 6,5 % frente a un retorno mensual en periodos no recesivos fue de 1 % y con desvío estándar muestral de 4,1 %. Estos últimos datos muestran que las contracciones del PIB impactan de forma importante en los retornos del mercado de capitales e incrementan el riesgo.

A simple vista, con los datos mencionados anteriormente, se puede determinar una estrategia de inversión que lograría mejores retornos y menor riesgo. Esta estrategia podría ser no participar en el mercado de capitales en periodos recesivos de la economía de Estados Unidos para evitar

---

<sup>1</sup><https://fred.stlouisfed.org/series/GDPC1>

<sup>2</sup><https://fred.stlouisfed.org/series/T10Y2Y>

<sup>3</sup>De acuerdo a National Bureau of Economic Research <https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions> la economía de Estados Unidos ha a travesado por 34 contracciones económicas desde 1854 hasta 2020.

encontrarse en momentos de incremento de la incertidumbre y los riesgos. Sin embargo, surge un problema no resuelto a la fecha, que pronosticar las recesiones económicas no es posible con precisión utilizando modelos actuales con diferentes metodologías.

Otra forma interesante es la propuesta de (Harvey, 2011) que mostró con evidencia empírica que cuando el spread <sup>4</sup> es negativo, anticipa una crisis económica. Aunque también carece de precisión, ya que calculando el tiempo en días este es muy heterogéneo (un desvío estándar muestral elevado) debido a que el promedio de tiempo es de 18 meses con un desvío estándar de 10 meses <sup>5</sup>. De este modo impide tener precisión de cuando inicia una recesión económica en Estados Unidos por el intervalo de confianza amplio de 2 meses y 37 meses suponiendo una distribución normal al 95 % de confianza para el tiempo transcurrido desde que el spread es negativo hasta que se observa la tasa de variación del PIB con un valor negativo.

Por último, la publicación del PIB tiene un trimestre de rezago, eso significa que es difícil saber con certeza si en el tiempo contemporáneo la economía norteamericana está en recesión. Por lo tanto, la estrategia tendría un rezago y los inversores no podrán aplicar esta estrategia de salir del mercado de capitales en momentos de crisis económica en el mismo periodo de tiempo. En definitiva, los inversores se encontrarían invirtiendo en el mercado de capitales de igual forma en periodos recesivos y no recesivos debido a que confirmarían este suceso un trimestre después.

Frente a todos estos hechos, resulta de interés indagar si existe alguna forma de construir una estrategia de inversión que permita a los inversionistas no encontrarse invirtiendo en el mercado de capitales de Estados Unidos en el momento que se esté atravesando una recesión económica, aunque este periodo de rezago sea de un mes como máximo. Para realizar este análisis se utilizó el modelo de valoración de activos financieros (CAPM <sup>6</sup>, por su siglas en inglés, (W. F. Sharpe, 1964); Lintner, 1969; Jarrow, 1988; A. Sharpe, 1999) y un marco metodológico con el modelo probit instrumental. Es importante destacar que la literatura sobre pronóstico de las recesiones económicas en Estados Unidos es amplia como se puede mencionar algunos trabajos de investigación de los siguientes autores (Del Negro y Schorfheide, 2011; Edge et al., 2010; Karlsson, 2013; Wieland y Wolters, 2011; Babii et al., 2022).

En lo referente a la verificación de la hipótesis se empleo el modelo de regresión lineal clásico

---

<sup>4</sup>La diferencia entre de los bonos del tesoro de Estados Unidos a 10 años con vencimiento constante menos los bonos del tesoro a 2 años con vencimiento constante.

<sup>5</sup>Para ver detalles ver Apéndice 1. Cabe aclarar que todos los cálculos se realizaron en Python.

<sup>6</sup>Capital Asset Pricing Model (CAPM).

estimado por máxima verosimilitud (Wooldridge, 2010) para verificar si la estrategia propuesta logra mejores resultados que el mercado de capitales de Estados Unidos (la variable que se utilizó como representativa al mercado fue el S&P500). Finalmente, en el marco referencial se destacan algunos trabajos al respecto debido a que de la misma manera que en los modelos de pronósticos existe una enorme cantidad de trabajos que plantean diferentes estrategias para la compra y venta de acciones en Estados Unidos, las mismas se pueden clasificar en dos inmensos grupos: las estrategias activas y pasivas. Estas referencias permiten clasificar esta estrategia de inversión propuesta por este trabajo de investigación como una estrategia híbrida debido a que plantea realizar inversiones a largo plazo, pero también con base en las series de tiempo macroeconómicas sugerir una salida del mercado de capitales como también el ingreso al mismo.

En lo que sigue, la sección 2 se plantea el tema de investigación; la sección 3 se explica cómo se planteó el problema de estudio; la sección 4 se establecen los objetivos, tanto el objetivo general como los objetivos específicos; la sección 5 ofrece el marco teórico del estudio; la sección 6 describe tanto la literatura para modelos de pronósticos como para estrategias de inversión; la sección 7 plantea la hipótesis de la investigación; la sección 8 desarrolla la metodología utilizada para el modelo de pronóstico y la estimación; la sección 9 muestra la procedencia de los datos; la sección 10 presenta los hallazgos del estudio para el pronóstico y la prueba de hipótesis con la evidencia empírica; y la sección 11 proporciona conclusiones basadas en los resultados de la investigación.

## 2. Tema de tesis

De acuerdo a lo planteado anteriormente el tema de tesis es:

**Estrategia secreta para ganarle al mercado de capitales de Estados Unidos basada en un modelo probit instrumental y estimaciones por máxima verosimilitud.**

## 3. Planteamiento del problema

El crecimiento de la economía de un país <sup>7</sup> (medido como la variación positiva de la cantidad de producción en un año) en su conjunto tiene un comportamiento cíclico. Eso quiere decir que

---

<sup>7</sup>El crecimiento de la economía se mide con el producto interno bruto (PIB) que es una medida que se obtiene de las cuentas nacionales en la economía de un país que registra la cantidad de producción en un año.

atraviesa por periodos de auge cuando el crecimiento de la producción registra valores positivos y recesión cuando el crecimiento de la producción registra valores negativos. Considerando este hecho, el trabajo de investigación encontró tres situaciones claves para resolver con el objetivo de obtener retornos superiores con menores riesgos en promedio a los que logra en conjunto el mercado de capitales de Estados Unidos, en otras palabras, para ganarle al mercado.

La primera situación es la dificultad que se tiene a la fecha en pronosticar las recesiones económicas. La capacidad de pronóstico en tiempo real para el crecimiento del producto de Estados Unidos con modelos lineales y no lineales estructurales y de forma reducida, y los pronósticos de juicio sobre el crecimiento de la producción (*Blue Chip*)<sup>8</sup> todos estos modelos y metodologías mostrarán tener mayor error de estimación en pronosticar el crecimiento del producto durante las recesiones que durante las expansiones. En conclusión, la combinación de modelos de pronóstico del crecimiento de la producción durante las expansiones es mas preciso, pero tiene un desempeño deficiente durante las recesiones. Esto se explica porque estos modelos no consideran el momento y la intensidad de las recesiones, particularmente, la volatilidad del crecimiento del PIB, que se acentúa particularmente durante las recesiones. (Ver Del Negro y Schorfheide, 2011, Edge et al., 2010, Karlsson, 2013 y Wieland y Wolters, 2011). De hecho, uno de los últimos métodos basados en *Machine Learning* muestran las mismas dificultades en pronosticar las recesiones con la utilización de algoritmos y una mayor cantidad de variables macroeconómicas, financieras y otras (Babii et al., 2022). En definitiva, es claro que no se puede pronosticar las recesiones económicas con precisión.

La segunda situación es que en la mayoría de los casos los retornos en promedio observados son negativos en el mercado de capitales en periodos de recesiones económicas frente a periodos de auge (expansión económica) donde en promedio los retornos son positivos. Es más, la incertidumbre, medida por la volatilidad del retorno de las acciones, tiende a crecer de manera importante. Por ejemplo, en la serie de tiempo de los retornos de las acciones (S&P500 se utilizó como referente del mercado en este trabajo de investigación) en el periodo de 1985 hasta 2022 de forma diaria, se observa que los periodos de crisis económica el retorno en promedio diario

---

<sup>8</sup>Los indicadores Blue Chip son una encuesta de alrededor de los 50 principales economistas de pronóstico de bancos, industrias manufactureras, firmas de corretaje y compañías de seguros. La encuesta se lleva a cabo desde 1976 y comprende varias series macro, incluido el crecimiento del PIB real. La encuesta contiene pronósticos de cada miembro y el promedio (o consenso) de sus pronósticos. También proporciona el promedio de los 10 pronósticos más altos y los 10 más bajos para cada variable, y el pronóstico mediano. Finalmente, también publica un índice de difusión que refleja los cambios en las expectativas que pueden ocurrir antes de cambios en el pronóstico de consenso.

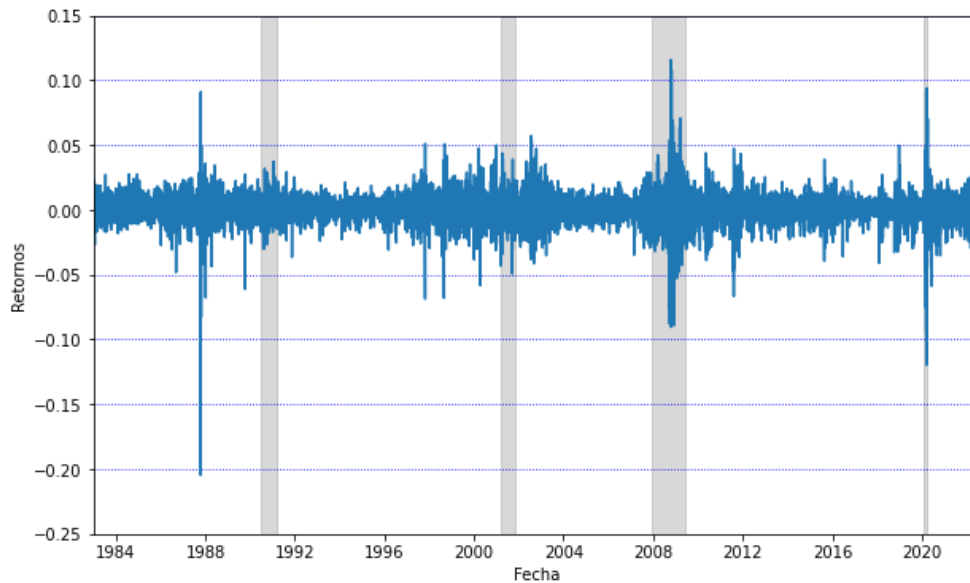


Figura 1: Retornos históricos S&P 500

es negativo alcanzando a -2,1% con un desvío estándar muestral de 6,5 %, mientras que en auge económico el retorno es positivo (1% promedio diario) con desvío estándar muestral de 4,1%. Todo esto evidencia que cuando la economía está en recesión los retornos del mercado de capitales generan pérdidas y aumentan los riesgos a los que se observa en periodos donde no se registran periodos de disminución de la producción (ver Figure 1). Adicionalmente, con datos mensuales se observa que los retornos anuales del S&P500 totalizan 9,8 % con un desvío estándar muestral de 15,2 %, mientras que cuando se excluye los periodos donde se registran recesiones económicas este retorno anual alcanza a 12 % y con un desvío de estándar muestral 13,5 %. Estos datos revelan que cuando se excluye los retornos del S&P500 en periodos recesivos el retorno se incrementa en 2,2 % y riesgo disminuye en 1,7 %. En consecuencia, no cabe la menor duda que en los periodos recesivos incorporan mayor riesgo al mercado de capitales por la incertidumbre y generan un menor retorno por la magnitud de las pérdidas observadas en estos periodos.

La tercera situación es el rezago que tiene la publicación de la tasa de crecimiento económico en Estados Unidos. Claramente, no es posible contar con toda la información del comportamiento de la economía en el mismo periodo de tiempo (contemporáneamente) por metodologías y oportunidad de la captura de información de acuerdo a manuales de cuentas nacionales <sup>9</sup>. Por ejemplo, en Estados Unidos la publicación del crecimiento de la economía (PIB) se realiza con un

<sup>9</sup><https://www.imf.org/external/pubs/ft/qna/>

trimestre de rezago <sup>10</sup>; aunque los datos se actualizan cada mes, pero esto es solo para corregir los datos preliminares publicados un trimestre anterior. Por otra parte, existe una gran cantidad de indicadores líderes<sup>11</sup> que buscan determinar de forma contemporánea y adelantada el comportamiento del PIB. Sin embargo, solo muestran si la economía de Estados Unidos se encuentra en una posible recesión o expansión económica sin mostrar la magnitud de la variación del PIB y mucho peor no señalan con precisión cuando inicia un periodo recesivo. De la misma manera la propuesta de (Harvey, 2011)<sup>12</sup> se podría clasificar como un indicador líder adelantado de una crisis económica, pero tiene el mismo problema que cualquier indicador líder: no da con precisión el inicio de una recesión económica en Estados Unidos. Particularmente en el periodo 1973 hasta 2020 (de acuerdo a disponibilidad de información) se encontró el tiempo que transcurre cuando el spread <sup>13</sup> se vuelve negativo en promedio en meses es de 18 meses con un desvío muestral de 10 meses donde el intervalo de confianza suponiendo una distribución normal al 95 % del tiempo es de 2 meses a 37 meses. Además, no muestran tampoco con claridad cuando acaba una recesión económica. Este resultado verifica la imprecisión del pronóstico de una recesión económica medida en meses. Por tanto, con lo observado anteriormente se puede concluir que existe un conjunto de señales de una recesión económica en Estados Unidos, pero ninguna muestra con exactitud su inicio ni el fin de la recesión; y que los datos publicados tienen un trimestre de rezago, imposibilitando establecer una estrategia de inversión que permita al inversionista encontrarse fuera del mercado de capitales en los momentos de aumento de la incertidumbre y riesgos.

### 3.1. Formulación del problema

De acuerdo a lo planteado en la anterior sección, el problema de investigación es:

**¿Cómo se puede construir un modelo de pronóstico de las recesiones económicas de Estados Unidos que permita contar una estrategia de inversión que logre mejores retornos y menor riesgo que el mercado de capitales?**

<sup>10</sup><https://www.bea.gov/data/gdp/gross-domestic-product>

<sup>11</sup><https://www.conference-board.org/topics/us-leading-indicators>

<sup>12</sup>Para detalles de la propuesta de (Harvey, 2011) ver apéndice 1.

<sup>13</sup>La diferencia entre de los bonos del tesoro de Estados Unidos a 10 años con vencimiento constante menos los bonos del tesoro a 2 años con vencimiento constante.

## 4. Objetivos

### 4.1. Objetivo general

Construir un modelo de pronóstico de la recesiones económicas de Estados Unidos que permita contar con una estrategia de inversión que logre mejores retornos y menor riesgo que el mercado de capitales.

### 4.2. Objetivos específicos

- Construir un modelo de pronóstico no lineales de las recesiones económicas en los Estados Unidos con el mayor grado de precisión posible.
- Plantear una estrategia con la variable indicadora que se obtiene del modelo de pronóstico que me muestre los periodos que se debe invertir o no en el mercado de capitales de Estados Unidos.
- Verificar si la estrategia alcanza retornos superiores y riesgos menores al mercado de capitales con un modelo de valoración de activos financieros.

## 5. Marco teórico

### 5.1. Modelo de valoración de activos financieros

Dada una serie de suposiciones que se pueden encontrar en las referencias (ver W. F. Sharpe, 1964; Black, 1972; Jensen, 1968; Lintner, 1969; Jarrow, 1988; A. Sharpe, 1999) el CAPM <sup>14</sup> concluye que el retorno de los activos tiene una relación lineal con los retorno del mercado. Específicamente, dado el retorno de todas las acciones que constituyen un mercado denotado como,  $M$ , y el de retorno de un activo libre de riesgo denominado,  $R_f$ . El CAPM establece que el retorno de cada activo,  $R(i)$ , en el mercado tiene la forma esperada:

$$E[R(i)] = R_f + \beta(i) * (E[M] - R_f) \quad (1)$$

<sup>14</sup>Capital Asset Pricing Model (CAPM), por su siglas en inglés, que es el modelo de valoración de activos financieros. El modelo de valoración de activos financieros donde sus bases fueron planteadas por (Markowitz, 1952) y (Tobin, 1958).

De la ecuación (1), para activos,  $i = 1, \dots, n$ , donde,  $\beta(i)$ , es un parámetro que especifica el grado de exposición entre un activo dado y el mercado subyacente. En palabras, el retorno esperado de cada activo es igual al retorno de un activo sin riesgo más un retorno de mercado esperado ajustado al riesgo neto del activo libre de riesgo. Los parámetros,  $\beta(1), \dots, \beta(n)$ , se denominan betas.

Teniendo en cuenta que la beta de un activo tiene la forma:

$$\beta(i) = \text{cov}(R(i), M) / \text{var}(M) \quad (2)$$

De la ecuación (2), donde,  $\beta$  es el cociente de la covarianza entre los retorno del activo y del mercado dividido por la varianza de los retorno del mercado. Por su parte, la interpretación de este parámetro es: si un activo tiene una beta igual a 1, dice que el,  $\beta$ , del activo se mueve con el mercado; si un activo tiene una beta mayor a 1, dice que el,  $\beta$ , del activo es más volátil que el mercado; y si un activo tiene una beta inferior a 1, se dice que el,  $\beta$ , del activo es menos volátil que el mercado.

## 5.2. Estimación del CAPM

La forma estándar del modelo CAPM para la estimación es un modelo lineal con parámetros adicionales para cada activo para caracterizar los errores residuales. Para cada uno de,  $n$ , activos con,  $m$ , muestras de retornos de activos observados,  $R(k, i)$ , retornos de mercado,  $M(k)$ , y retornos de activos sin riesgo,  $Rf(k)$ , el modelo de estimación tiene la forma:

$$R(k, i) = \alpha(i) + Rf(k) + \beta(i) * (M(k) - Rf(k)) + V(k, i) \quad (3)$$

De la ecuación (3), para muestras,  $k = 1, \dots, m$ , y activos,  $i = 1, \dots, n$ , donde,  $\alpha(i)$ , es un parámetro que especifica el retorno no sistemático de un activo,  $\beta(i)$ , es el activo beta, y,  $V(k, i)$ , es el error residual para cada activo con variable aleatoria asociada,  $V(i)$ .

Los parámetros,  $\alpha(1), \dots, \alpha(n)$ , se denominan activos alfa para medir el rendimiento anormal (Jensen, 1968). La forma estricta del CAPM especifica que los alfas deben ser cero y que las desviaciones de cero son el resultado de desequilibrios temporales. Sin embargo, en la práctica los activos pueden tener alfas distintas de cero, donde gran parte de la gestión activa de inversiones



se dedica a la búsqueda de activos con alfas distintos de cero explotables. Para permitir la posibilidad de alfas distintos de cero, el modelo de estimación generalmente busca estimar alfas y realizar pruebas para determinar si los alfas son estadísticamente iguales a cero. Por su parte, los errores residuales,  $V(i)$ , se supone que tienen momentos:

$$E[V(i)] = 0 \quad (4)$$

y

$$E[V(i) \times V(j)] = S(i, j) \quad (5)$$

De las ecuaciones (4 y 5), para activos,  $i, j = 1, \dots, n$ , donde los parámetros,  $S(1, 1), \dots, S(n, n)$ , se denominan varianzas y covarianzas residuales o no sistemáticas. La raíz cuadrada de la varianza residual de cada activo, es decir,  $\sqrt{S(i, i)}$ , para,  $i = 1, \dots, n$ , se dice que es el riesgo residual o no sistemático del activo, ya que caracteriza la variación residual en los precios de los activos que no puede ser explicada por variaciones en los precios de mercado.

Para finalizar, también hay trabajos de investigación que muestran la vigencia en la actualidad del modelo CAPM ( Markowski, 2020). Por su parte, (Wang et al., 2022 ) prueba que la relación lineal con una estimación no paramétrica es robusta entre el coeficiente de riesgo sistemático, beta y el rendimiento esperado de las acciones. En definitiva, la forma de corrección de la estimación que podría presentar sesgo en el beta se probó con estimaciones del método generalizado de momentos (GMM, por sus siglas en inglés) corrigiendo el error en variables (Jegadeesh et al., 2019). Por tanto, hay bastantes estudios empíricos que mantienen la vigencia del modelo CAPM.

## 6. Marco referencial

### 6.1. El modelo de pronóstico probit

De acuerdo a (Ng, 2012), utiliza un conjunto de factores de riesgo de recesión y los modelos probit dinámicos avanzados para pronosticar recesiones en EE. UU. Los factores de riesgo incluyen las expectativas del mercado financiero de una perspectiva económica sombría, el riesgo de liquidez en el economía general, el riesgo de un efecto riqueza negativo como resultado del estallido de burbujas de precios de activos, y señales de fundamentos macroeconómicos frágiles. Sus con-

clusiones fueron: los factores de riesgo propuestos combinados con las diversas especificaciones del modelo probit pueden pronosticar la duración de las recesiones con mayor precisión que los modelos estáticos ordinarios con solo el diferencial de retorno y el retorno del mercado de valores como predictores; y mientras que los modelos probit dinámicos avanzados superan al modelo estático en términos de predicción de la duración de las recesiones, el modelo probit estático con los factores de riesgo propuestos es tan bueno como los modelos probit dinámicos con respecto a la previsión de los picos de los ciclos económicos. Sin embargo, sus pronósticos fuera de muestra tienen a señalar de 1 a 3 meses antes de lo que se observa en los datos, mostrando dificultad de establecer un pronóstico con precisión.

(Xinyu y David, 2020) examinaron el poder predictivo de diferentes variables y utilizarlas para pronosticar la probabilidad de una futura recesión en Estados Unidos en tres horizontes de pronóstico diferentes. En total estudiaron 20 variables entre indicadores bursátiles, variables macroeconómicas, tasas de interés y spreads. Estas variables se usaron en un modelo de pronóstico de probabilidad llamado modelo probit para generar probabilidades de recesión pronosticadas, que evaluó con diferentes medidas de desempeño. Lamentablemente, sus pronósticos de su modelo probit establecen que una recesión de EE.UU. de 2020 sería en junio a agosto de 2020, aunque la observada fue en febrero y marzo de 2020.

En definitiva, hay bastante literatura sobre pronósticos de las recesiones de Estados Unidos con otras metodologías y modelos (ver, por ejemplo; Del Negro y Schorfheide, 2011; Edge et al., 2010; Karlsson, 2013; Wieland y Wolters, 2011 y Babii et al., 2022). Todas estas metodologías muestran falta de precisión en el momento de realizar el pronóstico fuera de la muestra y tienden a aproximar el evento sin tener exactitud de cuando inicia con precisión. Inclusive, los últimos trabajos como de (Wang et al., 2022) con redes neuronales profundas muestran resultados claves para aplicar la estrategia de que los inversores no participen en el mercado en momentos de recesión económica.

En lo que respecta la literatura de modelos de pronósticos de las recesiones de Estados Unidos se observa que a la fecha no se evidencia trabajos de modelo probit instrumental. Por tanto, el trabajo de investigación explorara esta alternativa para construir un modelo probit instrumental que calcule la probabilidad de las recesiones económicas en Estados Unidos. Luego, esta probabilidad se utilizará para construir una variable indicadora dicotómica <sup>15</sup>.

<sup>15</sup>Esto significa que toma dos valores, 1, cuando hay una tasa de variación del PIB positiva y, 0, cuando hay una tasa

## 6.2. Estrategías de inversión

A menudo, cuando las expectativas del futuro <sup>16</sup> de la economía marcan conductas de angustia en los inversionistas en el mercado de capitales de Estados Unidos en el corto plazo. Es más, la guerra entre Rusia y Ucrania no encuentra una solución a la fecha, provocando un sentimiento creciente de pesimismo e incertidumbre sobre la economía. De hecho, la pandemia de la COVID-19, que todavía está vigente, provocó una caída de la producción de Estados Unidos por dos meses en el 2020. Sin embargo, como lo menciona (Marks, 2022) en su memorándum explica que desde 1920 han existido 17 recesiones, una Guerra Mundial, una Gran Depresión, múltiples periodos de angustia por un cataclismo global y otros eventos difíciles, a pesar de estos eventos el S&P500 logra un retorno superior a 10 % anual promedio durante más de un siglo. Es claro, con esta posición, que la mejor estrategia es invertir a largo plazo.

No obstante, se cuenta con literatura abundante sobre estrategias de inversión, donde el análisis técnico es uno de los más empleados con el objetivo de lograr mayores retornos que el S&P500. Por ejemplo, en (Barry y Robert, 2013) muestran técnicas gráficas para realizar pronósticos del comportamiento futuro de los precios de las acciones. Muchas de estas estrategias buscan encontrar patrones en las series de tiempo, observando sus tendencias, máximos, mínimos y otros estadígrafos para establecer resistencias y soportes que son denominados como pisos y techos con el fin de predecir el futuro del comportamiento de los precios de las acciones. En definitiva, es bastante amplio el análisis de estas gráficas que además buscan encontrar alguna forma particular que forma la serie de tiempo, aunque no se tiene evidencia científica de los resultados alcanzados que mejoren el retorno del S&P500.

De acuerdo a (Kakushadze y Serur, 2018), en su capítulo de acciones, se puede encontrar 20 estrategias para realizar compra y venta de acciones. Dentro de estas estrategias se encuentran las estrategias de análisis técnico que se basan en promedios móviles, soporte y resistencia en horizontes de tiempo cortos en busca de realizar un arbitraje estadístico. Por otra parte, se tiene estrategias basadas en análisis fundamental donde las estrategias se basan en la percepción de los inversores sobre cómo deberían comportarse los precios y retornos de las acciones en función de los fundamentos de las empresas (los productos y servicios de las empresas, fuentes de ingresos, proveedores, competidores, socios, etc.). Sin embargo, no muestran cierto favoritismo sobre

---

de variación negativa del PIB.

<sup>16</sup>Son creencias de lo que podría suceder.

determinada estrategia calificándola como mejor o peor.

(Tweedy, 2021) establece que una estrategia de inversión es la infravaloración (comprar acciones por debajo del valor observado de acuerdo al análisis fundamental) que se asocia con un riesgo bajo, y también se asociaría con rendimientos satisfactorios. De acuerdo a sus resultados históricos de inversión de casi 50 años y amplios estudios, se ha confirmado empíricamente que el enfoque fundamental del análisis de valores desarrollado por Benjamin Graham produce atractivas tasas de retorno a largo plazo. En consecuencia, una estrategia de inversión recomendada es a largo plazo y con análisis fundamental.

En lo que respecta a la estrategia de inversión, es claro que el análisis fundamental es el que con la evidencia empírica mostró resultados superiores al mercado. Este tipo de estrategias de inversión son fundamentalmente de largo plazo. Por tanto, este trabajo de investigación considera este elemento muy importante, ya que no plantea en ningún momento una salida o entrada al mercado en cualquier momento, sino que se basa en una estrategia de largo plazo combinada con la observación del comportamiento de algunas series de tiempo macroeconómicas. Por último, se puede clasificar la estrategia planteada en este trabajo de investigación como híbrida ya que combina el elemento de inversión a largo plazo con el seguimiento de algunas variables macroeconómicas.

## 7. Planteamiento de la hipótesis

En lo que corresponde al planteamiento de la hipótesis se tiene que considerar dos partes: la primera, es construir la variable indicadora dicotómica de orientación del momento de inversión y salida del mercado de capitales; la segunda, es verificar con el modelo CAPM la significatividad de alfa con el objetivo de observar si con la estrategia logra mejores resultados que el mercado de capitales de Estados Unidos.

Dado el modelo probit:

La variable dependiente es dicotómica: recesión económica (crisis) y expansión económica (auge):

$$y_t = \begin{cases} 1 & \text{crisis} \\ 0 & \text{auge} \end{cases} \quad (6)$$

De la ecuación (6), donde el vector de regresores se denotan por,  $x_t$ , suponemos que influyen en el resultado de,  $y_t$ . Por tanto, el modelo probit es:

$$P[y_t = 1|x_t] = \Phi(\phi_1 + \phi_2 x_t) \quad (7)$$

De la ecuación (7), donde,  $P$ , denota la probabilidad  $y_t$ , y  $\Phi$ , es la función de distribución acumulada de la distribución normal estándar. Finalmente, los parámetros desconocidos son,  $\phi_1$  y  $\phi_2$ .

En consecuencia, si la,  $P$ , es superior al promedio de probabilidades calculado con el modelo probit,  $\bar{P}$ , se tiene como resultado  $I$  que es la variable indicadora que toma dos valores, 1, el momento de realizar la inversión (comprar el índice S&P500) y, 0, el momento de encontrarse fuera del mercado de capitales (mantener el dinero en efectivo).

Posteriormente, se multiplica el retorno del mercado de capitales,  $R_{M,t}$ , por la variable indicadora,  $I$ , y se obtiene como resultado,  $R_{E,t}$ , que es el retorno que alcanza la estrategia planteada.

Por su parte, dado el modelo CAPM e incorporando una variable estocástica:

$$R_{E,t} = \alpha + \beta \times R_{M,t} + \epsilon_t \quad (8)$$

De la ecuación (8), donde,  $R_{E,t}$ , es el retorno de la estrategia,  $R_{M,t}$ , es el retorno del mercado y el vector de parámetros desconocidos es  $[\alpha, \beta, \sigma_\epsilon]$  y  $\epsilon_t$  es ruido blanco.

## 7.1. Hipótesis

Se plantean las siguientes hipótesis:

- La hipótesis nula:  $H_0 : \alpha = 0$
- La hipótesis alternativa:  $H_1 : \alpha > 0$

**Por lo tanto, si se rechaza,  $H_0$ , significa que de acuerdo al modelo CAPM el,  $\alpha$ , no es un desequilibrio temporal, sino que es permanente, debido a que,  $\alpha$ , no es un retorno que responde a las fluctuaciones del mercado de capitales, de hecho, disminuye el riesgo garantizando un retorno derivado de la estrategia al no rechazar,  $H_1$ . demostrando que el retorno de la estrategia logró mejores retornos que los rendimientos del mercado de capitales.**

## 8. Marco metodológico

### 8.1. Modelo de pronóstico probit

#### 8.1.1. Especificación

De acuerdo a (Cameron y Trivedi, 2005), cuando suponemos que los errores del modelo en lugar de describirse con la función de distribución logística, pueden describirse con la función de distribución Normal estándar. Cuando representamos los errores con la distribución Normal, el modelo resultante recibe el nombre de modelo probit.

Una variable,  $z$ , se dice que tiene distribución normal cuando su función de distribución de probabilidad acumulada tiene la siguiente forma:

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad (9)$$

En términos del modelo:

$$\Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t) = \int_{-\infty}^{\hat{\phi}_1 + \hat{\phi}_2 x_t} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad (10)$$

Existe una relación no lineal en los parámetros del modelo y, por lo tanto, el método de mínimos cuadrados no puede usarse para estimar el modelo probit.

### 8.2. Modelo de pronóstico probit

#### 8.2.1. Estimación

Como lo mencionamos anteriormente el método de mínimos cuadrados no puede emplearse para la estimación del modelo probit, por lo tanto, debemos recurrir al método de máxima verosimilitud.

Como la función de probabilidad de los errores y sabemos que tenemos una muestra aleatoria (es decir, compuesta por variables aleatorias independientes), la función de verosimilitud es simplemente la multiplicación de las funciones de probabilidad para todas las observaciones que hay en la muestra.

En términos matemáticos, la función de verosimilitud del modelo probit es:

$$L(\hat{\phi}_1, \hat{\phi}_2; x_t) = \prod_{t=1}^T \left[ \Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t) \right]^{y_t} \times \left[ 1 - \Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t) \right]^{1-y_t} \quad (11)$$

El logaritmo natural de la función de verosimilitud (11) es:

$$l(\hat{\phi}_1, \hat{\phi}_2; x_t) = \sum_{t=1}^T \left[ y_t \ln(\Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t)) + (1 - y_t) \ln(1 - \Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t)) \right] \quad (12)$$

Las condiciones de primer orden para la maximización de esta función (12) son:

$$\frac{\partial l(\cdot)}{\partial \hat{\alpha}} = \sum_{t=1}^T \left[ \frac{y_t - \Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t)}{\Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t)[1 - \Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t)]} \phi(\hat{\phi}_1 + \hat{\phi}_2 x_t) \right] = 0 \quad (13)$$

$$\frac{\partial l(\cdot)}{\partial \hat{\alpha}} = \sum_{t=1}^T \left[ \frac{y_t - \Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t)}{\Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t)[1 - \Phi(\hat{\phi}_1 + \hat{\phi}_2 x_t)]} \phi(\hat{\phi}_1 + \hat{\phi}_2 x_t) \right] x_t = 0 \quad (14)$$

Como se puede observar en las condiciones de primer orden (13 y 14), las incógnitas de ambas ecuaciones ( $\hat{\phi}_1, \hat{\phi}_2$ ) entran en forma no lineal y por lo tanto no pueden resolverse por métodos lineales.

(Amemiya, 1985) demostró que la función de verosimilitud del modelo probit es globalmente cóncava por lo que las condiciones de segundo orden para un máximo se cumplen. Los estimadores de máxima verosimilitud del modelo probit son insesgados, consistentes y eficientes. La matriz de varianzas y covarianzas de los estimadores del modelo está dada por la inversa de la matriz de las condiciones de segundo orden con el signo opuesto. En la diagonal principal de esta matriz tenemos las varianzas de los coeficientes mientras que fuera de la diagonal principal tenemos las covarianzas entre los coeficientes. Una vez que los parámetros del modelo han sido estimados, pueden realizarse los contrastes de hipótesis habituales.

Por su parte, una medida alternativa de bondad del ajuste se conoce como  $R^2$  de McFadden y se define como:

$$R_{MF}^2 = 1 - \frac{l(\hat{\phi}_1, \hat{\phi}_2)}{l(\hat{\phi}_0)} \quad (15)$$

De la ecuación (15) donde,  $l(\hat{\phi}_1, \hat{\phi}_2)$  es el valor del logaritmo de la función de verosimilitud evaluada en los estimadores de MV y  $l(\hat{\phi}_0)$  es el valor del logaritmo de la función de verosimilitud de un modelo que tiene solo una constante.

El  $R_{MF}^2$  tiene la misma interpretación que el  $R^2$  común. Es decir, nos dice que porcentaje de la

Cuadro 1: Predicción y realización

Predicción		
Realización	$\hat{y}_t = 0$	$\hat{y}_t = 1$
$y_t = 0$	$p_{00}$	$p_{10}$
$y_t = 1$	$p_{01}$	$p_{11}$

variabilidad de la variable dependiente está explicado por la regresión y no sirve para comparar el ajuste de diferentes modelos con distinto número de variables explicativas. La medida de bondad del ajuste que se utiliza para comparar ajustes es el  $R_{MF}^2$  ajustado.

El  $R_{MF}^2$  ajustado se define como:

$$R_{MF}^2 = 1 - \frac{l(\hat{\phi}_1, \hat{\phi}_2) - k}{l(\hat{\phi}_0)} \quad (16)$$

De la ecuación (16) donde  $k$  es el número de parámetros a estimar.

Otra forma de medir la bondad del ajuste es observar como clasifica a las observaciones el modelo en comparación con los datos realmente observados.

Observamos como el modelo predice que  $y_t = 1$  (es decir, cuándo  $\hat{y}_t = 1$ ).

Como se tiene una estimación de la probabilidad de trabajar una regla de clasificación sencilla podría ser que siempre que la estimación de la probabilidad de recesión sea mayor a un valor  $c$  (por ejemplo 0,5) entonces clasifique a ese periodo como recesión. En términos matemáticos:

Si:

$$F(\hat{\phi}_1 + \hat{\phi}_2 x_t) = \frac{e^{\hat{\phi}_1 + \hat{\phi}_2 x_t}}{1 + e^{\hat{\phi}_1 + \hat{\phi}_2 x_t}} > c \Rightarrow \hat{y}_t = 1 \quad (17)$$

De la ecuación (17) en la práctica uno puede elegir cualquier valor para,  $c$ . Una posibilidad es emplear la probabilidad empírica:

$$c = \frac{\#y_t = 1}{t} \quad (18)$$

Una vez que clasificadas las observaciones se puede construir lo que se denomina la tabla de predicción-realización:

La fracción,  $p_{00} + p_{11}$ , se denomina tasa de aciertos. Cuanto mayor es la tasa de aciertos mejor es el ajuste del modelo.

La tasa de aciertos se define como la fracción de predicciones correctas en la muestra. Formalmente, si definimos a la variable aleatoria  $w_t$  como la indicadora de una predicción correcta



(esto es,  $w_t = 1$ , si,  $y_t = \hat{y}_t$  y  $w_t = 0$ , si,  $y_t \neq \hat{y}_t$ ), entonces la tasa de aciertos se define como,  $h = \frac{1}{t} \sum_{t=0}^T w_t$ .

En la población, la proporción de unos es,  $p$ . Supongamos que nosotros hacemos una predicción completamente aleatoria. Esto es, predecimos un 1 con probabilidad,  $p$ , y predecimos un, 0, con probabilidad,  $(1 - p)$ . En este escenario, la probabilidad de hacer una predicción correcta es  $q = p^2 + (1 - p)^2$ .

Usando las propiedades de la distribución binomial para el número de predicciones (hechas en forma aleatoria) correctas, la tasa de aciertos aleatoria,  $(h_a)$ , tiene,  $E(h_a) = q$  y  $Var(h_a) = q(1 - q)/n$ . Entonces la habilidad predictiva del modelo estimado puede evaluarse comparándola con esta tasa de aciertos aleatoria.

Se realiza una prueba de hipótesis, cuya hipótesis nula sea que las predicciones del modelo estimado no son mejores que las predicciones hechas en forma aleatoria; y cuya hipótesis alternativa sea que las predicciones del modelo son mejores que las hechas aleatoriamente.

Bajo la hipótesis nula, la tasa de aciertos,  $h$ , se distribuye normalmente con media,  $q$ , y varianza,  $q(1 - q)/n$ . Por lo tanto el estadístico de contraste es:

$$z = \frac{h - q}{\sqrt{q(1 - q)/n}} = \frac{nh - nq}{\sqrt{nq(1 - q)}} \quad (19)$$

En la práctica,  $q = p^2 + (1 - p)^2$ , se estima con,  $\hat{p}^2 + (1 - \hat{p})^2$ , con,  $\hat{p}$ , siendo la proporción de unos en la muestra.

La regla de decisión es rechazar la hipótesis nula siempre que el valor de probabilidad de,  $z$ , sea menor al nivel de significación de la prueba o siempre que,  $z$ , sea mayor al valor crítico de la normal estándar.

Otros índices de interés son los denominados sensibilidad y especificidad. Donde la sensibilidad es la probabilidad de predecir un éxito entre los éxitos:  $Pr(\hat{y}_t = 1|y_t = 1)$ . Y la especificidad es la probabilidad de predecir un fracaso entre los fracasos:  $Pr(\hat{y}_t = 0|y_t = 0)$ . Como resultado, la probabilidad de predecir un éxito falso o falso positivo es uno menos la especificidad:  $Pr(\hat{y}_t = 1|y_t = 0) = 1 - Pr(\hat{y}_t = 0|y_t = 0)$ . En definitiva, los falsos positivos corresponden a lo que llamamos error de tipo I. Es claro que una mejor bondad de ajuste se obtiene con una alta sensibilidad y especificidad.

En una tabla (2) de contingencia (similar a la de predicción-realización) se pueden ver estas

Cuadro 2: Tabla de contingencia

Predicción		
Realización	$\hat{y}_t = 0$	$\hat{y}_t = 1$
$y_t = 0$	$Pr(\hat{y}_t = 0 y_t = 0)$	$Pr(\hat{y}_t = 1 y_t = 0)$
$y_t = 1$	$Pr(\hat{y}_t = 0 y_t = 1)$	$Pr(\hat{y}_t = 1 y_t = 1)$

definiciones:

Una forma de resumir la bondad del ajuste con estas dos medidas es graficando la curva ROC (*Relative (Receiver) Operating Characteristic*). Donde, la curva ROC es una representación gráfica de la sensibilidad frente a (1 - especificidad) para un sistema clasificador binario según se varía el umbral de clasificación. Una forma de resumir la curva ROC es calcular el área bajo la curva (*AUC Area Under the Curve*). La interpretación es si el AUC es igual a 1, el valor de diagnóstico es perfecto y si AUC es 0,5 no se tiene valor diagnóstico.

### 8.3. Modelo de regresión lineal

De acuerdo a (Wooldridge, 2010), el modelo de regresión lineal simple y de acuerdo a la hipótesis de trabajo es que  $E(R_{E,t}/R_{M,t})$  es lineal en  $R_{M,t}$  del tipo:

$$E(R_{E,t}/R_{M,t}) = \alpha + \beta R_{M,t} \quad (20)$$

De la ecuación (20) donde,  $\alpha$  y  $\beta$ , son parámetros fijos desconocidos que se llaman coeficientes de regresión.

Por lo tanto, podemos expresar el desvío del retorno de la estrategia,  $R_{E,t}$ , alrededor de su valor esperado como:

$$u_t = R_{E,t} - E(R_{E,t}/R_{M,t}) \quad (21)$$

o

$$R_{E,t} = E(R_{E,t}/R_{M,t}) + u_t \quad (22)$$

De la ecuación (22) donde,  $u_t$ , es una variable aleatoria no observable que puede tomar valores positivos o negativos.

Reemplazando se tiene,

$$R_{E,t} = \alpha + \beta R_{M,t} + u_t \quad (23)$$

La ecuación (23) establece retorno de la estrategia planteada está linealmente relacionada con el retorno del mercado más un término aleatorio.

Para la estimación utilizando la máxima verosimilitud (MV), se supone que,  $u_t | R_{M,t} \sim N(\mu_t, \sigma_t^2)$ , donde,  $\mu_t = \alpha + \beta R_{M,t}$ . La función de verosimilitud es:

$$\prod_{t=1}^T f(R_{M,t}; R_{E,t}) = \prod_{t=1}^T f_{R_M}(R_{M,t}) \times f_{R_E|R_M}(R_{E,t}|R_{M,t}) = \mathcal{L}_1 \times \mathcal{L}_2 \quad (24)$$

De la ecuación (24) donde,  $\mathcal{L}_1 = f_{R_M}(R_{M,t})$ , y,  $\mathcal{L}_2 = f_{R_E|R_M}(R_{E,t}|R_{M,t})$ . El término,  $\mathcal{L}_1$ , no involucra los parámetros,  $\alpha$ ,  $\beta$ . Nos concentraremos en el segundo término,  $\mathcal{L}_2$ , que se denomina verosimilitud condicional, dada por:

$$\mathcal{L}_2 \equiv \mathcal{L}(\alpha, \beta, \sigma) = \prod_{t=1}^T f_{R_E|R_M}(R_{E,t}|R_{M,t}) \propto \sigma^{(-T)} \exp\left[-\frac{1}{2\sigma^2} \sum_t (R_{E,t} - \mu_t)^2\right] \quad (25)$$

La log-verosimilitud condicional de la función (25) es:

$$l(\alpha, \beta, \sigma) = -T \log \sigma - \frac{1}{2\sigma^2} \sum_{t=1}^T [R_{E,t} - (\alpha + \beta R_{M,t})]^2 \quad (26)$$

Para encontrar el estimador de MV de  $(\alpha, \beta)$  maximizamos  $l(\alpha, \beta, \sigma)$ . Es importante resaltar que bajo el supuesto de normalidad, el estimador de mínimos cuadrados clásicos es también un estimador de máxima verosimilitud. La diferencia es que al maximizar,  $l(\alpha, \beta, \sigma)$ , el parámetro  $\sigma$  por máxima verosimilitud es:

$$\hat{\sigma}^2 = \frac{1}{T} \sum_t \hat{u}_t^2 \quad (27)$$

Este estimador de la ecuación (27) es similar, pero no idéntico, al estimador insesgado de la estimación por mínimos cuadrados clásicos<sup>17</sup>.

Lo anterior muestra como es el retorno de la estrategia a nivel poblacional. Sin embargo, utili-

<sup>17</sup>En este caso particular no es tan importante la insesgades del estimador, de hecho, si la muestra es inmensa no tiene impacto en la varianza.

zaremos una muestra para aproximar esta relación poblacional a través de una regresión muestral:

$$\hat{R}_{E,t} = \hat{\alpha} + \hat{\beta}R_{M,t} \quad (28)$$

De la ecuación (28) donde,  $\hat{R}_{E,t}$ , es la estimación de,  $E(R_{E,t}/R_{M,t})$ , y  $(\hat{\alpha}$  y  $\hat{\beta})$ , son los estimadores de los parámetros poblacionales,  $\alpha$  y  $\beta$ , respectivamente.

De la misma forma en que expresamos a la función de regresión poblacional para una familia particular, podemos expresar a la función de regresión muestral como:

$$R_{E,t} = \hat{\alpha} + \hat{\beta}R_{M,t} + \hat{u}_t \quad (29)$$

Finalmente, para lograr una aproximación lo más cerca posible de la regresión poblacional, se realiza por el método de estimación de máxima verosimilitud (MV).

Es importante considerar que el modelo de regresión lineal incorpora supuestos:

1. Linealidad: Este supuesto implica que el efecto marginal no depende del nivel de los regresores del modelo.
2. Exogeneidad estricta:  $E(u_t/R_{M,t}) = 0, t = 1, 2, \dots, T$ , implica que por ley de esperanzas iteradas,  $E(u_t) = 0, \forall t$ , los regresores son ortogonales al error para todas las observaciones.
3. No singularidad: Este supuesto muestra que no existe relaciones linealmente dependientes de la variable explicativa.
4. Varianza del error esférica:  $E(u_t^2/R_{M,t}) = \sigma^2 > 0, t = 1, 2, \dots, T$  implica que el segundo momento condicional, que es una función no lineal de,  $R_{M,t}$ , es una constante. Asimismo, ausencia de correlación serial,  $Cov(u_t, u_v) = 0, \forall t \neq v$ , la covarianza condicional de los errores del modelo es cero.

Técnicamente, el método de MV consiste en maximizar la suma de residuos al cuadrado, contrariamente al método de mínimos cuadrados que minimiza la suma de residuos al cuadrado.

Ahora para determinar la bondad de ajuste se lo realiza a través del,  $R^2$ , no centrado que se encuentra entre,  $0 < R^2 < 1$ , y su interpretación es la fracción de variabilidad de la variable dependiente que es atribuible a la variación en las variables explicativas.

Finalmente, para contrastar la hipótesis se construye el estadístico,  $z$ , de la siguiente forma:

$$z = \frac{\hat{\alpha} - \alpha}{\sqrt{\sigma^2[X'X]^{-1}}} \quad (30)$$

De la ecuación (30) que se distribuye como una normal estándar,  $z$ , distribución,  $z$ . En consecuencia, un desvío inmenso de,  $z$ , de cero es un signo de la falla de la hipótesis nula. Luego, se encuentra el valor crítico,  $z_{\alpha/2}$ , en la distribución  $z$  como el intervalo alrededor de cero que deja,  $\alpha/2$  %, de cada lado de la distribución. Esto es,

$$Prob(-z_{\alpha/2} < z < z_{\alpha/2}) = 1 - \alpha \quad (31)$$

Donde la regla de decisión es: Rechace,  $H_0$ , si,  $z < -z_{\alpha/2}$  ó  $z > z_{\alpha/2}$ . Ahora esta regla de decisión de la prueba se puede formular en términos del p-valor, de la siguiente forma: Si,  $Prob(z > |z|) < \alpha$ , se rechaza,  $H_0$ .

## 9. Métodos y técnicas de recolección de datos

### 9.1. Datos para el modelo econométrico de pronóstico de las recesiones económicas

Los datos se seleccionaron de acuerdo a disponibilidad y extraídos del banco de la Reserva Federal de *Saint Louis* de Estados Unidos<sup>18</sup>. Para construir el modelo de pronóstico de las recesiones económicas el periodo de tiempo de las series es de enero de 1959 hasta agosto de 2022. Esta serie de tiempo se dividió en dos partes: la primera una serie de tiempo de periodo de estimación del modelo (enero de 1959 hasta diciembre de 2018) que son 719 observaciones y un periodo de tiempo de prueba de la efectividad en el pronóstico del modelo (enero 2019 hasta agosto de 2022) que son 44 observaciones<sup>19</sup>. Asimismo, se utilizaron un total de 128 variables macroeconómicas, financieras y otras.

Para limpiar la información se procedió a eliminar columnas con observaciones faltantes y

<sup>18</sup><https://fred.stlouisfed.org/series/>

<sup>19</sup>Es importante aclarar que no se sigue los criterios empíricos de periodos de evaluación dentro de muestra. Esto debido a que el objetivo del modelo no es predecir el crecimiento económico, sino las recesiones, que en promedio suceden cada 10 años. Entonces se tomó la decisión de ver si con el modelo podía identificar la última recesión económica en Estados Unidos de 2020.

filas con valores perdidos, se agregaron rezagos (desde 3 meses hasta 18 meses), se comprobó la estacionariedad (para evitar regresiones espurias), finalmente se estandarizaron las series de tiempo <sup>20</sup>. Con todo este procesamiento de información se construyó un total de 708 variables incluyendo todos los rezagos que fueron realizados con el objetivo de observar si estas variables podrían explicar de forma no contemporánea las recesiones económicas de Estados Unidos. Una variable dicotómica se añadió, que toma valor, 1, cuando hay recesión económica y, 0, cuando hay expansión económica. En conclusión, se tiene un total de 709 variables en la base de datos.

Posteriormente, se calcularon las correlaciones entre las 709 variables en el tiempo, donde la correlación analizada fue con la variable dicotómica de las recesiones económicas con el conjunto de otras variables. Cabe destacar que trabajar con una base de datos de 128 variables y hacer el seguimiento a las mismas dificulta el pronóstico debido a que muchas de estas variables en el tiempo tienden a discontinuar su publicación, en consecuencia, no permite tener pronósticos por falta de estos datos, sobre todo a futuro. Por lo tanto, se planteó dos criterios para seleccionar algunas variables que explican las recesiones económicas (variable dicotómica): el primero es las variables que tengan una correlación superior a 0,3 <sup>21</sup>; y el segundo en la estimación por el principio de simplificación parsimoniosa <sup>22</sup> las variables que tengan significatividad estadística luego de la estimación del modelo econométrico probit.

Con estos dos criterios: el resultado del primer criterio se limitó la base de datos a 50 variables; con el segundo criterio en la estimación del modelo econométrico probit se obtuvo 3 variables <sup>23</sup> que explican de manera significativa la variable dicotómica de recesiones económicas de Estados Unidos.

<sup>20</sup>Todos los cálculos se realizaron en el language de programación Python.

<sup>21</sup>Este criterio se seleccionó debido a que si bien la correlación superior a 0,7 asociación alta. Sin embargo, ninguna de las variables alcanzaba esta correlación respecto a las recesiones económicas, por tanto, se disminuyó el criterio a 0,5, asociación moderada, donde solo una variable con sus rezagos registraba este valor. En definitiva, se optó por un criterio de 0,3, asociación mediana, que generaba un conjunto de 50 variables para estimar el modelo probit (Ver Amat Rodrigo, 2016).

<sup>22</sup>Es importante recordar la palabras de Milton Friedman: Una hipótesis [modelo] es importante si esta 'explica' mucho, con poco.... Esto significa que se deben introducir en el modelo unas pocas variables claves que capturen la esencia del fenómeno bajo estudio relegando toda influencia menor y aleatoria al término de error.

<sup>23</sup>Las tres variables son: Moody's seasoned Aaa corporate bond Minus Federal Funds Rate (AAAFFM), All Employees, Total Nonfarm (PAYEMS) y 5-Year Treasury Constant Maturity Minus Federal Funds Rate (T5YFFM) con 6 rezagos mensuales

## 9.2. Datos para el modelo econométrico de estimación del modelo CAPM

Las series de tiempo para el modelo econométrico para la estimación del CAPM se recolectaron de *Yahoo Finance*<sup>24</sup>, donde la variable es el índice Standard & Poor's 500 (*Standard & Poor's 500 Index*). Esta variable también conocido como S&P500, la cual es uno de los índices bursátiles más importantes de Estados Unidos y se lo considera el índice más representativo de la condición del mercado. La forma de selección del periodo de análisis es de acuerdo a la disponibilidad de la misma en forma mensual<sup>25</sup>. Por tanto, la muestra inicia de enero de 1985 hasta agosto de 2022. El total de observaciones dentro la muestra es 453.

## 10. Resultados

### 10.1. Resultados del modelo de pronóstico probit

Esta sección modela las recesiones económicas de Estados Unidos.

La variable dependiente es dicotómica con: recesión (crisis) y expansión económica (auge),

$$y_t = \begin{cases} 1 & \text{crisis} \\ 0 & \text{auge} \end{cases} \quad (32)$$

Donde los valores, 1, y, 0, se eligen por simplicidad. Las variables explicativas son: AAAFFM, PAYEMS y T5YFFM6MLAG. Con las realizaciones de las variables explicativas y la estimación por máxima verosimilitud del modelo probit se estimó la probabilidad de crisis económica para el periodo enero 1959 a diciembre de 2008. Para una estimación del probit instrumental se utilizaron como instrumentos el resto de variables (un total de 47 variables) que fueron no consideradas por su falta de significatividad estadística<sup>26</sup>.

<sup>24</sup><https://finance.yahoo.com/>

<sup>25</sup>Se utilizó Python para la recolección y por medio de este software solo se descarga la serie S&P500 con frecuencia mensual desde enero de 1985.

<sup>26</sup>Las variables son: INDPRO, IPDMAT, IPMANSICS, UNRATE, USGOOD, USCONS, MANEMP, DMANEMP, SRVPRD, TB3SMFFM, TB6SMFFM, T1YFFM, T5YFFM, T10YFFM, HWIURATIO12MLAG, COMPAPFFX3MLAG, COMPAPFFX6MLAG, COMPAPFFX9MLAG, COMPAPFFX12MLAG, TB3SMFFM3MLAG, TB3SMFFM6MLAG, TB3SMFFM9MLAG, TB3SMFFM12MLAG, TB6SMFFM3MLAG, TB6SMFFM6MLAG, TB6SMFFM9MLAG, TB6SMFFM12MLAG, T1YFFM3MLAG, T1YFFM6MLAG, T1YFFM9MLAG, T1YFFM12MLAG, T5YFFM3MLAG, T5YFFM9MLAG, T5YFFM12MLAG, T10YFFM3MLAG, T10YFFM6MLAG, T10YFFM9MLAG, T10YFFM12MLAG, AAAFFM3MLAG, AAAFFM6MLAG, AAAFFM9MLAG, AAAFFM12MLAG, BAAFFM3MLAG, BAAFFM6MLAG, BAAFFM9MLAG y BAAFFM12MLAG.

Por tanto el modelo probit estimado es:

$$\hat{P}[y_t = 1|\cdot] = \hat{P}[y_t = 1|PAYEMS_t; AAAFFM_t; T5YFFM6MLAG_t] \quad (33)$$

Y,

$$\hat{P}[y_t = 1|\cdot] = \Phi(\hat{\phi}_1 + \hat{\phi}_2 PAYEMS_t + \hat{\phi}_3 AAAFFM_t + \hat{\phi}_4 T5YFFM6MLAG_t) \quad (34)$$

El resultado de la estimación del modelo probit instrumental <sup>27</sup> es:

$$\hat{P} = -1,835846 - 0,660103 PAYEMS - 0,5507298 AAAFFM - 0,4322547 T5YFFM6MLAG$$

(0,1234988)      (0,0875862)      (0,0993091)      (0,0938943)

$$T = 695 \quad \bar{R}^2 = 0,4439 \quad \hat{\sigma} = 0,26415$$

(Desviaciones típicas entre paréntesis)

También se puede observar gráficamente los resultados de la estimación de las probabilidades de una recesión económica en Estados Unidos (Ver Figura 2). Cabe destacar que en los periodos donde la probabilidad supera el umbral, se evidencia una recesión económica en Estados Unidos. Sin embargo, es importante señalar que el modelo probit instrumental no predice con exactitud y el umbral encontrado tampoco alcanza ese nivel de precisión dentro de muestra. De hecho el error cuadrático medio es diferente de cero (0,1353).

De acuerdo a la tabla de contingencia de predicción y realización se encuentra que el número de casos correctamente predichos es 630 (90,6%). Por su parte, con a la curva el área bajo la curva AUC es 0,9565, mostrando un valor muy próximo a 1, por tanto se puede afirmar que el ajuste del modelo es muy cerca a perfecto. Finalmente, con la bondad de ajuste medida por el  $\bar{R}^2$  de McFadden es 0,4439, esto significa las recesiones económicas de Estados Unidos son explicadas en un 44,39% con las tres variables dependientes <sup>28</sup>.

Por su parte, la probabilidad promedio dentro de muestra para generar la variable indicadora <sup>29</sup> (Ver Figura 2) es:

$$\bar{P} = 0,1155 \quad (35)$$

<sup>27</sup>Para ver detalles ir a primer Apéndice 2.

<sup>28</sup>Ver Apéndice 2 y 3.

<sup>29</sup>Es importante destacar que utilizar este umbral se lo realizó de acuerdo a Ng, 2012 donde encuentra que la probabilidad tiene una alta volatilidad.



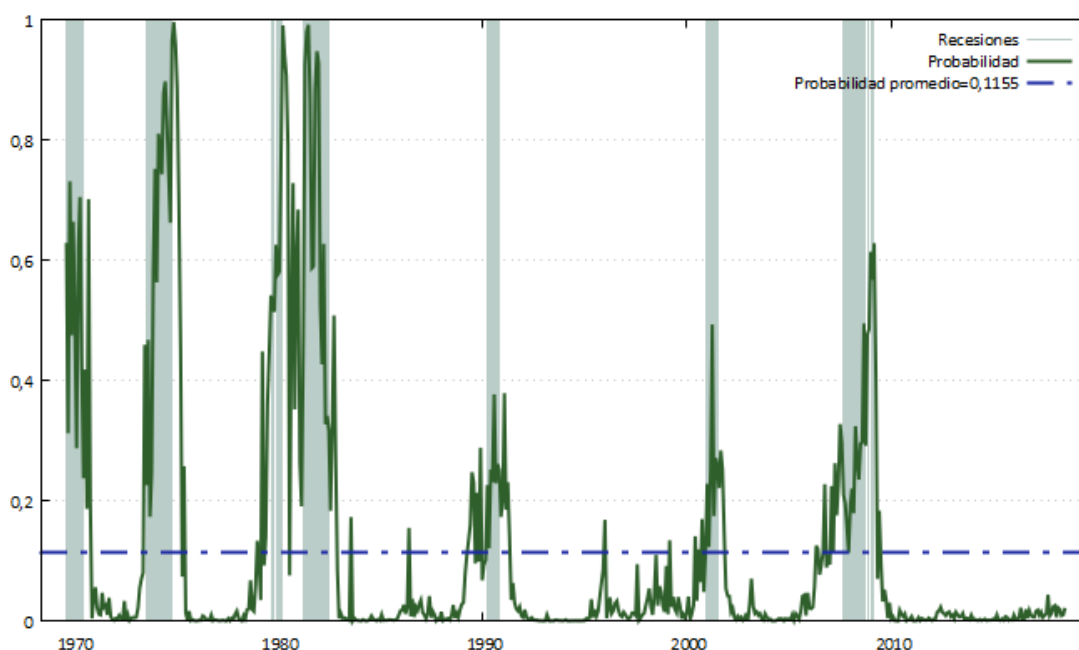


Figura 2: Recesiones y probabilidades dentro de la muestra

Por tanto, la variable indicadora que señala el momento de invertir es:

$$I_t = \begin{cases} 1 & \bar{P} < 0,1155 \\ 0 & \bar{P} \geq 0,1155 \end{cases} \quad (36)$$

Luego, con la observación de las tres variables dependientes se estimó las probabilidades fuera de muestra que se desarrolla a continuación.

#### 10.1.1. Evaluación fuera de muestra

La evaluación fuera de la muestra se realizó de agosto de 2018 hasta julio de 2022. De inicio se puede observar (Figure 3) un incremento de la probabilidad superior al 42 % en marzo y 100 % en abril de 2020, y comparando con la recesión observada en Estados Unidos esta se observó en febrero y marzo de 2020. Este resultado evidencia que fuera de muestra el modelo probit instrumental aproxima bastante bien el periodo recesivo. Y el error cuadrático medio es igual a 0,04167 si se considera la variable indicadora como umbral, este error cuadrático medio es inferior al valor dentro de muestra, aunque, no es cero. Sin embargo, aproxima bastante bien los meses recesivos en Estados Unidos considerando que la recesión se observó solo en dos meses (febrero y

marzo de 2020) como resultado de las medidas asumidas por este país para disminuir los contagios de la COVID-19 en el 2020.

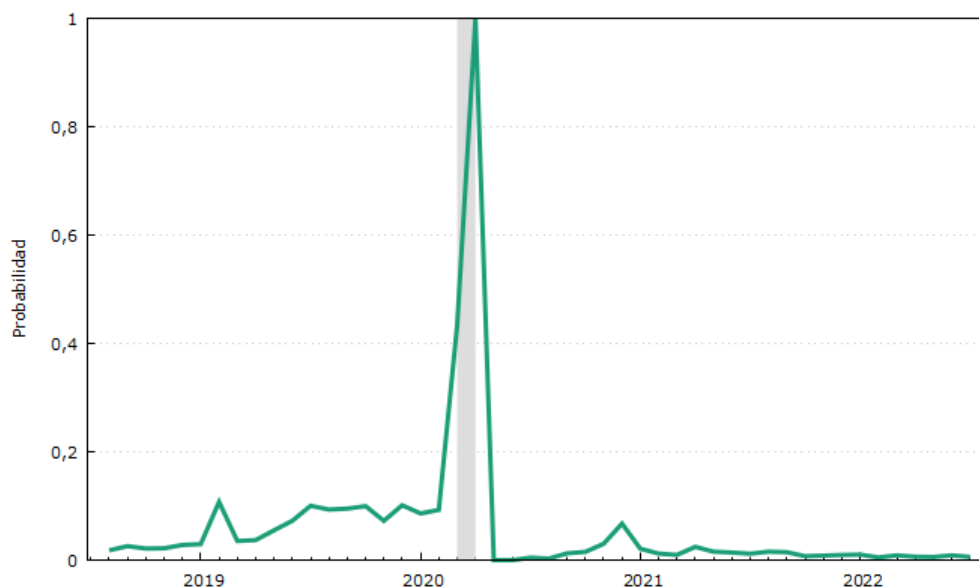


Figura 3: Recesiones y probabilidades fuera de muestra

Una situación no resuelta es que las variables explicativas que permiten calcular la probabilidad son contemporáneas, por tanto, para alcanzar estos resultados se requiere contar con AAAFFM, PAYEMS y T5YFFM6MLAG en el mismo periodo de tiempo. Lamentablemente, no se cuenta con estas variables en el mismo periodo de tiempo, sino con un mes de rezago. Por tanto, se rezagó un mes la variable indicadora de forma que en el periodo contemporáneo se tiene la orientación de invertir o no en el mercado de capitales de Estados Unidos.

Luego, se multiplicó la variable indicadora dicotómica por el retorno del S&P500 generando el retorno de la estrategia,  $R_E$ , planteada en este trabajo de investigación <sup>30</sup>. El retorno de la estrategia se verificará si alcanza un,  $\alpha$ , significativo que evidencia que la estrategia logra retornos que no dependen del mercado de capitales como se planteó en la hipótesis.

## 10.2. Resultados del modelo de regresión lineal simple

Por su parte, la estimación por máxima verosimilitud del modelo CAPM <sup>31</sup> es

<sup>30</sup>Es importante aclarar que se supone que tanto la entrada y salida del mercado de capitales en Estados Unidos es libre y no existe comisiones, impuestos u otros gastos para aplicarlo en la realidad.

<sup>31</sup>Para más detalles de la estimación ver Apéndice 4. Por su parte, el criterio de información de Akaike (AIC) (Akaike, 1974) y el criterio de información bayesiano (BIC) (Bhat y Kumar, 2010) se muestran en los resultados.

$$\widehat{R}_E = 0,0023331 + 0,7380859 R_M$$

(0,0009194)      (0,0206865)

$$T = 451 \quad \ln L = 1143,372233 \quad \hat{\sigma} = 0,019172 \quad AIC = -5,061518 \quad BIC = -2743,883$$

(Desviaciones típicas entre paréntesis)

### 10.2.1. Constrastación de la hipótesis

De inicio, calculando el estadístico,  $z$ , se tiene que es estadísticamente significativo al 5% de nivel de significación. Es más, el p-valor es 0,011 que significa que esa es la probabilidad de no rechazar la hipótesis nula. Por tanto, de forma individual,  $\hat{\alpha}$ , no es cero, como también,  $\hat{\alpha}$  se encuentra en el intervalo al 95% de nivel de confianza.

$$z = \frac{\hat{\alpha} - \alpha}{\sigma_{\hat{\alpha}}} = \frac{0,0023331 - 0}{0,0009194} = 2,54 \quad (37)$$

y el intervalo,

$$Prob(0,0005311 < \hat{\alpha} < 0,0041351) = 0,95 \quad (38)$$

Por otra parte, también utilizó el test de Wald para contrastar la hipótesis. Donde el estadístico de contraste es:

$$\chi^2 = \frac{(\hat{\alpha} - \alpha)^2}{\sigma_{\hat{\alpha}}^2} = \frac{(0,0023331 - 0)^2}{0,0009194^2} = 6,4396 \quad (39)$$

Por su parte,  $\alpha = 0,05$ , el valor para la distribución Chi-cuadrado ( $\chi^2_{(1)}$ ) es 3,8415. Dado que el valor 3,8415 es menor que 6,4396, rechazamos la hipótesis nula. El p-valor es:

$$p - \text{valor} = \mathbf{Prob}(\chi^2_{(1)}) = 0,0112 \quad (40)$$

De acuerdo a las dos pruebas de hipótesis que se realizaron se encontraron: primero, que el p-valor se encuentra entre 0,01 y 0,05 muestra robusta evidencia (Wasserman, 2004) en contra de la hipótesis nula ( $H_0$ ); y segundo, el intervalo de confianza contiene al 95% de nivel de confianza el valor estimado de,  $\hat{\alpha}$ , con valores positivos. Por lo tanto, los datos contradicen la teoría de valoración de activos (CAPM) sobre el valor,  $\alpha$ , sean un desequilibrio temporal, sino que este es

Cuadro 3: Comparación de resultados: febrero de 1985 a julio de 2022

Variables	Indicadores financieros (En porcentaje)	
	Retorno	Riesgo
S&P500	9,8	15,2
Estrategia	10,2	13,0

permanente y se debe a la estrategia planteada en este trabajo de investigación. En otras palabras, se tiene evidencia que la estrategia alcanzó un retorno que no se explica por el comportamiento del mercado de capitales de Estados Unidos, sino son propios de la estrategia.

### 10.2.2. Retornos anuales

Para finalizar, se calculó los retornos de forma anual desde febrero de 1985 hasta julio de 2022 (aproximadamente 4 décadas). El retorno anual alcanzado por el S&P500 es 9,8 % con un desvío estándar muestral de 15,2 %, mientras que el retorno anual de la estrategia es 10,2 % con un desvío muestral de 13,0 %. Estos resultados muestran que la estrategia en casi 4 décadas alcanzó un mayor retorno y con menor riesgo que el mercado de capitales de Estados Unidos (Ver Cuadro 3).

## 11. Conclusiones

En conclusión, se logró alcanzar el objetivo de esta investigación al construir una estrategia que alcance mejores retornos y menor riesgo que el mercado de capitales de Estados Unidos.

Hay que destacar que para realizar este trabajo se eligió el modelo CAPM como modelo teórico y el modelo probit instrumental de forma metodológica. El modelo teórico sirvió para plantear la hipótesis que muestra que existe un parámetro independiente del mercado que es un desequilibrio permanente. Por su parte, el modelo econométrico estimado por máxima verosimilitud probó la estrategia de inversión.

Los resultados de la investigación empírica permiten señalar las siguientes conclusiones:

- Es evidente que los periodos de recesión de la economía norteamericana los retornos en promedio muestran ser negativos, mientras que el riesgo (medido con el desvío estándar

muestral) es mayor en comparación a periodos donde el crecimiento de la economía es positivo.

- El pronóstico de una recesión económica todavía está en desarrollo, debido a que no se encuentra a la fecha un modelo que realice la estimación de este evento con precisión. Sin embargo, el desarrollo de la estrategia con el uso de un modelo probit instrumental evidencia una muy buena aproximación para el pronóstico de las recesiones económicas de Estados Unidos. De hecho, estos resultados permitirán construir una variable indicadora (dicotómica que toma valor 0 cuando la probabilidad de recesión supera el umbral observado y 1 cuando se encuentra por debajo de mencionado umbral) que marca la estrategia de inversión.
- Finalmente, se verificó la hipótesis de los resultados alcanzados por la estrategia son superiores a los resultados que logra el mercado de capitales con estimaciones de máxima verosimilitud. En definitiva, se evidenció que la estrategia desarrollada en este trabajo de investigación le gana al mercado de capitales de Estados Unidos durante 4 décadas.

Queda pendiente aplicar esta metodología al conjunto de mercados de capitales, no solo a los países desarrollados, sino también, para países en vías de desarrollo. Debido a que con la globalización de los mercados de capitales, el inversionista que busca diversificar su portafolio de inversión tenga una orientación que le permita lograr mejores resultados que el mercado de capitales de todo el mundo con mayor precisión y una estrategia ganadora.

## Referencias

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- Amat Rodrigo, J. (2016). Correlación lineal y Regresión lineal simple. *Ciencia de Datos*.
- Amemiya, Y. (1985). Instrumental variable estimator for the nonlinear errors-in-variables model. *Journal of Econometrics*, 28(3), 273-289.
- Babii, A., Ghysels, E., & Striaukas, J. (2022). Machine learning time series regressions with an application to nowcasting. *Journal of Business & Economic Statistics*, 40(3), 1094-1106.
- Barry, S., & Robert, S. (2013). *Technical analysis: Modern perspectives*. CFA Institute Research Foundation.
- Bhat, H. S., & Kumar, N. (2010). On the derivation of the bayesian information criterion. *School of Natural Sciences, University of California*, 99.
- Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of business*, 45(3), 444-455.
- Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.
- Del Negro, M., & Schorfheide, F. (2011). Bayesian macroeconometrics. *The Oxford handbook of Bayesian econometrics*, 293, 389.
- Edge, R. M., Gürkaynak, R. S., Reis, R., & Sims, C. A. (2010). How useful are estimated dsge model forecasts for central bankers?[with comments and discussion]. *Brookings Papers on Economic Activity*, 209-259.
- Harvey, C. R. (2011). Yield Curve Inversions and Future Economic Growth. *Diakses dari [https://faculty.fuqua.duke.edu/~charvey/Research/Professional\\_Materials/Term\\_Structure\\_May\\_17\\_2011.pdf](https://faculty.fuqua.duke.edu/~charvey/Research/Professional_Materials/Term_Structure_May_17_2011.pdf)*.
- Jarrow, R. A. (1988). *Finance theory*. Prentice Hall.
- Jegadeesh, N., Noh, J., Pukthuanthong, K., Roll, R., & Wang, J. (2019). Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation. *Journal of Financial Economics*, 133(2), 273-298.
- Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of finance*, 23(2), 389-416.

- Kakushadze, Z., & Serur, J. A. (2018). 151 Trading Strategies. Z. Kakushadze and JA Serur, 151.
- Karlsson, S. (2013). Forecasting with Bayesian vector autoregression. *Handbook of economic forecasting*, 2, 791-897.
- Lintner, J. (1969). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets: A reply. *The review of economics and statistics*, 222-224.
- Markowitz, H. (1952). Portfolio selection. *The Journal of finance*, 7(3), 77-91.
- Markowski, L. (2020). Further evidence on the validity of CAPM: The Warsaw Stock Exchange application. *Journal of Economics & Management*, 39, 82-104.
- Marks, H. (2022). Oaktree Clients. En *I Beg to Differ*. 2022 Oaktree Capital Management, L.P.
- Ng, E. C. (2012). Forecasting US recessions with various risk factors and dynamic probit models. *Journal of Macroeconomics*, 34(1), 112-125.
- Sharpe, A. (1999). *Bailey: Investments*.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3), 425-442.
- Tobin, J. (1958). Liquidity preference as behavior towards risk. *The review of economic studies*, 25(2), 65-86.
- Tweedy, B. C. L. (2021). *What Has Worked in Investing: Studies of Investment Approaches and Characteristics Associated with Exceptional Returns*.
- Wang, Z., Li, K., Xia, S. Q., & Liu, H. (2022). Economic recession prediction using deep neural network. *The Journal of Financial Data Science*, 4(3), 108-127.
- Wasserman, L. (2004). *All of statistics: a concise course in statistical inference* (Vol. 26). Springer.
- Wieland, V., & Wolters, M. H. (2011). The diversity of forecasts from macroeconomic models of the US economy. *Economic Theory*, 47(2), 247-292.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
- Xinyu, W., & David, C. (2020). Forecasting U.S. Recessions by using the Probit Model. *Master Thesis in COPENHAGEN BUSINESS SCHOOL*.

## 12. Apéndice 1: Cálculo de la anticipación de la propuesta de (Harvey, 2011)

La Figura 4 muestra en línea roja la fecha en la que la variable de spread se vuelve negativa y las recesiones económicas con color plomo, adicionalmente se añadió el S&P500.

La siguiente Figura 5 muestra en días después de que la variable spread se volvió negativa y se

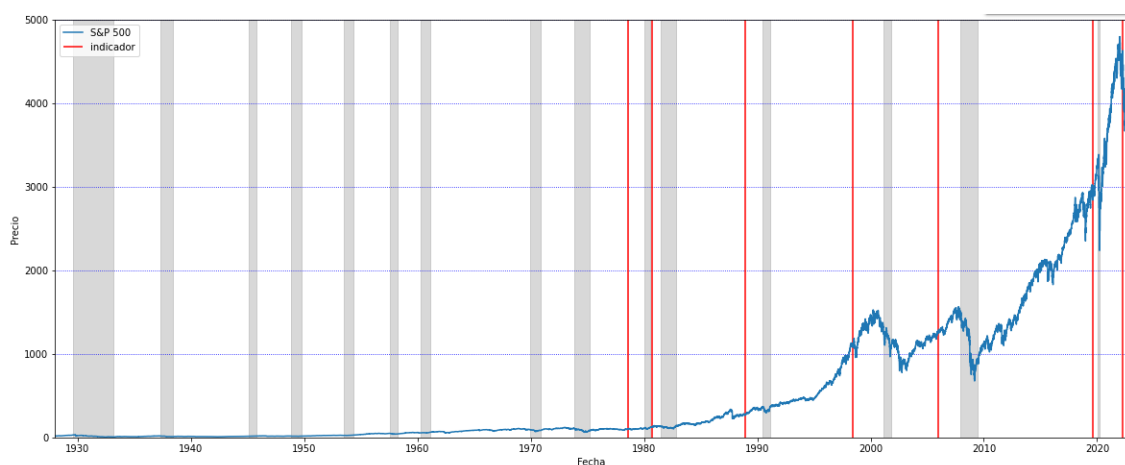


Figura 4: Fechas antes que se observe una recesión económica

observó una recesión económica en Estados Unidos.

En promedio de días suponiendo que tienen distribución normal es 536 días (18 meses) y con un

cima	crisis	indicador	días después del indicador
1973-11-01	1975-03-01	1978-08-18	NaT
1980-01-01	1980-07-01	1980-09-12	501 days
1981-07-01	1982-11-01	1988-12-13	292 days
1990-07-01	1991-03-01	1998-06-09	565 days
2001-03-01	2001-11-01	2005-12-29	996 days
2007-12-01	2009-06-01	2019-08-27	702 days
2020-02-01	2020-04-01	2022-04-01	158 days

Figura 5: Días antes que se observe una recesión económica en Estados Unidos

desvió estándar muestral de 298 días (10 meses). Finalmente, suponiendo que desde la última recesión fue 158 días de que el spread fue negativo la próxima recesión económica iniciaría el 06 de septiembre de 2022.



### 13. Apéndice 2: Resultado de estimación del modelo probit

Todas las estimaciones se realizaron en el lenguaje de programación Gretl.

Modelo 2: Probit, usando las observaciones 1960:09–2018:07 ( $T = 695$ )

Variable dependiente: y

Desviaciones típicas QML

	Coefficiente	Desv. Típica	$z$	Pendiente*
const	-1,83591	0,105666	-17,37	
payems	-0,652482	0,104062	-6,270	-0,0482577
aaaffm	-0,489009	0,116508	-4,197	-0,0361672
t5yffm6mlag	-0,475092	0,114800	-4,138	-0,0351379
Media de la vble. dep.	0,122302	D.T. de la vble. dep.	0,327870	
$R^2$ de McFadden	0,459358	$R^2$ corregido	0,443866	
Log-verosimilitud	-139,5847	Criterio de Akaike	287,1694	
Criterio de Schwarz	305,3450	Hannan-Quinn	294,1977	

\*Evaluado en la media

Número de casos 'correctamente predichos' = 630 (90,6 por ciento)

Contraste de razón de verosimilitudes:  $\chi^2(3) = 237,197$  [0,0000]

Contraste de normalidad de los residuos -

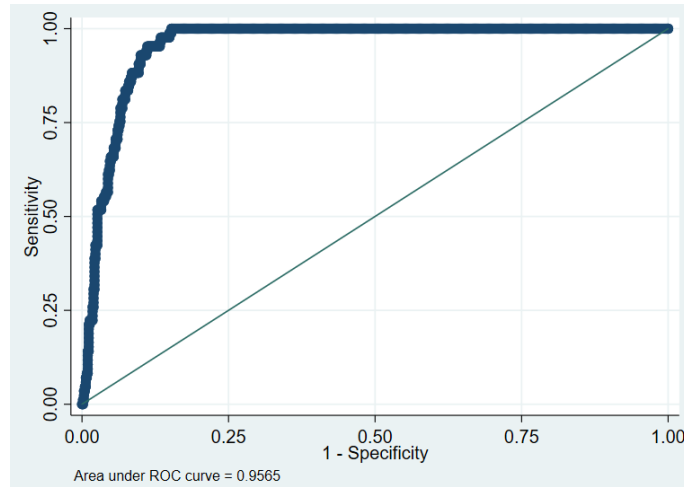
Hipótesis nula: [El error tiene distribución Normal]

Estadístico de contraste:  $\chi^2(2) = 239,101$

con valor p = 1,20178e-52

## 14. Apéndice 3: Curva ROC y Bondad de Ajuste

Se puede observar que la curva AUC es 0,9565 y una correcta clasificación de 90,65 %, mostrando un muy buen ajuste del modelo Probit.



Classified	True		Total
	D	~D	
+	36	16	52
-	49	594	643
Total	85	610	695

Classified + if predicted  $\Pr(D) \geq .5$   
 True D defined as  $y \neq 0$

Sensitivity	$\Pr(+ D)$	42.35%
Specificity	$\Pr(- \sim D)$	97.38%
Positive predictive value	$\Pr(D +)$	69.23%
Negative predictive value	$\Pr(\sim D -)$	92.38%
False + rate for true ~D	$\Pr(+ \sim D)$	2.62%
False - rate for true D	$\Pr(- D)$	57.65%
False + rate for classified +	$\Pr(\sim D +)$	30.77%
False - rate for classified -	$\Pr(D -)$	7.62%
Correctly classified		90.65%

## 15. Apéndice 4: Resultado de estimación del modelo CAPM

Todas las estimaciones se realizaron en el lenguaje de programación Gretl.

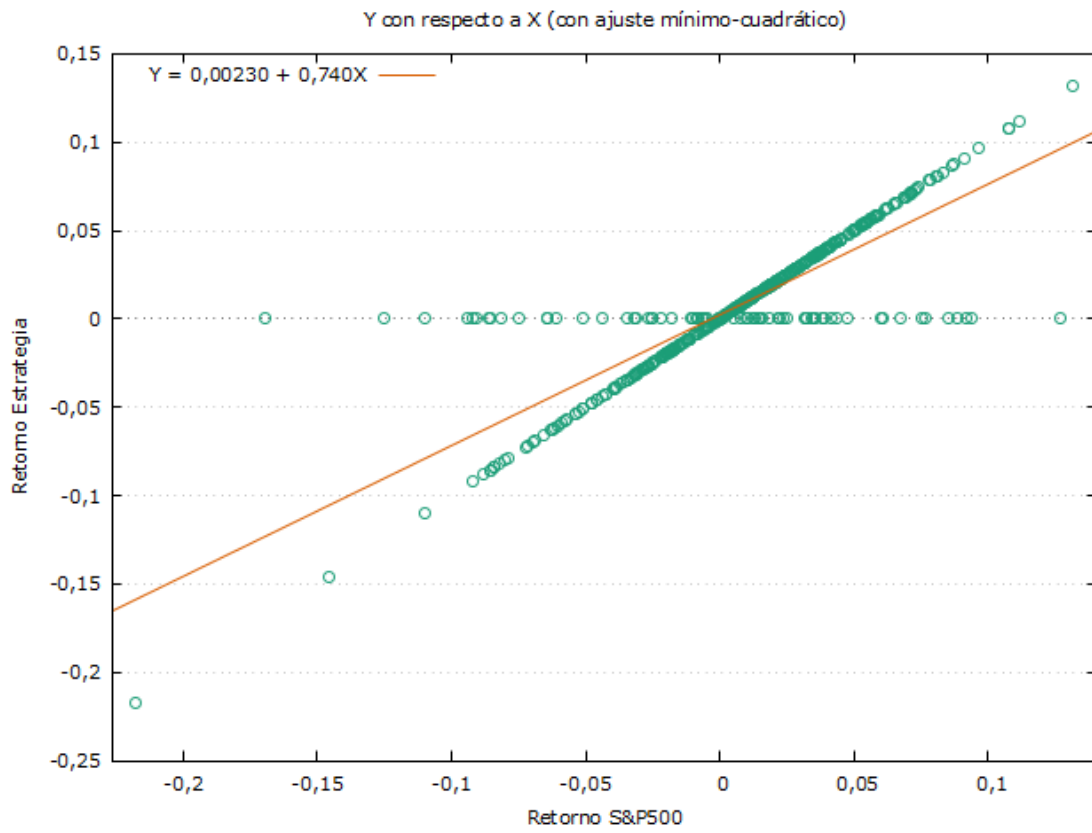
Modelo 1: MV, usando las observaciones 1985:02–2022:08 ( $T = 451$ )

Variable dependiente:  $R_E$

Mediante independientes:  $R_M$

	Coeficiente	Desv. Típica	$z$	valor p	$L_I$	$L_S$
$\hat{\alpha}$	0,0023331	0,0009194	2,54	0,011	0,0005311	0,0041351
$R_M$	0,7380859	0,0206865	35,68	0,000	0,6975411	0,7786308

Media de la vble. dep.	0,008112	D.T. de la vble. dep.	0,037649
Suma de cuad. residuos	0,165044	D.T. de la regresión	0,019172
Chi-cuadrado(1)	1273,0313	valor p	0,000000
Log-verosimilitud	1143,372233	Criterio de Akaike	-5,061518
Criterio de Schwarz	-2743,883		



## 16. Apéndice 5: Matriz de correlaciones

Solo se observan algunas etiquetas debido a que son 709 variables, cabe destacar que la variable dicotómica es RPI (Las correlaciones se calcularon con el language de programación Python).



## Autorización para publicar los trabajos finales

Completar cada punto con SI o NO:

- **Repositorio Institucional** (completar con SI o NO):

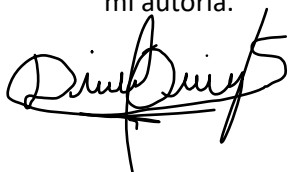
  SI   autorizo a la Universidad del CEMA a publicar y difundir en el **Repositorio Institucional** de la Universidad de la Biblioteca con fines exclusivamente académicos y didácticos el Trabajo Final de mi autoría.

- **Catálogo en línea** (completar con SI o NO):

  SI   autorizo a la Universidad del CEMA a publicar y difundir en el **Catálogo en línea** (acceso con usuario y contraseña) de la Biblioteca con fines exclusivamente académicos y didácticos el Trabajo Final de mi autoría.

- **Página web UCEMA** (completar con SI o NO):

  SI   autorizo a la Universidad del CEMA a publicar y difundir en la **página web de la Universidad** como Trabajo destacado, si el mismo obtuviese la distinción correspondiente, con fines exclusivamente académicos y didácticos el Trabajo Final de mi autoría.



Firma digital:

Aclaración: David Quiroz Sillo

DNI: 4755522