EL COLEGIO DE MÉXICO
CENTRO DE ESTUDIOS ECONÓMICOS

MAESTRÍA EN ECONOMÍA

TRABAJO DE INVESTIGACIÓN PARA OBTENER EL GRADO DE MAESTRO EN ECONOMÍA

UNA PRUEBA DEL CAPM CONDICIONAL PARA MÉXICO CON TÉCNICAS NO-PARAMÉTRICAS

JORGE HUMBERTO DEL CASTILLO SPÍNDOLA

PROMOCIÓN 1998 - 2000

ASESOR:

MTRO. ENEAS CALDIÑO GARCÍA

2005

Una Prueba del CAPM Condicional para México con Técnicas No-paramétricas

Jorge Humberto Del Castillo

Agradecimientos

Primero que nada agradezco a el Profr. Eneas Caldiño García toda su ayuda y gran paciencia en la realización de este trabajo, asi como su apoyo y grandes clases durante mis estudios en El Colegio de México. Muchas gracias Eneas.

Quiero agradecer al Centro de Estudios Económicos de El Colegio de México por todo el apoyo que he recibido desde mi ingreso a la institución. En particular a los diferentes coordinadores de la maestría con los que he tenido oportunidad de estar hasta la fecha: al Dr. Jaime Sempere (gracias por el gran apoyo, confianza y ayuda durante la etapa final de este trabajo), al Dr. Gerardo Esquivel (gracias por la ayuda, el apoyo y los consejos a lo largo de esta etapa) y al Dr. Alejandro Castañeda (gracias por el apoyo y confianza durante la primera etapa de mis estudios). Asimismo, agradezco el apoyo y confianza del Director del Centro. Dr. Horacio Sobarzo.

Un agradecimiento muy fuerte merecen todos aquellos profesores cuyas clases y/ó amistad influyeron determinantemente en mi formación. Con mucho cariño a Eneas Caldiño, Oscar Fernández, Gerardo Esquivel, Carlos Urzúa, Jaime Sempere, Graciela Márquez, Jorge Fernández, Subir, Francisco Venegas y José Antonio Nuñez.

De igual forma un agradecimiento a Mercedes por su ayuda y en especial a Monica Vargas por su gran ayuda y amistad durante estos años, ambas del Centro de Estudios Económicos.

Finalmente, agradezco a los que siempre han estado ahí dando todo su apoyo: con todo mi amor a mi familia.

Dedicatoria

Dedico este trabajo a Karla Salas. Por toda su ayuda en la elaboración y por todo lo que me ha enseñado en este tiempo.

Contenido

Introducción ix

1 El CAPM y el CAPM Condicional 1

2 Marco Teórico e Instrumentación Empírica 7

3 Los Datos y Resultados 15

4 Apéndice 19

Bibliografía 27

INTRODUCCIÓN

El Modelo de Valuación de Activos de Capital, CAPM de su nombre en inglés (Capital Asset Pricing Model) fué desarrollado inicialmente por Treynor (1961) y Sharpe (1964) y posteriormente por Lintner (1965), Mossin (1966) y Black (1972) y es una extensión inmediata de la teoría de Markowitz, con la ventaja de ser empíricamente más manejable.
El CAPM determina el precio de los activos considerando que los inversionistas han tomado decisiones óptimas cuando el mercado está en equilibrio. De esto se deriva que los activos riesgosos sólo pagan por el riesgo sistemático o no diversificable.
Supuestos del CAPM

1) Los inversionistas son aversos al riesgo y buscan maximizar su utilidad, la cual está en función de su riqueza terminal.

2) Las decisiones de portafolio se toman de acuerdo a criterios de media y varianza.

3) Los inversionistas tienen expectativas homogéneas.

4) La información está disponible para todos.

5) No hay costos de transacción ni impuestos.

6) Todos los activos están en el mercado y son divisibles.

7) Existe un portafolio de mercado que consiste en la tenencia de todos los activos en la proporción correspondiente a su valor de mercado relativo.

De manera opcional, algunas versiones (la nuestra es una) suponen que existe una tasa libre de riesgo con la cual se presta y se pide prestado infinitamente. Esta tasa tiene varianza cero, covarianza cero con los otros activos y da un rendimiento positivo.
El CAPM se construye examinando el comportamiento de los inversionistas en una economía modelo hipotética de un solo periodo. Entonces, para la examinación empírica es necesario hacer supuestos. Uno de los supuestos comunes es que las betas de los activos permanecen constantes a lo largo del tiempo. esto no es muy razonable ya que el riesgo relativo del flujo de efectivo de una empresa puede variar durante el ciclo de negocio. Por ejemplo, en Estados Unidos, el famoso artículo de Fama y French (1992) examina empíricamente la versión del CAPM de Black (1972), encontrando que el valor estimado del coeficiente de la beta es casi cero. Ellos interpretan esta relación plana entre el retorno promedio y beta como evidencia fuerte contra el CAPM. Sin embargo, esta no es necesariamente evidencia en contra de la versión condicional del CAPM, la cual toma en cuenta la naturaleza de la información disponible en cada momento.
Aún cuando los retornos esperados son lineales en las betas para cada periodo de tiempo, basado en la información disponible en el momento, la relación entre los retornos esperados no condicionales y la beta no condicional podría ser plana. En el capítulo 1 se expone un

ejemplo que permite verificar esto y el cual ayuda a ilustrar las fallas contenidas en un estudio empírico del CAPM que no tome en cuenta la variación en el tiempo de las betas.
En consecuencia, las betas y los retornos esperados dependerán en general de la naturaleza de la información disponible en cualquier punto del tiempo dado y variarán a lo largo de este.
Por todo esto, supondremos que el CAPM condicional se cumple, es decir, el retorno esperado de un activo basado en la información disponible en cualquier punto del tiempo dado es lineal en su beta condicional.
Ahora justifiquemos el porque del uso de técnicas de estadística no-paramétrica.
Los modelos condicionales de valuación de activos imponen restricciones en momentos condicionales no observados de los retornos de los activos, pero no describen exactamente como estos momentos varían a lo largo del tiempo. Por ejemplo, el CAPM condicional establece que los retornos en exceso esperados condicionales de los activos en cualquier punto del tiempo son lineales en sus covarianzas condicionales con el mercado, pero no dice nada acerca de la variación en el tiempo de los retornos y covarianzas esperados condicionales. Sin embargo, para realizar pruebas empíricas se necesita evaluar algunos momentos condicionales o funciones de estos en su variación a lo largo del tiempo.
Lo que se hace actualmente en la literatura es suponer formas funcionales acerca de los momentos condicionales de los retornos de los activos y/o en funciones de estos momentos para obtener series de tiempo y construir las pruebas. Sin embargo, errores de especificación tienen gran impacto en la inferencia y diferentes especificaciones pueden facilmente producir resultados empíricos diferentes.
No hay un consenso general acerca de que momentos condicionales de los retornos modelar y como serán especificados. De hecho, debido a que las especificaciones se escogen por conveniencia o facilidad de uso, estas son arbitrarias y mutuamente excluyentes.
También, se ha observado que existe un problema de hipótesis conjunta. Esto es, si la prueba de un modelo se basa en un modelo estadístico para algunos momentos condicionales de los retornos, se tiene simultaneamente una prueba de la predicción de la valuación y del modelo auxiliar. Aún cuando el modelo de valuación sea correcto, este tipo de prueba puede producir un rechazo y estimaciones grandes de errores de valuación simplemente por suponer una forma funcional pobre acerca de los momentos condicionales. Este problema es relevante a todas las metodologias propuestas en la literatura.
Desafortunadamente, el problema de la hipótesis conjunta no es el único. Problemas de potencia pueden facilmente surgir en muestras finitas debido simplemente a ciertas especificaciones.
Una implicación de los efectos de especificación es que uno puede cambiar de manera significativa los resultados de la prueba con solo alterar la especificación empírica.
Sin información detallada de los efectos que causa la especificación que se escoge, cómo podemos decidir si un resultado empírico es evidencia confiable o el simple producto de cierta especificación?
En este trabajo, siguiendo a Qing Wang (1997) se construye una prueba que evita los efectos de especificación asociados con los metodos existentes. Se prueba la eficiencia media-varianza condicional de un portafolio de mercado dado. Esta prueba se basa en la siguiente idea, primero se tiene un modelo de regresión lineal para los errores de los retornos esperados condicionales, el cual siempre es consistente con la eficiencia media-varianza condicional del

portafolio de mercado. A través de un factor de descuento no-paramétrico, se obtiene un estimador de mínimos cuadrados ponderados para los coeficientes de regresión. Entonces, se prueba si los coeficientes son cero usando el estimador y sus propiedades asintóticas para obtener una estadística de prueba. Con esta metodologia se evita el problema de la hipótesis conjunta y la prueba está libre del impacto de errores de especificación del factor de descuento.

En el primer capítulo, se revisan el CAPM y el CAPM condicional, así como se da una justificación del uso de un factor de descuento no-paramétrico observando lo que pasaría con el metodo generalizado de momentos.

En el capítulo 2 se desarrolla todo el marco teórico necesario para construir las pruebas y se detalla toda la instrumentación empírica.

En el capítulo 3 se detalla todos los datos a utilizar, sus propiedades estadísticas básicas, el porque de su elección y el resultado y conclusiones de la prueba.

Se incluye un apéndice con el programa desglosado para calcular las pruebas elaborado en MATLAB.

Capítulo 1

El CAPM y el CAPM Condicional

La mayoría de los estudios del CAPM estático (no condicional) suponen que las betas permanecen constantes a lo largo del tiempo y que el retorno en el portafolio de valores ponderados de todos los activos es una aproximación para el retorno en la riqueza agregada. Suponer que el CAPM se cumple en un sentido condicional es suponer que las betas y el premio de riesgo de mercado varían a lo largo del tiempo.
Gran parte de la investigación en finanzas es dedicada a entender como los inversionistas valuan los flujos de efectivo riesgosos. En general se esta de acuerdo en que los inversionistas demandan un mayor retorno esperado cuando invierten en activos mas riesgosos. Sin embargo, aún no se comprende del todo como los inversionistas manejan el riesgo del flujo de efectivo en un activo y como determinan que premio al riesgo demandar. Se han sugerido muchos modelos que describen como los inversionistas manejan el riesgo y valuan los flujos de efectivo riesgosos. Entre ellos el mas usado es el CAPM de Sharpe-Lintner-Black.
De acuerdo al CAPM:

i) La relación entre el retorno esperado requerido y beta es lineal.

Ahora bien, sabemos que el riesgo de un activo es medido por la beta del flujo de efectivo con respecto al retorno en el portafolio de mercado de todos los bienes en la economía.
Se han hecho muchos estudios para examinar empíricamente el desempeño del CAPM estático en explicar el efecto cruzado de los retornos promedio realizados. Los resultados obtenidos permiten ver que es posible construir un conjunto de portafolios donde el CAPM estático no es capaz de explicar la variación del efecto cruzado en los retornos promedio entre ellos. [1]

El CAPM se construye examinando el comportamiento de los inversionistas en una economía hipotética donde viven solo un periodo. Esto, obviamente no es cierto en el mundo real. Entonces, para la examinaciíon empírica del modelo con datos reales es necesario hacer ciertos supuestos. Uno de los supuestos comunes es que la beta de los activos permanecen constantes a lo largo del tiempo. Esto no es muy razonable ya que el riesgo relativo del flujo de efectivo de una empresa puede variar durante el ciclo de negocio. Por ejemplo, durante una recesión el apalancamiento financiero de empresas con problemas puede incrementarse abruptamente relativo al de otras empresas ocasionando que las betas de sus activos se incrementen. También, la participación relativa de diferentes sectores en la economía fluctua, induciendo cambios en las betas de las firmas en estos sectores. En consecuencia, las betas y

[1] Ver Banz(1981), Reinganum(1981), Gibbons(1982), Basu(1983), Chan, Chen y Hsieh(1985), Shanken(1985), y Bhandari(1988).

los retornos esperados dependerán en general de la naturaleza de la información disponible en cualquier punto del tiempo dado y variarán a lo largo de este.

Por todo esto, supondremos que el CAPM condicional se cumple, i.e., el retorno esperado de un activo basado en la información disponible en cualquier punto del tiempo dado es lineal en su beta condicional.

CAPM ESTATICO (Sharpe-Lintner-Black)

Denotemos por R_i al retorno en el activo i y por R_p al retorno en el portafolio de mercado de todos los activos en la economía.

La versión del CAPM de Black (1972) es:

$$E[R_i] = \gamma_0 + \gamma_1 \beta_i \tag{1.1}$$

donde

$$\beta_i = \frac{Cov(R_i, R_p)}{Var(R_p)}$$

En Estados Unidos, Fama y French (1992) examinan empíricamente esta versión, encontrando que el valor estimado de γ_1 es casi cero. Ellos interpretan esta relación plana entre el retorno promedio y beta como evidencia fuerte contra el CAPM. Sin embargo, esta no es necesariamente evidencia en contra del CAPM condicional.

Aún cuando los retornos esperados son lineales en las betas para cada periodo de tiempo, basado en la información disponible en el momento, la relación entre los retornos esperados no condicionales y la beta no condicional podría ser plana.[2]

El siguiente ejemplo nos permite ver esto:

Considérese una economía hipotética en donde se cumple el CAPM periodo por periodo. Supóngase que un econometrista considera solo dos activos y que hay solo dos periodos. Las betas del primer activo son 0.5 para el primer periodo y 1.25 para el segundo, dando una beta promedio de 0.875. Para el segundo activo son de 1.5 para el primer periodo y 0.75 para el segundo, dando una beta promedio de 1.125. Supóngase que el premio de riesgo esperado en el mercado es de 10 por ciento en el primer periodo y 20 por ciento en el segundo. Entonces, si el CAPM se cumple en cada periodo, el premio de riesgo esperado para el primer activo será 5 por ciento en el primer periodo (10×0.5) y 25 por ciento en el segundo (1.25×20). El premio de riesgo esperado para el segundo activo será 15 por ciento en ambos periodos.

En consecuencia, un econometrista que ignore el hecho de que las betas y los premios de riesgo varían en el tiempo, erroneamente concluirá que el CAPM no se cumple, ya que los dos activos ganan un premio de riesgo promedio de 15 por ciento pero sus betas promedio difieren.

Aún cuando los números de este ejemplo son extremos y poco realistas, nos ayudan a ilustrar las fallas contenidas en un estudio empírico del CAPM que no tome en cuenta la variación en el tiempo de las betas.

[2]Esto es debido a que un activo que está en la frontera de media-varianza condicional no necesariamente está en la frontera no condicional. Ver por ejemplo Dybvig y Ross(1985), y Hansen y Richard(1987)

CAPM CONDICIONAL

Denotaremos por $R_{i,t+1}$ el retorno bruto (i.e., 1 mas la tasa de retorno) del activo i en el periodo $t+1$. Análogamente $R_{p,t+1}$ será el retorno bruto del portafolio de riqueza agregada de todos los activos en la economía en el periodo $t+1$. Nos referiremos a $R_{p,t+1}$ como el retorno de mercado.

Sea I_t el conjunto de información común de los inversionistas al final del periodo t.

Supondremos que todas las series de tiempo usadas son covarianza- estacionarias y que todos los momentos condicionales y no condicionales existen.

Cuando se tienen inversionistas racionales aversos al riesgo en una economía dinámica, estos típicamente se anticiparán y cubriran contra la posibilidad de que las oportunidades de inversión en el futuro pudiesen cambiar adversamente. Debido a esta necesidad de cubrirse que surge en una economía dinámica, el retorno esperado condicional en un activo será de manera típica conjuntamente lineal en la beta de mercado condicional y las betas del portafolio para cubrirse. [3]

Sin embargo, siguiendo el desarrollo de Merton (1980), supondremos que los motivos para cubrirse no son suficientemente importantes y en consecuencia que el CAPM se cumplirá de manera condicional como sigue:

Para cada activo i y en cada periodo $t+1$,

$$E[R_{i,t+1}|I_t] = \gamma_{0,t} + \gamma_{1,t}\beta_{i,t} \tag{1.2}$$

donde $\beta_{i,t}$ es la beta condicional del activo i definida como

$$\beta_{i,t} = \frac{Cov(R_{i,t+1}, R_{p,t+1}|I_t)}{Var(R_{p,t+1}|I_t)} \tag{1.3}$$

$\gamma_{0,t}$ es el retorno esperado condicional en un portafolio cero-beta y $\gamma_{1,t}$ es el premio de riesgo de mercado condicional.

Ahora tomemos la esperanza no condicional en ambos lados de (2) para obtener:

$$E[E[R_{i,t+1}|I_t]] = E[\gamma_{0,t} + \gamma_{1,t}\beta_{i,t}]$$

de donde obtenemos

$$E[R_{i,t+1}] = E[\gamma_{0,t}] + E[\gamma_{1,t}]E[\beta_{i,t}] + Cov(\gamma_{1,t}, \beta_{i,t})$$

definiendo $\gamma_0 = E[\gamma_{0,t}]$, $\gamma_1 = E[\gamma_{1,t}]$ y $\overline{\beta_i} = E[\beta_{i,t}]$ obtenemos finalmente

$$E[R_{i,t+1}] = \gamma_0 + \gamma_1\overline{\beta_i} + Cov(\gamma_{1,t}, \beta_{i,t}) \tag{1.4}$$

γ_1 es el premio de riesgo de mercado esperado y $\overline{\beta_i}$ es la beta esperada.

Si la covarianza en (1.4) es cero ó una función lineal de la beta esperada para cada activo i escogido arbitrariamente, entonces este se asemeja al CAPM estático, i.e., el retorno esperado es una función lineal de la beta esperada. Sin embargo, en general, el premio de riesgo condicional en el mercado y las betas condicionales estan correlacionados.

[3] Ver Merton (1973) y Long (1974)

Por estudios previos se sabe que el premio de riesgo esperado en el mercado, así como las betas condicionales no son constantes[4] y varían a lo largo del ciclo económico.[5]

En consecuencia, el último término en (4) en general no es cero y el retorno esperado no condicional no es una función lineal de la beta esperada solamente.
Una vez que hemos visto esto, analizaremos el CAPM condicional a partir del factor de descuento. En la siguiente sección se argumenta el porque del uso de un factor de tipo no-paramétrico.

UN FACTOR DE DESCUENTO NO-PARAMETRICO

Utilicemos el factor de descuento del método generalizado de momentos[6] para discutir los efectos que causa el tipo de especificación que se seleccione.
Para probar la hipótesis de que un portafolio benchmark p dado es eficiente media-varianza condicionalmente con el factor de descuento del método generalizado de momentos, primero se especifica un factor de descuento lineal en $r_{p,t+1}$

$$m_{t+1}(\theta) = a(X_t, \theta) - b(X_t, \theta) r_{p,t+1}$$

donde
$r_{p,t+1}$ es retorno del portafolio p menos la tasa libre de riesgo.
X_t es un vector de variables de estado.
θ es un vector de parámetros.

Ahora, se aplica el procedimiento del método generalizado de momentos para probar si se cumple la condición

$$E[m_{t+1}(\theta_0) r_{t+1}] = 0$$

Para algún valor del vector de parámetros θ_0, donde r_{t+1} es el vector de retornos menos la tasa libre de riesgo.
Como saber si un factor de descuento está especificado de manera correcta?
La hipótesis de eficiencia condicional predice un factor de descuento lineal en $r_{p,t+1}$, pero no dice nada acerca de la forma funcional de a y de b. Sin embargo, aparentemente no tiene sentido que cualquier a y b sirvan para una prueba apropiada de la hipótesis.
La hipótesis de eficiencia media-varianza condicional no deja sin restricción a a y b.
Es directo ver que la ecuación condicional beta-precio

$$E[r_{i,t+1}|I_t] = E[r_{p,t+1}|I_t] \frac{Cov(r_{i,t+1}, r_{p,t+1}|I_t)}{Var(r_{p,t+1}|I_t)}$$

tiene una representación del factor de descuento

$$E[m_{t+1} r_{t+1}] = 0$$

[4] Keim y Stambaugh(1986), Breen, Glosten y Jagannathan(1989)
[5] Fama y French(1989), Chen(1991) y Ferson y Harvey(1991)
[6] ya que es muy frecuente su uso en los estudios actuales

donde

$$m_{t+1} = E[r^2_{p,t+1}|X_t] - E[r_{p,t+1}|X_t]r_{p,t+1}$$

De aquí se desprende que las especificaciones correctas de a y b ó un factor de descuento especificado correctamente deben cumplir

$$\frac{b(X_t,\theta_0)}{a(X_t,\theta_0)} = \frac{E[r_{p,t+1}|X_t]}{E[r^2_{p,t+1}|X_t]}$$

para algún θ_0.

Especificaciones incorrectas de a y b dan origen a un problema de hipótesis conjunta.

El procedimiento del factor de descuento del método generalizado de momentos prueba de manera simultánea la eficiencia media-varianza condicional del benchmark y los supuestos de la forma funcional de a y b. Errores de especificación pueden facilmente resultar en violaciones de la condición $E[m_{t+1}(\theta_0)r_{t+1}] = 0$ (el objetivo de la prueba) aún cuando el benchmark este en la frontera de eficiencia condicional.

El problema de la hipótesis conjunta no es la única posible consecuencia del error de especificación. Especificaciones incorrectas de a y b pueden afectar el tamaño de prueba, la potencia de la prueba, el sesgo de estimación y el ruido de estimación.

Los efectos de especificación originan serias dudas acerca del uso de factores de descuento paramétricos injustificados.

Capítulo 2

Marco Teórico e Instrumentación Empírica

UN ESTIMADOR DE MINIMOS CUADRADOS PONDERADO

Considérese un marco donde hay un activo condicionalmente sin riesgo. Igual que antes $r_{p,t+1}$ es el retorno del portafolio p menos la tasa libre de riesgo.
$r_{i,t+1}$ es el retorno del activo i menos la tasa libre de riesgo. $i = 1, \dots, n$.
Sea X_t un vector de tamaño k de variables de estado tal que

$$E[r_{p,t+1}|I_t] = E[r_{p,t+1}|X_t] \tag{2.1}$$
$$E[r^2_{p,t+1}|I_t] = E[r^2_{p,t+1}|X_t] \tag{2.2}$$

donde I_t es el conjunto de información de los inversionistas[1] al tiempo t.
Se utilizará el supuesto de que $r_{p,t+1}$, $r_{i,t+1}$ y las variables de estado son estrictamente estacionarias.
Si el portafolio benchmark p es eficiente media-varianza condicionalmente, entonces

$$E[r_{i,t+1}|I_t] = E[r_{p,t+1}|I_t]\frac{Cov(r_{i,t+1}, r_{p,t+1}|I_t)}{Var(r_{p,t+1}|I_t)} \tag{2.3}$$

o de manera equivalente

$$E[r_{i,t+1}|I_t] = E[r_{p,t+1}|I_t]\frac{E[r_{i,t+1}r_{p,t+1}|I_t]}{E[r^2_{p,t+1}|I_t]} \tag{2.4}$$

para $i = 1, \dots, n$.
La representación por medio de la covarianza (2.3) es la usual ecuación beta-precio. La representación de (2.4) es la de momento cruzado.
Hagamos

$$\begin{aligned} g_p(X_t) &= E[r_{p,t+1}|X_t] \\ g_{pp}(X_t) &= E[r^2_{p,t+1}|X_t] \\ b(X_t) &= \frac{g_p(X_t)}{g_{pp}(X_t)} \end{aligned}$$

[1]Nótese que estas ecuaciones son solo para el portafolio p. No se requiere que X_t sea una caracterización completa del conjunto de información I_t, solo que sea suficiente para desarrolllar la prueba no-paramétrica.

Si (2.1) y (2.2) se cumplen, los errores de los retornos esperados condicionales de (4) se pueden expresar como

$$E[r_{i,t+1}|I_t] - E[r_{p,t+1}|I_t]\frac{E[r_{i,t+1}r_{p,t+1}|I_t]}{E[r^2_{p,t+1}|I_t]} = E[m_{t+1}r_{i,t+1}|I_t]$$

donde

$$m_{t+1} = 1 - b(X_t)r_{p,t+1}$$

y en consecuencia (2.4) es equivalente a

$$E[m_{t+1}r_{i,t+1}|I_t] = 0 \qquad (2.5)$$

Ahora bien, sea $e_{i,t+1} = m_{t+1}r_{i,t+1}$ y Z_t un vector de tamaño q de variables estacionarias observadas en I_t.
Si se pudiese observar el factor de descuento m_{t+1}, una manera natural de probar la condición $E[e_{i,t+1}|I_t] = 0$ sería hacer una regresión de $e_{i,t+1}$ en Z_t y probar si los coeficientes son cero. Esto es debido a que las regresiones siguientes

$$e_{i,t+1} = Z_t'\delta_i + u_{i,t+1} \qquad (2.6)$$

donde $E[u_{i,t+1}|I_t] = 0$ para $i = 1, \ldots, n$ son siempre consistentes con (2.5).
Obviamente (2.5) implica que (2.6) se cumple con $\delta = 0$ donde $\delta = (\delta_1'\delta_2' \ldots \delta_n')'$.
Para implementar esta idea, siguiendo el desarrollo de Wang (1997) se reemplazará m_{t+1} con un factor de descuento no-paramétrico $\widehat{m}_{t+1}$ y se estimará el vector de parámetros δ_i con

$$\widehat{\delta}_i = \left(\frac{1}{N}\sum_{t=1}^{N}\widehat{w}_t Z_t Z_t'\right)^{-1}\left(\frac{1}{N}\sum_{t=1}^{N}\widehat{w}_t Z_t \widehat{e}_{i,t+1}\right) \qquad (2.7)$$

para $i = 1, \ldots, n$ donde

$$\begin{aligned}
\widehat{e}_{i,t+1} &= \widehat{m}_{t+1}r_{i,t+1} \quad \text{y} \\
\widehat{m}_{t+1} &= 1 - \widehat{b}(X_t)r_{p,t+1} \quad \text{con} \\
\widehat{b}(X) &= \frac{\widehat{g}_p(X)}{\widehat{g}_{pp}(X)}
\end{aligned}$$

La función ponderadora se escoge como

$$\widehat{w}_t = \widehat{f}(X_t)\widehat{g}_{pp}(X_t)$$

$\widehat{f}$, $\widehat{g}_p$ y $\widehat{g}_{pp}$ son kernels definidos como

$$\begin{aligned}
\widehat{f}(X) &= N^{-1}h^{-k}\sum_{s=1}^{N}K\left(\frac{X - X_s}{h}\right) \\
\widehat{g}_p(X) &= N^{-1}h^{-k}\widehat{f}(X)^{-1}\sum_{s=1}^{N}K\left(\frac{X - X_s}{h}\right)r_{p,s+1}
\end{aligned}$$

$$\widehat{g}_{pp}(X) \;=\; N^{-1}h^{-k}\widehat{f}(X)^{-1}\sum_{s=1}^{N}K\left(\frac{X-X_s}{h}\right)r^2_{p,s+1} \qquad (2.8)$$

$\widehat{f}$ es el estimador no-paramétrico de densidad Rosenblatt-Parzen, con kernel $K(\cdot)$ y banda de ancho h. $\widehat{g}_p$ y $\widehat{g}_{pp}$ son estimadores de función de regresión Nadaraya-Watson.
La función ponderadora $\widehat{w}_t$ se escogió de tal forma que $\widehat{w}_t Z_t Z_t'$ y $\widehat{w}_t Z_t \widehat{e}_{i,t+1}$ pueden ser expresadas como estadísticas U generalizadas de segundo orden, permitiendo analizar propiedades de $\widehat{\delta}_i$ para muestras grandes.
La prueba que se propone está basada en el estimador de mínimos cuadrados ponderado $\widehat{\delta}_N$

$$\widehat{\delta}_N = \left(\widehat{\delta}_1 \quad \widehat{\delta}_2 \cdots \widehat{\delta}_n\right)'$$

Intuitivamente, $\widehat{\delta}_N$ converge a cero si el benchmark p es eficiente media-varianza condicionalmente. De otro modo, el estimador converge a un límite diferente de cero[2]. Entonces, se puede hacer una prueba de eficiencia condicional checando que tanto dista $\widehat{\delta}_N$ de cero usando teoría de la distribución asintótica para considerar los errores de muestreo.
(2.6) también puede ser visto como un modelo para los errores de valuación; i.e., $Z_t'\delta_i$ serviría como aproximación para $E[e_{i,t+1}|I_t]$, los errores de retorno esperado condicional en (2.4). Por esto es que no se ha puesto ninguna restricción para escoger Z_t.
Nótese que el modelo de regresión (2.6) no ocasiona el problema de hipótesis conjunta, sin importar la elección que se haga de Z_t. Esto es debido a que el modelo siempre es correcto cuando el benchmark es eficiente media-varianza condicional.

COMPORTAMIENTO ASINTOTICO DEL ESTIMADOR Y LA ESTADISTICA DE PRUEBA

Sea r_{t+1} un vector de excesos de retorno (retorno en el activo menos la tasa libre de riesgo) $(r_{1,t+1}\cdots r_{n,t+1})' \otimes Z_t$ y sea $Y_{t+1} = (X_t' Z_t' r_{p,t+1} r_{t+1}')'$, donde $\otimes$ es el operador de Kronecker. Denotemos

$$\begin{aligned} w_t &= f(X_t)g_{pp}(X_t) \\ A &= \mathbb{I}_n \otimes E[w_t Z_t Z_t'] \\ \widehat{A_N} &= \mathbb{I}_n \otimes N^{-1}\sum_{t=1}^{N}\widehat{w}_t Z_t Z_t' \end{aligned}$$

donde $\mathbb{I}_n$ es la matriz identidad de tamaño n.
Sea $\delta = (\delta_1' \cdots \delta_n')'$ con

$$\delta_i = \left(E[w_t Z_t Z_t']\right)^{-1} E[w_t Z_t e_{i,t+1}]$$

[2]a menos que $e_{i,t+1}$ sea ortogonal a todos los componentes de Z_t para $i = 1,\ldots,n$

Definamos

$$\begin{aligned}
\gamma(Y_{t+1}) &= \eta(Y_{t+1}) - [\mathbb{I}_n \otimes a(Y_{t+1})]\delta & (2.9)\\
\eta(Y_{t+1}) &= f(X_t)[g_{pp}(X_t)r_{t+1} - g_p(X_t)r_{p,t+1}r_{t+1} + g_r(X_t)r^2_{p,t+1} - g_{pr}(X_t)r_{p,t+1}] & (2.10)\\
a(Y_{t+1}) &= f(X_t)[g_{pp}(X_t)Z_tZ_t' + r^2_{p,t+1}g_{zz}(X_t)] & (2.11)
\end{aligned}$$

donde

$$\begin{aligned}
g_r(X_t) &= E[r_{t+1}|X_t]\\
g_{pr}(X_t) &= E[r_{p,t+1}r_{t+1}|X_t] \quad \text{y}\\
g_{zz}(X_t) &= E[Z_tZ_t'|X_t]
\end{aligned}$$

Resultado 1. Tenemos que

i) $\widehat{A}_N$ converge en probabilidad a A.

ii) La distribución límite de $\sqrt{N}\widehat{A}_N(\widehat{\delta}_N - \delta)$ es idéntica a la de $N^{-\frac{1}{2}}\sum_{t=1}^{N}\gamma(Y_{t+1})$.

iii) $E[\gamma(Y_{t+1})] = 0$.

Ahora, enunciaremos algunos supuestos técnicos que permiten obtener resultados acerca de la distribución límite del estimador $\widehat{\delta}_N$ y del estadístico de prueba que se dará mas adelante.

SUPUESTO 1.- La sucesión de datos $\{Y_{t+1}\}$ es un proceso estrictamente estacionario del tipo β-mixing y el subvector X_t tiene distribución absolutamente continua con densidad $f(X_t)$ y para alguna $\rho > 2$ los números $\beta_n \quad n = 1, 2, \ldots$ satisfacen $\sum_{n=1}^{\infty} n\beta_n^{\frac{(\rho-2)}{\rho}} < \infty$.

SUPUESTO 2.-

i) $r^2_{p,t+1}$, r_{t+1} y $r_{p,t+1}r_{t+1}$ tienen primer momento finito.

ii) $||w_1(t,s)||_\rho < \infty$ y $||w_2(t,s)||_\rho < \infty \quad \forall \quad t < s$ donde

$$\begin{aligned}
w_1(t,s) &\equiv (r^2_{p,s+1} - r_{p,s+1}r_{p,t+1})r_{t+1} + (r^2_{p,t+1} - r_{p,t+1}r_{p,s+1})r_{s+1}\\
w_2(t,s) &\equiv r^2_{p,s+1}Z_tZ_t' + r^2_{p,t+1}Z_sZ_s'
\end{aligned}$$

y $||\cdot||_\rho$ denota la norma [3] ρ.

iii) $||\eta(Y_{t+1})||_\rho < \infty$ y $||a(Y_{t+1})||_\rho < \infty$.

[3] $||X_{ij}||_\rho \equiv \left(E[X_{ij}^\rho]\right)^{\frac{1}{\rho}}$

SUPUESTO 3.- fg_p, fg_{pp}, fg_r, fg_{pr} y fg_{zz} satisfacen la condición local de Lipschitz para alguna función $m(X)$; donde $m(X_t)r_{t+1}$, $m(X_t)r_{p,t+1}r_{t+1}$, $m(X_t)r_{p,t+1}$, $m(X_t)r^2_{p,t+1}$ y $m(X_t)Z_tZ_t'$ tienen norma ρ finita.[4]

SUPUESTO 4.- El kernel K es una función simétrica acotada que satisface

i) $\int K(u)du = 1$

ii) $\int |u|^j |K(u)| du < \infty$ si $0 \leq j \leq k+1$

iii) $\int u_1^{l_1} \cdots u_k^{l_k} K(u)du = 0$ si $0 < l_1 + \cdots + l_k < k+1$ donde u_j es el j-ésimo elemento del vector u. Es decir, el nucleo K es de orden $k+1$

SUPUESTO 5.-

i) La j-ésima derivada parcial de fg_p, fg_{pp}, fg_r, fg_{pr} y fg_{zz} existen para toda $j \leq k+1$.

ii) Las esperanzas $E[g_{pr}\nabla_{l_1,\ldots,l_j}(fg_p)]$, $E[g_r\nabla_{l_1,\ldots,l_j}(fg_{pp})]$, $E[g_{pp}\nabla_{l_1,\ldots,l_j}(fg_r)]$, $E[g_p\nabla_{l_1,\ldots,l_j}(fg_{pr})]$ y $E[g_{pp}\nabla_{l_1,\ldots,l_j}(fg_{zz})]$ existen para toda $j \leq k+1$, donde las funciones y las derivadas parciales[5] están evaluadas en X_t.

SUPUESTO 6.- Las matrices A y Γ_0 son no-singulares. ($\Gamma_0 = E[\gamma(Y_{t+1})\gamma(Y_{t+1})']$).

La condición mixing en el supuesto 1 restringe el grado de dependencia permitida en la sucesión de datos, esto es entre otras cosas para poder aplicar un teorema de límite central. Las condiciones que requieren que β_n se anule como una potencia de n no suelen ser restrictivas para la mayoría de series financieras y son usuales en la literatura. Mientras mas fuertes son las restricciones para la existencia de los momentos (una mayor ρ) mas dependencia es permitida. Esta relación inversa no es poco común para establecer resultados asintóticos para datos serialmente correlacionados. El supuesto 5 es una condición de regularidad para corrección asintótica de sesgo, a través del uso de un nucleo de orden mayor.
Para las condiciones en los momentos hechas en el supuesto 2, primero nótese que ρ puede escogerse arbitrariamente cercana a 2 si β_n decae exponencialmente; en seguida, es directo verificar que el supuesto 2 se cumple $\forall \quad \rho > 2$ si la distribución conjunta de todas las variables es normal o lognormal.
Una vez dados los supuestos técnicos, observamos que debido a que $N^{-1}\sum_{t=1}^{N}\gamma(Y_{t+1})$ es un promedio simple de vectores aleatorios estacionarios, la aplicación de un teorema de límite central[6] nos da el siguiente resultado.

---

[4]Se dice que una función $h(X)$ satisface la condición local de Lipschitz para alguna función $m(X)$ si $|h(X+Y) - h(X)| < m(X)||Y||$

[5]$\nabla(h)$ denota el vector gradiente de la función h

[6]Doukhan et al. (1994) han mejorado los teoremas de límite central clásicos de Ibragimov y Linnik (1971) y probaron que si $2 < \rho < \infty$, $E[|X|^\rho] < \infty$ y $\sum_{n=1}^{\infty} \alpha_n n^{\frac{2}{(\rho-2)}} < \infty$ entonces $n^{-\frac{1}{2}}\sum_{i=1}^{n}(X_i - E[X_i])$ converge a un vector aleatorio normal centrado.

Teorema 2.1. Dados los supuestos 1-6, si $h \to 0$, $Nh^{2k} \to \infty$ y $Nh^{2k+2} \to 0$, entonces el estimador de mínimos cuadrados ponderado $\widehat{\delta}_N$ es tal que $\sqrt{N}(\widehat{\delta}_N - \delta)$ tiene una distribución límite normal multivariada con media 0 y matriz de varianza-covarianza Ω, donde $\Omega = A^{-1}\Gamma A^{-1}$, $\Gamma = \sum_{-\infty}^{\infty}\Gamma_j$ y $\Gamma_j = E[\gamma(Y_{t+1})\gamma(Y_{t+j+1})']$.

Este teorema muestra que $\widehat{\delta}_N$ tiene las propiedades límite estandar, consistencia-$\sqrt{N}$ y normalidad asintótica de los estimadores paramétricos.
Este resultado no se basa en (2.1) y (2.2), ni requiere que las ecuaciones de (2.6) esten especificadas correctamente. Hay que notar que las condiciones para h son diferentes de aquellas para estimadores de kernel puntuales. Las condiciones $Nh^{2k} \to \infty$ y $Nh^{2k+2} \to 0$ dan las cotas superior e inferior de la tasa en que h converge a 0 para que $\widehat{\delta}_N$ exhiba el comportamiento asintótico deseado. La condición $Nh^{2k+2} \to 0$ se debe al uso de un kernel de orden $k+1$ y en consecuencia el rango admisible para la tasa puede relajarse usando un kernel de orden mayor a $k+1$.
Ahora, para construir una prueba de eficiencia condicional usando la distribución de $\widehat{\delta}_N$ necesitamos un estimador para la matriz de covarianza Ω. Para esto, primero veamos la estimación de $\gamma(Y_{t+1})$. Reemplazando en (2.10) y (2.11) $f(X)$, $g_p(X)$, $g_{pp}(X)$, $g_r(X)$, $g_{pr}(X)$ y $g_{zz}(X)$ por kernels estandar y reemplazando en (2.9) δ por $\widehat{\delta}_N$, una aproximación natural para $\gamma(Y_{t+1})$ sería

$$\begin{aligned}
\widehat{\gamma}_N(Y_{t+1}) &= \widehat{\eta}_N(Y_{t+1}) - [\mathbb{I}_n \otimes \widehat{a}_N(Y_{t+1})]\widehat{\delta}_N && (2.12)\\
\widehat{\eta}_N(Y_{t+1}) &= \widehat{f}(X_t)[\widehat{g}_{pp}(X_t)r_{t+1} - \widehat{g}_p(X_t)r_{p,t+1}r_{t+1} + \widehat{g}_r(X_t)r_{p,t+1}^2 - \widehat{g}_{pr}(X_t)r_{p,t+1}] && (2.13)\\
\widehat{a}_N(Y_{t+1}) &= \widehat{f}(X_t)[\widehat{g}_{pp}(X_t)Z_tZ_t' + r_{p,t+1}^2\widehat{g}_{zz}(X_t)] && (2.14)
\end{aligned}$$

con $\widehat{f}$, $\widehat{g}_p$ y $\widehat{g}_{pp}$ como se definieron previamente[7] y

$$\begin{aligned}
\widehat{g}_r(X) &= N^{-1}h^{-k}\widehat{f}(X)^{-1}\sum_{s=1}^{N} K\left(\frac{X-X_s}{h}\right)r_{s+1}\\
\widehat{g}_{pr}(X) &= N^{-1}h^{-k}\widehat{f}(X)^{-1}\sum_{s=1}^{N} K\left(\frac{X-X_s}{h}\right)r_{p,s+1}r_{s+1}\\
\widehat{g}_{zz}(X) &= N^{-1}h^{-k}\widehat{f}(X)^{-1}\sum_{s=1}^{N} K\left(\frac{X-X_s}{h}\right)Z_sZ_s'
\end{aligned}$$

Se puede mostrar que el estimador

$$\widehat{\Gamma}_j = N^{-1}\sum_{t=1}^{N-j}\widehat{\gamma}_N(Y_{t+1})\widehat{\gamma}_N(Y_{t+j+1})'$$

es consistente para Γ_j. También puede verse que dado (2.1) y (2.2) $\Gamma_j = 0 \quad \forall \quad j \neq 0$ cuando las ecuaciones en (2.6) se cumplen. Es por esto que se propone el estadístico de prueba

$$\widehat{T}_\delta = N\widehat{\delta}_N'\widehat{\Omega}_N^{-1}\widehat{\delta}_N \qquad (2.15)$$

[7]Hay que notar que $g_{zz}(X_t) = Z_tZ_t'$ cuando Z_t es una transformación fija de X_t, por ejemplo $Z_t = (1X_t')'$ y en dado caso no es necesario utilizar el nucleo de estimación $\widehat{g}_{zz}(X_t)$ en (14). Simplemente habría que reemplazar $g_{zz}(X_t)$ por Z_tZ_t' lo que da $\widehat{a}_N(Y_{t+1}) = \widehat{f}(X_t)[\widehat{g}_{pp}(X_t) + r_{p,t+1}^2]Z_tZ_t'$.

donde $\widehat{\Omega}_N = \widehat{A}_N^{-1}\widehat{\Gamma}_0\widehat{A}_N^{-1}$ para probar la eficiencia media-varianza condicional.

Teorema 2.2. Si se cumplen las condiciones del Teorema 1

i) Dados (2.1) y (2.2), si el portafolio p es eficiente media-varianza condicional, entonces el estadístico de prueba $\widehat{T}_\delta$ tiene una distribución límite χ^2 con $q \times n$ grados de libertad.

ii) $\widehat{\Gamma}_j$ es un estimador consistente de Γ_j para cualquier j fija.

Vale la pena notar que la prueba no tendrá potencia si se escoge algún vector Z_t que sea ortogonal[8] a $e_{i,t+1}$. Sin embargo, la prueba tendrá potencia si un componente de Z_t puede pronosticar significativamente $e_{i,t+1}$, sin importar si el modelo de regresión (2.6) es especificado de manera correcta o no.

EL KERNEL K Y LA BANDA h

El kernel K utilizado es una función de densidad normal multivariada independiente

$$K(u) = \prod_{i=1}^{k} \phi_i(u_i)$$

donde ϕ_i es la densidad normal univariada con media cero y varianza σ_i^2 (σ_i es la desviación estandar de la i-ésima variable de estado). Para el computo, σ_i fué reemplazada por la desviación estandar muestral.
Para escoger la banda h óptima no hay un consenso teórico. Una opción práctica que toma en cuenta las condiciones de la tasa de convergencia para los resultados de distribución límite del Teorema 1 es

$$h = N^{-\frac{1}{2k+1}}$$

SELECCION DE LAS VARIABLES DE ESTADO

Es muy importante la selección del vector X_t debido a que no se quisiera perder variables predictoras importantes en el conjunto de información I_t con el cual se condiciona.
En el extremo, una prueba no va a ser eficiente condicionalmente si no usa ninguna variable de estado. Por otro lado, no se quisiera incluir variables redundantes que pudieran tener un impacto importante en la potencia de la prueba.
Lo que haremos es probar (2.1) y (2.2) para un candidato X_t dado. Para esto, se propondrá una prueba no-paramétrica.
Sea Z_t un vector de tamaño $q_1 \times 1$ en el conjunto I_t, cuyas componentes son diferentes a las de X_t. La idea para poder hacer la selección es verificar si Z_t puede pronosticar los residuales

[8] i.e., $E[Z_t e_{i,t+1}|X_t] = 0 \quad i = 1, \ldots, n$

$r_{p,t+1} - g_p(X_t)$ y $r^2_{p,t+1} - g_{pp}(X_t)$.
Si (2.1) se cumple (i.e., $E[r_{p,t+1}|I_t] = g_p(X_t)$) entonces

$$r_{p,t+1} - g_p(X_t) = Z_t'\mu + \epsilon_{t+1}$$

con $E[\epsilon_{t+1}|I_t] = 0$ y $\mu = 0$.
Bajo ciertas condiciones de regularidad, si $Nh^{2k} \to \infty$ y $Nh^{2k+2} \to 0$, entonces se va a tener que el estimador de μ

$$\widehat{\mu} = \left(\frac{1}{N}\sum_{t=1}^{N} \widehat{f}(X_t)Z_tZ_t'\right)^{-1}\left(\frac{1}{N}\sum_{t=1}^{N} \widehat{f}(X_t)Z_t[r_{p,t+1} - \widehat{g}_p(X_t)]\right)$$

es tal que

$$\sqrt{N}(\widehat{\mu} - \mu) \xrightarrow{D} N(0, \Omega_\mu)$$

Como en el Teorema 2, se puede construir un estimador de matriz de covarianzas

$$\widehat{\Omega}_\mu = \widehat{A}_\mu^{-1}\widehat{\Gamma}_\mu\widehat{A}_\mu^{-1}$$

donde

$$\begin{aligned}\widehat{A}_\mu &= N^{-1}\sum_{t=1}^{N} \widehat{f}(X_t)Z_tZ_t' \\ \widehat{\Gamma}_\mu &= N^{-1}\sum_{t=1}^{N} \widehat{\gamma}_{\mu,t+1}\widehat{\gamma}_{\mu,t+1}'\end{aligned}$$

con

$$\widehat{\gamma}_{\mu,t+1} = \widehat{f}(X_t)[Z_tr_{p,t+1} - Z_t\widehat{g}_p(X_t) - r_{p,t+1}\widehat{g}_z(X_t) + \widehat{g}_{zp}(X_t) - \widehat{g}_{zz}(X_t)\widehat{\mu} - Z_tZ_t'\widehat{\mu}]$$

$\widehat{g}_{zz}(X)$ como se definió previamente y

$$\begin{aligned}\widehat{g}_z(X) &= N^{-1}h^{-k}\widehat{f}(X)^{-1}\sum_{s=1}^{N} K\left(\frac{X - X_s}{h}\right)Z_s \\ \widehat{g}_{zp}(X) &= N^{-1}h^{-k}\widehat{f}(X)^{-1}\sum_{s=1}^{N} K\left(\frac{X - X_s}{h}\right)Z_sr_{p,s+1}\end{aligned}$$

Esto da un estadístico de prueba

$$Q = N\widehat{\mu}'\widehat{\Omega}_\mu^{-1}\widehat{\mu}$$

con distribución límite bajo (2.1) de $\chi^2(q_1)$.
La prueba para (2.2) se construye de manera similar unicamente reemplazando $r_{p,t+1}$ y $r_{p,s+1}$ con $r^2_{p,t+1}$ y $r^2_{p,s+1}$ respectivamente.

Capítulo 3

Los Datos y Resultados

SELECCION DE DATOS

Comencemos primero por describir los datos que representaran los retornos en exceso $r_{p,t+1}$ del portafolio de mercado p y a los que representaran los retornos en exceso $r_{i,t+1}$ $i = 1, \ldots, n$ de los activos ó portafolio que se usan para pronosticar.
La tasa libre de riesgo utilizada es la tasa mensual porcentual de los cetes a 28 días, la cual denotaremos λ. Para el portafolio p se utilizó el crecimiento porcentual del Indice de Precios y Cotizaciones menos la tasa libre de riesgo como sigue

$$r_{p,t+1} = \left(\frac{IPC_{t+1} - IPC_t}{IPC_t} \right) \times 100 - \lambda$$

Ahora bien, como el modelo está ubicado en un contexto macroeconómico y siguiendo a Qing Wang (1997) quien usa los deciles para tamaño impares del value-weighted portfolio of New York Stock Exchange, nosotros utilizaremos el crecimiento porcentual del IPC para los sectores

1) Transformación.

2) Construcción.

3) Comercio.

4) Transporte.

5) Servicios.

menos la tasa libre de riesgo como sigue

$$r_{i,t+1} = \left(\frac{IPC_{i,t+1} - IPC_{i,t}}{IPC_{i,t}} \right) \times 100 - \lambda \quad i = 1, \ldots, 5$$

de esto podemos observar que en nuestro estudio la n del capítulo 2 vale 5.
El análisis estadístico básico de estos datos es

portafolio	media
1	0.068965459
2	1.832799235
3	1.206859485
4	6.42690945
5	0.602764374
p	0.525386306

portafolio	desv. est.
1	9.15240845
2	21.85293699
3	10.60173723
4	66.83048961
5	14.7918615
p	10.05646583

$$\mathbf{C} = \begin{pmatrix} 1 & & & & & \\ 0.7505 & 1 & & & & \\ 0.7439 & 0.5501 & 1 & & & \\ 0.0155 & -0.0772 & -0.0001 & 1 & & \\ 0.6019 & 0.4838 & 0.5734 & -0.0527 & 1 & \\ 0.6432 & 0.4148 & 0.5597 & 0.1040 & 0.7328 & 1 \end{pmatrix}$$

Donde $\mathbf{C}$ es la matriz de correlaciones.
Sigamos con la obtención del vector X_t, el cual representará las variables dentro del conjunto de información de los inversionistas I_t con las cuales será suficiente para tomar decisiones de inversión.
El total de variables que fueron escogidas y que pudiesen conformar el vector X_t son las siguientes

1) Indice de Volumen Físico de la Actividad Industrial en crecimiento porcentual. Se tomó como medida del crecimiento de la economía. Lo denotaremos a.

2) Indice Nacional de Precios al Consumidor (INPC) en crecimiento porcentual. Se tomó como medida de la inflación. Lo denotaremos b.

3) Agregado Monetario M4 en crecimiento porcentual. Se tomó para captar la riqueza en la economía considerando la liquidez. Lo denotaremos c.

4) Tipo de cambio peso-dólar en crecimiento porcentual. Lo denotaremos d.

5) Diferencia entre la tasa mensual del papel comercial y λ. Se tomó como el premio de incumplimiento. Lo denotaremos e.

6) Indice Dow Jones en crecimiento porcentual. Se tomó por considerar las principales acciones que cotizan en el New York Stock Exchange, mercado donde se encuentran listadas las principales acciones mexicanas. Lo denotaremos f.

Las desviaciones estándar correspondientes σ_i utilizadas en el kernel K de las muestras son las siguientes

i	σ_i
1	4.489683122
2	2.339468186
3	2.357376626
4	6.402002774
5	0.388932972
6	4.535023512

Los datos utilizados en este estudio son mensuales y van de febrero de 1987 a diciembre de 2001 para el caso de las variables de estado. En cuanto a los retornos, van de marzo de 1987 a enero de 2002 (recorrido un periodo debido a su uso en $t+1$ y la forma en que se realizaron los programas computacionales). De esto vemos que $N = 179$.
Siguiendo el procedimiento para la selección de variables de estado descrito en el capítulo 2, se conformaron diferentes vectores z_t de tamaño 4 (i.e., q_1 del capítulo 2 vale 4) y para cada uno de estos se calculo el valor del estadístico de prueba

$$Q = N\widehat{\mu}'\widehat{\Omega}_{\mu}^{-1}\widehat{\mu}$$

con distribución límite bajo (2.1) de $\chi^2(q_1)$.
Se conformaron 5 vectores z, donde por ejemplo z_{1236} indica que se utilizaron las variables 1,2,3 y 6 (i.e., Indice de Volumen Físico de la Actividad Industrial, INPC, M4 y Dow Jones). Con esto se obtuvieron los siguientes estadísticos de prueba Q

z	Q
z_{1234}	$Q_{1234} = 1.4378$
z_{1236}	$Q_{1236} = 6.3058$
z_{1246}	$Q_{1246} = 6.1003$
z_{1245}	$Q_{1245} = 1.4048$
z_{1256}	$Q_{1256} = 6.9765$

Hay que notar por supuesto que el número de combinaciones posibles para formar los z a partir de las 6 variables es de 15 que es el número de combinaciones de 4 elementos tomados de un conjunto de 6 elementos, pero aquí consideramos sólo estas 5 posibilidades. Aún mas, es posible formar los z de tamaño 5 ó 3 lo que da mas posibilidades.
También observemos que aquí solo tenemos un total de 6 variables de estado para escoger, pero es posible hacer las pruebas con mas variables y con los correspondientes vectores z que se puedan construir a partir de estas. Para esto, se anexa un apéndice con el programa[1] para ejecutar todas las pruebas propuestas en este trabajo, el cual fué realizado en el lenguaje MATLAB[2].

[1]Consultar el apéndice para ver como el programa recibe los inputs (las series) de forma matricial

[2]Cualquier duda relacionada con el programa, su uso, los comandos ó las series de datos utilizadas, favor de contactarme a través del Centro de Estudios Económicos del Colegio de México.

Consultando los cuantiles que dejan una probabilidad a la derecha de una χ^2 de 5 y 10 por ciento, observamos que todas estas elecciones son aceptadas. Aún cuando debiese de escogerse a X_t como

$$X_t = (X_t^1 = a_t \quad X_t^2 = b_t \quad X_t^3 = d_t \quad X_t^4 = e_t)$$

debido a que $Q_{1245} = 1.4048$ es el valor que se acepta a mayor confianza, por completitud y para efectos de comparación realizaremos la prueba de eficiencia media-varianza condicional para el portafolio p para los 5 casos de X_t debido a que estas 5 pruebas de z fueron aceptadas (i.e., se acepto que los 5 z pueden pronosticar[3] X).

LA PRUEBA DE EFICIENCIA DEL PORTAFOLIO p

Para todas estas pruebas, Z_t se tomará como $Z_t = (1 \quad X_t')'$ (importante no confundir con el z que se acaba de utilizar arriba).
A continuación se reporta para cada uno de los casos el valor del estadístico de prueba

$$\widehat{T}_\delta \; = \; N\widehat{\delta}_N'\widehat{\Omega}_N^{-1}\widehat{\delta}_N \tag{3.1}$$

cuaya distribución es $\chi^2_{(q\times n)}$

z	Q
X_{1234}	$T_{1234} = 34.9501$
X_{1236}	$T_{1236} = 11.1104$
X_{1246}	$T_{1246} = 10.4051$
X_{1245}	$T_{1245} = 22.4002$
X_{1256}	$T_{1256} = 14.1639$
X_{123456}	$T_{123456} = 11.6842$

Donde para los primeros 5 estadísticos utilizamos $\chi^2_{(5\times5)}$ y para el último $\chi^2_{(7\times5)}$.
Los cuantiles que dejan una probabilidad a la derecha de una χ^2 con 25 grados de libertad de 5 y 10 por ciento son respectivamente 37.65 y 34.38, lo que nos permite concluir que p es eficiente media-varianza condicionalmente para todos los casos excepto para X_{1234} al 10 por ciento. Es importante observar que aún cuando una de las mejores elecciones de variables era X_{1234} debido a que Q_{1234} era de los valores mas bajos, con esta elección se rechaza para 10 por ciento la hipótesis de eficiencia. Aún mas, para una χ^2 con 35 grados de libertad, observamos que usando todas las variables de estado listadas anteriormente (i.e., para el caso X_{123456}) la hipótesis de eficiencia de p es aceptada ampliamente.
La información anterior permite concluir que con los datos utilizados hay evidencia a favor del modelo CAPM en su versión condicional para el caso de México.

[3] Ver la sección de Selección de las Variables de Estado del capítulo 2

Capítulo 4

Apéndice

En este apéndice se incluyen las instrucciones necesarias para poder utilizar los programas elaborados en MATLAB y poder correr las pruebas desarrolladas en este trabajo.
El hecho de poder capturar todas las series utilizadas en forma matricial y vectorial, así como el poder programar las pruebas con operaciones matriciales de manera simple, fué el motivo que inspiró la decisión del uso de MATLAB; el cual es un lenguaje con un soporte muy amplio y poderoso en el uso de matrices y las operaciones que envuelven a éstas.
Al final de este apéndice se encuentra el como elaborar los programas en MATLAB para que el lector pueda copiarlos (si desea realizar algunas pruebas con nuestros programas).

Para comenzar, se utilizarán las matrices A, Z y R, así como los vectores p y sigma y el valor numérico h.
La matriz A se conforma de los vectores X_t como columnas. Es decir, la columna 1 de la matriz A es el vector columna de variables de estado X_1 y así sucesivamente hasta la columna X_N (recordemos que N es nuestro horizonte temporal, i.e., el número de meses que deseamos utilizar en el estudio empírico). Observemos que A es de tamaño $k \times N$
Una vez que se está en la pantalla de comandos en MATLAB, aparece el prompt y se escribirá simplemente A=[ ]; . Dentro de los corchetes se incluirán los valeres de la matriz, los cuales pueden ser transportados de manera sencilla de EXCEL con un simple copiar-pegar (El punto y coma después del corchete permite que la matriz no sea desplegada y simplemente quede en memoria).
La matriz Z es igual que la matriz A con la salvedad de que le aumentamos un primer renglón de 1's, lo cual refleja que estamos utilizando $Z_t = (1 \quad X_t')'$. Recordemos que Z_t es de tamaño $q \times 1$ (en nuestro caso $q = k + 1$). La forma de capturar Z es análoga a la de A.
La matriz R contiene la información de los $r_{i,t+1}$ con $i = 1, \ldots, n$ de la siguiente manera: el renglón 1 hace referencia a $i = 1$ y así hasta el renglón n. Esto nos da una matriz de tamaño $n \times N$. Para no lidiar con el hecho de tener los retornos en el tiempo $t + 1$ en la programación, se utilizó t; por lo cual hay que ser cuidadosos y meter los datos de cada renglón recorridos una fecha. Por ejemplo, si en la matriz A se comenzó con enero del 87 y se terminó en diciembre de 2001, para R se comenzará en febrero del 87 y se terminará en enero de 2002.
El vector p contiene la información de los $r_{p,t+1}$ e igual se tiene que recorrer la captura en una fecha.
La captura en MATLAB de un vector es igual a la de una matriz, simplemente se escribe a continuación del prompt p=[ ]; y en medio de los corchetes un copiar-pegar directo de

EXCEL.
El vector sigma de tamaño $k \times 1$ recoge las correspondientes desviaciones estándar de los renglones de A, i.e., las desviaciones estándar muestrales de las variables de estado. Estas desviaciones se utilizan en las densidades normales que conforman el nucleo K.
De igual manera simplemente delante del prompt se escribe sigma=[ ]; es importante meter estas variables de esta forma para que los programas corran.
Para el valor numérico h simplemente se escribe delante del prompt $h = N^{-\frac{1}{2k+1}}$, claro utilizando los correspondientes valores de N y k en cuestión (Para elevar un número a una potencia en MATLAB se utiliza el símbolo ^ y para dividir /).
Ahora bien, para los programas utilizados en la selección de las variables de estado recordemos que utilizamos los vectores z_t (minúsculas, no confundir con los Z_t), con los cuales conformamos una matriz z, cuyos renglones son un subconjunto de los renglones de A en diferentes combinaciones posibles (ver capítulo 3 donde se escogió 5 combinaciones diferentes cada una de 4 renglones de A). La matriz z se captura en MATLAB de forma análoga al resto de matrices.

Cada programa debe salvarse como un archivo diferente por cuestiones de optimización computacional. El orden en el que los programas corren es el siguiente, aunque solo hay que ejecutar algunos de ellos (mas adelante se especificará cuales).

1) Función K. Calcula todos los valores del kernel para los posibles valores de t y s. Arroja una matriz de $N \times N$.

2) Función K1. Da un valor del kernel específico.

3) Función fgorro. Calcula $\widehat{f}(X) = N^{-1}h^{-k}\sum_{s=1}^{N} K\left(\frac{X-X_s}{h}\right)$.

4) Función erret. Calcula r_t.

5) Función ggorro1. Calcula $\widehat{g}_p$, $\widehat{g}_{pp}$, $\widehat{g}_{pr}$ y $\widehat{g}_r$.

6) Función bmew. Calcula $\widehat{b}$, $\widehat{m}$, $\widehat{w}_t$ y $\widehat{e}$.

7) Función deltaNgorro. Calcula $\widehat{\delta}_N$.

8) Función gorroN. Calcula $\widehat{a}_N$, $\widehat{\eta}_N$ y $\widehat{\gamma}_N$.

9) Función Tdelta. Calcula $\widehat{\Gamma}_0$, $\widehat{A}_N$, $\widehat{\Omega}_N$ y el estadístico de prueba $\widehat{T}_\delta$.

10) Función mugorro. Calcula $\widehat{\mu}$.

11) Función ggorro3. Calcula $\widehat{g}_z$, $\widehat{g}_{zp}$ y $\widehat{\gamma}_\mu$.

12) Función pfinal. Calcula $\widehat{A}_\mu$, $\widehat{\Gamma}_\mu$, $\widehat{\Omega}_\mu$ y Q.

Las únicas funciones que deben ejecutarse para obtener la información necesaria de las pruebas propuestas en este trabajo son

i) Función K.

ii) Función deltaNgorro.

iii) Función Tdelta.

iv) Función mugorro.

v) Función pfinal.

Al grabar cada función en un archivo de MATLAB, este es guardado en la ventana del editor. Para ejecutar la función deseada, en la ventana de comando se escribe simplemente el nombre de la función delante del prompt (por ejemplo $\gg Kts = K(A, sigma, h)$) y se oprime la tecla enter.
Una vez que se tienen todas las posibles variables de estado, se construyen la matriz A,Z,. . .,z. Se ejecutan los cinco programas anteriores para diferentes opciones de z. Una vez que se haya escogido la mejor z que pronostique los X_t, se substituirá la matriz A con esta elección de z y en consecuencia también tendrá que ser substituida la matriz Z de acuerdo al procedimiento descrito arriba (aumentar un renglón de $1^{'}$s. . .). Una vez hecho esto se ejecutarán de nuevo *i*), *ii*) y *iii*) de la lista de arriba, con lo que se encontrará el estadístico de prueba $\widehat{T_\delta}$ definitivo, el cual permitirá decidir si el portafolio p es o no eficiente media-varianza condicional.
Los programas:

Función K

```
function Kts=K(A,sigma,h)
sigma=sigma(:);
[k,N]=size(A);
Kts=ones(N,N);
for t=1:N
   for s=1:N
      Kts1=1;
      for i=1:k
          c2=(1/sqrt(2*pi*sigma(i)))*exp(-((((A(i,t)-A(i,s))/h).^2)/(2*sigma(i).^2)));
          Kts1=Kts1*c2;
      end
      Kts(t,s)=Kts1;
   end
end
```

Función K1

```
function Kts1=K1(A,t,s,sigma,h)
sigma=sigma(:);
[k,N]=size(A);
Kts1=1;
for i=1:k
    c2=(1/sqrt(2*pi*sigma(i)))*exp(-((((A(i,t)-A(i,s))/h).^2)/(2*sigma(i).^2)));
    Kts1=Kts1*c2;
end
```

Función fgorro

```
function f=fgorro(A,t,sigma,h,Kts)
[k,N]=size(A);
sum=Kts(t,:)*ones(N,1);
f=N^-1*h^-k*sum;
```

Función erret

```
function rt=erret(R,Z,t)
q=size(Z,1);
n=size(R,1);
rt=zeros(q*n,1);
rt=kron(R(:,t),Z(:,t));
```

Función ggorro1

```
function [gp,gpp,gr,gpr]=ggorro1(A,t,sigma,h,p,R,Z,Kts)
p=p(:);
k=size(A,1);
q=size(Z,1);
[n,N]=size(R);
gpr=zeros(N,1);
gr=zeros(N,1);
Kts1=Kts(t,:)*p;
Kts2=Kts(t,:)*p.^2;
ft=fgorro(A,t,sigma,h,Kts);
gp=N^-1*h^-k*ft^-1*Kts1;
gpp=N^-1*h^-k*ft^-1*Kts2;
sum1=0;
sum2=0;
for s=1:N
   rs=erret(R,Z,s);
   Krt1=Kts(t,s)*rs;
   Krt2=Kts(t,s)*p(s)*rs;
   sum1=sum1+Krt1;
   sum2=sum2+Krt2;
end
ft=fgorro(A,t,sigma,h,Kts);
gr=N^-1*h^-k*ft^-1*sum1;
gpr=N^-1*h^-k*ft^-1*sum2;
```

Función bmew

```
function [wt,egorro]=bmew(A,t,sigma,h,p,R,Z,Kts)
n=size(R,1);
N=size(A,2);
egorro=zeros(n,1);
ft=fgorro(A,t,sigma,h,Kts);
[gp,gpp]=ggorro1(A,t,sigma,h,p,R,Z,Kts);
bgorrot=gp/gpp;
mgorrot=1-bgorrot*p(t);
for i=1:n
   egorro(i)=mgorrot*R(i,t);
end
wt=ft*gpp;
```

Función deltaNgorro

```
function dN=deltaNgorro(A,sigma,h,p,R,Z,Kts)
q=size(Z,1);
n=size(R,1);
[k,N]=size(A);
dN=[];
sum=0;
sum1=0;
for i=1:n
   for t=1:N
      [wt,egorro]=bmew(A,t,sigma,h,p,R,Z,Kts);
      res1=wt*Z(:,t)*Z(:,t)';
      res2=wt*Z(:,t)*egorro(i);
      sum=sum+res1;
      sum1=sum1+res2;
   end
   deltai=(inv((1/N)*sum)*((1/N)*sum1))';
   dN=[dN,deltai];
end
dN=dN';
```

Función gorroN

```
function gammaN=gorroN(A,t,sigma,h,p,R,Z,Kts,dN)
n=size(R,1);
q=size(Z,1);
In=eye(n);
ft=fgorro(A,t,sigma,h,Kts);
rt=erret(R,Z,t);
[gp,gpp,gr,gpr]=ggorro1(A,t,sigma,h,p,R,Z,Kts);
aN=ft*(gpp*(Z(:,t)*Z(:,t)')+p(t).^2*(Z(:,t)*Z(:,t)'));
etaN=ft*(gpp*rt-gp*p(t)*rt+gr*p(t).^2-gpr*p(t));
gammaN=etaN-kron(In,aN)*dN;
```

Función Tdelta

```
function Td=Tdelta(A,sigma,h,p,R,Z,Kts,dN)
[n,N]=size(R);
sum=0;
for t=1:N
   gammaN=gorroN(A,t,sigma,h,p,R,Z,Kts,dN);
   Krt=gammaN*gammaN';
   sum=sum+Krt;
end
sigma0=N^-1*sum;
sum1=0;
for t=1:N
   [wt,egorro]=bmew(A,t,sigma,h,p,R,Z,Kts);
   Krt=wt*(Z(:,t)*Z(:,t)');
   sum1=sum1+Krt;
end
AN=kron(eye(n),N^-1*sum1);
omN=inv(AN)*sigma0*inv(AN);
Td=N*dN'*inv(omN)*dN;
```

Función mugorro

```
function mu=mugorro(A,sigma,h,p,R,Z,z,Kts,dN,Td)
[n,N]=size(R);
sum2=0;
sum3=0;
for t=1:N
   ft=fgorro(A,t,sigma,h,Kts);
   gp=ggorro1(A,t,sigma,h,p,R,Z,Kts);
   Krt=ft*(z(:,t)*z(:,t)');
   Krt1=ft*z(:,t)*(p(t)-gp);
   sum2=sum2+Krt;
   sum3=sum3+Krt1;
end
mu=inv(N^-1*sum2)*(N^-1*sum3);
```

Función ggorro3

```
function gammamu=ggorro3(A,t,sigma,h,p,R,Z,z,Kts,dN,Td,mu)
k=size(A,1);
[n,N]=size(R);
sum=0;
sum1=0;
for s=1:N
   Krt=K1(A,t,s,sigma,h)*z(:,s);
   Krt1=K1(A,t,s,sigma,h)*z(:,s)*p(s);
   sum=sum+Krt;
   sum1=sum1+Krt1;
end
ft=fgorro(A,t,sigma,h,Kts);
gz=N^-1*h^-k*ft^-1*sum;
gzp=N^-1*h^-k*ft^-1*sum1;
[gp,gpp,gr,gpr]=ggorro1(A,t,sigma,h,p,R,Z,Kts);
gammamu=ft*(z(:,t)*p(t)-z(:,t)*gp-p(t)*gz+gzp-(z(:,t)*z(:,t)')*mu-(z(:,t)*z(:,t)')*mu);
```

Función pfinal

```
function Q=pfinal(A,sigma,h,p,R,Z,z,Kts,dN,Td,mu)
[k,N]=size(A);
sum=0;
sum1=0;
for t=1:N
   ft=fgorro(A,t,sigma,h,Kts);
   gammamu=ggorro3(A,t,sigma,h,p,R,Z,z,Kts,dN,Td,mu);
   Krt=ft*(z(:,t)*z(:,t)');
   Krt1=gammamu*gammamu';
   sum=sum+Krt;
   sum1=sum1+Krt1;
end
Amu=N^-1*sum;
Sigmamu=N^-1*sum1;
Omegamu=inv(Amu)*Sigmamu*inv(Amu);
Q=N*mu'*inv(Omegamu)*mu;
```

BIBLIOGRAFÍA

[1] Arcones, Miguel A. 1995. *On the Central Limit Theorem for U-statistics Under Absolute Regularity.* Statistics and Probability Letters 24. 245-249.

[2] Banz, Rolf W. 1981. *The relationship between return and market value of common stocks.* Journal of Financial Economics 9. 3-18.

[3] Basu,Sanjoy. 1983. *The relationship between earnings yield, market value, and return for NYSE common stocks: Further evidence.* Journal of Financial Economics 12. 51-74.

[4] Bhandari, Laxmi Chand. 1988. *Debt/equity ratio and expected common stock returns: Empirical evidence.* Journal of Finance 43. 507-528.

[5] Bierens, Herman J. 1987. *Kernel estimators of regression functions.* Chapter 3 in T.F. Bewley ed. Advances in Econometrics. 99-144.

[6] Breen, William J., Larry, R. Glosten, and Ravi Jagannathan. 1989. *Economic significance of predictable variations in stock index returns.* Journal of Finance 44. 1177-1190.

[7] Chan, K. C., Nai-fu Chen, and David A. Hsieh. 1985. *An exploratory investigation of the firm size effect.* Journal of Financial Economics 14. 451-471.

[8] Chen, Nai-fu. 1991. *Financial investment opportunities and the macroeconomy.* Journal of Finance 46. 529-554.

[9] Doukhan, P., P. Massart, and E. Rio. 1994. *The functional central limit theorem for strongly mixing processes.* Ann. Inst. H. Poincaré Probab. Statist. 30. 63-82.

[10] Dybvig, P.H., and Stephen A. Ross. 1985. *Differential information and performance measurement using a security market line.* Journal of Finance 40. 383-400.

[11] Fama, Eugene F., and Kenneth R. French. 1989. *Business conditions and the expected returns on bonds and stocks.* Journal of Financial Economics 25. 23-50.

[12] Fama, Eugene F., and Kenneth R. French. 1992. *The cross-section of expected stock returns.* Journal of Finance 47. 427-466.

[13] Fan, J., and Gijbels, I. 1996. *Local polynomial modelling and its applications.* Chapman and Hall.

[14] Ferson, Wayne E., and Campbell R. Harvey. 1991. *The variation of economic risk premiums.* Journal of Political Economy 99. 385-415.

[15] Gibbons, Jean D., and Chakraborti, Subhabrata. 1992. *Non-parametric Statistical Inference.* Dekker.

[16] Gibbons, Michael R. 1982. *Multivariate tests of financial models: A new approach.* Journal of Financial Economics 10. 3-27.

[17] Hansen, Lars Peter, and Scott F. Richard. 1987. *The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models.* Econometrica 55. 587-613.

[18] Jagannathan, Ravi, and Zhenyu Wang. 1996. *The conditional CAPM and the cross-section of expected returns.* Journal of Finance 51. 3-53.

[19] Keim, Donald B., and Robert F. Stambaugh. 1986. *Predicting returns in the stock and bond markets.* Journal of Financial Economics 17. 357-390.

[20] Long, J. B. 1974. *Stock prices, inflation, and the term structure of interest rates.* Journal of Financial Economics 1. 131-170.

[21] Merton, Robert C. 1973. *An intertemporal capital asset pricing model.* Econometrica 41. 867-887.

[22] Merton, Robert C. 1980. *On estimating the expected return on the market: An exploratory investigation.* Journal of Financial Economics 8. 323-361.

[23] Reinganum, Mark R. 1981. *Misspecification of capital asset pricing: empirical anomalies based on earnings yields and market values.* Journal of Financial Economics 9. 19-46.

[24] Shanken, Jay. 1985. *Multivariate tests of the zero-beta CAPM.* Journal of Financial Economics 14. 327-348.

[25] Wang, Qing. 1997. *Tests of conditional asset pricing models.*