

A NON-PARAMETRIC TEST OF THE CONDITIONAL CAPM FOR THE MEXICAN ECONOMY

Jorge H. del Castillo-Spíndola

*Universidad Nacional Autónoma de México**

Resumen: Se han sugerido muchos modelos para describir cómo los inversionistas valúan flujos de efectivo riesgosos. El más usado es el Modelo de Valuación de Activos de Capital (CAPM por sus siglas en inglés) de Sharpe-Lintner-Black. Sin embargo, se han observado muchas anomalías y fuerte evidencia en su contra. Suponer que el CAPM se cumple en una versión condicional es suponer que las betas y el premio de riesgo de mercado varían a lo largo del tiempo. Presentamos una prueba de la versión condicional del CAPM para el caso de México, basada en una metodología no-paramétrica sugerida por Wang que evita el problema asociado a la especificación de la forma funcional de las betas de los activos.

Abstract: Many models have been suggested to describe how investors manage risk and value risky cash flows. Among them, the most widely used is the Sharpe-Lintner-Black Capital Asset Pricing Model (CAPM). However many anomalies and evidence against this version have been presented. To assume that the CAPM holds in a conditional sense is to assume that the betas and the market risk premium vary along time. We present a test of the conditional version of the CAPM for the Mexican economy, that uses a non-parametric methodology suggested by Wang that avoids the problem of functional form misspecification of the betas of the assets.

Clasificación JEL: C12, C13, C14, G12, G13, G14

Palabras clave: CAPM, conditional mean-variance efficiency, betas of the assets, market risk premium, non-parametric testing, eficiencia media-varianza condicional, betas de los activos, premio de riesgo de mercado, prueba no-paramétrica.

Fecha de recepción: 16 XII 2005

Fecha de aceptación: 27 VI 2006

* Useful comments and suggestions were made by Eneas Caldiño, Onésimo Hernández, Pablo Padilla, Karina Caballero and two anonymous referees. I also thank to the seminar in stochastic control and games participants at Cinvestav. Useful help with the data in an earlier draft was provided by Karla Salas. Any errors remain my responsibility. hdelcastillo@gmail.com

1. Introduction

The Capital Asset Pricing Model (CAPM) appeared for the first time in the early 60's as a natural and immediate extension of the theory developed by Markowitz. The first works were made by Sharpe (1964) and Lintner (1965). Further developments were made by Mossin (1966) and Black (1972). Basically, the CAPM determines asset prices by considering that optimal investment decisions have been taken while the market is in equilibrium. From this, risky assets only pay for the non-diversifiable risk.

The CAPM is built by analyzing the investors' behavior in a hypothetical economy of a single period. Thus, assumptions are necessary for empirical testing. The most common assumption is that the betas of the assets remain constant along time. This assumption does not seem to be very reasonable since the risk associated with a company's cash flow may vary during the business cycle.

Much work has been devoted to the study of the CAPM; among them, the work of Fama and French (1992) empirically analyzes the Black version. They find that the estimated value for the coefficient of beta is almost zero. This result is presented as strong evidence against the CAPM. This finding motivated the consideration of the available information at each moment, which is the main feature of the conditional version of the CAPM. Put differently, in the conditional CAPM the expected return of an asset based on the available information in a given period of time is linear in its conditional beta.

The evidence against the CAPM is not necessarily against the conditional version. It may well happen that even when the expected returns are linear in the betas for each period of time based on the available information at that time, the relation between the non-conditional expected returns and the non-conditional beta is flat.¹ It is, of course, not sufficient to say that the evidence against the CAPM is not evidence against the conditional version in order to work with this. Stronger, theoretical arguments are needed. The work of Hansen and Richard (1987) shows that the conditional CAPM can hold even when the CAPM presents anomalies.

The purpose of this work is to test the conditional CAPM for the Mexican economy. To achieve this, we still face another problem: According to the conditional CAPM, the conditional expected excess returns on the assets in a given time, should be linear in their conditional covariances with the market, but the model says nothing

¹ Wang (2002) provides an illustrative example.

about how these conditional expected returns and covariances vary over time. How then can we test empirically? Usually, functional forms on these conditional moments are assumed to obtain time series. However, it has turned out that different specifications can easily produce different empirical results. Moreover, one can easily change the results of a test by simply altering the empirical specification.

What is then the criteria for deciding know whether an empirical result is reliable evidence or just the effect of a certain specification? Unfortunately this is not the only problem: By assuming a certain functional form we are not only testing the pricing prediction but also the auxiliary model (joint hypothesis problem). Thus, it may happen that even when the pricing model is correct, the test could produce a rejection and large pricing error estimates because of a poor functional form. There are methodologies which avoid the specification effects.

Here, we choose to follow a methodology proposed by Wang (1998, 2002, 2003) and based on the following idea. First we have a linear regression model for the errors in the conditional expected returns, which is always consistent with the conditional mean-variance efficiency of the market portfolio. Then, through a non-parametric discount factor we get a weighted least squares estimator for the regression coefficients. We test whether the coefficients of the regression are zero by using the estimator and its asymptotical properties to obtain a test statistic. By considering a non-parametric discount factor the test avoids the joint hypothesis problem and the test is actually free of the impact of specification errors.

2. The CAPM and the Conditional CAPM

2.1. CAPM (*Sharpe-Lintner-Black*)

Let us denote R_i the return in the asset i and R_p the return in the market portfolio of all assets in the economy. Black's version of the CAPM (1972) is:

$$E[R_i] = \gamma_0 + \gamma_1 \beta_i \quad (1)$$

where

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

Fama and French (1992) found that the estimated value of γ_1 for the US economy is almost zero. They interpret this flat relation

between the average return and beta as evidence against the CAPM. However, as mentioned earlier, this is not necessarily evidence against the conditional version.

Even when the expected returns are linear in the betas for each period of time based on the available information at the time, the relation between the non-conditional expected returns and the non-conditional beta could be flat.²

2.2. Conditional CAPM

Let us denote $R_{i,t+1}$ the gross return (i.e., 1 plus the rate of return) of asset i on period $t + 1$. Analogously, $R_{p,t+1}$ will be the gross return of the portfolio of aggregated wealth of all assets in the economy in period $t + 1$. $R_{p,t+1}$ will be known as the market return. We also denote I_t the investors common information set at the end of period t .

We assume all used time series are covariance-stationary as well as the existence of all conditional and non-conditional moments.

Rational, risk-averse investors in a dynamical economy will certainly anticipate and hedge under the possibility of future adverse investment conditions. Thus, the conditional expected return on an asset will typically be jointly linear in the market beta and in the betas of the hedging portfolio.³ However, following Merton's development (1980), we will assume that hedging considerations are not sufficiently important and so the CAPM will hold conditionally as follows:

For each asset i and period $t + 1$,

$$E[R_{i,t+1}|I_t] = \gamma_{0,t} + \gamma_{1,t}\beta_{i,t} \quad (2)$$

where $\beta_{i,t}$ is the conditional beta of asset i defined as:

$$\beta_{i,t} = \frac{Cov(R_{i,t+1}, R_{p,t+1}|I_t)}{Var(R_{p,t+1}|I_t)}. \quad (3)$$

We have that $\gamma_{0,t}$ is the conditional expected return in a zero-beta portfolio and $\gamma_{1,t}$ is the conditional market risk premium.

² This is because an asset located in the conditional mean-variance frontier is not necessarily located in the non-conditional frontier. See Dybvig and Ross(1985), and Hansen and Richard(1987).

³ See Merton (1973) and Long (1974).

Now we take the unconditional expectation in (2):

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

to get

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

by defining $\gamma_0 = E[\gamma_{0,t}]$, $\gamma_1 = E[\gamma_{1,t}]$ and $\bar{\beta}_i = E[\beta_{i,t}]$ we finally obtain

$$E[R_{i,t+1}] = \gamma_0 + \gamma_1\bar{\beta}_i + Cov(\gamma_{1,t}, \beta_{i,t}) \quad (4)$$

Here γ_1 is the expected market risk premium and $\bar{\beta}_i$ is the expected beta.

Observe that if the covariance in (4) is zero or a linear function of the expected beta for each arbitrary asset i , then this actually resembles the static CAPM; i.e., the expected return is a linear function of the expected beta. However, in general, the conditional risk premium in the market and the conditional betas are correlated. Some authors have shown that the conditional risk premium in the market, as well as the conditional betas, are not constant⁴ and vary along the business cycle.⁵ Thus, the last term in (4) is in general not zero and so the unconditional expected return is not only a linear function of the expected beta.

3. Theoretical Framework and Empirical Instrumentation

The path we follow to analyze the Conditional CAPM is through considering a discount factor.

3.1. *Why a Non-parametric Discount Factor?*

As mentioned earlier, a wrong specification of this factor may lead to a joint hypothesis problem. Moreover, the test power, and the estimation bias can also be affected.⁶ To see this, we present an example given in Wang (1998). To test the hypothesis that a given

⁴ Keim and Stambaugh(1986), Breen, Glosten, and Jagannathan(1989).

⁵ Fama and French(1989), Chen(1991) and Ferson and Harvey(1991).

⁶ See Wang (1998).

benchmark portfolio p is conditionally mean-variance efficient, Wang specifies a discount factor linear in $r_{p,t+1}$:

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

where

$r_{p,t+1}$ is the return of p minus the risk-free rate,

X_t is a vector of state variables,

θ is a vector of parameters.

Then, the procedure of the Generalized Method of Moments is applied to test if the condition

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

holds for some value θ_0 of the vector of parameters, where r_{t+1} is the vector of returns minus the risk-free rate.

How to know whether a discount factor is correctly specified? The conditional efficiency hypothesis predicts a discount factor linear in $r_{p,t+1}$ but says nothing of the functional forms of a and b . However, the conditional mean-variance efficiency hypothesis does not leave a and b unrestricted. It is immediate by clear that the conditional beta-price equation:

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

has a representation of the discount factor

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

where

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

From here, the correct specifications of a and b (a correctly specified discount factor) must satisfy:

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

for some θ_0 .

Wrong specifications of a and b give origin to a joint hypothesis problem. The procedure of the GMM discount factor simultaneously

tests the conditional mean-variance efficiency of the benchmark portfolio and the functional form assumptions of a and b . Specification errors can easily lead to violations of the condition $E[m_{t+1}(\theta_0)r_{t+1}] = 0$ even when the benchmark portfolio is in the conditional efficiency frontier. These problems motivate the use of a non-parametric discount factor, allowing the data to self-adjust instead of assuming a priori a functional form.

3.2. A Weighted Least Squares Estimator

Let us consider a framework where there exists a conditionally riskless asset. As before:

- $r_{p,t+1}$ is the return of portfolio p minus the risk-free rate,
- $r_{i,t+1}$ is the return of asset i minus the risk-free rate. $i = 1, \dots, n$.

Let X_t be a k -dimensional vector of state variables such that

$$E[r_{p,t+1}|I_t] = E[r_{p,t+1}|X_t] \tag{5}$$

$$E[r_{p,t+1}^2|I_t] = E[r_{p,t+1}^2|X_t] \tag{6}$$

where I_t is the investors information set⁷ at time t .

We assume that $r_{p,t+1}$, $r_{i,t+1}$ and the state variables are strictly stationary. If the benchmark portfolio p is conditionally mean-variance efficient, we then have

$$E[r_{i,t+1}|I_t] = E[r_{p,t+1}|I_t] \frac{Cov(r_{i,t+1}, r_{p,t+1}|I_t)}{Var(r_{p,t+1}|I_t)} \tag{7}$$

or equivalently

$$E[r_{i,t+1}|I_t] = E[r_{p,t+1}|I_t] \frac{E[r_{i,t+1}r_{p,t+1}|I_t]}{E[r_{p,t+1}^2|I_t]} \tag{8}$$

for $i = 1, \dots, n$.

The covariance representation (7) is the usual beta-price equation. Representation (8) is that of cross moment. Set

$$g_p(X_t) = E[r_{p,t+1}|X_t]$$

⁷ Notice that these equations are just for portfolio p . It is not required that X_t fully characterizes I_t , just that it be sufficient to develop the non-parametric test.

$$g_{pp}(X_t) = E[r_{p,t+1}^2|X_t]$$

$$b(X_t) = \frac{g_p(X_t)}{g_{pp}(X_t)}$$

If (5) and (6) hold, the errors of the conditional expected returns of (8) can be expressed as

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

where

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

Then, (8) is equivalent to

$$E[m_{t+1}r_{i,t+1}|I_t] = 0. \tag{9}$$

Now, set $e_{i,t+1} = m_{t+1}r_{i,t+1}$ and let Z_t be a q -dimensional vector of observed stationary variables in I_t .

If the discount factor m_{t+1} were observable, a natural way to test the condition $E[e_{i,t+1}|I_t] = 0$ would be to run a regression of $e_{i,t+1}$ in Z_t and test if the coefficients are zero. This is because the following regressions

$$e_{i,t+1} = Z_t' \delta_i + u_{i,t+1} \tag{10}$$

where $E[u_{i,t+1}|I_t] = 0$ for $i = 1, \dots, n$, are always consistent with (9).

Clearly (9) implies that (10) holds with $\delta = 0$ where $\delta = (\delta'_1, \delta'_2 \dots \delta'_n)'$.

Following the development by Wang (1998, 2002, 2003) to implement this idea, m_{t+1} will be substituted with a non-parametric discount factor \widehat{m}_{t+1} and the vector of parameters δ will be estimated with

$$\widehat{\delta}_i = \left(\frac{1}{N} \sum_{t=1}^N \widehat{w}_t Z_t Z_t' \right)^{-1} \left(\frac{1}{N} \sum_{t=1}^N \widehat{w}_t Z_t \widehat{e}_{i,t+1} \right) \tag{11}$$

for $i = 1, \dots, n$ where

$$\widehat{e}_{i,t+1} = \widehat{m}_{t+1} r_{i,t+1} \quad \text{and}$$

$$\widehat{m}_{t+1} = 1 - \widehat{b}(X_t)r_{p,t+1} \quad \text{with}$$

$$\widehat{b}(X) = \frac{\widehat{g}_p(X)}{\widehat{g}_{pp}(X)}.$$

The weighting function is chosen as

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

Here \widehat{f} , \widehat{g}_p and \widehat{g}_{pp} are kernels defined as

$$\begin{aligned} \widehat{f}(X) &= N^{-1}h^{-k} \sum_{s=1}^N K\left(\frac{X - X_s}{h}\right) \\ \widehat{g}_p(X) &= N^{-1}h^{-k}\widehat{f}(X)^{-1} \sum_{s=1}^N K\left(\frac{X - X_s}{h}\right)r_{p,s+1} \\ \widehat{g}_{pp}(X) &= N^{-1}h^{-k}\widehat{f}(X)^{-1} \sum_{s=1}^N K\left(\frac{X - X_s}{h}\right)r_{p,s+1}^2 \end{aligned} \quad (12)$$

where \widehat{f} is the Rosenblatt-Parzen non-parametric density estimator with kernel $K(\cdot)$ and bandwidth h , and \widehat{g}_p and \widehat{g}_{pp} are Nadaraya-Watson regression function estimators.

The weighting function \widehat{w}_t was chosen so that $\widehat{w}_t Z_t Z_t'$ and $\widehat{w}_t Z_t \widehat{e}_{i,t+1}$ could be expressed as generalized U-statistics of second order, allowing us to analyze the properties of $\widehat{\delta}_i$ for large samples.

The proposed test is based in the weighted least squares estimator $\widehat{\delta}_N$:

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

Intuitively, $\widehat{\delta}_N$ converges to zero if the benchmark portfolio p is conditionally mean-variance efficient. Otherwise, the estimator converges to a non-zero limit.⁸ Then, a conditional efficiency test can be constructed by checking how far $\widehat{\delta}_N$ is from zero using asymptotic distribution theory to consider the sampling errors.

⁸ unless $e_{i,t+1}$ is orthogonal to all components of Z_t for $i = 1, \dots, n$.

We can also see (10) as a model for the pricing errors; i.e., $Z_t' \delta_i$ would serve as a proxy for $E[e_{i,t+1}|I_t]$, the conditional expected return errors in (8). This is why no restriction has been placed in choosing Z_t .

Notice that the regression model (10) does not give origin to the joint hypothesis problem, regardless of the chosen Z_t . This is due to the fact that the model is correct as long as the benchmark portfolio is conditionally mean-variance efficient.

3.3. The Estimator's Asymptotic Behavior and the Test Statistic

Let r_{t+1} be a vector of excess returns (return of the asset minus the risk-free asset)

$$r_{t+1} = (r_{1,t+1} \cdots r_{n,t+1})' \otimes Z_t$$

and let

$$Y_{t+1} = (X_t' Z_t' r_{p,t+1} r_{t+1})'$$

where \otimes is Kronecker's operator.

Let us denote

$$w_t = f(X_t) g_{pp}(X_t)$$

$$A = \mathbb{I}_n \otimes E[w_t Z_t Z_t']$$

$$\widehat{A}_N = \mathbb{I}_n \otimes N^{-1} \sum_{t=1}^N \widehat{w}_t Z_t Z_t'$$

where \mathbb{I}_n is the identity matrix of size n .

Set $\delta = (\delta_1' \cdots \delta_n')'$ with

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

Define

$$\gamma(Y_{t+1}) = \eta(Y_{t+1}) - [\mathbb{I}_n \otimes a(Y_{t+1})] \delta \tag{13}$$

$$\eta(Y_{t+1}) = f(X_t) [g_{pp}(X_t) r_{t+1} - g_p(X_t) r_{p,t+1} r_{t+1}] \tag{14}$$

$$\begin{aligned}
 &+g_r(X_t)r_{p,t+1}^2 - g_{pr}(X_t)r_{p,t+1}] \\
 a(Y_{t+1}) &= f(X_t)[g_{pp}(X_t)Z_tZ_t' + r_{p,t+1}^2g_{zz}(X_t)] \tag{15}
 \end{aligned}$$

where

$$g_r(X_t) = E[r_{t+1}|X_t]$$

$$g_{pr}(X_t) = E[r_{p,t+1}r_{t+1}|X_t] \text{ and}$$

$$g_{zz}(X_t) = E[Z_tZ_t'|X_t].$$

We have a first result for the limit behavior:

THEOREM 3.1. *We have that*

- i) \widehat{A}_N converges in probability to A ,
- ii) The limit distribution of $\sqrt{N}\widehat{A}_N(\widehat{\delta}_N - \delta)$ is identical to that of $N^{-\frac{1}{2}}\sum_{t=1}^N\gamma(Y_{t+1})$,
- iii) $E[\gamma(Y_{t+1})] = 0$.

The next step is to get the limiting distribution of the estimator, then to construct the test statistic and finally to obtain its limiting distribution. For these two results some technical assumptions are needed. These assumptions are presented in the appendix.

We now observe that since $N^{-1}\sum_{t=1}^N\gamma(Y_{t+1})$ is a simple average of stationary random vectors, one application of a central limit theorem⁹ gives the next result.

THEOREM 3.2. *Given assumptions 1-6 in the appendix, and if $h \rightarrow 0$, $Nh^{2k} \rightarrow \infty$ and $Nh^{2k+2} \rightarrow 0$, then the weighted least squares estimator $\widehat{\delta}_N$ is such that $\sqrt{N}(\widehat{\delta}_N - \delta)$ has a multivariate Gaussian limit distribution with mean 0 and covariance matrix Ω , where $\Omega = A^{-1}\Gamma A^{-1}$, $\Gamma = \sum_{-\infty}^{\infty}\Gamma_j$ and $\Gamma_j = E[\gamma(Y_{t+1})\gamma(Y_{t+j+1})']$.*

This theorem shows that $\widehat{\delta}_N$ has the standard limiting properties, namely \sqrt{N} -consistency and asymptotic normality of parametric estimators.

⁹ Doukhan, Massart and Rio (1994) have improved the classical central limit theorems of Ibragimov and Linnik (1971) and proved that if $2 < \rho < \infty$, then $E[|X|^\rho] < \infty$ and $\sum_{n=1}^{\infty}\alpha_n n^{\frac{2}{(\rho-2)}} < \infty$. It then follows that $n^{-\frac{1}{2}}\sum_{i=1}^n(X_i - E[X_i])$ converges to a centered Gaussian random vector.

This result is based neither on (5) and (6), nor requires that equations (10) be correctly specified. We have to notice that the conditions for h differ from those used for pointwise kernel estimators. The conditions $Nh^{2k} \rightarrow \infty$ and $Nh^{2k+2} \rightarrow 0$ give the upper and lower bounds of the rate at which h converges to 0 for $\widehat{\delta}_N$ to exhibit the wanted asymptotic behavior. The condition $Nh^{2k+2} \rightarrow 0$ is due to the use of a kernel of order $k + 1$ and consequently the admissible range for the rate can be relaxed by using a kernel of order higher than $k + 1$.

Now, to construct a conditional efficiency test using the distribution of $\widehat{\delta}_N$ we need an estimator for the covariance matrix Ω . For this, we first estimate $\gamma(Y_{t+1})$. We substitute $f(X)$, $g_p(X)$, $g_{pp}(X)$, $g_r(X)$, $g_{pr}(X)$ and $g_{zz}(X)$ in (14) and (15) by standard kernels and δ in (13) by $\widehat{\delta}_N$. A natural approximation for $\gamma(Y_{t+1})$ would be

$$\widehat{\gamma}_N(Y_{t+1}) = \widehat{\eta}_N(Y_{t+1}) - [\mathbf{II}_n \otimes \widehat{a}_N(Y_{t+1})]\widehat{\delta}_N \tag{16}$$

$$\begin{aligned} \widehat{\eta}_N(Y_{t+1}) = & \widehat{f}(X_t)[\widehat{g}_{pp}(X_t)r_{t+1} - \widehat{g}_p(X_t)r_{p,t+1}r_{t+1} \\ & + \widehat{g}_r(X_t)r_{p,t+1}^2 - \widehat{g}_{pr}(X_t)r_{p,t+1}] \end{aligned} \tag{17}$$

$$\widehat{a}_N(Y_{t+1}) = \widehat{f}(X_t)[\widehat{g}_{pp}(X_t)Z_tZ_t' + r_{p,t+1}^2\widehat{g}_{zz}(X_t)] \tag{18}$$

with \widehat{f} , \widehat{g}_p and \widehat{g}_{pp} as previously defined¹⁰ and

$$\widehat{g}_r(X) = N^{-1}h^{-k}\widehat{f}(X)^{-1} \sum_{s=1}^N K\left(\frac{X - X_s}{h}\right)r_{s+1}$$

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

$$\widehat{g}_{zz}(X) = N^{-1}h^{-k}\widehat{f}(X)^{-1} \sum_{s=1}^N K\left(\frac{X - X_s}{h}\right)Z_sZ_s'$$

¹⁰ Note that $g_{zz}(X_t)=Z_tZ_t'$ when Z_t is a fixed transformation of X_t . For example $Z_t=(1X_t)'$. In such a case, it is not necessary to use the estimation kernel $\widehat{g}_{zz}(X_t)$ en (18): we simply substitute $g_{zz}(X_t)$ by Z_tZ_t' which gives $\widehat{a}_N(Y_{t+1})=\widehat{f}(X_t)[\widehat{g}_{pp}(X_t)+r_{p,t+1}^2]Z_tZ_t'$.

It can be shown that

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

is a consistent estimator of Γ_j . Also, given (5) and (6) we have that $\Gamma_j = 0$ for all $j \neq 0$ when the equations (10) hold. In consequence, to test the conditional mean-variance efficiency, the proposed test statistic is

$$\widehat{T}_\delta = N \widetilde{\delta}_N' \widehat{\Omega}_N^{-1} \widetilde{\delta}_N \tag{19}$$

where $\widehat{\Omega}_N = \widehat{A}_N^{-1} \widehat{\Gamma}_0 \widehat{A}_N^{-1}$.

THEOREM 3.3. *If the conditions of theorem 3.2 hold, then*

- i) Given (5) and (6), if the portfolio p is conditionally mean-variance efficient, then the test statistic \widehat{T}_δ has a limit distribution χ^2 with $q \times n$ degrees of freedom,*
- ii) $\widehat{\Gamma}_j$ is a consistent estimator of Γ_j for each fixed j .*

It is worth noting that the test will have no power if a Z_t is chosen orthogonal¹¹ to $e_{i,t+1}$. However, the test will have power if a component of Z_t is able to significantly forecast $e_{i,t+1}$, regardless whether the regression model (10) is correctly specified or not.

3.4. The Kernel K and the Bandwidth h

One simple selection of the kernel K that fulfills the required characteristics is an independent multivariate Gaussian density function

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

where ϕ_i is the one-dimensional Gaussian density with mean 0 and variance σ_i^2 (σ_i is the standard deviation of the i -th state variable). For the empirical study, σ_i was substituted by the sample standard deviation.

¹¹ i.e., $E[Z_t e_{i,t+1} | X_t] = 0 \quad i=1, \dots, n$.

There is no theoretical consensus for choosing h . A practical choice, taking into account the convergence rate conditions of theorem 3.2., is

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

3.5. State Variables Selection

We do not want to ignore important forecasting variables in the conditioning set I_t . This is why the selection of the vector X_t is very important.

In one extreme, a test will not be conditionally efficient if none of the state variables are used. In the other, the inclusion of redundant variables may have a significant impact on the test power.

Our procedure will be to test (5) and (6) for a given X_t . For this, a non-parametric test is proposed.

Let z_t be a q_1 -dimensional vector in the set I_t , whose components are different from those of X_t . The idea for being able to make the selection is to verify if z_t can forecast the residuals $r_{p,t+1} - g_p(X_t)$ and $r_{p,t+1}^2 - g_{pp}(X_t)$.

If (5) holds (i.e., $E[r_{p,t+1}|I_t] = g_p(X_t)$), then

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

with $E[\epsilon_{t+1}|I_t] = 0$ and $\mu = 0$.

Under certain regularity conditions, if $Nh^{2k} \rightarrow \infty$ and $Nh^{2k+2} \rightarrow 0$, then the estimator

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

of μ is such that

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

As in theorem 3.3., an estimator of the covariance matrix can be constructed. Namely,

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

where

$$\widehat{A}_\mu = N^{-1} \sum_{t=1}^N \widehat{f}(X_t) z_t z_t'$$

$$\widehat{\Gamma}_\mu = N^{-1} \sum_{t=1}^N \widehat{\gamma}_{\mu,t+1} \widehat{\gamma}'_{\mu,t+1}$$

with

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

We have $\widehat{g}_{zz}(X)$ as previously defined and

$$\widehat{g}_z(X) = N^{-1} h^{-k} \widehat{f}(X)^{-1} \sum_{s=1}^N K\left(\frac{X - X_s}{h}\right) z_s$$

$$\widehat{g}_{zp}(X) = N^{-1} h^{-k} \widehat{f}(X)^{-1} \sum_{s=1}^N K\left(\frac{X - X_s}{h}\right) z_s r_{p,s+1}$$

This gives a test statistic

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

with a limit distribution under (5) of $\chi^2(q_1)$.

The corresponding test for (6) is constructed in a similar way, only substituting $r_{p,t+1}$ and $r_{p,s+1}$ with $r_{p,t+1}^2$ and $r_{p,s+1}^2$ respectively.

4. Data and Results

4.1. Data Selection

We start by describing the data representing the excess returns $r_{p,t+1}$ of portfolio p and those representing the excess returns $r_{i,t+1}$ $i = 1, \dots, n$ of the assets or portfolios used to forecast.

For the risk-free rate we consider the monthly rate of *Cetes* 28-days, denoted by λ . For the portfolio p we chose the percentage growth of the *índice de precios y cotizaciones* (IPC) minus the risk-free rate as follows

dad industrial. Considered as a measure of the growth in the economy and denoted by a ,

2) Percentage growth of the *índice nacional de precios al consumidor* (INPC). Considered as a measure of inflation and denoted by b ,

3) Percentage growth of the *agregado monetario*, M4. Taken to capture the wealth in the economy (considering liquidity) and denoted by c ,

4) Percentage growth in the peso-dollar exchange rate. Denoted by d ,

5) Spread between the monthly rate of commercial paper and λ . Considered as the default premium and denoted by e ,

6) Percentage growth of the Dow Jones index. Denoted by f .

The sample standard deviations σ_i used in the kernel K are the following

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

We are considering monthly data. The returns go from March 1987 to January 2002. The state variables are lagged one period¹² (February 1987 to December 2001). We then see that $N = 179$.

Following the procedure to select the state variables described in section 3, different vectors z_t of size 4 (i.e., $q_1 = 4$) were constructed. For each of these the value of the test statistic was computed.

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

with limit distribution under (5) $\chi^2(q_1)$.

We constructed 5 vectors of z , where, for example, z_{1236} indicates that variables 1, 2, 3 and 6 were used (i.e., *índice del volumen físico de la actividad industrial*, INPC, M4 and Dow Jones). The following values for the test statistic Q were obtained:

¹² See section 3 for the theoretical and empirical construction of the tests.

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}	$Q_{1256} = 6.9765$

All the above selections are accepted at 5 and 10 percent of confidence. Even though X_t should be chosen as

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

since $Q_{1245} = 1.4048$ is the value accepted at the highest confidence, for the sake of completeness and comparison effects we will perform the conditional mean-variance efficiency test of p for the 5 possibilities of X_t since these 5 tests of z were accepted.

4.2. *The Efficiency Test for the Portfolio p*

For all these tests, Z_t will be taken¹³ as $Z_t = (1 \quad X_t')'$. We next present the value of the test statistic

$$\widehat{T}_\delta = N \widehat{\delta}'_N \widehat{\Omega}_N^{-1} \widehat{\delta}_N \tag{20}$$

whose distribution is $\chi^2_{(q \times n)}$, for each case:

X	T
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}	$T_{123456} = 11.6842$

Where for the first 5 test statistics a $\chi^2_{(5 \times 5)}$ was used, and for the last test statistic a $\chi^2_{(7 \times 5)}$ was used.

¹³ not to be confused with the lowercase z used just before.

The quantiles which left a probability of .05 and .10 to the right of a $\chi^2_{(5 \times 5)}$ are 37.65 and 34.38 respectively. Thus, we conclude that p is conditionally mean-variance efficient for all cases but X_{1234} at 10 per cent confidence. There is a important point here to notice. Even though X_{1234} was one of the best choices for the state variables since Q_{1234} was one of the smaller values, with this selection the efficiency hypothesis is rejected at 10 per cent confidence.

We also find that by using all proposed state variables, (i.e., for the case X_{123456} using a $\chi^2_{(7 \times 5)}$) the efficiency hypothesis is largely accepted.

This information allows us to conclude that under the chosen variables and their corresponding data, there is evidence supporting the validity of the conditional CAPM for the Mexican economy.

5. Conclusions

Testing the CAPM, or equivalently, the mean-variance efficiency of the market portfolio has been a crucial point in financial economics. The risk associated with an asset is measured by the beta of the cash flow with respect to the market portfolio of all assets in the economy. Thus, its importance lies in helping to understand how investors manage the cash flow risk of an asset and how they determine which risk premium to demand. To achieve this goal, the importance of considering the available information at each time has been argued. This information is represented by the time series of some state variables chosen in the economy. In order to avoid the problems associated with the functional forms specification of certain conditional moments, a non-parametric methodology has been suggested. This methodology not only avoids the problems of misspecification, but also has the advantage of keeping the same rate of convergence as the parametric ones.¹⁴

We have presented results supporting the validity of the conditional CAPM for Mexican data. By no means is it implied that this evidence is conclusive. If we accept that the state variables chosen are a fairly good representative of the available economical information for the investors' decisions, and if we accept that the IPC sectoral data

¹⁴ Some non-parametric methodologies present lower rates of convergence than those of the parametric models.

we choose is the economic data best suited to represent the portfolios to forecast,¹⁵ then these results actually say that the conditional CAPM is a reasonable model to consider for the Mexican economy.

References

- Arcones, Miguel A. (1995). On the Central Limit Theorem for U-statistics under Absolute Regularity, *Statistics and Probability Letters*, 24, 245-249.
- Banz, Rolf W. (1981). The Relationship between Return and Market Value of Common Stocks, *Journal of Financial Economics*, 9, 3-18.
- 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.
- Bhandari, Laxmi Chand (1988). Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence, *Journal of Finance*, 43, 507-528.
- Bierens, Herman J. (1987). Kernel Estimators of Regression Functions, in T. F. Bewley (Ed.), *Advances in Econometrics*, 99-144.
- Black, F., Jensen, M., and M. Scholes (1972). The Capital Asset Pricing Model: Some Empirical Tests, in M. Jensen (Ed.), *Studies in the theory of capital markets*, Praeger Publishers, Inc.
- Breen, W. J., L. R. Glosten, and Ravi Jagannathan (1989). Economic Significance of Predictable Variations in Stock Index Returns, *Journal of Finance*, 44, 1177-1190.
- Caldino, Eneas (1996). On the Mean-standard Deviation Frontier, *Estudios Económicos*, vol. 11, núm. 2, 297-319.
- 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.
- Chen, Nai-fu (1991). Financial Investment Opportunities and the Macroeconomy, *Journal of Finance*, 46, 529-554.
- Doukhan, P., P. Massart, and E. Rio (1994). The Functional Central Limit Theorem for Strongly Mixing Processes, *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 30, 63-82.
- Dybvig, P. H., and S. A. Ross (1985). Differential Information and Performance Measurement Using a Security Market Line, *Journal of Finance*, 40, 383-400.
- Fama, Eugene F., and K. R. French (1992). The Cross-section of Expected Stock Returns, *Journal of Finance*, 47, 427-466.
- (1989). Business Conditions and the Expected Returns on Bonds and Stocks, *Journal of Financial Economics*, 25, 23-50.

¹⁵ The programs to run the tests are available upon request.

- Fan, J., and I. Gijbels (1996). *Local Polynomial Modelling and its Applications*, Chapman and Hall.
- Ferson, W. E., and R. H. Campbell (1991). The Variation of Economic Risk Premiums, *Journal of Political Economy*, 99, 385-415.
- Gibbons, Jean D., and Chakraborti, Subhabrata (1992). *Non-parametric Statistical Inference*, NY, Marcel Dekker.
- Gibbons, Michael R. (1982). Multivariate Tests of Financial Models: A New Approach, *Journal of Financial Economics*, 10, 3-27.
- Hansen, Lars Peter, and F. R. Scott (1987). The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models, *Econometrica*, 55, 587-613.
- Jagannathan, Ravi, and Zhenyu Wang (1996). The Conditional CAPM and the Cross-section of Expected Returns, *Journal of Finance*, 51, 3-53.
- Keim, Donald B., and R. F. Stambaugh (1986). Predicting Returns in the Stock and Bond Markets, *Journal of Financial Economics*, 17, 357-390.
- Lintner, J. (1965). Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *Review of Economics and Statistics*, 47, 13-37.
- Long, J. B. (1974). Stock Prices, Inflation, and the Term Structure of Interest Rates, *Journal of Financial Economics*, 1, 131-170.
- Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investment*, John Wiley.
- Merton, Robert C. (1980). On Estimating the Expected Return on the Market: An Exploratory Investigation, *Journal of Financial Economics* 8, 323-361.
- (1973). An Intertemporal Capital Asset Pricing Model, *Econometrica*, 41, 867-887.
- Mossin, J. (1968). Equilibrium in a Capital Asset Market, *Econometrica*, 34, 768-783.
- 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.
- Shanken, Jay (1985). Multivariate Tests of the Zero-beta CAPM, *Journal of Financial Economics*, 14, 327-348.
- Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance*, 19, 425-442.
- Wang, K. Q. (2003). Asset Pricing with Conditioning Information: A New Test, *Journal of Finance*, vol. LVIII, 1, 161-196.
- (2002). Non-parametric Tests of Conditional Mean-variance Efficiency of a Benchmark Portfolio, *Journal of Empirical Finance*, 9, 133-169.
- (1998). *Tests of Conditional Asset Pricing Models: A New Approach*, Ph.D. dissertation, Graduate School of Business, University of Chicago.

Appendix

The technical assumptions needed to obtain theorems 3.2 and 3.3 are the following:

- 1) The data sequence $\{Y_{t+1}\}$ is a strictly stationary process of the β -mixing type, X_t has an absolutely continuous distribution with density $f(X_t)$, and for some $\rho > 2$ the numbers $\beta_n \quad n = 1, 2, \dots$ satisfy

$$\sum_{n=1}^{\infty} n \beta_n^{\frac{\rho-2}{\rho}} < \infty$$

- 2) *i)* $r_{p,t+1}^2, r_{t+1}$ and $r_{p,t+1}r_{t+1}$ have finite first moments,
ii) $\|w_1(t, s)\|_{\rho} < \infty$ and $\|w_2(t, s)\|_{\rho} < \infty \quad \forall \quad t < s$ where

$$w_1(t, s) \equiv (r_{p,s+1}^2 - r_{p,s+1}r_{p,t+1})r_{t+1} + (r_{p,t+1}^2 - r_{p,t+1}r_{p,s+1})r_{s+1}$$

$$w_2(t, s) \equiv r_{p,s+1}^2 Z_t Z_t' + r_{p,t+1}^2 Z_s Z_s'$$

and $\|\cdot\|_{\rho}$ denotes¹⁶ the ρ -norm,

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

- 3) $fg_p, fg_{pp}, fg_r, fg_{pr}$ and fg_{zz} satisfy the Lipschitz local condition for some function $m(X)$, where $m(X_t)r_{t+1}, m(X_t)r_{p,t+1}r_{t+1}, m(X_t)r_{p,t+1}, m(X_t)r_{p,t+1}^2$ and $m(X_t)Z_t Z_t'$ have finite ρ -norm.¹⁷

- 4) The kernel K is a bounded symmetric function satisfying

- i)* $\int K(u)du = 1,$

- ii)* $\int |u|^j |K(u)|du < \infty$ if $0 \leq j \leq k + 1,$

- iii)* $\int u_1^{l_1} \dots u_k^{l_k} K(u)du = 0$ if $0 < l_1 + \dots + l_k < k + 1,$ where u_j is the j -th element of the vector u . In other words, the kernel K is of order $k + 1$.

- 5) *i)* The j -th partial derivative of $fg_p, fg_{pp}, fg_r, fg_{pr}$ and fg_{zz} exist for all $j \leq k + 1,$
ii) The expectations $E[g_{pr} \nabla_{l_1, \dots, l_j}(fg_p)], E[g_r \nabla_{l_1, \dots, l_j}(fg_{pp})],$

¹⁶ $\|X_{ij}\|_{\rho} \equiv E[X_{ij}^{\rho}]^{\frac{1}{\rho}}.$

¹⁷ $h(X)$ satisfy the Lipschitz local condition for some function $m(X)$ if $|h(X+Y) - h(X)| < m(X)\|Y\|.$

$E[g_{pp}\nabla_{l_1,\dots,l_j}(fg_r)]$, $E[g_p\nabla_{l_1,\dots,l_j}(fg_{pr})]$ and $E[g_{pp}\nabla_{l_1,\dots,l_j}(fg_{zz})]$ exist for all $j \leq k + 1$, where the functions and partial derivatives¹⁸ are evaluated at X_t .

- 6) The matrices A and Γ_0 are non-singular. ($\Gamma_0 = E[\gamma(Y_{t+1})\gamma(Y_{t+1})']$).

The mixing condition in 1 restricts the allowed degree of dependence in the data sequence. Among other things, this allows an application of a central limit theorem. The conditions requiring that β_n vanishes as a power of n are, for most financial series, not too restrictive and their use is common practice in the literature. The stronger the restrictions for the existence of the moments (a bigger ρ), the more dependence is allowed. This trade-off is common to establish asymptotic results for serially correlated data. Assumption 5 is a regularity condition for bias asymptotic correction through the use of a kernel of higher order.

For the moment conditions in 2 *i*), we first note that ρ can be chosen arbitrarily close to 2 if β_n decays exponentially. Then it is easy to see that assumption 2 holds for all $\rho > 2$ if the joint distribution of all variables is normal or log-normal.

¹⁸ $\nabla(h)$ denotes the gradient of h .