

DOCUMENTOS DE TRABAJO

BILTOKI

D.T. 2006.03

Nonparametric estimation betas in the Market Model.

M^a Victoria Esteban González y Susan Orbe Mandaluniz

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

Facultad de Ciencias Económicas.
Avda. Lehendakari Aguirre, 83
48015 BILBAO.

Documento de Trabajo BILTOKI DT2006.03

Editado por el Departamento de Economía Aplicada III (Econometría y Estadística)
de la Universidad del País Vasco.

Depósito Legal No.: BI-3005-06

ISSN: 1134-8984

Nonparametric estimation betas in the Market Model*

María Victoria Esteban[†]

Susan Orbe-Mandaluniz[‡]

Abstract

In this study an alternative nonparametric estimator to the Fama and MacBeth approach for the CAPM estimation is proposed. Betas and risk premiums are estimated simultaneously in order to increase the explanatory power of the proxy for betas. A data driven method is proposed for selecting the smoothness degrees, which are directly related to the subsample sizes. Based on this relation, the traditional estimator is obtained as a particular case. Contrary to the results obtained in other studies our empirical evidence for Spanish market data is favorable to the CAPM.

Running title: The smoothed rolling estimator.

*Financial aid from G. C. 9/UPV 00038.321-13631/2001, Ministerio de Educación y Ciencia (SEJ-2005-05549/ECON) and from FBBVA (1/BBVA00038.16421/2004) is gratefully acknowledged.

[†]Departamento de Econometría y Estadística. Universidad del País Vasco. Avenida Lehendakari Agirre, 83; 48015 Bilbao (Spain). E-mail: mvictoria.esteban@ehu.es. Author for correspondence.

[‡]Departamento de Econometría y Estadística. Universidad del País Vasco. Avenida Lehendakari Agirre, 83; 48015 Bilbao (Spain). E-mail: susan.orbe@ehu.es.

1 Introduction

One of the most widely studied topics in financial economics is the description of the trade-off between risk and expected return. The principal result concerning this question is the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965), where the first problem to be solved is the estimation of the betas. Based on empirical finance betas are not stable over time. Evidence on beta instability dates back to the early 1970s (Blume (1971), Gonedes (1973), Baesel (1974), etc.). More recent evidence can be found in Bos & Newbold (1984) and González-Rivera (1997) among others.

The sequence of these unknown time-varying betas is usually estimated through either maximum likelihood, the generalized method of moments or least squares techniques. However, direct estimation by any of these techniques without further restrictions proves unfeasible. Thus, in order to solve this estimation problem some assumptions must be established concerning the structure of time-varying betas. A traditional solution to this estimation problem consists of introducing smoothness constraints on time-varying betas. One alternative is to assume that the sequence of coefficients is random (Cooley & Prescott (1976)), where smoothness restrictions can be introduced in several ways; using prior distributions (Spall (1989)), likelihood procedures in the state space form (Aoki (1987)) or generalized flexible least squares (Lütkepohl & Herwatz (1996)). In CAPM context, random coefficient models have been used by Fabozzi & Francis (1978), and Kalman Filter procedures (see Harvey (1990)) for estimating CAPM betas have been applied by Black, Fraser & Power (1992), Wells (1994) and Brooks, Faff & Josev (1997). The main disadvantage of the random approach is that all a priori distributions involved in the estimation as well as the initial values have to be determined in advance.

Another alternative is to assume that the sequence of betas is a deterministic function of time. In this framework, the estimation of time-varying coefficients is reduced to the estimation of an underlying function where the crucial point is the specification of this unknown function that relates the sequence of betas with the time index. Most approaches made in this direction consider that betas vary in a deterministic way across subsets of

observations within the sample. The estimation method used is generally ordinary or rolling least squares for a prefixed subsample.

In this paper we are interested in this last alternative: Deterministic time-varying betas. We assume that betas are unknown smooth functions of the time index and in order to avoid misspecification problems we propose using semiparametric estimation techniques. The semiparametric estimator proposed is based on the estimator described in Robinson (1989), modified to reach consistency according to the characteristics of the estimation framework corresponding to the CAPM.

The goal of this paper is to propose a flexible semiparametric estimation method that generalizes the traditional rolling least squares estimator. We present the traditional estimator as a particular case and relate the selection of the prefixed subsample to the selection of the bandwidth or smoothness parameter, which is chosen using a data driven method. A comparison of empirical results based on variable significance and expected signs, between the proposed and the traditional rolling estimators is made using Spanish stock market data. We conclude that the empirical evidence is favorable to the CAPM in terms of absence of intercept and a positive risk market premium statistically significant.

The rest of the paper is organized as follows. Section 2 presents the theoretical model and the proposed estimation procedure, which generalizes the traditional one. Section 3 presents an illustration to compare empirical results. Section 4 concludes.

2 Theoretical model and estimation methodology

The CAPM implies that the expected return of an asset must be linearly related to the covariance between its return and the return of the market portfolio. Sharpe (1964) and Lintner (1965) derive the CAPM assuming the possibility of lending and borrowing at a risk free rate of interest. This version of CAPM is generally expressed for the expected return of the asset i through the following equation:

$$E(R_i) - r_f = \beta_i [E(R_m) - r_f] \tag{1}$$

where R_i is the rate of return on the i -th risky asset, R_m is the rate of return on the market portfolio and r_f is the return on the riskfree asset. Equation 1 indicates that the expected return on the i -th risky asset must be the return on the riskfree asset plus a risk premium. The parameter β_i is interpreted as the contribution to market risk made by the i -th asset. This coefficient is the ratio of the covariance, between its return and the return of the market portfolio, and the market variance:

$$\beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)}. \quad (2)$$

The Sharpe-Lintner version can be expressed in terms of returns in excess of the riskfree rate, “excess returns”, so defining $Z_i = R_i - r_f$ as the excess of return for the i -th asset in excess of the riskfree rate and Z_m as the excess of return on the market portfolio, Equations 1 and 2 can be rewritten as:

$$E(Z_i) = \beta_i E(Z_m) \quad (3)$$

$$\beta_i = \frac{Cov(Z_i, Z_m)}{Var(Z_m)}. \quad (4)$$

In this setting, Equation 3 has three direct implications. First, in this relation there is no intercept. Second, the parameter β_i completely captures the cross-sectional variation of expected excess returns. And third, the market risk premium, $E(Z_m)$, must be positive.

Since the CAPM is a single-period model, that is, Equation 3 does not have a time dimension, Fama & MacBeth (1973) estimate the CAPM using a cross-sectional approach in order to test the second and third implications derived from Equation 3. The cross-sectional regression model for a given time t is given by:

$$Z_{it} = \gamma_{0t} + \gamma_{1t}\beta_{it} + \eta_{it} \quad i = 1, \dots, N \quad (5)$$

where the sample size N is determined by the number of portfolios, the coefficients γ_{1t} (the market risk premium) and γ_{0t} (the expected zero-beta portfolio return with respect to the market) are unknown, Z_{it} is the excess return of asset (portfolio) i at time t , β_{it} is the explanatory variable and η_{it} is the error term. In this context the main estimation problem

comes from the feature that the coefficients $(\gamma_{0t}, \gamma_{1t})$ and the explanatory variable (β_{it}) are all unknown. Therefore, in order to estimate the gammas, the unknown betas must be estimated. The main idea of the estimation method is to run the excess returns over the proxy of the betas for each cross section and aggregate the estimates of the premium in the time dimension. Thus the final model to be fitted is:

$$Z_{it} = \gamma_{0t} + \gamma_{1t}\hat{\beta}_{it} + \eta_{it} \quad i = 1, \dots, N. \quad (6)$$

Since the CAPM must be tested in Model 6 above the procurement of an adequate proxy for the betas is crucial. Next we propose a flexible semiparametric estimation method for estimating the proxy and risk premium simultaneously in Model 5 taking into account the characteristics of the estimation setting. Before we present the whole estimation procedure, we will focus our attention on obtaining the proxy for the unknown explanatory variable, which must satisfy two conditions. First, it must be closely correlated to the real variable β_{it} , so in consequence it must be able to explain the excess of returns, Z_{it} . And second, in order to obtain consistency it must be temporarily uncorrelated to the error term (η_{it}) , that is $E(\sum_{i=1}^N \hat{\beta}_{it}\eta_{it}) = 0$.

We approach the first condition for the proxy in the same way as Fama & MacBeth (1973). The proxy of β_{it} for a given portfolio i and a fixed time estimation moment t is obtained as the slope coefficient in the excess-return market model for a given subsample. That is, the beta coefficient in the following regression equation:

$$Z_{ij} = \alpha_{itm} + \beta_{itm}Z_{mj} + \epsilon_{ij} \quad j \in s(t, h_i, T) \quad (7)$$

where Z_{mj} is the excess return on the market portfolio at time j . The subscript m in the coefficients, indicating market portfolio, will be dropped from now on in order to simplify the notation (thus $\beta_{it} \equiv \beta_{itm}$). The subsample used, $s(t, h_i, T)$, depends on the moment in time for which we are estimating the equation (t), the parameter that regulates the subsample or window size (h_i) and the total number of observations available in the time dimension (T). In fact each subsample goes from $t-1$, i.e. the previous observation of the time estimation moment, to $s(t, h_i, T)$ which varies according the values of t , h_i and T .

Running Equation 7 over the time dimension we obtain the values $\{\hat{\beta}_{it}\}_{t=1}^T$ and repeating it independently for each portfolio $i \in [1, N]$ we get all values required to estimate the premiums in (6).

Nonetheless, the empirical results in the relevant literature are not satisfactory because neither the sign nor the estimated premiums are those expected. We think that one reason for these results might be the use of an inappropriate proxy variable. Therefore, in order to improve the explanatory power of the proxy we use a semiparametric estimator to estimate time-varying betas which is a modified version of the estimator in Robinson (1989). The modification is introduced in order to satisfy the second condition for the proxy, which is easily reached if the estimation of the proxy for β_{it} only uses past observations ($j < t$) in Equation 7. Thus we propose to estimate the proxy by minimizing the following smoothed sum of squared residuals, given portfolio i and smoothness degree h_i :

$$\min_{(\alpha_{it}, \beta_{it})} \sum_{j=t-1}^{t-Th_i} K_{h_i,tj} (Z_{ij} - \alpha_{it} - \beta_{it} Z_{mj})^2, \quad (8)$$

where $K_{h_i,tj} = h_i^{-1} K((t/T - j/T)/h_i)$ is a symmetric second order kernel with compact support $[-1, 1]$ and h_i , called bandwidth, determines the smoothness degree imposed and therefore regulates the window size. So the subsample size used at each estimation time t , given by $[t - Th_i, t - 1]$, is the same when estimating the betas for the i -th portfolio but can be different when we estimate the betas corresponding to another portfolio. Note that the size of all the subsamples in the i -th portfolio is determined by the smoothing parameter h_i and the number of observations in them is Th_i . Large values of h_i impose higher smoothness, implying a larger subsample size so that more past observations are employed at each local estimation and vice versa. In this sense, choosing the bandwidth (h_i) implies selecting the subsample size. This semiparametric technique for estimating the proxy is based on the assumption that betas are somehow smooth over time. This means that each sequence of coefficients, $\{\beta_{it}\}_{t=1}^T$, lies on an unknown function of the time index, that is $\{\beta_{it} = f_i(t/T)\}_{t=1}^T$ is a smooth function such that $f_i(t/T) \in C^2[0, 1]$ for all $i \in [1, N]$. The advantage of this estimator is that it does not need to specify the unknown function $f_i(t/T)$ to determine how coefficients behave in time. It also allows for

linear and nonlinear specifications of time index.

The expression of the “*smoothed rolling estimator*” for the betas derived from the normal equations from (8), once the smoothness degree h_i is fixed, is given by:

$$(\hat{\alpha}_{it} \quad \hat{\beta}_{it})'_{SR} = \left(\sum_{j=t-1}^{t-Th_i} K_{h_i,tj} X_j X_j' \right)^{-1} \sum_{j=t-1}^{t-Th_i} K_{h_i,tj} X_j Z_{ij} \quad (9)$$

where $X_j = (1 \ Z_{mj})'$ is the j -th observation of the explanatory variables and the subscript SR denotes “smoothed rolling estimator”. The closed form expression for the estimator ensures that no iterative methods are needed in order to calculate the estimations. And assuming that the matrix to invert in (9) is not singular, it is the unique solution to the system of normal equations from (8). Note that the smoothness assumption is made over the coefficients. So a small bandwidth parameter provides very rough coefficients, usually with no reasonable interpretations, and leads to an estimated response variable equal to its past value ($\hat{Z}_{it} = Z_{it-1}$). By contrast, with high degree of smoothness, little variability is allowed and the estimations tend to be very similar over time. As usual in a semiparametric setting (see Eubank (1988) and Härdle (1990) among others) the selection of the bandwidth (h_i) is crucial. The bandwidth cannot be very small because the estimations will not be interpretable, but it cannot be very large because in that case the values of the explanatory variable in the regression Equation 6 will be nearly the same for all time periods.

Note that the traditional rolling estimator used by Fama & MacBeth (1973) can be obtained as a particular case from (9). Recall that they estimate the series of betas for each portfolio i by repeating the estimation for subsamples of several years prior to each estimation moment t . The window size used, which determines the length of the subsample, is set to be the same throughout the period and for all the portfolios under study. The selection of the window size is based on the assumption that the joint distribution of Z_{it} and Z_{mt} is stationary over time. Fisher (1970) and Gonedes (1973) among other authors, find empirically that when using monthly data the optimal subsample is between four and seven years, so traditionally the window size has been taken as five years: Each subsample begins at $t - 60$ and ends at $t - 1$. Therefore the

estimator used by Fama & MacBeth (1973) for the proxy derives from minimizing the local sum of squared residuals corresponding to Model 7 at each moment of time using sixty past monthly observations:

$$\min_{(\alpha_{it}, \beta_{it})} \sum_{j=t-1}^{t-60} (Z_{ij} - \alpha_{it} - \beta_{it} Z_{mj})^2 \quad (10)$$

and the expression of the “rolling estimator” for the betas derived from the normal equations from (10) is given by:

$$(\hat{\alpha}_{it} \ \hat{\beta}_{it})'_R = \left(\sum_{j=t-1}^{t-60} X_j X'_j \right)^{-1} \sum_{j=t-1}^{t-60} X_j Z_{ij} \quad (11)$$

where $X_j = (1 \ Z_{mj})'$ is the j -th observation of the explanatory variables and the subscript R denotes “rolling estimator”. Comparing expressions (10) and (11) with (8) and (9) respectively, it is easy to observe that the estimator proposed by Fama & MacBeth (1973) is obtained when the kernel used is uniform, that is, all observations in the subsample are given the same weight, and the bandwidths for all portfolios are chosen as $h_i = 60/T$, which implies that the number of observations used at each local estimation is sixty.

If our aim in this paper were only to generalize the estimator for the proxy, maintaining the rest of the methodology of Fama and MacBeth, then we should pick the t -th observation from the second element of (9) for each portfolio, build the proxy and estimate the premiums in Model 6. But if the final goal is the estimation of the regression coefficients γ_{0t} and γ_{1t} in Model 5 then we have to recall that we need a good proxy for the cross-sectional regression model. That is, we are interested on estimating as well as possible the slope values corresponding to all portfolios at the same moment in time t that are contained in the proxy, $\{\hat{\beta}_{it}\}_{i=1}^N$. Hence for the proxy, besides a high correlation with the real variable to replace, we want to maximize dispersion because this increases the precision in the estimator of the premiums. Thus, with the values of the smoothness degrees ($\{h_i\}_{i=1}^N$) fixed we propose to estimate proxies and coefficients simultaneously by minimizing the following smoothed sum of squared residuals:

$$\min_{(\gamma_{0t}, \gamma_{1t}, \hat{\beta}_{h_i, it})} \sum_{t=1}^T \sum_{i=1}^N (Z_{it} - \gamma_{0t} - \gamma_{1t} \hat{\beta}_{h_i, it})^2 \quad (12)$$

subject to

$$\hat{\beta}_{h_i, it} = \frac{\sum_{j=t-1}^{t-Th_i} K_{h_i, tj} \sum_{j=t-1}^{t-Th_i} K_{h_i, tj} Z_{mj} Z_{ij} - \sum_{j=t-1}^{t-Th_i} K_{h_i, tj} Z_{mj} \sum_{j=t-1}^{t-Th_i} K_{h_i, tj} Z_{ij}}{\sum_{j=t-1}^{t-Th_i} K_{h_i, tj} \sum_{j=t-1}^{t-Th_i} K_{h_i, tj} Z_{mj}^2 - \left(\sum_{j=t-1}^{t-Th_i} K_{h_i, tj} Z_{mj} \right)^2} \quad (13)$$

which is the slope semiparametric estimator given by the second element of (9). Usually the bandwidth is selected using a data driven method, mainly cross-validation, the penalized sum of squared residuals or plug-in methods. But since in this estimation procedure the observation corresponding to a given moment in time t is not included in the subsample when estimating at that point, the minimization of the smoothed sum of squared residuals in (8) does not lead to the typical selection of a null smoothness degree. Therefore, we propose to select the bandwidth by minimizing the sum of squared residuals in the main regression, Model 5. The bandwidth chosen by this minimization procedure is such that the dispersion of the proxy, $(\{\hat{\beta}_{h_i, it}\}_{i=1}^N)$, increases, leading to a greater explanatory power for estimating the risk premium in Model 5. Under the assumptions usually taken in nonparametric settings, this estimator is consistent and asymptotically normally distributed. Note that the convergence rate is lower than optimal because the order of the bias is $O(h)$ due to the kernel employed, equivalent to a one-sided kernel that only takes into account past observations¹.

The above minimization problem generalizes the one used by Fama & MacBeth (1973) because no decisions have to be taken about the subsamples to be used. First, there is no need of a subjective choice of subsample size because it is controlled by the bandwidth, which is selected by a data driven method. Second, once the subsample is fixed, it allows for observations to be weighted differently, giving higher weights to those that are closer in time. Finally it allows the information about the portfolios together to be used for obtaining a proxy with greater explanatory power.

¹Similar nonparametric asymptotic results are obtained by Cline & Hart (1991) for density estimation, Müller (1992) for regression and Chen & Liu (1993) for autoregressive models.

3 Empirical results: An illustration

In this section we illustrate the empirical results obtained using monthly excess returns with the appropriate adjustments for stock dividends and capital changes for 150 assets during the period January 1968 through December 2000 for the Spanish capital market. Ten portfolios are formed according to the capital stock exchange (size) of each firm at the end of each year. The portfolios have an approximately similar number of assets. The assets that make up the portfolios might change from year to year but the size remains constant. All the assets in a given portfolio receive the same weight in calculating the aggregate return.

The return of the risk free asset is the monthly percentage rate offered by one year Spanish Treasury Bills in the secondary market. Before 1982 the rates of the loans granted by financial institutions were used. Although these rates cannot be considered as real rates of interest of risk free assets, they are considered to be a reasonable approximation. The proxy for the market portfolio is the value-weighted stock market index in excess of the risk free asset formed with the sample available. Table 1 presents the summary statistics for the excess returns for the size portfolios and the index. The results are similar to other studies using Spanish data. As usual the portfolio that contains the smallest assets obtains higher mean returns than the rest.

Table 2 reports the summary statistics for the time series of estimated betas from the market model, Equation 7, using the traditional rolling estimator for the period 1973:1-2000:12. Each beta, $\{\hat{\beta}_{it}\}_{t=1}^T$, is estimated as the slope coefficient in a time-series regression using as its explanatory variable the market portfolio, the value weighted index. The subsamples employed at each estimation moment t for each portfolio contain the sixty previous observations (months), which are equally weighted. Columns two and three present the mean and standard deviations of the time series of estimated betas for each decile portfolio. Columns four and five show the minimum and maximum values of the R -squared, which increase from C1 to C10 as expected due to their construction. Columns six and seven show the minimum and maximum values of the t -statistics for testing the

significance of the variable ($H_0 : \beta_i = 0$). Also as expected, all the betas are statistically different from zero at a significance level of 5% for all the portfolios. So the value weighted market index used to approach the true market portfolio adequately explains the excess return of portfolios formed by size.

Tables 3 and 4 report the summary statistics for the estimated betas from the same market model and sample period 1973:1-2000:12, but using the smoothed rolling estimator with two alternative kernels, the Uniform and the Epanechnikov kernels, respectively. As before, each beta, $\{\hat{\beta}_{it}\}_{t=1}^T$, is estimated as the slope coefficient in a time-series regression using the value weighted index as the explanatory variable. The advantage of, and therefore the difference in, the proposed smoothed rolling estimator with respect to the traditional rolling estimator is that it uses different subsample sizes across portfolios and offers the possibility of weighting observations within subsamples differently. These smoothing parameters are selected according to the data driven method based on the minimization of the sum squared residuals in (5), where the proxy and the risk premium are estimated through (12) using the corresponding kernel function (uniform or Epanechnikov). Once the smoothing parameters are chosen, and thus the subsample sizes fixed, the weights are determined by the kernel function employed. When a uniform kernel ($K(u) = (1/2)I(|u| \leq 1)$) is applied, the observations in the subsample are given the same weight. By contrast, if an Epanechnikov ($K(u) = (3/4)(1 - u^2)I(|u| \leq 1)$) kernel is used, the observations in the subsample are not weighted equally. In this last case the observations nearest the estimation time t are given higher weights than those farthest from it, that is, weights decline as the time lag increases. In consequence the differences between the results of these two tables are due to the use of different kernels. In this way, columns two to seven present the same statistics as Table 2, but referring to the smoothed estimators. Column three shows an increase in the variability in the time series estimated betas, the standard deviation is greater for all portfolios than for the rolling estimator. Columns four and five present more extreme values for R -squared. Columns six and seven show the minimum and maximum values of the t-statistics for testing the significance of

the market portfolio ($H_0 : \beta_i = 0$). Although for portfolios C5, C6, C8 and C10 the market portfolio is always relevant, for the remaining portfolios the minimum value for some t-statistics indicate no rejection of the null hypothesis, but the percentage with this result is negligible. These results must not be discouraged: we prefer to renounce the best fitting when estimating the betas in order to increase the variability for obtaining a greater precision in the estimated premiums. With regard to Table 4 similar results are obtained.

In order to check whether the explanatory power of the proxy increases when we use the proposed estimation method, we estimate Model 5 and calculate the corresponding sample variance for the explanatory variable (the proxy) and the sum of squared residuals (SSR) for each period of time from January 1973 to December 2000. The columns in Table 5 present the maximum and minimum values of these statistics when the proxy is estimated using the traditional rolling estimator given in (11) and the smoothed rolling estimator in (9) for Uniform and Epanechnikov kernels. It can be observed that the range of variability for the proxy variable increases when we use the smoothed rolling estimators (different subsample sizes across portfolios) instead of the traditional rolling estimator (same subsample size for all portfolios). Furthermore when we use the smoothed rolling estimator, the variability is greater when the observations in the subsamples are weighted differently than when they are weighted equally. Thus, according to the fact that an explanatory variable with a greater variance is preferred to increase its explanatory power, the third proxy will perform better. In relation to the sum of squared residuals, the result it is similar: the range of values for the SSR is smaller for the smoothed rolling estimator with the Epanechnikov kernel. The explained variability of the cross section of the excess returns of portfolios formed by size is greater for this estimator.

Table 6 presents some results for the estimated risk premiums, $\hat{\gamma}_{0t}$ and $\hat{\gamma}_{1t}$. Each horizontal panel contains the estimated mean premium, the t-ratio for testing the hypothesis that the mean premium over time is equal to zero (in round brackets) and the t-ratio calculated with the Shanken (1992) correction (in square brackets). The results in the

first column are for the rolling estimator. It can be observed that when the proxy is estimated using the rolling estimator, the average premium associated with the constant is statistically different from zero at 5%. Thus the implication of absence of intercept in the model is not supported by the data. In this same framework, the average premium for the market is negative but not statistically different from zero at 5%, so the sensitivity to market risk does not explain the excess of return when portfolios are formed by size. In other words, the market does not price risk. Similar results, contrary to what is expected by theory, are obtained by Esteban (1997) and Nieto & Rodríguez (2005) for the Spanish stock market². Columns two and three present the results for the smoothed rolling estimator using the two kernels considered. Independently of the kernel used the results are more satisfactory: we find absence of intercept and that the market prices risk so there is favorable evidence for CAPM. Clearly, the statistical results are better for the Epanechnikov kernel than for the Uniform kernel. Recall that the Epanechnikov kernel weights observations differently within the subsample, hence the estimator for the proxy (Equation 9) can be interpreted as a kind of generalized least squares estimator where observations are penalized according to their distance from the estimation time. The improvement comes from the way the available information is used, the weights decrease as the observation goes back in time such that recent information is more highly valued.

4 Conclusions

Applying the procedure of Fama & MacBeth (1973) for the estimation of the CAPM using Spanish stock market data produces a negative estimated market risk premium which is not statistically significant, which implies that the market does not price risk. These results are contrary to the CAPM because, among other things, they require that the market risk premium to be positive and statistically different from zero in order to support the theoretical implications. There may be several reasons for this statistical

²For the American stock market Chen, Roll & Ross (1986) or Ferson & Harvey (1991), among others, obtain similar contrary results.

evidence against the CAPM: Improper selection of the proxy for the market portfolio, inappropriate subsamples for the local estimation of betas or incorrect assumptions about the returns distribution. We are inclined to attribute it to the first two reasons and we try to overcome the problem by using a more flexible estimator.

Our alternative to the Fama and MacBeth approach is to estimate betas and risk premiums simultaneously using nonparametric techniques. This estimator generalizes the estimator for the betas used by Fama & MacBeth (1973) including a data driven method for selecting the subsample sizes. However this particular case does not minimize the sum of squared residuals of Model 5 so it has not been chosen. Thus one possible reason for the evidence against the CAPM is the poor explanatory power of the proxy, leading to a reduction in the precision in the estimator of the risk premiums and consequently to the CAPM not being supported by the data. We think that the problem with the proxy might be due to the following points. On the one hand the same subsample size they use for all portfolios does not seem to be optimal, and all past observations are given the same weight. On the other hand there is no connection between the temporal and cross-sectional dimensions in the estimation.

The advantage of the estimation procedure depends on its flexibility: it is able to generalize the above points. We think that the data structure corresponding to different portfolios is not necessarily common so the same subsample size for all portfolios may not be adequate. And giving the same weight to all past observations presupposes that no structural changes have occurred in previous years. It is more reasonable to give higher weights to observations that are closer in time. We are convinced that estimating Model 5 employing all the information simultaneously leads to a more efficient estimator because it takes into account the time and cross relations between the market, sensitivities and risk premiums. We recall that the empirical results from the proposed estimation procedure, considering the above generalizations, support the CAPM for the Spanish stock market in terms of absence of intercept and a positive risk market premium which is statistically significant. Finally, further analysis could consist of accommodating additional risk measures

beyond the CAPM beta and estimating these models using this estimation procedure to check whether the proposed estimation procedure provides the same satisfactory results.

5 Acknowledgements

Financial aid from G. C. 9/UPV 00038.321-13631/2001, Ministerio de Educación y Ciencia (SEJ-2005-05549/ECON) and from FBBVA (1/BBVA00038.16421/2004) is gratefully acknowledged.

Table 1: Summary statistics for asset excess returns

Portfolio	Mean	S.D.	Min.	Max.
<i>C1</i>	1.31	9.62	-46.02	53.56
<i>C2</i>	1.28	7.28	-28.01	37.47
<i>C3</i>	0.97	7.73	-30.79	43.27
<i>C4</i>	1.29	7.19	-26.13	30.98
<i>C5</i>	0.99	6.79	-34.53	37.69
<i>C6</i>	0.92	6.27	-30.55	29.85
<i>C7</i>	0.59	6.46	-23.90	31.40
<i>C8</i>	0.70	6.64	-33.44	45.12
<i>C9</i>	0.63	6.05	-31.68	27.61
<i>C10</i>	0.74	5.85	-23.18	25.77
<i>VW</i>	0.61	5.88	-29.74	21.14
<i>RFR</i>	0.56	0.29	0.14	1.19

The table presents the summary statistics for ten portfolios formed by size, VW is the value-weighted stock market index in excess of the risk free asset, RFR. Period 1968:1 2000:12 with 396 observations. Data in percentages per month.

Table 2: Summary statistics for betas estimated using the rolling estimator.

Portfolio	$\bar{\hat{\beta}}$	$S_{\hat{\beta}}$	Min R^2	Max R^2	Min $t_{\hat{\beta}}$	Max $t_{\hat{\beta}}$
<i>C1</i>	1.11	0.265	26.50	76.42	4.57	13.71
<i>C2</i>	0.83	0.226	18.12	68.35	3.58	11.19
<i>C3</i>	0.97	0.176	31.01	79.61	5.10	15.04
<i>C4</i>	0.93	0.143	31.14	77.16	5.12	13.99
<i>C5</i>	0.96	0.101	52.37	83.28	7.98	16.99
<i>C6</i>	0.83	0.136	46.72	83.08	7.13	16.87
<i>C7</i>	0.91	0.208	55.04	83.36	8.42	17.04
<i>C8</i>	0.95	0.172	51.17	88.22	7.79	20.84
<i>C9</i>	0.96	0.093	56.82	90.53	8.73	23.55
<i>C10</i>	0.98	0.093	79.38	97.89	14.94	51.96

The table presents the summary statistics for the betas estimated in $Z_{it} = \alpha_{im} + \beta_{im}Z_{mt} + \epsilon_{it}$, the market model, for 336 monthly observations corresponding to the period 1973:1-2000:12. For each portfolio i the betas are estimated using the rolling estimator with a window size of 60, that is, all subsamples start at $t - 60$ and end at $t - 1$:

$$(\hat{\alpha}_{it} \ \hat{\beta}_{it})'_R = \left(\sum_{j=t-60}^{t-1} X_j X_j' \right)^{-1} \sum_{j=t-60}^{t-1} X_j Z_{ij}$$

where $X_j = (1 \ Z_{mj})'$ is the j -th observation of the explanatory variables and Z_i is the excess return in the i -th portfolio formed by size, Z_m is the value-weighted market return (in excess). $\bar{\hat{\beta}}$ is the average of the time series of estimated betas. $S_{\hat{\beta}}$ is the standard error of the respective time series. Min R^2 and Max R^2 are the minimum and maximum values of the respective R^2 time series. Min $t_{\hat{\beta}}$ and Max $t_{\hat{\beta}}$ are the minimum and maximum values of the t-statistics time series for testing $H_0 : \beta_i = 0 \quad i = 1, \dots, 10$.

Table 3: Summary statistics for the time series of betas estimated using the smoothed rolling estimator with a Uniform kernel.

Portfolio	$\bar{\hat{\beta}}_i$	$S_{(\hat{\beta}_i)}$	Min R^2	Max R^2	Min $t_{\hat{\beta}}$	Max $t_{\hat{\beta}}$
<i>C1</i>	1.06	0.578	0.050	95.80	-0.97	12.65
<i>C2</i>	0.77	0.483	0.002	96.78	-0.86	13.44
<i>C3</i>	0.97	0.569	0.003	99.09	-0.37	20.93
<i>C4</i>	0.88	0.293	4.36	85.96	0.82	9.58
<i>C5</i>	0.93	0.149	35.93	89.60	3.81	14.97
<i>C6</i>	0.81	0.188	31.22	92.58	3.43	18.02
<i>C7</i>	0.88	0.388	10.18	96.83	0.95	15.63
<i>C8</i>	0.91	0.281	13.35	95.34	1.35	15.66
<i>C9</i>	0.92	0.210	10.30	96.12	1.17	17.24
<i>C10</i>	0.99	0.146	67.19	98.81	6.07	38.77

The table presents the summary statistics for the time series of betas estimated in $Z_{it} = \alpha_{im} + \beta_{im}Z_{mt} + \epsilon_{it}$, the market model, for 336 monthly observations corresponding to the period 1973:1-2000:12. For each portfolio i the betas are estimated using the smoothed rolling estimator with a uniform kernel ($K(u) = (1/2)I(|u| \leq 1)$) and corresponding smoothing parameter h_i . So, each window size is Th_i , that is, the subsamples start at $t - Th_i$ and end at $t - 1$:

$$(\hat{\alpha}_{it} \quad \hat{\beta}_{it})'_{SR} = \left(\sum_{j=t-1}^{t-Th_i} K_{h_i,tj} X_j X_j' \right)^{-1} \sum_{j=t-1}^{t-Th_i} K_{h_i,tj} X_j Z_{ij}$$

where $X_j = (1 \quad Z_{mj})'$ is the j -th observation of the explanatory variables and Z_i is the excess return in the i -th portfolio formed by size, Z_m is the value-weighted market return (in excess). $\bar{\hat{\beta}}$ is the average of the time series of estimated betas. $S_{\hat{\beta}}$ is the standard error of the respective time series. Min R^2 and Max R^2 are the minimum and maximum values of the respective R^2 time series. Min $t_{\hat{\beta}}$ and Max $t_{\hat{\beta}}$ are the minimum and maximum values of the t-statistics time series for testing $H_0 : \beta_i = 0 \quad i = 1, \dots, 10$.

Table 4: Summary statistics for the time series of betas estimated using the smoothed rolling estimator with an Epanechnikov kernel.

Portfolio	$\bar{\hat{\beta}}_i$	$S_{\hat{\beta}_i}$	Min R^2	Max R^2	Min $t_{\hat{\beta}}$	Max $t_{\hat{\beta}}$
<i>C1</i>	1.06	0.545	7.4e-004	9.5e+001	-1.04	13.40
<i>C2</i>	0.78	0.573	7.7e-006	9.8e+001	-1.06	19.44
<i>C3</i>	0.97	0.569	3.0e-005	99.09	-0.37	20.93
<i>C4</i>	0.88	0.293	4.36	85.96	0.82	9.58
<i>C5</i>	0.93	0.149	35.93	89.60	3.81	14.97
<i>C6</i>	0.81	0.167	31.22	92.58	3.43	18.02
<i>C7</i>	0.88	0.372	9.18	93.91	0.95	11.78
<i>C8</i>	0.92	0.264	23.85	94.66	2.23	16.84
<i>C9</i>	0.93	0.190	10.09	94.89	1.34	17.24
<i>C10</i>	0.99	0.127	67.04	98.98	6.69	46.21

The table presents the summary statistics for the time series of betas estimated in $Z_{it} = \alpha_{im} + \beta_{im}Z_{mt} + \epsilon_{it}$, the market model, for 336 monthly observations corresponding to the period 1973:1-2000:12. For each portfolio i the betas are estimated using the smoothed rolling estimator with an Epanechnikov kernel ($K(u) = (3/4)(1 - u^2)I(|u| \leq 1)$) and corresponding smoothing parameter h_i . So, each window size is Th_i , that is, the subsamples start at $t - Th_i$ and end at $t - 1$:

$$(\hat{\alpha}_{it} \quad \hat{\beta}_{it})'_{SR} = \left(\sum_{j=t-1}^{t-Th_i} K_{h_i,tj} X_j X_j' \right)^{-1} \sum_{j=t-1}^{t-Th_i} K_{h_i,tj} X_j Z_{ij}$$

where $X_j = (1 \quad Z_{mj})'$ is the j -th observation of the explanatory variables and Z_i is the excess return in the i -th portfolio formed by size, Z_m is the value-weighted market return (in excess). $\bar{\hat{\beta}}$ is the average of the time series of estimated betas. $S_{\hat{\beta}}$ is the standard error of the respective time series. Min R^2 and Max R^2 are the minimum and maximum values of the respective R^2 time series. Min $t_{\hat{\beta}}$ and Max $t_{\hat{\beta}}$ are the minimum and maximum values of the t-statistics time series for testing $H_0 : \beta_i = 0 \quad i = 1, \dots, 10$.

Table 5: Summary statistics for the proxy variable $\hat{\beta}$ estimated using different estimators.

	Traditional Rolling		Smoothed Rolling (Uniform kernel)		Smoothed Rolling (Epanechnikov kernel)	
	Min	Max	Min	Max	Min	Max
$Var(\hat{\beta})$	0.0069	0.0507	0.0075	0.7983	0.0124	1.550
SSR	3.41e-004	1.72e-001	3.37e-004	1.43e-001	2.98e-004	1.33e-001

The table presents the summary statistics for the estimated proxy variable to be used as an explanatory variable in the main model $Z_{it} = \gamma_{0t} + \gamma_{1t}\hat{\beta}_{it} + \eta_{it}$ in order to estimate the market risk premium. The columns present the maximum and minimum values of the sample variance for the explanatory variable and the sum of squared residuals (SSR) of the T cross-sectional series of order N .

Table 6: Average risk premium

	Traditional Rolling	Smoothed Rolling (Uniform kernel)	Smoothed Rolling (Epanechnikov kernel)
$\bar{\hat{\gamma}}_1$	-0.0028 (-0.44756) [-0.44539]	0.0079 (1.90786) [1.90039]	0.0081 (2.03633) [2.03084]
$\bar{\hat{\gamma}}_0$	0.0104 (1.76310) [1.75450]	0.00049 (0.09780) [0.09761]	-0.00018 (-0.03785) [-0.03758]

The estimation of the model $Z_{it} = \gamma_{0t} + \gamma_{1t}\beta_{it} + \eta_{it}$ $i = 1, \dots, N$ produces the risk premium estimates $\hat{\gamma}_{1t}$. The mean of the time series risk premium estimated is called $\bar{\hat{\gamma}}_1$. The t-statistics for the average of the risk premiums over time are in parentheses. This statistic, $\bar{\hat{\gamma}}_1/\hat{\sigma}_{\hat{\gamma}}_1$, is asymptotically normally distributed. The second parenthesis is the t-statistics with the correction proposed by Shanken (1992) for the biases introduced by the errors-in-variables problem. Estimation results for the intercept are also presented. Period 1973:1-2000:12

References

- Aoki, M. (1987), *State Space Modelling of Time Series*, Springer Verlag: Berlin.
- Baesel, J. (1974), ‘On the assessment of risk: Some further considerations’, *Journal of Finance* **29**, 1491–1494.
- Black, A., Fraser, P. & Power, D. (1992), ‘UK unit trust performance, 1980-1989: A passive time-varying approach’, *Journal of Banking and Finance* **16**, 1015–1033.
- Blume, M. (1971), ‘On the assesment of risk’, *Journal of Finance* **26**, 1–10.
- Bos, T. & Newbold, P. (1984), ‘An empirical investigation of the possibility of stochastic systematic risk in the market model’, *Journal of Business* **57**, 34–41.
- Brooks, R., Faff, R. & Josev, T. (1997), ‘Beta stability and monthly seasonal effects: Evidence from the Australian capital markets’, *Applied Economics Letters* **4**, 563–566.
- Chen, N., Roll, R. & Ross, S. (1986), ‘Economic forces and stock market’, *Journal of Business* **61**, 409–425.
- Chen, R. & Liu, L. (1993), ‘Functional coefficient autoregressive models: estimation and tests of hypotheses’, *Journal of Time Series Analysis* **22**, 151–173.
- Cline, D. & Hart, J. (1991), ‘Kernel estimation of densities with discontinuities or discontinuous derivatives’, *Statistics* **22**, 69–84.
- Cooley, T. & Prescott, E. (1976), ‘Estimation in the presence of stochastic parameter variation’, *Econometrica* **44**, 167–184.
- Esteban, M. (1997), ‘Variabilidad predecible en los rendimientos de los activos. evidencia e implicaciones.’, *Investigaciones Económicas* **21**, 523–542.
- Eubank, R. (1988), *Spline smoothing and nonparametric regression*, Marcel Dekker, Inc., New York.
- Fabozzi, F. & Francis, J. (1978), ‘Beta as a random coefficient’, *Journal of Financial and Quantitative Analysis* **13**, 101–116.

- Fama, E. & MacBeth, J. (1973), 'Risk, return, and equilibrium: Empirical test', *Journal of Political Economy* **81**, 607–636.
- Ferson, W. & Harvey, C. (1991), 'The variation of economic risk premiums', *Journal of Political Economy* **99**, 385–415.
- Fisher, L. (1970), The estimation of systematic risk: Some new findings, in 'Proceedings of the Seminar on the Analysis of Security Price', University of Chicago, May.
- Gonedes, N. (1973), 'Evidence on the information content of accounting numbers: Accounting-based and market-based estimates of systematic risk', *Journal of Financial and Quantitative Analysis* **8**, 407–443.
- González-Rivera, G. (1997), 'The pricing of time-varying beta', *Empirical Economics* **22**, 345–363.
- Härdle, W. (1990), *Applied Nonparametric Regression*, Cambridge University Press, Cambridge.
- Harvey, A. (1990), *Forecasting Structural Time Series and the Kalman Filter*, Cambridge University, Cambridge.
- Lintner, J. (1965), 'Valuation of risk assets and the selection of risk investments in stock portfolios and capital budgets', *Review of Economics and Statistics* **47**, 13–37.
- Lütkepohl, H. & Herwatz, H. (1996), 'Specification of varying coefficients time series models via generalized flexible least squares', *Journal of Econometrics* **70**, 261–290.
- Müller, H. (1992), 'Change-points in nonparametric regression analysis', *The Annals of Statistics* **20**, 737–761.
- Nieto, B. & Rodríguez, R. (2005), 'Modelos de valoración de activos condicionales: Un panorama comparativo', *Investigaciones Económicas* **29**, 33–71.
- Robinson, P. (1989), Nonparametric estimation of time varying parameters, in P. Hackl, ed., 'Statistical Analysis and Forecasting Economic Structural Change', Springer-Verlag, New York, pp. 253–264.

- Shanken, J. (1992), 'On the estimation of beta-pricing models', *The Review of Financial Studies* **5**, 1–34.
- Sharpe, W. (1964), 'Capital asset prices: A theory of market equilibrium', *Journal of Finance* **19**, 425–442.
- Spall, J. (1989), *Bayesian Analysis of Time Series and Dynamic Models*, Decker, New York.
- Wells, C. (1994), 'Variable betas on the Stockholm exchange', *Applied Economics* **4**, 74–92.