SORT 38 (1) January-June 2014, 13-42

Time-varying market beta: does the estimation methodology matter?

Belén Nieto[1], Susan Orbe[2] and Ainhoa Zarraga[3]

Abstract

This paper compares the performance of nine time-varying beta estimates taken from three different methodologies never previously compared: least-square estimators including nonparametric weights, GARCH-based estimators and Kalman filter estimators. The analysis is applied to the Mexican stock market (2003-2009) because of the high dispersion in betas. The comparison between estimators relies on their financial applications: asset pricing and portfolio management. Results show that Kalman filter estimators with random coefficients outperform the others in capturing both the time series of market risk and their cross-sectional relation with mean returns, while more volatile estimators are better for diversification purposes.

MSC: 62M10, 62J15, 62G08, 91G70.

Keywords: Time-varying beta, nonparametric estimator, GARCH-based beta estimator, Kalman filter.

1. Introduction

Precise estimates for market betas are crucial in many financial applications, including asset pricing, corporate finance and risk management. From a pricing perspective, the empirical failure of the unconditional Capital Asset Pricing Model (CAPM) has led to two possible ways of relaxing restrictive assumptions under the model being considered: the first is the use of an intertemporal framework, as in Merton (1973), that implies multiple sources of systematic risk. The ad-hoc three-factor model of

[1] Department of Finance. University of Alicante. Ap. 99. E-03080 Alicante. Spain. belen.nieto@ua.es

[2] Department of Applied Economics III, BETS. University of the Basque Country (UPV/EHU). Avda. Lehendakari Aguirre, 83. 48015 Bilbao. Spain. susan.orbe@ehu.es

[3] Department of Applied Economics III, BETS. University of the Basque Country (UPV/EHU). Avda. Lehendakari Aguirre, 83. 48015 Bilbao. Spain. ainhoa.zarraga@ehu.es
Received: November 2012
Accepted: September 2013

Fama and French (1993) and the four-factor model of Carhart (1997) are successful examples of multifactor models. The second is to eliminate the static context in the relationship between expected return and risk by allowing time variation in both factors and loadings. In that sense, Jagannathan and Wang (1996), Lettau and Ludvigson (2001) and Petkova and Zhang (2005) find that betas of assets with different characteristics move differently over the business cycle and Campbell and Vuolteenaho (2004), Fama and French (1997) and Ferson and Harvey (1999) show that time-variation in betas helps to explain anomalies such as value, industry and size. However, this conditional time-varying framework does not seem to be enough to improve the weak fit of the CAPM, as shown by Lewellen and Nagel (2006). The main problem in beta dynamics literature is that the investor's set of conditioning information is unobservable and consequently some assumptions have to be made. There are two main alternatives: making assumptions about the dynamics of the betas and making assumptions about the conditional covariance matrix of the returns.

For the first alternative, many different structures have been considered. There are studies that estimate the dynamics of betas by Kalman filter assuming standard stochastic processes such as random walk, autoregressive, mean reverting and switching models driving those dynamics. Some examples can be found in Wells (1994), Moonis and Shah (2003) and Mergner and Bulla (2008). Other studies use parametric approaches based on Shanken (1990), in which betas are modelled as a function of state variables or firm characteristics as in Jagannathan and Wang (1996) and in Lettau and Ludvigson (2001). A nonparametric version of this approach can be found in Ferreira, Gil and Orbe (2011). Betas have also been assumed as a function of time, with both linear and parabolic functional forms, as in Lin, Chen and Boot (1992) and Lin and Lin (2000). Nonetheless neither empirical estimation nor simulation results can produce a clear conclusion about the best way to model betas. If no parametric functions are specified and no additional conditions are assumed except that betas vary smoothly over time, then the seminal work of Fama and MacBeth (1973) suggests the use of a rolling window ordinary least squares (OLS) estimation of the market model. This data-driven approach has the advantage of no parameterization but requires prior selection of the window length. More recently, other estimators from the family of rolling least squares have been considered. In this sense, based on the nonparametric time-varying estimator proposed by Robinson (1989), time-varying conditional betas have been nonparametrically estimated by Esteban and Orbe (2010), Li and Yang (2011) and Ang and Kristensen (2012) assuming that betas vary smoothly over time and possibly nonlinearly. The flexibility of this nonparametric setting avoids the problem of misspecification derived from selecting a functional form but it also requires that window length be selected.

The second alternative, consisting of making assumptions about the conditional covariance matrix of the returns, relies on the parametric approach of ARCH-class models. In this context the assumptions under multivariate GARCH (MGARCH) models make it possible to estimate time-varying betas. In fact, the transmission of volatility between assets is captured by a time-varying conditional covariance matrix whose elements are used to calculate the beta as a ratio of covariance to variance. As the conditional covariance matrix is time dependent, the beta obtained will also be time dependent. There has been a great proliferation of multivariate models with GARCH structures in the last few decades, see Bauwens, Laurent and Rombouts (2006) or Silvennoinen and Teräsvirta (2009) for a survey. Some examples of the use of MGARCH models to estimate time-varying betas can be found in Bollerslev, Engle and Wooldridge (1988), Ng (1991), De Santis and Gérard (1998) and more recently in Choudhry (2005) and Choudhry and Wu (2008), among others.

Given the wide variety of time-varying beta estimates, some papers compare different approaches. The most common comparison is between GARCH-based estimators and Kalman filter approaches. In general, results indicate that the latter class of estimators performs better in terms of forecasting ability (Faff, Hillier and Hillier (2000) and Choudhry and Wu (2008)). However, there is no agreement about the best process assumption for beta dynamics. Moreover, when the Kalman filter is compared with estimators in the class of least squares, as in Ebner and Neumann (2005), the latter outperform the former.

In this paper three different methodologies for estimating time varying betas are compared: least-squares-based estimators, including the well-known rolling window OLS and the nonparametric time-varying estimator proposed in Esteban and Orbe (2010), beta estimators based on GARCH processes for the conditional covariance matrix of returns, including also asymmetric versions, and dynamic beta estimators obtained by the Kalman filter method. The main theoretical difference between the OLS and nonparametric estimators is that the latter have guaranteed consistency if the bandwidth is optimally chosen. In practice, there is an advantage in using the nonparametric estimator since there are many data-driven window selection criteria while the OLS estimator uses the rule of the thumb. The GARCH-based beta estimator does not rely on a smoothness assumption but has the advantage of taking into account the potential conditional heteroscedasticity of the returns. Finally, the Kalman filter method, unlike the other estimators, imposes assumptions about the specific functional form of beta dynamics. To the best of our knowledge, this is the first paper to compare these different methodologies simultaneously for the specific estimation of market risk.

Specifically, the OLS, the nonparametric estimator with both a uniform and a Gaussian kernel, the bivariate BEKK (after Baba, Engle, Kraft and Kroner) and the bivariate dynamic conditional correlation (DCC) structures together with their corresponding asymmetric versions, and random walk and random coefficient structures for the dynamic of the betas under the Kalman filter estimation are considered. The analysis is applied to daily returns for the Mexican stock market between 2003 and 2009. This market is selected because of the high cross-sectional dispersion in the sensitivity of individual returns to market returns in terms of both level and variability. Thus, grouping stocks into portfolios on the basis of trading volume provides high dispersion in time series and cross-sectionally which allows the performance of the beta estimates covering very different patterns to be analysed. The sample period also contributes to the aim of the paper because it includes the recent financial and economic crisis, ensuring enough time variation in betas potentially related, in this case, to the business cycle. Finally, the data frequency selection seeks to exploit the benefits of using high-frequency data in measuring systematic risk while avoiding problems of errors in variables that stem from nonsynchronous trading effects.

A second distinctive feature of the paper is the way in which the different estimates are compared. Instead of using only standard statistical measures based on the standard errors of the estimates or on the fit of the simple market model, the accuracy of the estimators is also determined by financial criteria. Specifically, the estimators are compared in terms of their usefulness for asset pricing or portfolio management purposes. On the one hand, the CAPM fit in both time series (pricing errors) and cross-sectional (risk premia) frameworks is analysed. On the other hand, the power in achieving the next period out of the sample minimum variance portfolio based on the use of each estimate is also compared.

Interesting results are found. The time-series analysis reveals that the wide time fluctuation combined with the moderate dispersion of the Kalman filter estimate assuming a random coefficient makes this the best beta estimator for reducing the adjustment errors in both the market model and the CAPM when daily frequency returns are used. At the same time, this estimator also produces a positive and significant risk premium in the cross-sectional estimation of the CAPM with monthly frequency data. This good fit between betas and mean returns is also obtained when the two nonparametric beta estimators are used. On the other hand, for the purpose of risk diversification, beta estimators with high volatility are more appropriate. The Kalman filter with random walk estimator and the GARCH-based beta estimators do a good job of estimating the composition of the portfolio with the minimum risk.

The rest of the paper is structured as follows. Section 2 presents the estimation methodologies. Section 3 describes the data. Section 4 compares beta estimates descriptively. Section 5 provides the empirical results for the comparison of the beta estimators in two frameworks: asset pricing and mean-variance portfolio analysis. Section 6 concludes and the Appendix contains the data information.

2. Methodology

The Capital Asset Pricing Model due to Sharpe (1964) and Lintner (1965) relates the expected return on an asset to its systematic market risk or beta. This beta is the sensitivity of the asset return to changes in the return on the market portfolio. That is, the beta is the slope of the market model:

$$R_{it} = \alpha_i + \beta_i R_{mt} + u_{it}, \qquad i = 1, \ldots, N, \quad t = 1, \ldots, T, \tag{1}$$

where R_{it} and R_{mt} are the return on asset or portfolio i and on the market portfolio at time t, respectively. Commonly, the unknown coefficients in (1) are estimated by OLS applied to the linear regression for each portfolio.

If it is assumed that these coefficients vary with time, model (1) must be rewritten as:

$$R_{it} = \alpha_{it} + \beta_{it} R_{mt} + u_{it}, \qquad i = 1, \ldots, N, \quad t = 1, \ldots, T \tag{2}$$

2.1. Least-squares-based time-varying beta estimators

As proposed by Fama and MacBeth (1973), one simple way to obtain time series estimates of betas is by a rolling OLS estimation of the market model. This consists of minimising a local sum of squared residuals for each portfolio i:

$$\min_{(\alpha_{it}, \beta_{it})} \sum_{j=t-1}^{t-r} (R_{ij} - \alpha_{it} - \beta_{it} R_{mj})^2, \tag{3}$$

where r indicates the amount of past observations to be considered at each estimation point. From the first order conditions of the optimisation problem (3), the rolling OLS estimator is obtained as:

$$(\hat{\alpha}_{it} \;\; \hat{\beta}_{it})'_{ROLL} = \left(\sum_{j=t-1}^{t-r} \mathbf{X}_j \mathbf{X}'_j \right)^{-1} \sum_{j=t-1}^{t-r} \mathbf{X}_j R_{ij}, \qquad i = 1, \ldots, N,$$

where $\mathbf{X}_j = (1 \;\; R_{mj})'$ is the jth observation of the data matrix, the subscript *ROLL* denotes the OLS rolling estimator and $'$ denotes matrix and vector transpose.

In the empirical application of this estimator, a window of 120 observations for data with daily frequency is used. The sampling frequency is selected based on the findings of Bollerslev and Zhang (2003) and Ghysels and Jacquier (2006), who show that high-frequency data result in a more effective measure of betas than the commonly used monthly returns. Since, in general, stocks in the Mexican market are not continuously traded, intraday data are discarded in order to avoid nonsynchronicity effects on beta estimates. Regarding the window length, an alternative number of observations was also considered but it did not alter the main conclusions of the paper.[1]

The nonparametric time-varying beta estimator can be considered within the family of rolling least-squares estimators. It relies on the assumption that the unknown time-varying coefficients, α_{it} and β_{it}, are smooth functions (linear or nonlinear) of the time

1. Specifically, windows of 90 and 400 days were analysed. Results are available upon request.

index. It is derived from minimising a smoothed sum of squared residuals for a given portfolio i and for a pre-selected smoothness degree h_i:

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

where $K_{h_i,tj} = h_i^{-1} K((t/T - j/T)/h_i)$ is a weight function and $K(\cdot)$ is a symmetric second order kernel. The shape of this kernel determines how past observations are to be weighted. If a uniform kernel is used all past observations selected are equally weighted but if the Epanechnikov or Gaussian kernels are used, higher weights are given to those observations closer to the estimation time point and lower weights to those farther away in time. The parameter h_i is the bandwidth that controls the amount of smoothness imposed on the coefficients associated with the ith portfolio. Solving the first-order conditions, the estimator has the following expression:

$$(\hat{\alpha}_{it} \;\; \hat{\beta}_{it})'_{NP} = \left(\sum_{j=t-1}^{t-Th_i} K_{h_i,tj} \mathbf{X}_j \mathbf{X}'_j\right)^{-1} \sum_{j=t-1}^{t-Th_i} K_{h_i,tj} \mathbf{X}_j R_{ij}, \qquad i = 1, \ldots, N,$$

where all elements are already defined and the subscript NP indicates the nonparametric estimator.

Once the smoothness degree h_i is set, the estimator obtained is consistent with the standard rate of convergence in nonparametric settings and has a closed form. Since the role of the bandwidth is to determine the amount of smoothness imposed on the betas and therefore the number of relevant past observations to be taken into account when estimating those betas, it is crucial to select it adequately in advance. If the bandwidth is large, the sub-sample of significantly weighted observations is larger, that is, more past observations are considered relevant in each local estimation. This results in a time series of estimated betas with little variability due to the high degree of smoothness. But if the bandwidth is small the estimation sub-sample is narrowed and the estimated betas have more dispersion. Different bandwidths (h_i) are allowed for the portfolios in order to capture different possible variations and curvatures of the betas. In consequence, the sub-sample size used at any estimation time point is the same when estimating the betas for a given portfolio but may be different for betas from another portfolio.

In regard to the practical issues of choosing the kernel and the bandwidths, it is well known that all kernels are asymptotically equivalent but that this is not the case for the bandwidth value. An optimal bandwidth is such that it minimises an error criterion in order to reach a tradeoff between the squared bias and the variance of the beta estimator. In the context of conditional factor models Ang and Kristensen (2012) and Li and Yang (2011) propose a bandwidth selection criterion for two-sided kernels, considering symmetric sub-samples that take into account not only past observations but also future observations. In this paper, only past observations are taken into account for estimating conditional betas and the data-driven method considered for selecting the bandwidths simultaneously is based on the proposal of Esteban and Orbe (2010), where the sum of squared residuals for all regressions is minimised together in order to take into account any possible relationships between portfolios.

Finally, note that this nonparametric estimator generalises the rolling OLS estimator since it can be derived as a particular case. If a uniform kernel that weights past observations equally is considered and $h_i = r/T$ is imposed instead of the smoothness degree being selected via a data-driven method, then the estimations obtained by the two estimators match.

2.2. The time-varying beta estimator based on multivariate GARCH models

The literature on financial econometric volatility has provided evidence of fluctuations and high persistence in conditional variance of asset returns and conditional covariance with the market return. Since market betas are ratios of estimated conditional covariances and variances, $\hat{\beta}_{it} = \widehat{cov}_t(R_i, R_m)/\widehat{var}_t(R_m)$, if these second moments are adequately estimated by an MGARCH, then betas are also expected to be accurate estimators.

The estimation procedure for MGARCH models involves maximising the following log-likelihood function for each portfolio i:

$$lnL(\boldsymbol{\theta}_i) = -\frac{1}{2}\sum_{t=1}^{T} ln|\mathbf{H}_{it}| - \frac{1}{2}\sum_{t=1}^{T} \mathbf{y}_{it}'\mathbf{H}_{it}^{-1}\mathbf{y}_{it},$$

where $\boldsymbol{\theta}_i$ is the vector of parameters to be estimated and $\mathbf{y}_{it} = (R_{it}\ R_{mt})'$ is the vector of dependent variables in the mean equation, expressed as $\mathbf{y}_{it} = \boldsymbol{\delta}_i + \boldsymbol{\varepsilon}_{it}$. $\boldsymbol{\delta}_i = (\delta_{i1}\ \delta_{i2})'$ is a bivariate vector of constants and $\boldsymbol{\varepsilon}_{it}$ is a bivariate vector given by $\boldsymbol{\varepsilon}_{it} = \boldsymbol{\mu}_{it}\mathbf{H}_{it}^{1/2}$, where $\boldsymbol{\mu}_{it}$ is a bidimensional i.i.d. normally distributed process with mean zero and identity covariance matrix. The specification of the conditional covariance matrix ($\mathbf{H}_{it}$) depends on the MGARCH structure considered.

This analysis considers two different MGARCH structures widely used in financial literature: BEKK and DCC. The former is the bivariate BEKK (1,1,1) due to Engle and Kroner (1995), which has the advantage that the positive-definite constraint of the conditional covariance matrix is guaranteed by construction. This matrix takes the form:

$$\mathbf{H}_{it} = \mathbf{C}_i'\mathbf{C}_i + \mathbf{A}_i'\boldsymbol{\varepsilon}_{it-1}\boldsymbol{\varepsilon}_{it-1}'\mathbf{A}_i + \mathbf{B}_i'\mathbf{H}_{it-1}\mathbf{B}_i, \tag{4}$$

where $\mathbf{C}_i$ is a (2×2) lower triangular coefficient matrix and $\mathbf{A}_i$ and $\mathbf{B}_i$ are (2×2) coefficient matrices. The latter, DCC, is the bivariate dynamic conditional correlation specification proposed by Engle (2002), where the conditional covariance matrix is decomposed into time-varying correlations and conditional standard deviations, ie.:

$$\mathbf{H}_{it} = \mathbf{D}_{it}\mathbf{R}_{it}\mathbf{D}_{it},$$

where $\mathbf{D}_{it}$ is a (2×2) diagonal matrix containing the conditional standard deviation of each process ε_{it}, obtained from univariate GARCH(1,1) models, $\sigma^2_{it} = \alpha_{i0} + \alpha_{i1}\varepsilon^2_{it-1} + \alpha_{i2}\sigma^2_{it-1}$, and the conditional correlation matrix can be written as:

$$\mathbf{R}_{it} = diag(q_{i11,t}^{-1/2}\, q_{i22,t}^{-1/2})\,\mathbf{Q}_{it}\, diag(q_{i11,t}^{-1/2}\, q_{i22,t}^{-1/2})$$

The (2×2) matrix $\mathbf{Q}_{it} = (q_{ijk,t})$ is given by:

$$\mathbf{Q}_{it} = \mathbf{S}_i(1 - \phi_{i1} - \phi_{i2}) + \phi_{i1}(\boldsymbol{\mu}_{it-1}\boldsymbol{\mu}'_{it-1}) + \phi_{i2}\mathbf{Q}_{it-1},$$

where $\mathbf{S}_i$ is the unconditional correlation matrix of $\boldsymbol{\mu}_{it}$.

The empirical evidence that negative shocks have a larger effect on the volatility of returns than positive shocks is also taken into account in this paper by the estimation of the asymmetric versions of the BEKK and DCC models, denoted by BEKK-A and DCC-A, respectively. In the case of the BEKK-A, the conditional covariance matrix is that of the BEKK model, equation (4), with the following term added:

$$\mathbf{E}'_i \boldsymbol{\nu}_{it-1}\boldsymbol{\nu}'_{it-1}\mathbf{E}_i,$$

where $\boldsymbol{\nu}_{it-1} = \boldsymbol{\varepsilon}_{it-1} \odot I_{\varepsilon-1}$, $\odot$ denotes the Hadamard product, $\mathbf{E}_i$ is a (2×2) coefficient matrix, and $I_{\varepsilon-1}$ is an indicator function which takes a value of one for negative residuals, ε_{t-1}, and zero otherwise. In the case of the DCC-A model, the term $e_i\varepsilon^2_{it-1}I_{t-1}$, with e_i being a coefficient, is added to each of the univariate GARCH(1,1) models that govern the variance of ε_{it}.

Once the conditional covariance matrix is estimated, the time-varying GARCH based beta for portfolio i is calculated as $\hat{\boldsymbol{\beta}}^l_{it} = \hat{\mathbf{H}}^l_{i12t}/\hat{\mathbf{H}}^l_{i22t}$, where $\hat{\mathbf{H}}^l_{i12t}$ is the estimated conditional covariance between the ith portfolio returns and the market returns and $\hat{\mathbf{H}}^l_{i22t}$ is the estimated conditional variance of the market return for l = BEKK, DCC, BEKK-A, DCC-A conditional covariance matrix structures.

2.3. The Kalman filter time-varying beta estimator

The state-space representation of the market model as in equation (2) enables time-varying coefficients to be estimated through the Kalman filter. The measurement equation is the market model and the transition equations that complete the state-space representation determine the changes in the coefficients over time. Therefore, some assumptions about the stochastic behaviour of the conditional betas are needed. Two of the most widely used characterisations of the dynamics of betas are used: the random walk (KF-RW) and the random coefficient (KF-RC).

Under the random walk assumption, betas vary smoothly so their current value is determined by their own previous value plus an error term: $\alpha_{it} = \alpha_{it-1} + \eta_{1it}$ and $\beta_{it} = \beta_{it-1} + \eta_{2it}$. Large variances in the error terms indicate that there is no persistence, so the current beta may be completely different from the previous one. As the variance of the error term of the transition equations decreases less variability is allowed and the betas become more stable. When the variances tend to zero constant betas are obtained.

In the random coefficient model, betas are assumed to vary randomly around a fixed value with some variance: $\alpha_{it} = \alpha_i + \eta_{1it}$ and $\beta_{it} = \beta_i + \eta_{2it}$. The smaller the variance in the error terms of the transition equations, the lower the variations in the betas are, and when the variances tend to zero constant betas are also obtained. As the variances increase more jumps are permitted. In contrast to the random walk, level shifts are not allowed.

The Kalman filter estimation method requires a distribution to be assumed for all stochastic terms in the measurement and transition equations. All errors $(u_{it}, \eta_{1it}, \eta_{2it})$ involved in the estimation process are assumed to be normally distributed with zero mean, to have constant variance and to be uncorrelated from one another. In order to overcome the practical issues of selecting initial parameters, the OLS estimate for each portfolio using the whole sample is chosen for the initial value of the coefficients. Moreover, large enough variances in the error terms in the transition equations are allowed.

3. Data

This analysis uses daily logarithms of returns on 42 stocks traded on the Mexican Stock Exchange between January 2, 2003 and December 31, 2009. The data series have been computed from close daily prices taking into account dividends and splits. The sample is selected on the basis of representative criteria in terms of both market capitalisation and trading volume. The sample basically coincides with the 35 firms included in the reference index, "Índice de Precios y Cotizaciones" (IPC, hereafter). As the composition of this market index is revised annually, this gives a total of 42 firms in the sample period. The proxy for the risk-free asset is the 28-day maturity Treasury Certificate (TC) and data for this proxy are collected from the Banco de Mexico.

To show the representativeness of the selected sample, the table in the Appendix provides the names of the firms selected, their industrial classifications and the percentage of the total trading volume in pesos on the Mexican Stock Exchange at the end of 2009 accounted for by each stock. At that time the market comprised stocks issued by 85 firms, five of which were non domestic companies. Although the sample only contains half of the firms extant, it accounts for 95% of the market in terms of trading volume in pesos in 2009, as can be seen by adding the weights in the last column of the table in the Appendix.[2] Moreover, the firms selected represent all the different industrial categories.

The individual stocks are sorted and grouped into portfolios. Since the aim of this work is to analyse the appropriateness of alternative beta estimators, it is important that the portfolios in the final sample produce different beta patterns. In that sense, individual betas could be used for sorting and locating stocks in portfolios. However, this would imply, on the one hand, selecting a beta estimation methodology first to analyse the appropriateness of each estimator. On the other hand, in subsequent sections asset pricing tests are used for comparing beta estimators and the results would be subject to the concerns raised by Lewellen, Nagel and Shanken (2010). This is why stocks have been sorted by individual money trading volume. This sorting produces sufficiently different portfolio betas in terms of both level and volatility. The composition of the portfolios is updated monthly by using the volume in pesos of the total trades for each stock during the month and the return of the portfolio is computed daily as the equally weighted average of the returns on stocks in the portfolio. Thus, Portfolio 1 contains the least liquid stocks while the most frequently traded stocks are in Portfolio 6.[3]

Table 1 reports the summary statistics for the returns on the six portfolios, on the market index and on the risk-free asset covering the whole sample period. The mean and the standard deviation are expressed on an annual basis. The beta estimator for each portfolio and its standard error come from the OLS estimation of the market model using the full sample period. Finally, the last row reports the average in time and across stocks within each portfolio of the monthly trading volume in millions of pesos. As can be seen, major differences in trading volume are observed; Portfolio 6 concentrates a large part of the market trading and its stocks have a trading volume 70 times greater than those of Portfolio 1. These liquidity differences do not imply differences in portfolio return volatilities, since standard deviation is similar for all six portfolios, but curiously they do produce increasing mean returns ranging from 14% for Portfolio 1 to 29% for Portfolio 6. Thus, this market seems to not show an illiquidity premium. More importantly, betas are monotonously increasing from Portfolio 1 to Portfolio 6 and also have different levels of standard errors. Therefore, the portfolio formation criterion produces the desirable dispersion in portfolio betas. The distribution of the returns is negatively skewed for the risk-free asset and all portfolios except the fifth and the market index, for which the returns' distributions are symmetric at the 5% significance level. Regarding the kurtosis coefficient, there is a significant positive excess of kurtosis for all cases except for the risk-free asset, for which the coefficient is negative.

2. The same calculation using trading volumes for other years in the sample period gives similar percentages of representativeness.

3. The classification has also been drawn up using trading volume in terms of number of shares and the characteristics of the resulting portfolios are very similar.

Table 1: *Summary Statistics of Returns.*

	Port. 1	Port. 2	Port. 3	Port. 4	Port. 5	Port. 6	IPC	TC
Mean	0.1401	0.2062	0.3263	0.2053	0.2404	0.2914	0.2367	0.0498
Standard deviation	0.2634	0.2182	0.2370	0.2662	0.2579	0.2719	0.2310	0.0006
Skewness	−0.5419	−0.5357	−0.1317	−1.4389	−0.0451	−0.1264	0.1024	−0.2892
Excess Kurtosis	5.2145	3.6191	7.5339	29.5636	5.6640	5.0603	5.3426	−0.4961
Beta	0.6523	0.6949	0.7958	0.9027	0.9667	1.1059		
Standard error	0.0223	0.0152	0.0154	0.0171	0.0133	0.0096		
Volume (millions)	119.87	298.96	497.89	817.24	1682.89	8280.50		

4. Conditional beta estimates

In this section descriptive statistics regarding the nine time series beta estimates obtained by the three methodologies considered are presented and compared. Rolling window OLS is obtained with subsamples of 120 previous observations for all portfolios and denoted by ROLL. The nonparametric estimator uses two alternative kernels: the uniform (NP-U) and the Gaussian (NP-G). The selected bandwidth is 0.1279 for Portfolios 1, 2, 3 and 6 and 0.0896 for Portfolios 4 and 5 when the uniform kernel is used, while for the Gaussian kernel the selected bandwidth is 0.0591 for all portfolios except the fifth, for which is 0.0398. Therefore, although bandwidths are allowed to vary with portfolios, the data-driven values selected indicate that betas have the same smoothness degree for most portfolios and hence the number of relevant past observations is the same. The alternative GARCH specifications produce time series of beta estimates that are denoted as BEKK and DCC for the symmetric versions of the BEKK and DCC models, respectively, and BEKK-A and DCC-A for the corresponding asymmetric versions. This estimation method does not weight the observations according to their temporal neighbourhood but according to the conditional heteroscedasticity structure. Finally, the Kalman filter method is applied with the assumption that the betas follow two alternative stochastic processes: random walk (KF-RW) and random coefficient (KF-RC). In the GARCH and Kalman filter context the total sample information is used, so that series of 1764 daily betas are produced. However, in order to provide a homogeneous, comparable context, the sample of beta estimates is restricted to the period between 17th October, 2003 and 31st December, 2009, with a total of 1564 daily beta estimates for each estimator.

Table 2 presents the mean and the standard deviation of the time series of estimated betas for each portfolio and for all the options considered. The general conclusion is that all estimation methods produce conditional beta series that have very similar mean values, smaller than the point beta estimate from the full sample (see Table 1). If there is any point worth commenting on, it is that KF-RC produces slightly higher mean betas for four out of the six portfolios. Differences between estimates are in standard deviations. The least volatile estimates are NP-U (for three portfolios) and KF-RC (for two portfolios, including the portfolio 1) while the most volatile estimates are KF-RW (for three portfolios) and BEKK (for two portfolios). The major difference in the volatility pattern of the two Kalman filter beta estimates is therefore noteworthy. Some differences are also found between the statistics of the symmetric and asymmetric GARCH estimates. For example, BEKK-A and DCC-A estimates have slightly higher means than the corresponding symmetric versions and BEKK-A estimates are less volatile than BEKK estimates while DCC-A estimates are more volatile than DCC ones.

The results in Table 2 are confirmed in Figure 1, which shows the time series beta estimates for the two extreme portfolios. Subfigures (a.1) and (a.2) compare ROLL and the two NP estimates, Subfigures (b.1) and (b.2) compare GARCH based estimates, and Subfigures (c.1) and (c.2) compare the beta estimates based on the Kalman filter methodology. All betas move around the same long term mean, the NP methods produce the smoothest betas and changes in the short term are much more pronounced in estimates from GARCH structures and from the Kalman filter method. Subfigures (c.1) and (c.2) show the high short term fluctuation of the estimate from the Kalman filter with random coefficient contrasting with the random walk assumption. However, the rank of this short term dispersion is lower for KF-RC than for KF-RW or for GARCH-based estimates, as the standard deviations in Table 2 indicate. In addition, independently of the estimation methodology, mean betas increase and standard deviations of betas decrease, almost monotonously, from the portfolio containing the least liquid stocks to the portfolio containing the most liquid stocks. Since the most liquid stocks have the highest correlation with the market index, a beta closer to one with lower variability is expected.

Table 2: *Summary Statistics of Beta Estimates.*

Portfolio	Statistic	ROLL	NP-U	NP-G	BEKK	DCC	BEKK-A	DCC-A	KF-RW	KF-RC
1	Mean	0.6431	0.6442	0.6436	0.6321	0.6386	0.6389	0.6479	0.6403	0.6486
	Std. Dv.	0.2154	0.1573	0.1733	0.2910	0.2275	0.2316	0.2453	0.2615	0.1386
2	Mean	0.6979	0.6872	0.6915	0.7115	0.7332	0.7170	0.7389	0.7113	0.7095
	Std. Dv.	0.1379	0.0929	0.1053	0.1566	0.1340	0.1151	0.1399	0.1716	0.0983
3	Mean	0.7543	0.7591	0.7594	0.7492	0.7837	0.7597	0.7909	0.7626	0.7792
	Std. Dv.	0.1179	0.0923	0.1017	0.1366	0.1207	0.1119	0.1272	0.1549	0.0975
4	Mean	0.8273	0.8288	0.8288	0.8058	0.8489	0.8181	0.8581	0.8229	0.8635
	Std. Dv.	0.1518	0.1477	0.1386	0.1643	0.1555	0.1064	0.1634	0.1711	0.1239
5	Mean	0.9180	0.9195	0.9195	0.9108	0.9299	0.9085	0.9285	0.9184	0.9544
	Std. Dv.	0.1125	0.1073	0.1130	0.1438	0.1233	0.1026	0.1200	0.1361	0.0850
6	Mean	1.0742	1.0752	1.0753	1.0722	1.0749	1.0779	1.0795	1.0720	1.1002
	Std. Dv.	0.0671	0.0582	0.0602	0.0688	0.0804	0.0700	0.0882	0.0865	0.0625

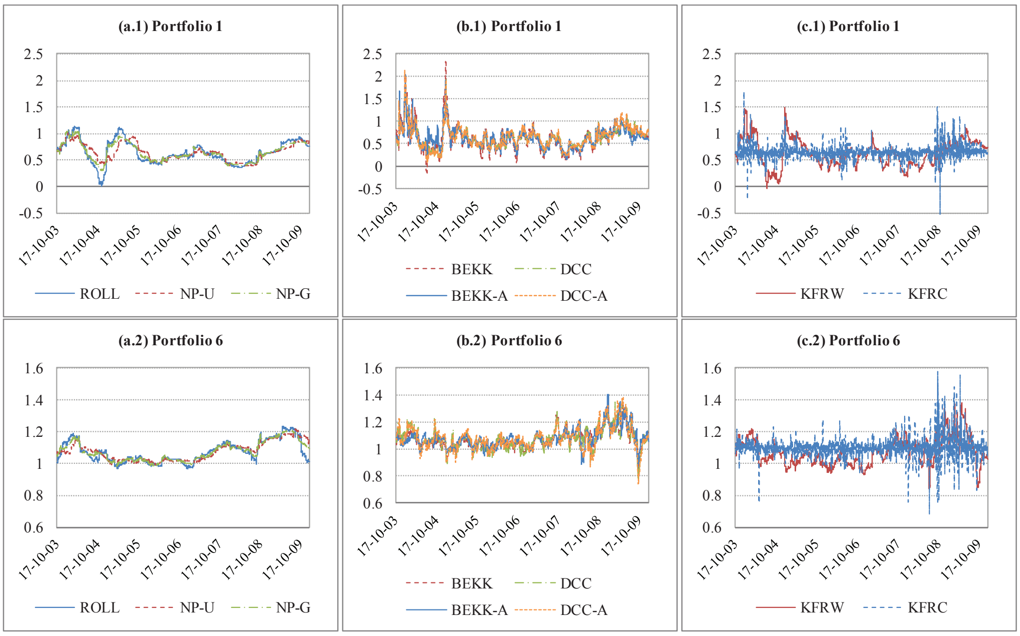

Figure 1: *Beta Estimates from Alternative Methodologies.*

A more formal comparison between the different beta estimates is carried out using the Kruskal-Wallis test. This is a non-parametric test based on ranked data that returns the p-value for the null hypothesis that two or more samples are drawn from the same population. For each portfolio, the Kruskall-Wallis test is applied to compare the different estimates all together on the one hand, and the estimates in each group on the other hand. The results are consistent for all six portfolios and indicate that the null is rejected when the nine estimates are compared simultaneously. The null is also rejected when the four GARCH based estimates are compared and when the two Kalman filter estimates are compared separately. Only for the group of least-squares-based estimates can the null not be rejected.[4]

In order to gain insight into the similarities of different time-varying beta estimates the correlations between pairs of conditional beta estimates are computed. Table 3 reports the correlations for each portfolio.[5] The results indicate that the pattern is very similar for beta estimates based on minimising some kind of least squares on the one hand, and for beta estimates from GARCH specifications on the other. However, the correlation between any of the estimated betas from each of these groups is much smaller. The different structures assumed for the beta dynamics in the Kalman filter

4. Results are available from the authors upon request.

5. Port.*i*/Port.*j* indicates that correlations for portfolio *j* are in the upper triangular panel while those for portfolio *i* are in the lower triangular panel.

method produce a lower correlation between the estimates. Moreover, the correlation between KF-RW beta estimates and those based on minimising least squares or GARCH structures are high, while the lowest correlations are those between KF-RC estimates and any other. This finding shows that the beta estimation method selected affects the resulting estimates.[6]

Table 3: *Correlations of Alternative Beta Estimates.*

	ROLL	NP-U	NP-G	BEKK	DCC	BEKK-A	DCC-A	KF-RW	KF-RC
Port. 2/Port. 1									
ROLL		0.7980	0.9628	0.4756	0.6081	0.4256	0.5783	0.7770	0.0712
NP-U	0.8213		0.8889	0.3434	0.4207	0.3194	0.4009	0.5228	0.0373
NP-G	0.9581	0.9130		0.5214	0.6395	0.4714	0.6098	0.7867	0.0694
BEKK	0.3530	0.2060	0.3670		0.9376	0.9186	0.9146	0.8242	0.1447
DCC	0.2889	0.1999	0.3073	0.6332		0.8476	0.9742	0.9118	0.1379
BEKK-A	0.3367	0.1932	0.3461	0.7234	0.8010		0.8580	0.7270	0.1568
DCC-A	0.4624	0.3214	0.4950	0.7171	0.8235	0.9069		0.8847	0.1341
KF-RW	0.7370	0.5139	0.7460	0.6349	0.4736	0.6158	0.7302		0.2548
KF-RC	0.0862	0.0345	0.0813	0.0950	0.0883	0.1223	0.1214	0.2773	
Port. 4/Port. 3									
ROLL		0.8323	0.9497	0.5399	0.5816	0.4781	0.5952	0.7330	0.0886
NP-U	0.9720		0.9279	0.4262	0.5287	0.4126	0.5713	0.6181	0.0760
NP-G	0.9686	0.9783		0.5716	0.6487	0.5174	0.6769	0.7781	0.1001
BEKK	0.5317	0.5251	0.5443		0.8092	0.9155	0.8032	0.8777	0.1427
DCC	0.5057	0.4960	0.5221	0.7944		0.7481	0.9119	0.8742	0.1244
BEKK-A	0.5329	0.5202	0.5444	0.8181	0.6872		0.7870	0.7905	0.1576
DCC-A	0.4746	0.4756	0.4981	0.7481	0.9626	0.6920		0.8397	0.1372
KF-RW	0.8193	0.7959	0.8279	0.7376	0.7652	0.6704	0.7181		0.2998
KF-RC	0.1254	0.1293	0.1265	0.1871	0.2005	0.1375	0.1858	0.3030	
Port. 6/Port. 5									
ROLL		0.9593	0.9481	0.3763	0.5037	0.1710	0.4689	0.6926	0.1094
NP-U	0.8479		0.9220	0.3526	0.4789	0.1486	0.4551	0.6634	0.1075
NP-G	0.9549	0.9481		0.4938	0.6416	0.2481	0.6130	0.8372	0.1196
BEKK	0.6086	0.5497	0.6526		0.9068	0.7574	0.8513	0.7485	0.1487
DCC	0.6181	0.5172	0.6485	0.9583		0.6078	0.9404	0.8867	0.1463
BEKK-A	0.5880	0.5646	0.6370	0.7525	0.7144		0.6077	0.4579	0.1575
DCC-A	0.5537	0.4307	0.5736	0.8740	0.9117	0.8211		0.8480	0.1523
KF-RW	0.7236	0.5766	0.7317	0.8807	0.9212	0.7320	0.8655		0.2908
KF-RC	0.1058	0.0942	0.1091	0.1352	0.1218	0.1421	0.1368	0.3031	

6. Similar results are obtained in Faff, Hillier and Hillier (2000) when comparing Kalman filter and GARCH-based beta estimators.

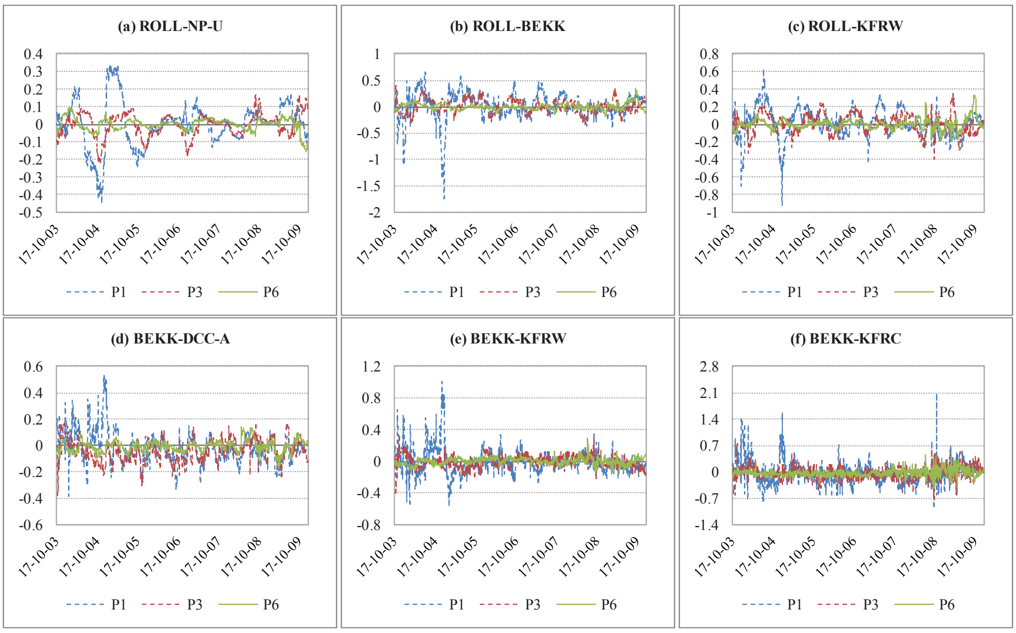

Figure 2: *Differences between Alternative Beta Estimates for Portfolios 1, 3 and 6.*

Figure 2 illustrates the high or low correlation between different estimates by showing the series of the differences between alternative pairs of beta estimates for Portfolios 1, 3 and 6.[7] As mentioned above, the volatility of beta estimates decreases and the mean increases from Portfolios 1 to 6 for all the estimation methodologies, so the largest differences are found between beta estimates in Portfolio 1. The most similar patterns correspond to the rolling and nonparametric estimates (Subfigure (a)) on the one hand, and to the BEKK and DCC-A (Subfigure (d)) on the other. However, major differences arise when beta estimates are obtained using methodologies based on different assumptions. For instance, Subfigure (f) shows that the largest difference is found when BEKK and KF-RC estimates are compared for Portfolio 1. These results are consistent with the correlation coefficients shown in Table 3; the higher the correlation between two beta estimate series the smaller the difference between them.

5. Beta estimator comparison

In this section the accuracy of the different estimators is compared in terms of the utility of time-varying beta estimates for two important financial applications: asset pricing and portfolio management.

7. The differences have been computed and plotted for all pairs of estimates and for all six portfolios. In order to save space, we only provide the most noteworthy cases.

5.1. The asset pricing perspective

This subsection analyses how systematic risk may be assessed more accurately through the use of one beta estimation methodology or another. For this purpose the simplest asset pricing framework is considered: the CAPM. It must be pointed out that this exercise does not set out to test the CAPM and that the analysis presented here could easily be extended to a multi-factor asset pricing model. However, this model offers a simple way of looking at the expected positive relationship between returns and systematic risk that any underlying investor's preferences would imply. In that sense, a beta estimate is more accurate if it is able to improve this relationship.

Next, two different settings for the comparison are considered. The first is based on time series analysis and the second on cross-section analysis.

5.1.1. Time series analysis

The first comparison between beta estimates relies on the appropriateness of the factor model representation. That is, for each portfolio the different beta estimates are compared in terms of fit for the market model. Since time-varying coefficients are estimated, R-squared statistics are not necessarily bounded and they cannot be comparable. Instead, unconditional variance ratios are studied as in Harvey, Solnik and Zhou (2002), among others. Specifically, the proportion of the unconditional variance of returns fitted by the market model, $VR1 = var(\hat{R}_i)/var(R_i)$, is used as a measure of goodness of fit, and the proportion of the unconditional variance of returns that the model fails to explain, $VR2 = var(\hat{u}_i)/var(R_i)$, as a measure of the estimation error. It must be pointed out that computing $\hat{R}_{it}$ and $\hat{u}_{it}$ requires estimates for parameter α_{it} and GARCH models do not provide them. In these cases, an estimation of α_{it} is obtained from the average of the market model where the time variation comes from each daily beta estimate:

$$\hat{\alpha}_{it}^{l} = \bar{R}_i - \hat{\beta}_{it}^{l}\bar{R}_m, \quad i = 1,\ldots,N, \quad t = 1,\ldots,T, \quad l = \text{BEKK, DCC, BEKK-A, DCC-A},$$

where $\bar{R}_i$ and $\bar{R}_m$ are the mean returns on portfolio i and on the market portfolio, respectively.

Table 4 shows the values of the $VR1$ and $VR2$ criteria for each portfolio and each estimator. The results for the two measures are very similar when the ROLL and NP estimators are compared, since they are both based on the use of rolling least squares. In general, ROLL estimates show a larger fit (larger $VR1$) but also a larger estimation error (larger $VR2$). This could be due to the bandwidth sizes selected. Since the numbers of relevant past observations selected by the data-driven method for the nonparametric estimator are smaller than for the rolling OLS, the smoothness degree imposed is lower and in consequence the estimated betas have a smaller bias but a larger variance. However, all least squares-based methods produce mostly lower values for $VR1$ and similar or higher values for $VR2$ than the rest of estimates. In general terms, according to the measure of variance explained, the market model is better adjusted when beta estimates come from GARCH structures (especially the asymmetric versions) and only slightly lower values are obtained when using the Kalman filter method with the random coefficient assumption. Moreover, the estimate that produces the lowest adjustment errors is clearly the Kalman filter with random coefficient for all portfolios. Therefore, it seems that the high daily fluctuation of the beta series from this estimation method benefits the time series adjustment of the market model.

Table 4: *Model Fit Criteria.*

Portfolio	Criteria	ROLL	NP-U	NP-G	BEKK	DCC	BEKK-A	DCC-A	KF-RW	KF-RC
1	VR1	0.3531	0.3207	0.3360	0.4107	0.3982	0.4208	0.4165	0.3900	0.4113
	VR2	0.6492	0.6542	0.6465	0.6380	0.6368	0.6310	0.6374	0.5456	0.4462
2	VR1	0.5616	0.5110	0.5329	0.6160	0.5832	0.6184	0.6166	0.5696	0.6068
	VR2	0.4464	0.4498	0.4437	0.4630	0.4592	0.4502	0.4539	0.3657	0.2923
3	VR1	0.6085	0.5730	0.5985	0.6370	0.6413	0.6332	0.6522	0.6392	0.6584
	VR2	0.3848	0.3883	0.3828	0.3725	0.3761	0.3682	0.3739	0.3163	0.2602
4	VR1	0.6383	0.6349	0.6238	0.6651	0.7241	0.6093	0.7293	0.6471	0.6592
	VR2	0.3758	0.3750	0.3740	0.3552	0.3689	0.3678	0.3726	0.3103	0.2334
5	VR1	0.7817	0.7734	0.7827	0.8059	0.7966	0.7634	0.7977	0.7734	0.7796
	VR2	0.2456	0.2451	0.2441	0.2401	0.2391	0.2366	0.2390	0.2060	0.1713
6	VR1	0.8850	0.8729	0.8836	0.8899	0.8947	0.8884	0.9038	0.8967	0.8995
	VR2	0.1109	0.1101	0.1103	0.1096	0.1111	0.1077	0.1101	0.0939	0.0739

The second comparison within this time-series framework employs Jensen's alpha as a measure of the error adjustment of the model: the difference between the observed return and the estimated return. Assuming the CAPM, the Jensen's alpha associated with each beta estimator is computed for each portfolio and period as:

$$\widehat{\alpha}_{it}^{J} = (R_{it} - R_{ft}) - \widehat{\beta}_{it}\,(R_{mt} - R_{ft}), \quad i = 1,\ldots,N, \quad t = 1,\ldots,T,$$

where R_{ft} represents the risk-free rate.

The quadratic sum of these alphas is calculated as a measure of the model misspecification, which allows a comparison to be made between different estimation methods. A large value of the quadratic sum of alphas indicates a poor specification of the model since the estimated returns differ greatly from the observed returns. Table 5 reports this measure. The bottom row shows the total sum of alphas for all portfolios. As expected, though with the exception of Portfolio 4, the quadratic sum of alphas decreases from Portfolio 1 to Portfolio 6 whatever the method used in the estimation of betas. Comparing the different estimates, the misspecification is similar for estimates based on least squares and on GARCH assumptions. The lowest values are obtained for Kalman filter methods and, specifically, for the case of random coefficient. The quadratic sum of alphas is considerably lower for KF-RC than for the other estimates and for all six portfolios. Consequently, the overall misspecification is also lower for this method.

Table 5: *Quadratic Sum of Jensen's Alphas.*

Portfolio	ROLL	NP-U	NP-G	BEKK	DCC	BEKK-A	DCC-A	KF-RW	KF-RC
1	0.2787	0.2801	0.2777	0.2740	0.2734	0.2707	0.2736	0.2478	0.1936
2	0.1381	0.1391	0.1375	0.1442	0.1428	0.1399	0.1410	0.1201	0.0920
3	0.1406	0.1415	0.1400	0.1370	0.1382	0.1354	0.1373	0.1229	0.0970
4	0.1775	0.1771	0.1774	0.1696	0.1763	0.1756	0.1779	0.1557	0.1124
5	0.1074	0.1074	0.1068	0.1063	0.1058	0.1047	0.1057	0.0957	0.0765
6	0.0542	0.0541	0.0540	0.0542	0.0549	0.0532	0.0544	0.0485	0.0369
Sum	0.8964	0.8993	0.8934	0.8854	0.8913	0.8795	0.8899	0.7907	0.6084

In order to learn whether the differences observed in Table 5 are relevant, a pairwise comparison of Jensen's alphas, in absolute values, associated with two beta estimators is conducted using the Wilcoxon signed rank test. Table 6 reports the median difference between the two series of alphas expressed on an annual basis. Each panel refers to a different portfolio and reports the median difference between the absolute value of alphas from the beta estimate indicated in the first column and the absolute value of alphas from the beta estimate indicated in the first row. For example, in the comparison of ROLL and NP-U for Portfolio 1, -0.0032 indicates that the pricing error is 0.32% lower when the ROLL beta estimate is used. Asterisks indicate that the null that this median difference is zero is rejected. Again, consistent with the evidence in Table 5, the results indicate that lower Jensen's alphas are obtained when betas are estimated by the Kalman filter for all six portfolios. And these time series errors are still lower when the random coefficient structure is assumed. Finally, although not for all the portfolios, some degree of relevance of the asymmetric BEKK estimator is shown when it is compared to the OLS or the non-parametric beta estimators.

Therefore, Tables 4, 5 and 6 provide a consistent conclusion: the lowest adjustment errors for both the market model and the CAPM are obtained when betas are estimated by the Kalman filter and a random coefficient model is assumed. It seems that the variability due to the random coefficient together with the dynamics incorporated into the estimation method are able to produce accurate beta estimates from the time series perspective.

5.1.2. Cross-sectional analysis

In this subsection the estimators are compared in terms of the market risk premium implied by the different estimated betas. Under rational expectations there should be a positive relationship between expected returns and systematic risk. For this purpose, the

Table 6: *Comparison of Jensen's Alphas in Absolute Values. Median Test.*

Port. 1	NP-U	NP-G	BEKK	DCC	BEKK-A	DCC-A	KF-RW	KF-RC
ROLL	−0.0032	−0.0001	0.0075	0.0115∗	0.0114∗	0.0087∗	0.0258∗	0.0813∗
NP-U		0.0046∗	0.0075∗	0.0113∗	0.0163∗	0.0066∗	0.0270∗	0.0899∗
NP-G			0.0045	0.0072∗	0.0106	0.0054	0.0220∗	0.0799∗
BEKK				0.0033∗	0.0015	−0.0006	0.0181∗	0.0634∗
DCC					−0.0021	−0.0022∗	0.0131∗	0.0613∗
BEKK-A						0.0010	0.0146∗	0.0623∗
DCC-A							0.0185∗	0.0561∗
KF-RW								0.0439∗
Port. 2	**NP-U**	**NP-G**	**BEKK**	**DCC**	**BEKK-A**	**DCC-A**	**KF-RW**	**KF-RC**
ROLL	−0.0013	0.0007	0.0055	0.0032	0.0076	0.0040	0.0202∗	0.0578∗
NP-U		0.0023∗	0.0074	0.0031	0.0057∗	0.0014	0.0207∗	0.0623∗
NP-G			0.0048	0.0037	0.0048	0.0024	0.0167∗	0.0580∗
BEKK				−0.0020	0.0015	−0.0036	0.0118∗	0.0524∗
DCC					0.0029	−0.0009	0.0173∗	0.0520∗
BEKK-A						−0.0045	0.0113∗	0.0522∗
DCC-A							0.0170∗	0.0467∗
KF-RW								0.0397∗
Port. 3	**NP-U**	**NP-G**	**BEKK**	**DCC**	**BEKK-A**	**DCC-A**	**KF-RW**	**KF-RC**
ROLL	0.0030	0.0014∗	0.0042∗	0.0020	0.0055∗	0.0025	0.0173∗	0.0683∗
NP-U		0.0002	0.0036	0.0022	0.0055∗	0.0008	0.0180∗	0.0642∗
NP-G			0.0021	0.0007	0.0049∗	−0.0013	0.0157∗	0.0598∗
BEKK				0.0005	0.0007	−0.0019	0.0121∗	0.0541∗
DCC					0.0003	0.0000	0.0141∗	0.0534∗
BEKK-A						−0.0026	0.0077∗	0.0627∗
DCC-A							0.0188∗	0.0544∗
KF-RW								0.0408∗
Port. 4	**NP-U**	**NP-G**	**BEKK**	**DCC**	**BEKK-A**	**DCC-A**	**KF-RW**	**KF-RC**
ROLL	0.0009∗	0.0012∗	−0.0009	−0.0016	0.0017	0.0005	0.0155∗	0.0546∗
NP-U		0.0005	−0.0005	0.0008	0.0017	0.0007	0.0139∗	0.0482∗
NP-G			−0.0006	−0.0031	0.0011	0.0009	0.0135∗	0.0480∗
BEKK				0.0021	0.0014	0.0008	0.0215∗	0.0503∗
DCC					0.0006	0.0004	0.0197∗	0.0490∗
BEKK-A						−0.0022	0.0096∗	0.0543∗
DCC-A							0.0183∗	0.0421∗
KF-RW								0.0306∗

Port. 5	NP-U	NP-G	BEKK	DCC	BEKK-A	DCC-A	KF-RW	KF-RC
ROLL	0.0003	0.0012*	0.0030	0.0074	0.0083*	0.0037	0.0179*	0.0591*
NP-U		0.0002	0.0000	0.0045	0.0085*	0.0032	0.0165*	0.0574*
NP-G			0.0000	0.0049	0.0045*	0.0035	0.0172*	0.0571*
BEKK				0.0011	0.0031	0.0000	0.0144*	0.0521*
DCC					0.0040	−0.0001	0.0132*	0.0524*
BEKK-A						−0.0030	0.0060*	0.0575*
DCC-A							0.0101*	0.0534*
KF-RW								0.0404*

Port. 6	NP-U	NP-G	BEKK	DCC	BEKK-A	DCC-A	KF-RW	KF-RC
ROLL	0.0006	0.0000	−0.0037	−0.0026*	−0.0009	−0.0031	0.0038*	0.0359*
NP-U		0.0002	−0.0014	−0.0025	0.0006	−0.0010	0.0036*	0.0297*
NP-G			−0.0018	−0.0023*	−0.0008	−0.0018	0.0037*	0.0331*
BEKK				−0.0019*	0.0002	−0.0005	0.0059*	0.0339*
DCC					0.0021*	0.0005	0.0093*	0.0346*
BEKK-A						−0.0011*	0.0036*	0.0312*
DCC-A							0.0086*	0.0318*
KF-RW								0.0264*

simple CAPM framework is used, which assumes only one source of systematic risk: the market beta.

Using the Fama and MacBeth (1973) methodology, the following cross-sectional regression is estimated for each day in the sample period:

$$R_{it} - R_{ft} = \gamma_{0t} + \gamma_{1t}\hat{\beta}_{it} + e_{it}, \quad i = 1, \ldots, N, \tag{5}$$

where the beta represents one of the nine alternative estimates. A reasonable beta estimator should produce a positive and significant market risk premium and the more precise the above cross-sectional relationship is, the more accurate the beta estimator is. Additionally, since excess returns are used as dependent variable, an intercept statistically equal to zero indicates a good model fit.

The results from the Fama-MacBeth estimation of the model are presented in Table 7. This table reports the estimates of the intercept and the market risk premium ($\times 10^2$), their t-statistics for individual significance and the corresponding Shanken (1992) adjusted t-statistics. Asterisks indicate that the risk premium is significantly different from zero using both t-statistics at the 5% level. The left panel of the table shows the results when daily portfolio returns and betas are used in the estimation of (5) and one regression is run each day. The right panel provides the results when monthly returns and the beta estimator corresponding to the last day of the previous month are

Table 7: *Cross-Sectional Risk Premium Estimation.*

		Daily frequency		Monthly frequency	
		γ_0	γ_1	γ_0	γ_1
	Estimate	0.0101	0.0821	0.4088	1.6655
ROLL	t-stat.	0.181	1.251	0.320	1.502
	Adj. t-stat.	0.181	1.250	0.311	1.459
	Estimate	−0.0045	0.1012	−0.2203	2.3191∗
NP-U	t-stat.	−0.081	1.568	−0.181	2.210
	Adj. t-stat.	−0.081	1.566	−0.176	2.147
	Estimate	−0.0054	0.1045	−0.0950	2.2120∗
NP-G	t-stat.	−0.098	1.601	−0.075	1.988
	Adj. t-stat.	−0.098	1.599	−0.073	1.930
	Estimate	0.0329	0.0566	0.9704	1.1540
BEKK	t-stat.	0.615	0.891	0.847	1.196
	Adj. t-stat.	0.614	0.890	0.823	1.161
	Estimate	0.0462	0.0436	0.4642	1.6817
DCC	t-stat.	0.852	0.671	0.410	1.682
	Adj. t-stat.	0.851	0.670	0.398	1.634
	Estimate	−0.0495	0.1498∗	0.9704	1.1540
BEKK-A	t-stat.	−0.911	2.242	0.729	1.054
	Adj. t-stat.	−0.910	2.239	0.708	1.024
	Estimate	0.0400	0.0512	1.2932	0.6088
DCC-A	t-stat.	0.735	0.788	1.101	0.564
	Adj. t-stat.	0.734	0.787	1.069	0.548
	Estimate	0.0283	0.0736	0.4793	1.6217
KF-RW	t-stat.	0.571	1.097	0.429	1.782
	Adj. t-stat.	0.570	1.096	0.417	1.730
	Estimate	0.0155	0.1181	−0.2193	2.3126∗
KF-RC	t-stat.	0.314	1.528	−0.155	2.000
	Adj. t-stat.	0.314	1.526	−0.150	1.942

used to reduce the excessive noise that daily observations could introduce into this cross-sectional analysis. In this case, the number of regressions is 75, which is the number of months in the period analysed.

The intercepts are non-statistically different from zero and market risk premia are positive for all beta estimates and for the two data frequencies. However, differences in the value and significance of the risk premia are observed for different beta estimators. At daily frequency, market risk premia are not significant in general. Only for the beta estimated from the asymmetric BEKK method is there a relevant cross-sectional relationship between returns and market betas. The results for the monthly frequency are better and more conclusive. The risk premia associated with betas from GARCH structures are similar and not significant. The cross-sectional relationship clearly improves when non-parametric betas or Kalman filter betas are employed. The risk premium estimate and the t-statistic are very similar when the two NP beta estimators or the KF-RC one are used. In these three cases risk premia are significant at the 5% level.

Thus, the results of this analysis indicate that the estimation of the risk premium depends on the characteristics of the beta estimator. Specifically, the three estimators with the lowest standard deviations are the ones that produce significant risk premia in the relationship between betas and returns at monthly frequency. On the one hand, comparing the standard OLS estimator with the non-parametric estimates, the results suggest that a correct size of the window and the use of weights decaying in time matter with a view to better capturing this cross-sectional relationship. Therefore, an optimal mechanism for choosing the bandwidth is important. On the other hand, the high variability that the Kalman filter produces (but with lower dispersion than GARCH-based methods) is also a good characteristic for having betas more closely related to the cross-section of returns.

5.2. Portfolio management analysis

An important application of betas is their use in portfolio management. Since individual betas are part of the variance of a portfolio, the power of prediction of the different beta estimators can be studied by analysing whether the purpose indicated in the portfolio construction criterion is achieved in the next period.

For each of the estimation methodologies considered, betas for all six portfolios are taken in order to obtain an estimate of the next period covariance matrix, which can then be used to obtain the composition of the overall minimum variance portfolio. Thus, the beta estimators are compared by analysing the variance of the resulting portfolio.

Specifically, according to the market model, for a given month s the covariance matrix of a set of N asset returns is:

$$\mathbf{\Sigma}_s = \sigma_{ms}^2 \mathbf{B}_s \mathbf{B}_s' + \mathbf{D}_s,$$

where σ_{ms}^2 is the variance of the market return, $\mathbf{B}_s$ is an N-vector of individual betas and $\mathbf{D}_s$ is an $N \times N$ matrix of the idiosyncratic variance-covariances, all of them measured in month s. The variance of the market return is estimated using daily returns within month

s; beta estimates on the last day of month $s-1$ are used as predictors of elements of $\mathbf{B}_s$; and $\mathbf{D}_s$ is estimated as the residual covariance matrix from the market model consistent with these beta estimates employing daily returns within month s:

$$\widehat{\mathbf{D}}_s = \frac{1}{T_d}\widehat{\mathbf{U}}_s'\widehat{\mathbf{U}}_s,$$

Table 8: *Out-of-Sample Comparison for the Prediction of the Global Minimum Variance Portfolio.*

x/y	NP-U	NP-G	BEKK	DCC	BEKK-A	DCC-A	KF-RW	KF-RC
ROLL	48	70.7	76	77.3	70.7	82.7	96	44
	−0.963	2.166∗	13.736∗	12.074∗	7.343∗	12.968∗	21.758∗	−1.731
NP-U		73.3	82.7	90.7	73.3	88	96	49.3
		2.706∗	18.189∗	9.889∗	10.984∗	12.144∗	20.717∗	−0.409
NP-G			78.7	78.7	70.7	82.7	97.3	37.3
			10.860∗	8.420∗	7.666∗	10.333∗	19.290∗	−3.976∗
BEKK				42.7	33.3	45.3	62.7	17.3
				−1.384	−4.339∗	−1.467	3.160∗	−17.213∗
DCC					44	57.3	74.7	18.7
					−2.808∗	0.342	5.541∗	−15.797∗
BEKK-A						65.3	80	17.3
						3.353∗	8.141∗	−16.040∗
DCC-A							77.3	13.3
							4.362∗	−16.368∗
KF-RW								6.7
								−23.543∗

where $\widehat{\mathbf{U}}_s$ is a $T_d \times N$ matrix containing the residuals $\widehat{u}_{isd} = R_{isd} - \widehat{\alpha}_{is-1} - \widehat{\beta}_{is-1}R_{msd}$ for $i = 1,\ldots,N$, $d = 1,\ldots T_d$, where T_d is the number of days in month s and $s = 1,\ldots,S$ with S being the number of months in the sample.

The portfolio formation criterion consists of investing in the minimum variance portfolio, which implies choosing the portfolio weights ($\boldsymbol{\omega}_s$) that solve the following problem:

$$\begin{aligned} &Min\ \boldsymbol{\omega}_s' \, \boldsymbol{\Sigma}_s \, \boldsymbol{\omega}_s \\ &s.t. \quad \boldsymbol{\omega}_s' \mathbf{1} = 1 \end{aligned}$$

This optimisation problem is solved for each month and each beta estimate, then the daily return of the minimum variance portfolio is computed for all the days in the month and its variance is recorded. The most successful beta estimator should lead to portfolios with the lowest variance.

Table 8 provides the results for the comparisons of pairs of series of the variance of the minimum variance portfolio conducted via the Wilcoxon median test. For each comparison x/y, the first number is the percentage of cases in which beta estimation method x produces higher variance than beta estimation method y. Below, the median difference ($\times 10^4$) is reported and an asterisk indicates that the null of equal medians is rejected at the 5% significance level.

The results are quite conclusive: the beta estimate that produces the lowest variance for the next period minimum variance portfolio is the Kalman filter with the random walk assumption. This is the case in the 96/97% of the out-of-sample predictions when it is compared to any least-squares-based estimates and in between 62% and 80% of the predictions when it is compared to GARCH-based estimates. It is also better than the other Kalman filter estimate in 6.7% of the predictions. Moreover, the difference between medians is larger in the cases when the KF-RW estimates is one of the beta estimates in the pair. By contrast, the beta estimated from the Kalman filter with the random coefficient produces the highest variance. On the other hand, GARCH beta estimators are superior to least-squares-based estimators for the purpose of risk hedging in portfolio decisions. Finally, when ROLL and NP estimators are compared the differences in the resulting variance portfolio are not so big but NP-G is significantly better than rolling OLS with both the standard selection of the window size and the optimal window size.

6. Conclusions

This paper compares the performances of three methodologies in estimating time-varying market betas: dynamic estimators based on least squares, time-varying estimators coming from GARCH structures for the conditional variance of the errors of the market model, and Kalman filter estimators. These three methodologies have never previously been compared with one another homogenously.

Specifically, three estimators in the group of least squares are selected: a rolling window OLS and two nonparametric estimators that use uniform and Gaussian kernels, respectively. The advantage of the nonparametric estimators is that they allow the optimal window length to be chosen. In the group of GARCH-based estimators standard DCC and BEKK models and their corresponding asymmetric versions are considered. In this case the potential benefits of taking into account the returns' conditional heteroscedasticity are examined. Finally, the Kalman filter estimator considering two different specifications for the transition equations is included in the comparison: one imposing a random walk process and the other assuming a random coefficient structure for the dynamics of the beta. Therefore, nine beta estimates are obtained for each of the six portfolios of daily returns for the Mexican stock market in the period 2003-2009. All the descriptive analyses indicate that the time pattern of these nine estimates are substantially different. The distribution of the estimates shows different sample moments for different estimates, especially regarding the standard deviation, and these differences are corroborated by an analysis of the correlations between them and by using the Kruskal-Wallis test.

The accuracy of the estimates is compared under two frameworks: an asset pricing perspective that assumes the CAPM and the mean-variance space for returns for portfolio management purposes. In the first case beta estimates are compared using different measures of the time-series fit of the model and looking at the cross-sectional relationship between mean returns and market betas. In the mean-variance context, the out-of-sample forecasting power of different beta estimates is obtained by comparing the results of the minimum variance portfolio.

The time-series analysis clearly concludes that the Kalman filter estimator that assumes a random coefficient is the best at reducing the adjustment errors in both the market model and the CAPM; moreover this is true for all six portfolios analysed. This estimate has the characteristic of presenting a very high time fluctuation, as GARCH-based estimates do, but a low standard deviation, as the smoothed nonparametric estimates do. This combination seems to be the reason for the good time series adjustment in the daily frequency sample used here.

The Kalman filter with the random coefficient estimate also produces a good fit for the CAPM cross-sectionally. In this case, this estimate and the two nonparametric estimates are the ones for which the relationship between betas and returns are positive and statistically significant. The high volatility in GARCH-based beta estimates has a negative effect on the stability of the relationship between systematic risk and mean returns. Consequently, in estimating the price of risk, dynamic methodologies that produce low dispersion are more appropriate for the prior estimation of systematic risk.

However, highly volatile market betas are appropriate in terms of risk diversification. The Kalman filter with a random walk estimate and the GARCH-based beta estimates are both better than estimates with lower volatility for estimating the composition of the portfolio with the minimum risk.

Given that different conclusions are obtained depending on whether betas or risk premia are estimated, one possible improvement along these lines could be to propose a new estimator that combines the advantages of these different estimators.

Acknowledgements

The authors thank Gonzalo Rubio and two anonymous referees for helpful comments and suggestions that substantially improved the contents of the paper. Financial support is acknowledged from Ministerio de Ciencia e Innovación under research grants ECO2012-35820, ECO2011-29268 and ECO2011-29751, from Generalitat Valenciana under the grant PROMETEO II/2013/015, and from Departamento de Educación, Universidades e Investigación del Gobierno Vasco under research grants IT-783-13 and IT-793-13.

Appendix: Individual stocks data information

Ticker	Firm Name	Sector	Trading Volume (Pesos %)
AMX-L	América Móvil	Telecomunications/Services	23.22
TELMEX-L	Teléfonos de Mexico	Telecomunications/Services	3.49
TELINT-L	Telmex Internacional	Telecomunications/Services	2.09
TELECOM-A1	Carso Global Telecom	Telecomunications/Services	1.89
AXTEL-CPO	Axtel	Telecomunications/Services	1.84
TLEVISA-CPO	Grupo Televisa	Telecomunications/Radio and Television	3.33
TVAZTCA-CPO	TV Azteca	Telecomunications/Radio and Television	1.07
ICH-B	Industrias CH	Materials/Steel	0.21
SIMEC-B	Grupo Simec	Materials/Steel	0.17
GMEXICO-B	Grupo Mexico	Materials/Metals and Mining	7.65
AUTLAND-B	Compañía minera Autland	Materials/Metals and Mining	0.12
CEMEX-CPO	Cemex	Materials/Construction	4.63
MEXCHEM	Mexichem	Materials/Chemical Products	0.93
ASUR-B	Grupo Aeroportuario del Sureste	Industrials/Transportation	0.87
GAP-B	Grupo Aeroportuario del Pacífico	Industrials/Transportation	0.50
OMA-B	Grupo Aeroportuario del Centro Norte	Industrials/Transportation	0.15
GEO-B	Corporación Geo	Industrials/Construction	1.73
URBI	Urbi Desarrollos Urbanos	Industrials/Construction	1.40
HOMEX	Desarrolladora Homex	Industrials/Construction	1.39
ICA	Empresas ICA	Industrials/Construction	1.33
IDEAL-B1	Impulsora del Desarrollo y el Empleo	Industrials/Construction	1.11
ARA	Consorcio Ara	Industrials/Construction	1.10
SARE-B	Sare Holding	Industrials/Construction	0.06
ALFA-A	Alfa	Industrials/Capital Goods	1.43
GCARSO-A1	Grupo Carso	Industrials/Capital Goods	1.02
LAB-B	Genomma Lab Internacional	Health/Medicine Distrib.	1.50
BOLSA-A	Bolsa Mexicana de Valores	Financial Services/Financial Markets	0.24
GFNORTE-O	Grupo Financiero Banorte	Financial Services/Financial Groups	2.04
GFINBUR-O	Grupo Financiero Inbursa	Financial Services/Financial Groups	1.07
COMPART-O	Banco Compartamos	Financial Services/Commercial Banks	0.79
WALMEX-V	Wal-Mart de Mexico	Consumer Staples/Hypermarkets	13.22
SORIANA-B	Organización Soriana	Consumer Staples/Hypermarkets	1.01
COMERCI-UBC	Controladora Comercial Mexicana	Consumer Staples/Hypermarkets	0.07
KIMBER-A	Kimberly-Clark Mexico	Consumer Staples/Household Products	1.06
BIMBO-A	Grupo Bimbo	Consumer Staples/Food	1.00
GRUMA-B	Gruma Sab de C.V.	Consumer Staples/Food	0.51
FEMSA-UBD	Fomento Económico Mexicano	Consumer Staples/Beverages	5.82
GMODELO-C	Grupo Modelo	Consumer Staples/Beverages	1.70
ARCA	Embotelladoras Arcas	Consumer Staples/Beverages	0.54
KOF-L	Coca-cola Femsa	Consumer Staples/Beverages	0.07
ELEKTRA	Grupo Elektra	Consumer Discret./Retails	1.28
GFAMSA-A	Grupo Famsa	Consumer Discret./Retails	0.50

References

Ang, A. and Kristensen, D. (2012). Testing conditional factor models. *Journal of Financial Economics*, 106, 132–156.

Bauwens, L., Laurent, S. and Rombouts, J. V. K. (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, 21, 79–109.

Bollerslev, T., Engle, R. and Wooldridge, J. (1998). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, 96, 116–131.

Bollerslev, T. and Zhang, B.Y.B. (2003). Measuring and modeling systematic risk in factor pricing models using high-frequency data. *Journal of Empirical Finance*, 10, 533–558.

Campbell, J. and Vuolteenaho, T. (2004). Bad beta, good beta. *American Economic Review*, 94, 1249–1275.

Carhart, M.M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52, 57–82.

Choudhry, T. (2005). Time-varying beta and the Asian financial crisis: Evidence from Malaysian and Taiwanese firms. *Pacific-Basin Finance Journal*, 13, 93–118.

Choudhry, T. and Wu, H. (2008). Forecasting ability of GARCH vs Kalman filter method: evidence from daily UK time-varying beta. *Journal of Forecasting*, 27, 670–689.

De Santis, G. and Gérard, B. (1998). How big is the premium for currency risk? *Journal of Financial Economics*, 49, 375–412.

Ebner, M. and Neumann, T. (2005). Time-varying betas of German stock returns. *Financial Markets and Portfolio Management*, 19, 29–46.

Engle, R. F. (2002). Dynamic conditional correlation-a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics*, 20, 339–350.

Engle, R. F. and Kroner, F. K. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, 11, 122–150.

Esteban, M. V. and Orbe, S. (2010). A nonparametric approach for estimating betas: the smoothed rolling estimator. *Applied Economics*, 42, 1269–1279.

Faff, R. W., Hillier, D. and Hillier, J. (2000). Time varying beta risk: an analysis of alternative modelling techniques. *Journal of Business Finance and Accounting*, 27, 523–554.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics*, 33, 3–56.

Fama, E. F. and French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics*, 43, 153–193.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return and equilibrium: empirical tests. *Journal of Political Economy*, 81, 607–636.

Ferreira, E., Gil, J. and Orbe, S. (2011). Conditional beta pricing models: a nonparametric approach. *Journal of Banking and Finance*, 35, 3362–3382.

Ferson, W. E. and Harvey, C. R. (1999). Conditioning variables and the cross-section of stock returns. *Journal of Finance*, 54, 1325–1360.

Ghysels, E. and Jacquier, E. (2006). Market beta dynamics and portfolio efficiency. Working Paper, Cornell University, CREF.

Harvey, C. R., Solnik, B. and Zhou, G. (2002). What determines expected international asset returns? *Annals of Economics and Finance*, 3, 249–298.

Jagannathan, R. and Wang, Z. (1996). The conditional CAPM and cross-section of expected returns. *Journal of Finance*, 51, 3–53.

Lettau, M. and Ludvigson, S. (2001). Resurrecting the (C)CAPM: a cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, 109, 1238–1287.

Lewellen, J. and Nagel, S. (2006). The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics*, 82, 289–314.

Lewellen, J., Nagel, S. and Shanken, J. (2010). A skeptical appraisal of asset-pricing tests. *Journal of Financial Economics*, 96, 175–194.

Li, Y. and Yang, L. (2011). Testing conditional factor models: a nonparametric approach. *Journal of Empirical Finance*, 18, 975–992.

Lin, W. T., Chen, Y. H. and Boot, J. C. (1992). The dynamic and stochastic instability of betas: Implications for forecasting stock returns. *Journal of Forecasting*, 11, 517–541.

Lin, H. J. and Lin, W. T. (2000). A dynamic and stochastic beta and its implications in global capital markets. *International Finance*, 3, 123–160.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47, 261–290.

Mergner, S. and Bulla, J. (2008). Time-varying beta risk of Pan-European industry portfolios: a comparison of alternative modeling techniques. *European Journal of Finance*, 14, 771–802.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41, 867–887.

Moonis, S. A. and Shah, A. (2003). Testing for time-variation in beta in India. *Journal of Emerging Market Finance*, 2, 163–180.

Ng, L. (1991). Tests of the CAPM with time-varying covariances: a multivariate GARCH approach. *The Journal of Finance*, 46, 1507–1520.

Petkova, R. and Zhang, L. (2005). Is value riskier than growth? *Journal of Financial Economics*, 78, 187–202.

Robinson, P. (1989). Nonparametric estimation of time varying parameters. Hackl, P. ed. *Statistical Analysis and Forecasting Economic Structural Change*. New York: Springer-Verlag, 253–264.

Shanken, J. (1990). Intertemporal asset pricing: an empirical investigation. *Journal of Econometrics*, 45, 99–120.

Shanken, J. (1992). On the estimation of beta-pricing models. *Review of Financial Studies*, 5, 1–33.

Sharpe, C. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, 19, 425–442.

Silvennoinen, A. and Teräsvirta, T. (2009). Multivariate GARCH models. Andersen, T. G., Davis, R. A., Kreiβ, J. and Mikosch, T. eds. *Handbook of Financial Time Series*. Berlin: Springer-Verlag, 201–229.

Wells, C. (1994). Variable betas on the Stockholm exchange 1971–1989. *Applied Financial Economics*, 4, 75–92.