DOCUMENTOS DE TRABAJO

BILTOKI

D.T. 2009.01

Predicting Betas: Two new methods.

Mª Victoria Esteban González y Fernando Tusell Palmer

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 DT2009.01
Editado por el Departamento de Economía Aplicada III (Econometría y Estadística) de la Universidad del País Vasco.

Depósito Legal No.: BI-1267-09
ISSN: 1134-8984
URN: RePEc:ehu:biltok:200901

Predicting Betas: Two new methods

M.V. Esteban* F. Tusell†

Abstract

Betas play a central role in modern finance. The estimation of betas from historical data and their extrapolation into the future is of considerable practical interest. We propose two new methods: the first is a direct generalization of the method in Blume (1975), and the second is based on Procrustes rotation in phase space. We compare their performance with various competitors and draw some conclusions.

keyword: risk prediction, systematic risk, time varying beta coefficients, Procrustes rotation

JEL Classification: G11, G12, G17.

*Departamento de Econometría y Estadística. Facultad de CC.EE. y Empresariales, Avda. Lehendakari Aguirre, 83; 48015 Bilbao. Spain. Phone:(34) 94 601 38 47. Fax(34) 94 601 37 54. E-mail: mvictoria.esteban@ehu.es

†Departamento de Estadística y Econometría. Facultad de CC.EE. y Empresariales, Avda. Lehendakari Aguirre, 83, 48015 Bilbao. Spain. Phone:(34) 94 601 37 33, Fax(34) 94 601 37 54 E-mail: fernando.tusell@ehu.es.

1 Introduction

The beta parameter as a measure of an asset's risk plays a central role in finance. Estimated betas and their predictions are used in asset pricing, cash flow valuation, risk management, making investment decisions or simply as a risk factor in models with more than one factor.

Because of their central role in portfolio theory, betas have been the object of enormous research interest. The traditional setting is one where beta risk is assumed to be constant. In this framework, betas can be estimated as the slope coefficient in a simple regression model fitted by ordinary least squares. However, empirical evidence in numerous studies suggest that betas are not constant over time.

Blume (1971) found that historical data for "individual securities and smaller portfolios have limited value in forecasting", whereas "larger portfolios are remarkably accurate in predicting future portfolios betas". In that paper, Blume documents the existence of a tendency for the betas of even well-diversified portfolios of extreme risk levels to regress towards the mean. Blume (1975) tested the reversion to the "grand mean" of the betas over time, namely 1, and attributed the existence of the reversion to the mean to the non-stationarity of the betas.

Furthermore, evidence on non-stability of beta coefficients is not tied to a particular market or type of asset. Gooding & O'Malley (1977), Fabozzi & Francis (1978), Bos & Newbold (1984) and Collins, Ledolter & Rayburn (1987) or Lee & Chen (1982) all present evidence for USA stocks. Faff, Lee & Fry (1992), Brooks, Faff & Lee (1992), Brooks, Faff & Lee (1994) do the same for the case of Australian stocks. Evidence for the German Stock Market is presented in Abberger (2004), Ebner & Neumann (2005), and more recently in Eisenbeiß, Kauermann & Semmler (2007); and many more references could be given covering different markets and assets.

The common finding of all this literature is that betas appear to be far from constant over time. Different models and estimation techniques have been used to account for non stationarity: state space models with different choices for the stochastic process of betas, the Hildreth & Houck (1968) random-coefficient model, the random walk coefficient or AR(1) coefficient model and non-parametric regression techniques among others: see for instance Fabozzi & Francis (1978) and Collins et al. (1987).

Lee & Chen (1982) detected some regression tendency of beta coefficients over time with a model that allows for coexistence of both beta instability and beta tendency. González-Rivera (1997) found evidence against the constancy of betas in favor of a random coefficient model, and concluded that the time variation of betas is due to non-systematic behavior of the firms.

Brooks et al. (1992), Brooks et al. (1994), Faff et al. (1992), Clare, Priestley & Thomas (1997) and Brooks, Faff, Gangemi & Lee (1997*b*) discuss different types of models for time varying betas. Abberger (2004) and Eisenbeiß et al. (2007), among others, use non-parametric estimation techniques to account for non-stationarity in betas of individual stocks. They treat beta coefficients as unspecified functions of time.

In the following we will consider estimation of betas, both using past and contemporaneous information (filtering) or just past information (prediction). However, betas are not directly observable. In order to test the effectiveness of the filtering and prediction methods discussed, we compute "target betas" in various ways, which use information in the future. Clearly, the performance of the filtering and prediction methods will be relative to the target betas used. For instance, very smooth target betas favour prediction methods with greater inertia. Since filtering and prediction methods have to be judged in association with the method used to estimate the target betas, we provide results for each combination.

The remaining of this paper is organized as follows. Section 2 provides a quick review of some methods proposed so far for the estimation of betas. Section 3 discusses the methods proposed in Vasicek (1973) and Blume (1975) to adjust (filter) and predict betas. It is against the background of this last paper that we motivate in Section 4 two new methods. Both can be seen as trying to endow Blume's method with a longer memory. Aside from its possible use in filtering or predicting betas, one of the methods generates as by-products several interesting descriptive measures, which provide insight on the behaviour of betas. In Section 5 we present our empirical results and Section 6 ends with some conclusions.

2 Estimating time varying betas

Sharpe (1964) and Lintner (1965) derive the Capital Asset Pricing Model (CAPM) assuming the existence of lending and borrowing at a risk-free rate of interest. The model implies that the expected return $E(R_i)$ of risky asset i (a single stock or a portfolio) must be linearly related to the covariance between its return and the return of the market portfolio and is usually expressed by the equation

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

where R is the return on the market portfolio and r_f is the return on the risk-free asset. The beta coefficient, β_i, is the coefficient of systematic risk of asset i expressed as the ratio of the covariance between its return and the return of the market portfolio and the market variance,

$$\beta_i = \frac{Cov(R_i, R)}{Var(R)}. \tag{2}$$

The Sharpe-Lintner CAPM can be expressed in "excess returns" as

$$E(Z_i) = \beta_i[E(Z)] \tag{3}$$

$$\beta_i = \frac{Cov(Z_i, Z)}{Var(Z)}, \tag{4}$$

where $Z_i = R_i - r_f$ is return of asset i in excess of the risk-free rate and Z is the excess return on the market portfolio of assets.

Usually, the beta is estimated by the slope in the market model,

$$Z_{it} = \alpha_i + \beta_i Z_t + \epsilon_{it}; \quad t = 1, \ldots, T. \tag{5}$$

The equation (5) suggests that a plausible estimate of β_i might be the estimated coefficient of a regression of excess asset returns on excess market returns, with or without intercept. There has been considerable evidence, though, that the beta stability assumption implied by (3) and (5) must be rejected and some scheme to accommodate time-varying betas should be adopted instead: see for instance Garbade & Rentzler (1981), Dotan & Ofer (1984), Elton & Gruber (1995), and references therein, in addition to papers cited in the Introduction. Attempts have been made to link beta's changes to seasonal effects, Brooks, Faff & Josev (1997*a*), market phase, Gooding & O'Malley (1977), and interest rates, Bildersee & Roberts (1981), among other factors.

Traditionally, a "rolling regression" beta estimator has been used, Fama & MacBeth (1973); the regression suggested by (5) is carried on overlapping sets of contiguous observations. The estimate of β_{it} (the i-th asset beta at time t) is thus given by the value $\hat{\beta}_{it}$ minimizing

$$\min_{\alpha_{it}, \beta_{it}} \sum_{j \in [t-q, t]} (Z_{ij} - \alpha_{it} - \beta_{it} Z_j)^2. \tag{6}$$

The length q of the block of contiguous observations used plays a fundamental role: the larger q, the smoother the variation of the $\hat{\beta}_{it}$. Traditionally, $q = 60$ (five years of monthly data) have been used[1].

Variations of the rolling approach exist in which (6) becomes a weighted regression problem:

$$\min_{\alpha_{it}, \beta_{it}} \sum_{j} K_{t,j} (Z_{ij} - \alpha_{it} - \beta_{it} Z_{mj})^2. \tag{7}$$

In (7), $K_{t,j}$ is a weight or kernel function, whose bandwidth (and therefore the implied smoothing in the estimated betas) can be chosen by a data driven method (see for instance Esteban & Orbe (to appear)).

The rolling regression approach is by no means the only possibility. As we are dealing with unobservable time-varying magnitudes, state-space models suggest themselves as an alternative. We can model $\{\beta_{it}\}$, $t = 1, T$ as a local level or local linear trend model, Harvey (1989), Durbin & Koopman (2001). Under the local level specification, $\{\beta_{it}\}$ follows the simple dynamics:

$$\beta_{it} = \beta_{i,t-1} + \eta_{it}. \tag{8}$$

[1] Fisher (1970) and Gonedes (1973) among other authors have found empirically that with monthly data the optimal subsample is between four and seven years: but there is no universal agreement on this issue, see also Meyers (1973) and Baesel (1974). Five years is a common choice.

The excess return of asset i at time t, Z_{it}, is:

$$Z_{it} = Z_t \beta_{it} + \epsilon_{it}. \tag{9}$$

Taken together, (8)–(9) define a state-space model, whose parameters (σ^2_η and σ^2_ϵ) can be estimated by maximum likelihood (assuming normality of the noises). The Kalman filter or smoother gives then estimates of the state β_{it} for each t, conditional on observations available at time $t-1$ (one step ahead prediction), at time t (filtering) or conditional on all observations, past and future (smoothing).

Other methods can also be used: Eisenbeiß et al. (2007) propose to minimize

$$\min_{\alpha_i, \beta_i, \tilde{\beta}_i(t)} \sum_t \left(Z_{it} - \alpha_i - \left[\beta_i + \tilde{\beta}_i(t) \right] Z_t \right)^2, \tag{10}$$

where $\tilde{\beta}_i(t)$ is an "smooth" function of time constrained by $\sum_t \tilde{\beta}_i(t) = 0$ (to preserve identifiability). The "total", time-varying, beta coefficient is thus estimated by $\hat{\beta}_i + \tilde{\beta}_i(t)$, with $\tilde{\beta}_i(t)$ capturing the time variation.

The "smooth" function $\tilde{\beta}_i(t)$ is often a spline, and the amount of smoothing can be set arbitrarily or chosen by cross-validation or other methods (see Hastie & Tibshirani (1991), Chapter 3 for instance).

As an illustration, Figure 1 shows the trajectories of estimated betas with monthly data for ten portfolios over the period January 1934-August 2007, using the rolling regression approach and a local level model with parameters estimated by maximum likelihood. (Data consists of 10 Industry portfolios, taken from Prof. K.R. French's Web page.)

There is good agreement between the two sets of estimates, and both show some compelling evidence of variation in the betas, with a noticeable downward trend in the betas for Utilities and Durables, partly reversed in the late nineties. There is also a sharp increase in the beta of the High Tech group around the year 2000. (Incidentally, the smoother evolution of betas for the portfolios of Manufactures, Durables, Health and Others also seems to occur in the German market: see Eisenbeiß et al. (2007).) We remark that at each time point, betas are estimated with past and future data: thus, the estimates are retrospective and cannot be computed at time t. However, they serve as targets to assess filtering and prediction methods realizable in real time, which we will examine in Section 5.

In Figure 2 we have used a generalized additive model (GAM, see for instance Hastie & Tibshirani (1991)), minimizing (10) with a spline in place of $\tilde{\beta}_i(t)$. The degree of smoothing has been chosen by generalized cross-validation. We have also used a local linear trend model,

$$\beta^{\ell}_{it} = \beta^{\ell}_{i,t-1} + \beta^{s}_{it} + \eta_{it} \tag{11}$$

$$\beta^{s}_{it} = \beta^{s}_{i,t-1} + \nu_{it} \tag{12}$$

$$Z_{it} = Z_t \beta^{\ell}_{it} + \epsilon_{it}, \tag{13}$$

Figure 1: Alternative estimates of betas using rolling regression over periods of $q = 60$ months and Kalman smoothing with local level dynamics for β_{it}. In both cases, past *and future* data is used when estimating β_{it} at time t. Monthly data for ten industrial portfolios in the period January 1934-August 2007.

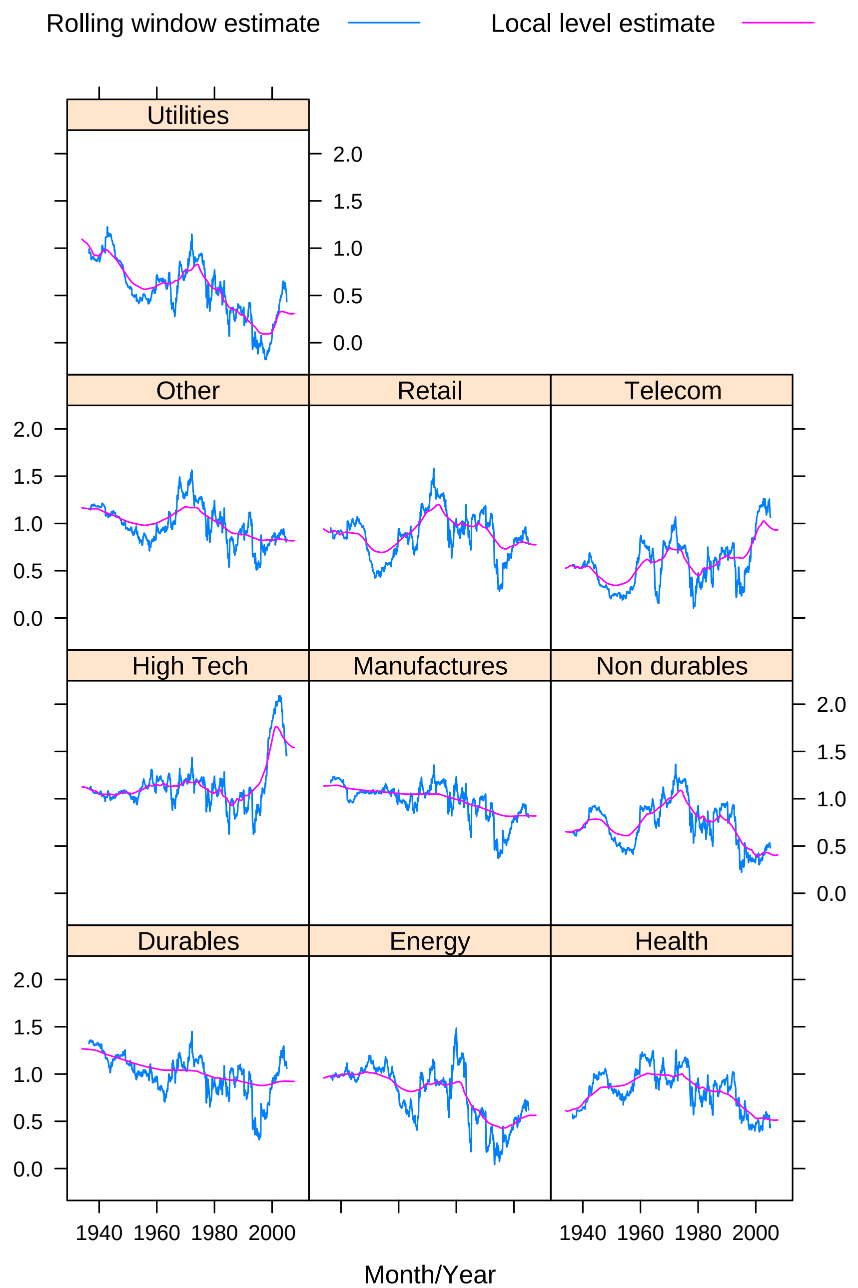
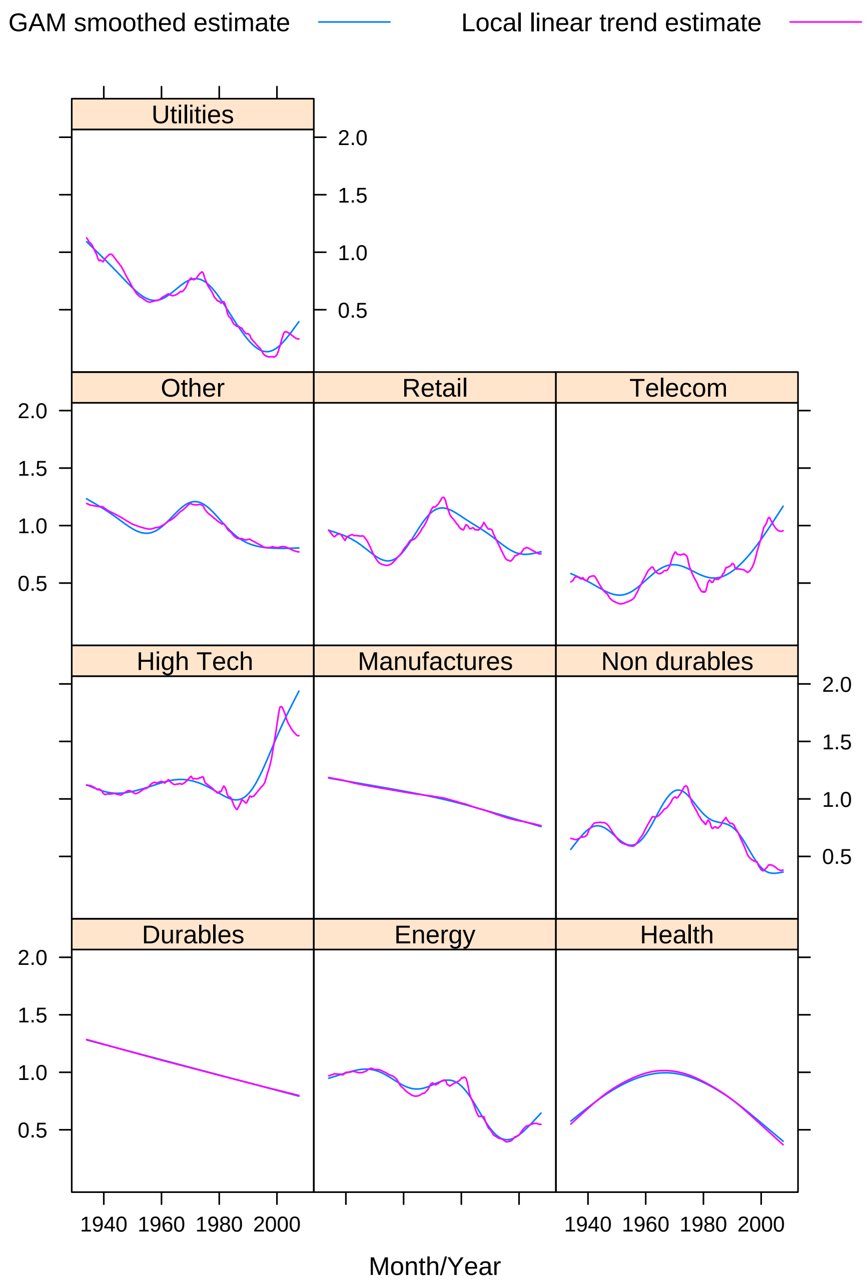


Figure 2: Alternative estimates of betas using a GAM with smoothing selected by generalized cross-validation and Kalman smoothing with local linear trend dynamics for β_{it}. In both cases, past *and future* data is used when estimating β_{it} at time t. Monthly data for ten industrial portfolios in the period January 1934-August 2007.



with parameters estimated by maximum likelihood. In (11)–(13) the betas β^{ℓ}_{it} follow a random walk with stochastic drift β^{s}_{it}, which itself follows a random walk. The local linear trend and GAM fits closely resemble each other: in fact, except for the amount of smoothing, both methods should be equivalent (Durbin & Koopman (2001), § 3.11).

The message of Figures 1 and 2 is that, while generally agreeing on their overall appearance, betas estimated by different methods and using different time spans may have quite different behaviour. In particular, with the span chosen, the rolling regression estimates in Figure 1 are jaggier than the smoothed estimates based on a local level model, and jaggier also than those in Figure 2.

As anticipated in the introduction, the performance of a filtering or prediction method will be quite dependent on what exactly we try to predict. Rather than using one single set of target betas as in e.g. Brooks & Faff (1997), we will compare our filtered and predicted betas with different target betas.

3 Predicting time-varying betas

Because single-index models call for estimates of betas in order to select assets for inclusion in a portfolio, considerable effort has been expended in trying to forecast these betas. We will review some of this work before we proceed to propose yet another two methods in the next Section.

In an early paper, Blume (1975) tried to assess how much association there is between betas in one period and their counterparts in the next period. He estimated betas for $i = 1, \ldots, n$ assets in two non-overlapping seven-year periods. Let β_{it}, $t = 1, 2$, denote the estimate of beta for asset i in period t. If we regress $\{\beta_{i2}\}$ on $\{\beta_{i1}\}$, $i = 1, \ldots, n$, fitting

$$\beta_{i2} = \delta_0 + \delta_1 \beta_{i1} + \epsilon_{i2}, \tag{14}$$

we can obtain estimates $\hat{\delta}_0$ and $\hat{\delta}_1$. These estimates can be used to obtain adjusted estimates of beta at $t = 2$

$$\hat{\beta}_{i2} = \hat{\delta}_0 + \hat{\delta}_1 \beta_{i1} \tag{15}$$

or predicted betas for the next period, i.e.

$$\hat{\beta}_{i3} = \hat{\delta}_0 + \hat{\delta}_1 \beta_{i2}. \tag{16}$$

It order to relate this procedure with our proposal in Section 4, it will help us to stack equation (16) for $i = 1, \ldots, n$ to form the single vector equation

$$\hat{\boldsymbol{\beta}}_{\mathbf{3}} = \hat{\delta}_0 \mathbf{1}_{\boldsymbol{n}} + \hat{\delta}_1 \boldsymbol{\beta}_{\mathbf{2}} \tag{17}$$

where $\boldsymbol{\beta}_{\mathbf{2}}{}^T = (\beta_{12}, \ldots, \beta_{n2})$ and $\mathbf{1}_{\boldsymbol{n}}$ is a vector of n "ones". The existing configuration of betas at time t can be thought of as a point $\boldsymbol{\beta}_{\boldsymbol{t}}$ in an n-dimensional space. The Blume forecast of the configuration of betas at time $t+1$, $\hat{\boldsymbol{\beta}}_{\boldsymbol{t+1}}$, is a linear combination of $\boldsymbol{\beta_t}$ and $\mathbf{1}_{\boldsymbol{n}}$. Quite often (as was indeed the case for the data used in Blume's original work) this will imply shrinkage of the vector of betas towards $\mathbf{1}_{\boldsymbol{n}}$.

Vasicek (1973) does something similar. The assets' betas, β_{i1}, and their variances, $\sigma^2_{\beta_{i1}}$, are computed for period 1. Let $\overline{\beta}_1$ be the average beta across the sample of assets (and $\sigma^2_{\overline{\beta}_1}$ its variance). The predicted beta for period 2 is computed as:

$$\hat{\beta}_{i2} = \frac{\sigma^2_{\beta_{i1}}}{\sigma^2_{\overline{\beta}_1} + \sigma^2_{\beta_{i1}}}\overline{\beta}_1 + \frac{\sigma^2_{\overline{\beta}_1}}{\sigma^2_{\overline{\beta}_1} + \sigma^2_{\beta_{i1}}}\beta_{i1} \tag{18}$$

This is similar in spirit to equation (16): each predicted beta is a (convex) linear combination of $\overline{\beta}_1$ and β_{i1}. However, the coefficients change with i, so we can no longer stack equation (18) for $i = 1, \ldots, n$ and write $\hat{\boldsymbol{\beta}}_{\boldsymbol{2}}$ as a linear combination of $\mathbf{1}_{\boldsymbol{n}}$ and $\boldsymbol{\beta_1}$, as in (17). It is true, though, that Vasicek's method relies, like Blume's, on shrinkage, this time towards $\overline{\beta}_1\mathbf{1}_{\boldsymbol{n}}$. Some variations exist, like the substitution of β_{i1} in (18) by β^*_{i1}, obtained by updating β_{i1} as in Blume's method (see Marín & Rubio (2001), p. 404). The idea of estimating betas by shrinking towards one or multiple points is reviewed and developed in Karolyi (1992).

4 Two new methods

We will look at the betas for the n assets considered at each time t as a vector $\boldsymbol{\beta_t}$ in R^n. The history of betas is then a trajectory in that "phase space". In our case, with ten portfolios, $n = 10$, so this trajectory wanders in 10-dimensional space. While we cannot visualize in more than three dimensions, it is useful to look at some two-dimensional projections to get a feeling of how betas evolve in phase space. (In order to develop some intuition on what is going on, we have found tools like `ggobi`, Cook, Swayne, Buja & Lang (2008), to be of invaluable help.) Figure 3 shows a subset of two-dimensional projections of four betas against another three. The betas are fitted using a GAM model (as represented in Figure 2 above).

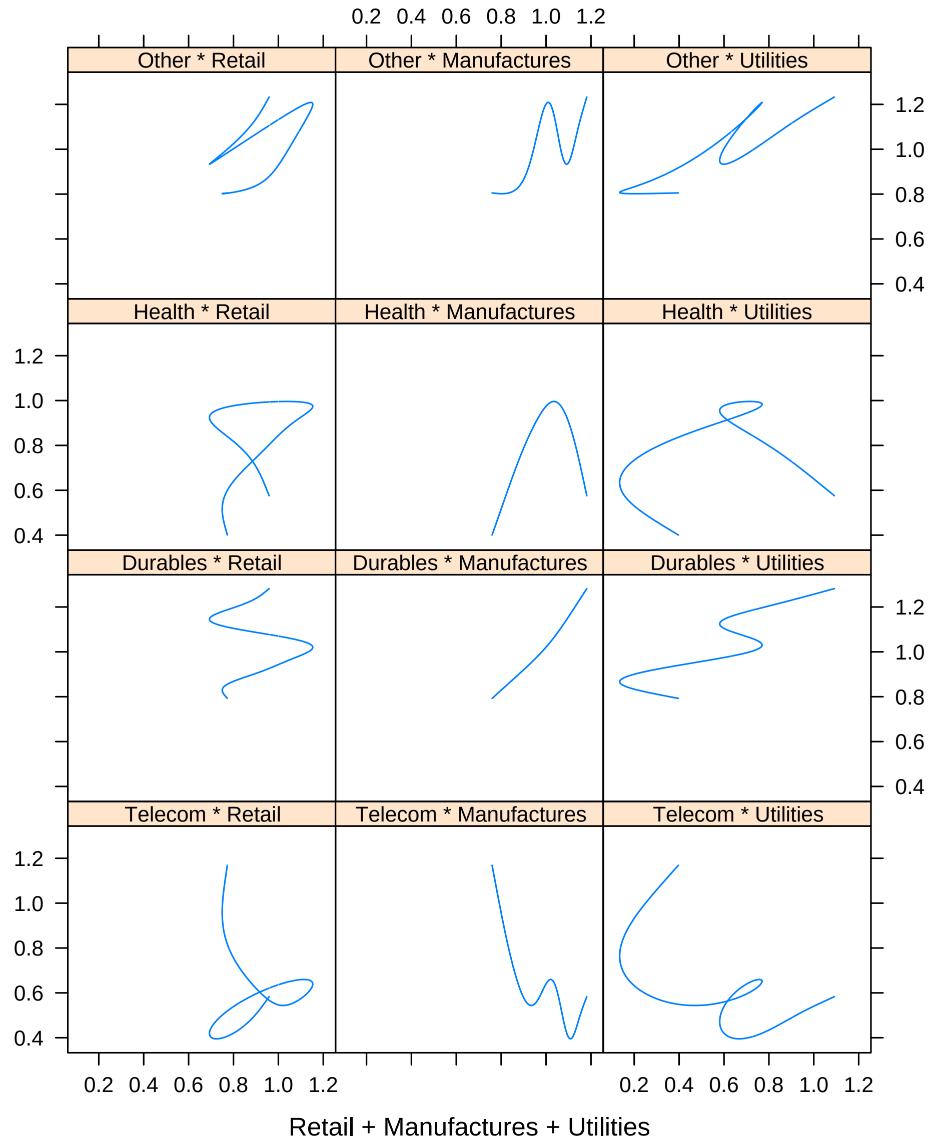
Blume's method can be seen as transforming the vector $\boldsymbol{\beta_t}$ of betas into a new

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{t+1}} = \hat{\delta}_0\mathbf{1}_{\boldsymbol{n}} + \hat{\delta}_1\boldsymbol{\beta_t}, \tag{19}$$

where $\hat{\delta}_0$ and $\hat{\delta}_1$ have been obtained by regressing $\boldsymbol{\beta_t}$ on $\boldsymbol{\beta_{t-1}}$. This seems a reasonable transformation in phase space for mapping one observed point to an approximation of the next.

One feature to notice is that $\hat{\delta}_0$ and $\hat{\delta}_1$ have been estimated in a cross-section regression with only $n-2$ degrees of freedom. Seen from another point of view, the transformation carrying $\boldsymbol{\beta_t}$ into $\hat{\boldsymbol{\beta}}_{\boldsymbol{t+1}}$ depends *only* on the evolution observed from $\boldsymbol{\beta_{t-1}}$ to $\boldsymbol{\beta_t}$. Looking at the long quasi-linear stretches in the phase space sections in Figure 3, one might conjecture that the change from $\boldsymbol{\beta_{t-1}}$ to $\boldsymbol{\beta_t}$ is likely to be quite similar to

Figure 3: Twelve two-dimensional sections of β_t in phase space. The betas shown are estimated using a GAM model. Monthly data for ten industrial portfolios in the period January 1934-August 2007. While some sharp turns can be seen, most of the time the trajectories evolve in quasi-linear stretches.



the changes at previous time steps. This suggests fitting the regression

$$\begin{bmatrix} \boldsymbol{\beta_t} \\ \boldsymbol{\beta_{t-1}} \\ \vdots \\ \boldsymbol{\beta_{t-k+1}} \end{bmatrix} = \delta_0 \begin{bmatrix} \mathbf{1}_{\boldsymbol{n}} \\ \mathbf{1}_{\boldsymbol{n}} \\ \vdots \\ \mathbf{1}_{\boldsymbol{n}} \end{bmatrix} + \delta_1 \begin{bmatrix} \boldsymbol{\beta_{t-1}} \\ \boldsymbol{\beta_{t-2}} \\ \vdots \\ \boldsymbol{\beta_{t-k}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon_t} \\ \boldsymbol{\epsilon_{t-1}} \\ \vdots \\ \boldsymbol{\epsilon_{t-k+1}} \end{bmatrix}, \tag{20}$$

i.e., stacking k cross-section regressions like (17) with the same parameters δ_0, δ_1. In other words, rather than seek a transformation carrying one point $\boldsymbol{\beta_{t-1}}$ in phase space into the next, $\boldsymbol{\beta_t}$, we seek a transformation carrying a group of k consecutive n-dimensional points (which we will call an "epoch") to the next. Doing so, we force smoothness in the evolution of predicted betas and use more observations to estimate δ_0, δ_1. The performance of this method, which we will name Blume-k, is examined in the sequel.

But we need not stop here: we may think of more general transformations carrying an epoch into the next. A rather general family of transformations is made of all rigid rotations in phase space, possibly followed by a dilation and translation.

Consider an epoch made of $\boldsymbol{\beta_t}, \boldsymbol{\beta_{t-1}}, \ldots, \boldsymbol{\beta_{t-k+1}}$ and stack all of its components as row vectors in the $k \times n$ matrix $\boldsymbol{B_t}$. Define likewise an epoch $\boldsymbol{B_{t+1}}$. We may look for a constant ρ, rotation matrix $\boldsymbol{G}$ and vector $\boldsymbol{c}$ such that:

$$\boldsymbol{B_{t+1}} \approx \rho \boldsymbol{B_t G} + \mathbf{1}_{\boldsymbol{k}} \boldsymbol{c}^T, \tag{21}$$

where $\boldsymbol{G}$ is restricted to be an orthogonal matrix, ρ is a non-negative constant and $\boldsymbol{c}$ is a n-dimensional translation vector. Note that ρ, $\boldsymbol{G}$ and $\boldsymbol{c}$ are all dependent on t, a dependency we drop in the notation for simplicity.

Consider initial and target $k \times n$ matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, and the problem of finding ρ, $\boldsymbol{G}$ and $\boldsymbol{c}$ such that

$$\boldsymbol{Y} \approx \rho \boldsymbol{X G} + \mathbf{1}_{\boldsymbol{k}} \boldsymbol{c}^T. \tag{22}$$

A reasonable criterion is to minimize trace$(\boldsymbol{Z^T Z})$ with $\boldsymbol{Z} = \boldsymbol{Y} - \rho \boldsymbol{XG} - \mathbf{1}_{\boldsymbol{k}} \boldsymbol{c}^T$. The resulting ρ, $\boldsymbol{G}$ and $\boldsymbol{c}$ are said to perform a Procrustean rotation carrying $\boldsymbol{X}$ into $\boldsymbol{Y}$, and can be readily obtained. Let $\tilde{\boldsymbol{X}}$ be the matrix $\boldsymbol{X}$ after centering its columns, and perform the singular value decomposition of $\boldsymbol{Y^T} \tilde{\boldsymbol{X}}$, so that let $\boldsymbol{UDV^T} = \boldsymbol{Y^T} \tilde{\boldsymbol{X}}$. It can be shown (see for instance Krzanowski (1988)) that the sought-for ρ, $\boldsymbol{G}$ and $\boldsymbol{c}$ are given by:

$$\begin{aligned} \boldsymbol{G} &= \boldsymbol{VU^T} && (23) \\ \rho &= \frac{\text{trace}(\boldsymbol{D})}{\text{trace}(\tilde{\boldsymbol{X}}^{\boldsymbol{T}} \tilde{\boldsymbol{X}})} && (24) \\ \boldsymbol{c^T} &= \mathbf{1}_{\boldsymbol{k}}^{\boldsymbol{T}} (\boldsymbol{Y} - \rho \boldsymbol{XG}). && (25) \end{aligned}$$

Use of these formulae will solve our problem. We will work in deviations from 1. Defining[2]

$$\boldsymbol{B}_{\boldsymbol{t}}^{*} = \boldsymbol{B}_{\boldsymbol{t}} - \boldsymbol{1}_{\boldsymbol{k},\boldsymbol{n}}, \tag{26}$$

where $\boldsymbol{1}_{\boldsymbol{k},\boldsymbol{n}}$ is a $k \times n$ matrix of ones, we will consider

$$\boldsymbol{B}_{\boldsymbol{t+1}}^{*} \approx \rho \boldsymbol{B}_{\boldsymbol{t}}^{*} \boldsymbol{G} + \boldsymbol{1}_{\boldsymbol{k}} \boldsymbol{c}^{T} \tag{27}$$

instead of (21) and use equations (23)–(25) with $\boldsymbol{B}_{\boldsymbol{t}}^{*}$ and $\boldsymbol{B}_{\boldsymbol{t+1}}^{*}$ in place of $\boldsymbol{X}$ and $\boldsymbol{Y}$ respectively.

The estimation of $\boldsymbol{G}$, ρ and $\boldsymbol{c}$ which approximately carry one epoch $\boldsymbol{B}_{\boldsymbol{t}}^{*}$ into $\boldsymbol{B}_{\boldsymbol{t+1}}^{*}$ provides a way to adjust and/or extrapolate betas, much in the same way as in our Blume-k generalization of Blume's method: given ρ, $\boldsymbol{G}$, $\boldsymbol{c}$ and $\boldsymbol{B}_{\boldsymbol{t}}^{*}$, we can predict $\boldsymbol{B}_{\boldsymbol{t+1}}^{*}$ by

$$\hat{\boldsymbol{B}}_{\boldsymbol{t+1}}^{*} = \rho \boldsymbol{B}_{\boldsymbol{t}}^{*} \boldsymbol{G} + \boldsymbol{1}_{\boldsymbol{k}} \boldsymbol{c}^{T}. \tag{28}$$

It is interesting to compare the last expression with Blume's method. Transposing (28) and picking the first column of the result, we have:

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{t+1}}^{*} = \rho \boldsymbol{G}^{T} \boldsymbol{\beta}_{\boldsymbol{t}}^{*} + \boldsymbol{c}, \tag{29}$$

which, since $\boldsymbol{\beta}_{\boldsymbol{t+1}}^{*} = \boldsymbol{\beta}_{\boldsymbol{t+1}} - \boldsymbol{1}$, can be written

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{t+1}} = \rho \boldsymbol{G}^{T} \boldsymbol{\beta}_{\boldsymbol{t}} + \boldsymbol{c} + (\boldsymbol{I}_{\boldsymbol{n}} - \rho \boldsymbol{G}^{T}) \boldsymbol{1}_{\boldsymbol{n}}. \tag{30}$$

In general, $\boldsymbol{c}$ will be close to zero: its only purpose is to shift the centroid of the transformed epoch so as to match the centroid of the target epoch, see (25). Both these centroids will be close to zero most of the time when we deal with betas expressed in deviations with respect to 1. We are left then with

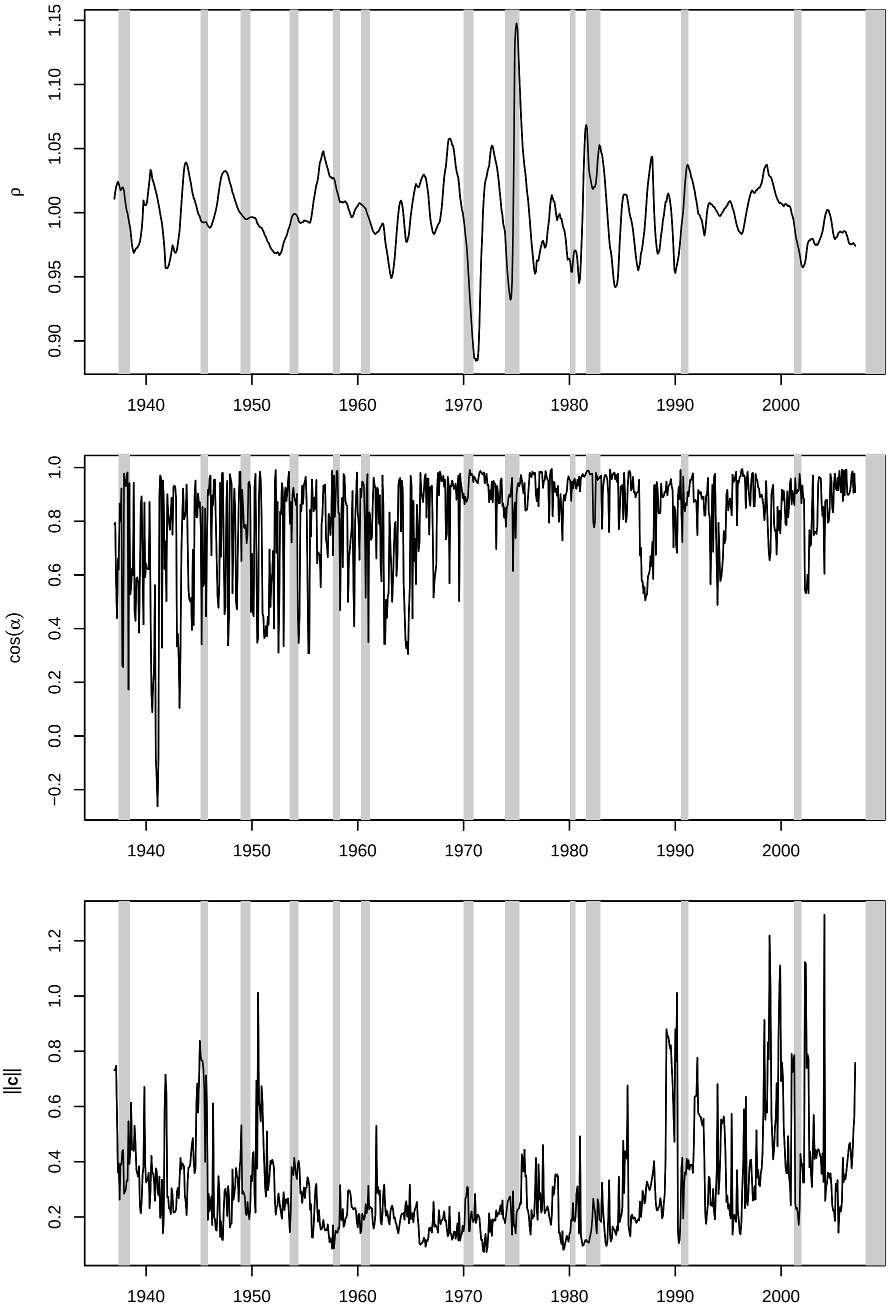
$$\hat{\boldsymbol{\beta}}_{\boldsymbol{t+1}} \approx \rho \boldsymbol{G}^{T} \boldsymbol{\beta}_{\boldsymbol{t}} + (\boldsymbol{I}_{\boldsymbol{n}} - \rho \boldsymbol{G}^{T}) \boldsymbol{1}_{\boldsymbol{n}}; \tag{31}$$

this gives $\hat{\boldsymbol{\beta}}_{\boldsymbol{t+1}}$ as a "weighted average" of $\boldsymbol{\beta}_{\boldsymbol{t}}$ and $\boldsymbol{1}_{\boldsymbol{n}}$, much as in expression (19). However, unlike in (19), the "weights" are matrices and unequal for different components of $\boldsymbol{\beta}_{\boldsymbol{t}}$.

As an example which may help to gain some intuition about the the performance of the rotation method, we have computed the Procrustes rotation on epochs of 24 consecutive months, $\boldsymbol{B}_{\boldsymbol{t}}$ and $\boldsymbol{B}_{\boldsymbol{t+1}}$. Thus, each epoch can be thought of as a set of 24 points in R^{n}, where $n = 10$ is the number of portfolios. Betas are expressed as deviations from 1, i.e. $\boldsymbol{\beta}^{*}$'s in the notation introduced in (26). The three panels of Figure 4

[2] The literature has reported a tendency of beta to revert to 1.0. Blume (1971) detected a tendency for the betas of well-diversified portfolios of extreme risk to regress toward the grand average of all betas. Blume (1975) tests that the mean is 1.0. Gooding & O'Malley (1977) confirms this tendency and assume that the causes are economic and not statistical. Goldberg (1981) and Garbade & Rentzler (1981) follow up along the same line.

Figure 4: Procrustes rotation statistics when rotating betas obtained by smoothing with a local linear trend model. The panels show from top to bottom the dilation coefficient ρ, $\cos(\alpha)$, where α is the angle of rotation associated to matrix $\boldsymbol{G}$ and $|\boldsymbol{c}|$, the modulus of the translation vector. Monthly data for the period January 1937-August 2007. Shaded bars identify recession periods as defined by the NBER.



show respectively the estimates of ρ defined in (24), the cosine of the angle α rotated by the orthogonal matrix $\boldsymbol{G}$ (computed as $\cos(\alpha) = k^{-1}\mathbf{1}_{\boldsymbol{k}}^{\boldsymbol{T}}\boldsymbol{G}\,\mathbf{1}_{\boldsymbol{k}}$) and the euclidean norm $\|c\|$ of the translation vector $\boldsymbol{c}$ defined in (25).

The results are interesting in themselves. A parameter ρ smaller than 1 can be interpreted as a regression of $\boldsymbol{\beta}^{*}$ towards $\mathbf{0}$ (and thus of $\boldsymbol{\beta}$ towards $\mathbf{1}$); this implies that the betas of the different assets become more alike. A parameter ρ greater than 1, on the other hand, implies increasing asset differentiation.

In Figure 4, ρ wanders around 1, rarely deviating more than 5% from it. There seems to be, though, a period with major beta convergence and then divergence: from the end of 1968 to the beginning of 1973. Interestingly, $\cos(\alpha)$ was rather stable and close to 1 for the whole of this period and $|\boldsymbol{c}|$ fairly small, implying that the realignment of betas was a phenomenon common to all of them. The opposite is true for the period through the early sixties, in which the orientation of $\boldsymbol{\beta}_{\boldsymbol{t}}^{*}$ seemed to change repeatedly, as evinced by the low $\cos(\alpha)$.

In an attempt to see if changes in $\boldsymbol{\beta}_{\boldsymbol{t}}$ are related to the economic cycle, we have shaded the periods which correspond to recessions, as defined by the NBER. It seems that some of the major swings in ρ, particularly in the seventies, are associated to recessions, but no clear pattern emerges. There are also a few, but relatively large, values of $\|c\|$ after the mid nineties, implying sudden changes in $\boldsymbol{\beta}_{\boldsymbol{t}}^{*}$ which cannot be accounted for by rotation and dilation.

5 An empirical investigation

In order to assess the relative merits of the different prediction and adjusting beta methods, we have conducted the following experiment. First, we have estimated target betas using four different methods described in Section 2, and illustrated in Figures 1 and 2. As mentioned there, these estimations are not realizable in real time, as they use observations in the future.

All computation was done in R (see R Development Core Team (2008)) and various packages built on R, notably Petris (2007), Sarkar (2008) and Zeileis & Grothendiek (2005). All code is available from the corresponding author.

All four tables presented in the following display mean square error (MSE) when target betas $\tilde{\boldsymbol{\beta}}_{\boldsymbol{it}}$ are approximated. Target betas are computed using past and future observations. The approximating betas $\hat{\boldsymbol{\beta}}_{\boldsymbol{it}}$ only use past or past and present information at time t, and are computed using various estimation methods, supplemented or not with a Blume-k or Procrustes transformation. Tables 1 and 2 show results for filtered betas, i.e. $\hat{\boldsymbol{\beta}}_{\boldsymbol{it}}$ are estimated with information up to time t. Tables 3 and 4 show results for predictions, when $\hat{\boldsymbol{\beta}}_{\boldsymbol{it}}$ are estimated with information up to time $t-1$.

The MSE figures have been computed as

$$MSE = N^{-1} \sum_i \sum_t (\hat{\beta}_{it} - \tilde{\beta}_{it})^2, \tag{32}$$

where the sums extend over all assets i and all N available months (N may differ from one estimate to the next, as the different methods force to discard a different number of observations at the start). Thus, the MSE pools results for all assets and times in a single figure for each combination of target and estimation method.

All tables have an "Unadjusted" panel, displaying the MSE for each combination of filtering/prediction method and target betas when no Blume-k or Procrustes adjustment is performed.

In Table 1 we see, under the "Unadjusted" heading, the performance of four filtering methods for target betas computed in different ways. For all filtering methods (except for the filtering based on the local linear trend), the MSE is lower for the smoother targets betas —the rolling regression method with a bilateral window extending over 60 months seems to produce poorly smoothed betas that are difficult to cope with by the filtering methods. Aside from that, filtering based on the local level model appears far better than the other filtering methods.

When betas go through a Procrustes transformation, results are largely unchanged, or even degraded, except for the local linear filtered betas, whose MSE drops markedly. This is true for all targets, except the RR60 variety, which all filtering methods appear to have trouble coping with. Overall, the Procrustes method does not shine in the comparison, although it does improve the local linear trend (LLT) estimates when the targets are LL, GAM and LLT.

More interesting patterns emerge when we use the Blume-k adjusting method. When $k = 1$ (i.e., for Blume's original method) the effect of the adjustment on the MSE is largely unnoticed *except* for the local linear filtered $\hat{\boldsymbol{\beta}}_{\boldsymbol{t}}$ and the smoother varietys (LL, GAM, LLT) of $\tilde{\boldsymbol{\beta}}_{\boldsymbol{t}}$. Filtering with the local linear model produces quite noisy betas and no matter what we do in introducing some constraints is useful.

The interesting part comes for the lower two panels: when using the Blume-k adjusting with epochs of $k = 24$, the improvement is striking, with reductions in the MSE of about 50% in some cases. This is even true with epochs of $k = 60$ months, although for such large k some figures already show signs of deterioration.

In order to gain some insight on the behaviour of the Procrustes and Blume-k method as the length k of the epoch changes, we have computed the MSE as a function of k for each combination of filtering method and target betas. The results can be seen in Figure 5.

It is clear that Blume's method (which corresponds to our Blume-k method in the particular case of $k = 1$) can be improved upon by increasing k. The optimal k seems to be for most combinations of method and target between $k = 30$ and $k = 60$. In contrast, the Procrustes adjusting method does not appear to help much except for the LLT variety of $\hat{\boldsymbol{\beta}}_{\boldsymbol{t}}$; and even then it seems dominated by the Blume-k method.

Aside from that, the estimation method based on the local level model (LL) seems clearly best for any k and most targets, and remarkably insensitive to the choice of k. When the targets are betas smoothed by rolling regression (bilateral, using both past and future data), the MSE becomes quite sensitive to the choice of k.

Tables 1 and 2 only tell part of the story. Both the Blume-k and Procrustes transformation method are geared towards prediction. When using information up to time t, the Blume-k method estimates the parameters δ_0, δ_1 fitting (20); let the estimates be $\hat{\delta}_0(t)$, $\hat{\delta}_1(t)$, where we have modified the notation to explicitly reflect the dependency of the estimates on t. Then, the adjusted betas for time t are given by

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{t}} = \hat{\delta}_0(t)\mathbf{1}_{\boldsymbol{n}} + \hat{\delta}_1(t)\hat{\boldsymbol{\beta}}_{\boldsymbol{t-1}} \tag{33}$$

and the one-step ahead predictions are computed as

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{t}} = \hat{\delta}_0(t-1)\mathbf{1}_{\boldsymbol{n}} + \hat{\delta}_1(t-1)\hat{\boldsymbol{\beta}}_{\boldsymbol{t-1}}. \tag{34}$$

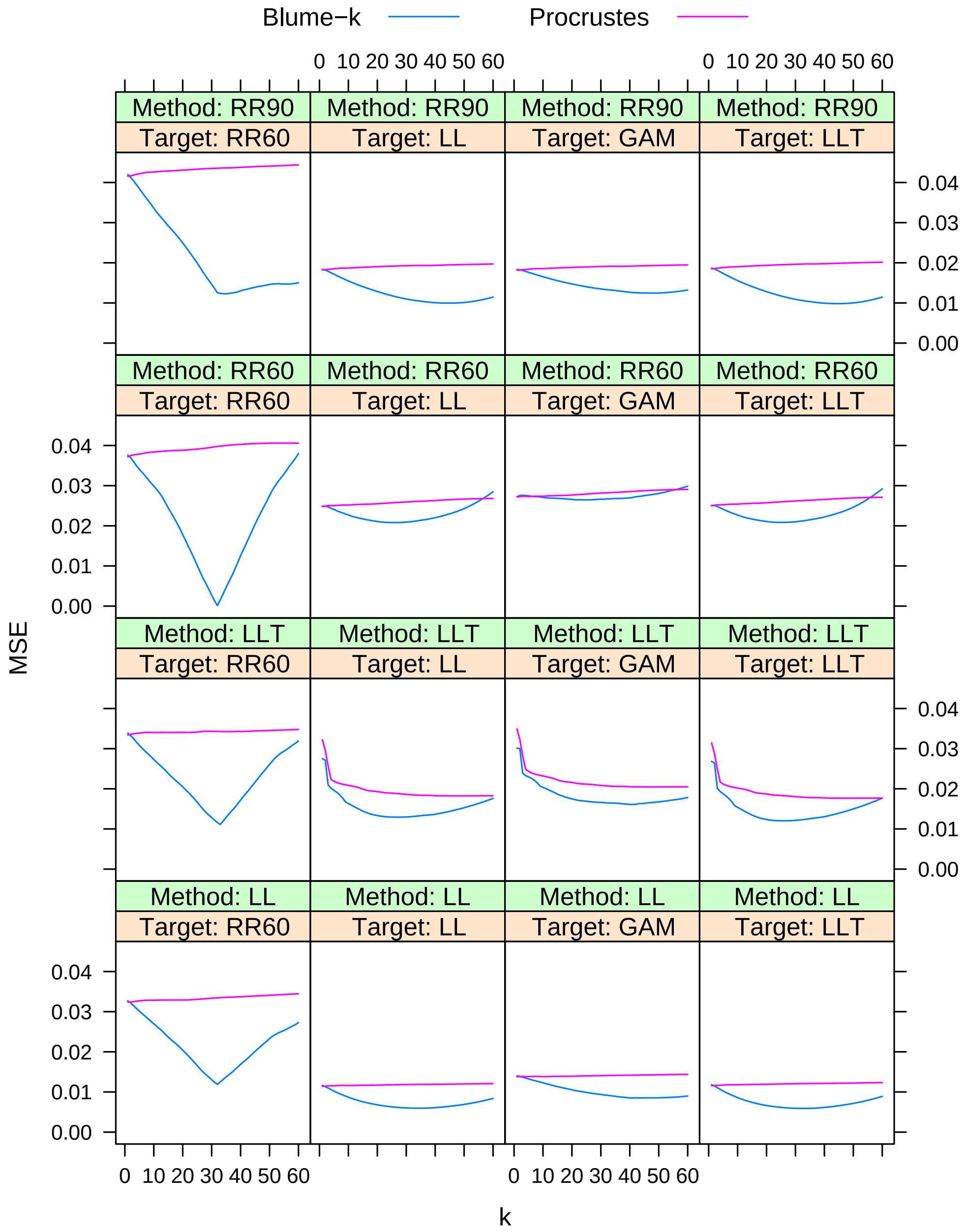
The adjusted estimates in Table 2 are basically one-step ahead predictions at time $t-1$ except for the fact that $\hat{\delta}_0(t)$, $\hat{\delta}_1(t)$ are estimated using the observation at time t, which is only used in the estimation of $\hat{\delta}_0(t)$ and $\hat{\delta}_1(t)$; similarly for the Procrustes estimator. This is in contrast to filtering using a local level or local linear trend model, which make fuller use of contemporaneous information. Therefore, one-step-ahead predictions may be a fairer basis for comparison of the merits of the different methods.

Tables 3 and 4 provide the analogues of Tables 1 and 2, using MSE of prediction one step ahead. Patterns mimic what we have seen in Tables 1 and 2. Most MSE figures are slightly larger, as one would expect given that we are using less information in our attempt to approximate $\hat{\boldsymbol{\beta}}_{\boldsymbol{t}}$. The local level model seems again the strongest performer, both before and after adjusting with the Blume-k method and a suitable k; of the values of k used in Table 4, $k = 24$ gives good results, while $k = 60$ is clearly too large. Not only is the local level method of prediction best, but it is also the one which stands to gain more from Blume-k adjusting.

6 Conclusions

Two new methods, Procrustes and Blume-k, have been introduced, aimed at the adjusting of raw betas estimated by several traditional methods. Both can be seen as mappings in phase space shrinking towards 1.0, and are thus similar in spirit to Blume's method. While the Procrustes method implements a rather flexible mapping which, as a by-product, produces useful descriptive statistics, it appears of limited value in terms of MSE reduction. The method that we have named Blume-k, on the other hand, implements a

Figure 5: MSE for different combinations of target betas and estimation method, using Blume-k and Procrustes adjusting with $k = 1, \ldots, 60$.



very simple mapping which exhibits considerable potential. Our main conclusion is that modifying Blume's method so as to have longer memory —as in the Blume-k method presented above— consistently improves beta estimates. The Procrustes method, on the other hand, provides an attractive decomposition of beta movements in phase space, but seems of lesser value as a predictive method.

Acknowledgements This research was supported partially by grant IT-321-07 from the Gobierno Vasco and ECO 2008-00777/ECON from Ministerio de Ciencia e Innovación.

Table 1: Mean square error of approximation of several target betas using different filtering methods. Targets are computed using past and future information. Methods listed in the left margin of the table use information up to and including the present observation. The figures in each cell are MSE. Details in text.

Estimation method for $\hat{\beta}_t$	**Target betas $\tilde{\beta}_t$**: **Rolling (RR60)**	**Local level (LL)**	**GAM model**	**Local linear (LLT)**
	UNADJUSTED			
RR (60 months)	0.037100	0.024729	0.027128	0.024917
RR (90 months)	0.041500	0.018204	0.018156	0.018475
Local level	0.032340	0.011472	0.013855	0.011610
Local linear	0.033437	0.033393	0.036142	0.032639
	PROCRUSTES ROTATED: EPOCH = 6 MONTHS			
RR (60 months)	0.037980	0.025075	0.027315	0.025292
RR (90 months)	0.042339	0.018574	0.018446	0.018870
Local level	0.032769	0.011588	0.013867	0.011762
Local linear	0.033968	0.021567	0.023978	0.020873
	PROCRUSTES ROTATED: EPOCH = 12 MONTHS			
RR (60 months)	0.038516	0.025269	0.027460	0.025501
RR (90 months)	0.042741	0.018780	0.018619	0.019102
Local level	0.032884	0.011632	0.013854	0.011827
Local linear	0.034049	0.020577	0.022888	0.019931
	PROCRUSTES ROTATED: EPOCH = 18 MONTHS			
RR (60 months)	0.038778	0.025421	0.027590	0.025666
RR (90 months)	0.042998	0.018985	0.018815	0.019314
Local level	0.032916	0.011692	0.013911	0.011897
Local linear	0.034062	0.019450	0.021735	0.018850

Table 2: Mean square error of approximation of several target betas using different filtering methods. Targets are computed using past and future information. Methods listed in the left margin of the table use information up to and including the present observation. The figures in each cell are MSE. Details in text.

	Target betas $\tilde{\beta}_t$			
Estimation method for $\hat{\beta}_t$	**Rolling (RR60)**	**Local level (LL)**	**GAM model**	**Local linear (LLT)**
	UNADJUSTED			
RR (60 months)	0.037100	0.024729	0.027128	0.024917
RR (90 months)	0.041500	0.018204	0.018156	0.018475
Local level	0.032340	0.011472	0.013855	0.011610
Local linear	0.033437	0.033393	0.036142	0.032639
	BLUME ADJUSTING: USING 1 MONTHS			
RR (60 months)	0.037611	0.024897	0.027240	0.025099
RR (90 months)	0.041984	0.018400	0.018327	0.018685
Local level	0.032720	0.011651	0.013982	0.011817
Local linear	0.033856	0.027499	0.030115	0.026838
	BLUME ADJUSTING: USING 24 MONTHS			
RR (60 months)	0.011806	0.020818	0.026441	0.020849
RR (90 months)	0.020985	0.011996	0.014158	0.011937
Local level	0.017391	0.006379	0.010010	0.006236
Local linear	0.017502	0.012989	0.016977	0.012051
	BLUME ADJUSTING: USING 60 MONTHS			
RR (60 months)	0.037985	0.028461	0.029850	0.029204
RR (90 months)	0.015019	0.011456	0.013214	0.011457
Local level	0.027277	0.008389	0.008988	0.008891
Local linear	0.031876	0.017619	0.017834	0.017683

Table 3: Mean square error of approximation of several target betas using different one step ahead prediction methods. Targets are computed using past and future information. Methods listed in the left margin predict $\tilde{\beta}_{t+1}$ using information up to and including time t. The figures in each cell are MSE. Details in text.

	Target betas $\tilde{\beta}_t$			
Estimation method for $\hat{\beta}_t$	**Rolling (RR60)**	**Local level (LL)**	**GAM model**	**Local linear (LLT)**
	UNADJUSTED			
RR (60 months)	0.038337	0.025201	0.027310	0.025440
RR (90 months)	0.042499	0.018652	0.018444	0.018953
Local level	0.033096	0.011909	0.014089	0.012090
Local linear	0.033610	0.022684	0.025184	0.021902
	PROCRUSTES ROTATED: EPOCH = 1 MONTHS			
RR (60 months)	0.040283	0.026936	0.029022	0.027170
RR (90 months)	0.043437	0.019501	0.019315	0.019805
Local level	0.033738	0.012405	0.014589	0.012601
Local linear	0.035549	0.023105	0.025382	0.022431
	PROCRUSTES ROTATED: EPOCH = 12 MONTHS			
RR (60 months)	0.039919	0.026299	0.028348	0.026549
RR (90 months)	0.043533	0.019278	0.019044	0.019615
Local level	0.033536	0.012126	0.014271	0.012335
Local linear	0.035116	0.021317	0.023503	0.020690
	PROCRUSTES ROTATED: EPOCH = 18 MONTHS			
RR (60 months)	0.039986	0.026336	0.028392	0.026595
RR (90 months)	0.043713	0.019438	0.019217	0.019788
Local level	0.033517	0.012098	0.014252	0.012319
Local linear	0.034992	0.020117	0.022303	0.019532

Table 4: Mean square error of approximation of several target betas using different one step ahead prediction methods. Targets are computed using past and future information. Methods listed in the left margin predict $\tilde{\beta}_{t+1}$ using information up to and including time t. The figures in each cell are MSE. Details in text.

	Target betas $\tilde{\beta}_t$			
Estimation method for $\hat{\beta}_t$	**Rolling (RR60)**	**Local level (LL)**	**GAM model**	**Local linear (LLT)**
	UNADJUSTED			
RR (60 months)	0.038337	0.025201	0.027310	0.025440
RR (90 months)	0.042499	0.018652	0.018444	0.018953
Local level	0.033096	0.011909	0.014089	0.012090
Local linear	0.033610	0.022684	0.025184	0.021902
	BLUME ADJUSTING: USING 1 MONTHS			
RR (60 months)	0.039666	0.026820	0.029177	0.027021
RR (90 months)	0.042859	0.019263	0.019201	0.019548
Local level	0.033523	0.012414	0.014766	0.012572
Local linear	0.035192	0.047748	0.050601	0.047096
	BLUME ADJUSTING: USING 24 MONTHS			
RR (60 months)	0.011933	0.020914	0.026551	0.020946
RR (90 months)	0.021020	0.012030	0.014199	0.011972
Local level	0.017452	0.006409	0.010050	0.006264
Local linear	0.017563	0.013038	0.017039	0.012098
	BLUME ADJUSTING: USING 60 MONTHS			
RR (60 months)	0.038048	0.028524	0.029920	0.029264
RR (90 months)	0.015046	0.011479	0.013240	0.011479
Local level	0.027310	0.008406	0.009011	0.008908
Local linear	0.031905	0.017640	0.017863	0.017704

References

Abberger, K. (2004), Conditionally parametric fits for CAPM betas, Technical report, Diskussionspapier des Center of Finance and Econometrics, Universität Konstanz.

Baesel, J. (1974), 'On the assessment of risk: Some further considerations', *Journal of Finance* **29**, 1491–1494.

Bildersee, J. & Roberts, G. (1981), 'Beta instability when interest rate levels change', *Journal of Financial and Quantitative Analysis* **XVI**, 375–380.

Blume, M. (1971), 'On the assessment of risk', *Journal of Finance* **XXVI**, 1–10.

Blume, M. (1975), 'Betas and their regression tendencies', *Journal of Finance* **XXX**(3), 785–795.

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. (1997), 'A note on beta forecasting', *Applied Economic Letters* **4**, 77–78.

Brooks, R., Faff, R. & Josev, T. (1997*a*), 'Beta stablity and monthly seasonal effects: evidence from the Australian capital market', *Applied Economics Letters* **4**, 563–566.

Brooks, R., Faff, R. & Lee, J. (1992), 'The form of time variation of systematic risk: some Australian evidence', *Applied Financial Economics* **2**, 191–198.

Brooks, R., Faff, R. & Lee, J. (1994), 'Beta stability and portfolio formation', *Pacific Basin Finance Journal* **2**, 463–479.

Brooks, R., Faff, R. W., Gangemi, M. & Lee, J. H. H. (1997*b*), 'A further examination of the effect of diversification on the stability of portfolio betas', *Applied Financial Economics* **7**, 9–14.

Chen, S.-N. & Lee, C. (1982), 'Bayesian and mixed estimators of time varying betas', *Journal of Economics and Business* **34**, 299–301.

Clare, A., Priestley, R. & Thomas, S. (1997), 'Is Beta dead? the role of alternative estimation methods', *Applied Economics Letters* **4**, 559–562.

Collins, D. W., Ledolter, J. & Rayburn, J. (1987), 'Some further evidence on the stochastic properties of systematic risk', *Journal of Business* **60**, 425–428.

Cook, D., Swayne, D. F., Buja, A. & Lang, D. T. (2008), *Interactive and Dynamic Graphics for Data Analysis: With R and GGobi.*, Springer-Verlag.

Dotan, A. & Ofer, A. (1984), 'Variable versus stationary beta in the market model', *Journal of Banking and Finance* **8**, 525–534.

Durbin, J. & Koopman, S. J. (2001), *Time Series Analysis by State Space Methods*, Oxford Univ. Press.

Ebner, M. & Neumann, T. (2005), 'Time-varying betas of German stock returns', *Financial Markets and Portfolio Management* **19**(1), 29–46.

Eisenbeiß, M., Kauermann, G. & Semmler, W. (2007), 'Estimating beta-coefficients of German stock data: A non-parametric approach', *The European Journal of Finance* **13**(6), 503–522.

Elton, E. & Gruber, M. (1995), *Modern Portfolio Theory and Investment Analysis*, John Wiley ans Sons.

Esteban, M. & Orbe, S. (to appear), ‘A nonparametric approach for estimating betas: The smoothed rolling estimator’, *Applied Economics*.

Fabozzi, F. J. & Francis, J. C. (1978), ‘Beta as a random coefficient’, *Journal of Financial and Quantitative Analysis* **13**, 101–116.

Faff, R. W., Lee, J. H. H. & Fry, T. R. L. (1992), ‘Time stationarity of systematic risk: some Australian evidence’, *Journal of Business Finance and Accounting* **19**, 253–270.

Fama, E. & MacBeth, J. (1973), ‘Risk, return and equilibrium: Empirical test’, *Journal of Political Economy* **81**, 607–636.

Fisher, L. (1970), The estimation of systematic risk: Some new findings. in: Proceedings of the seminar on the analysis of security price, Technical report, University of Chicago.

Garbade, K. & Rentzler, J. (1981), ‘Testing the hypothesis of beta stationarity’, *International Economic Review* **22**(3), 577–5867.

Goldberg, M. (1981), ‘An explanation of security β regression tendencies’, *Southern Economic Journal of Nonparametric Statistics* **47**(3), 809–815.

Gonedes, N. J. (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**(3), 407–443.

González-Rivera, G. (1997), ‘The pricing of time-varying beta’, *Empirical Economics* **22**, 345–363.

Gooding, A. & O’Malley, T. (1977), ‘Market phase and the stationarity of beta’, *Journal of Financial and Quantitative Analysis* pp. 833–857.

Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge Univ. Press.

Hastie, T. J. & Tibshirani, R. J. (1991), *Generalized Additive Models*, 2nd. edn, Chapman & Hall, London.

Hildreth, C. & Houck, J. P. (1968), ‘Some estimators for a linear model with random coefficients’, *Journal of the American Statistical Association* **63**, 584–595.

Karolyi, G. (1992), ‘Prediction risk: some new generalizations’, *Management Science* **38 (1)**, 57–74.

Krzanowski, W. J. (1988), *Principles of Multivariate Analysis: A User’s Perspective*. Oxford, signatura: 519.23 KRZ.

Lee, C. & Chen, C. (1982), ‘Beta stability and tendency. An application of a variable mean response regression model’, *Journal of Economics and Business* **34**, 201–206.

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.

Marín, J. & Rubio, G. (2001), *Economía Financiera*, Bosch.

Meyers, S. L. (1973), ‘The stationarity problem in the use of market model of security price behavior’, *The Accounting Review* **48**(2), 318–322.

Petris, G. (2007), *dlm: Bayesian and Likelihood Analysis of Dynamic Linear Models*. R package version 0.9-1.

Sarkar, D. (2008), *lattice: Lattice Graphics*. R package version 0.17-15.

Sharpe, W. (1964), 'Capital asset prices: A theory of market equilibrium', *Journal of Finance* **19**, 425–442.

Team, R. D. C. (2008), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Vasicek, O. (1973), 'A note on using cross-sectional information in Bayesian estimation of security betas', *Journal of Finance* **VIII**, 1233–1239.

Zeileis, A. & Grothendiek, G. (2005), 'Zoo: S3 infrastructure for regular and irregular time series', *Jss* **14**(6), 1–27.