Bayesian Robust Estimation of Systematic Risk Using Product Partition Models

Fernando A. Quintana* Pilar L. Iglesias† Manuel Galea-Rojas‡

March, 2005

Abstract

We consider Bayesian estimation of the systematic risk of a share using product partition models (PPM). The cluster structure of the PPM is used to derive a robust Bayes estimator of beta and also to identify outliers or clusters of observations. The procedure is implemented considering independent scale mixture of normals for the error terms. The results are illustrated with an application to a set of shares of companies in the Chilean stock market.

Key words: Bayesian inference, capital asset pricing model, outlier detection.

---

*Departamento de Estadística, Pontificia Universidad Católica de Chile. Corresponding author. Partially funded by Grant FONDECYT 1020712. e-mail: quintana@mat.puc.cl

†Departamento de Estadística, Pontificia Universidad Católica de Chile. Partially funded by Grant FONDECYT Líneas Complementarias 8000004. e-mail: pliz@mat.puc.cl

‡Departamento de Estadística, Universidad de Valparaíso, Chile. Partially funded by Grant FONDECYT 1000424. e-mail: manuel.galea@uv.cl

1 Introduction

Much interest in the field of financial economics is focused on efficient estimation of parameters of the return-generating process. The estimated beta (the slope in a simple linear regression model) from a capital asset pricing model (CAPM) is an important measure of risk for financial analysis and also for risk and portfolio managers. This parameter reflects the sensitivity of returns on an asset to movements in the market, and is very useful for calculating the capital of equities cost, a key quantity for the evaluation methods.

In finance literature, Bayesian approaches for the estimation of beta have received relatively little attention, as this is usually done within a classical framework. The beta parameter is typically estimated by the least squares method, which coincides with the maximum likelihood estimator under normality. However, such estimation method has at least two important limitations. First, no available prior information on the CAPM parameters is used. And secondly, it is well known (Chatterjee and Hadi 1988) that outliers (defined as a shift in the regression mean) and gross errors can distort the beta estimate.

The literature contains several references to problems such as the non-normality of returns and the lack of robustness of simple estimates (see, e.g., Fama 1965, Blattberg and Gonedes 1974, Zhou 1993). Lange et al. (1989) suggest the use of student-t distributions as a robust alternative to normality, illustrating its use in multivariate analysis and regression. More recently, Hodgson, Linton and Vorkink (2002) discuss estimation of linear asset pricing models, including the CAPM, under elliptical distributions. Cademartori et al. (2003) consider classical estimation of beta assuming independently student-t distributed share returns in the Chilean stock market. Brännäs and Nordman (2003) study conditional skewness modeling for stock returns. An alternative formulation using quantile regression under non-normal errors has been described in Min and Kim (2004). Bayesian solutions to related linear regression problems are considered in, e.g., Geweke (1993) and in Branco et al. (1998). Bayesian estimation of beta in the CAPM under normality has been considered by Harvey and Zhou (1990) and Polson and Tew (1999) in the context of portfolio selection.

We describe an alternative Bayesian method for estimating beta in the CAPM, which is robust to departures from normality. Specifically, we explore a product partition model (PPM) for scale mixture of normals regression. PPMs were introduced by Hartigan (1990) and Barry and Hartigan (1992) and we use them to achieve two goals: accommodating and identifying outliers, and detecting clusters of observations.

The article is organized as follows. Section 2 briefly reviews the CAPM and develops our method for the detection of influential returns in the estimator of systematic risk using PPM. In Section 3, the methodology is applied to a set of shares traded in the Santiago stock exchange market. Some final comments are stated in Section 4.

2 The capital asset pricing model (CAPM)

The CAPM states that the expected return on any asset or portfolio is related to the riskless rate of return and the expected market return via the expression:

$$E[R] = R_f + \beta(E[R_m] - R_f) \tag{1}$$

where R denotes the share return, R_f is the return on the risk free asset, β is the systematic risk of the asset under study, and R_m is the return on the market portfolio (given by an index). This model was independently derived by Sharpe (1964), Lintner (1965) and Mossin (1966) and can be useful in cases where a measure of expected stock returns is required. Applications include cost of capital estimation, portfolio performance evaluation, and event-study analysis. See Elton and Gruber (1995) and Campbell et al. (1997).

The usual estimator of beta of the equity is the OLS estimator of the slope coefficient in the excess-return market model, that is, beta in the regression equation

$$\begin{aligned} R_t - R_{ft} &= \alpha + \beta(R_{mt} - R_{ft}) + \epsilon_t \\ y_t &= \alpha + \beta x_t + \epsilon_t \end{aligned} \tag{2}$$

where R_t, R_{mt} and R_{ft} are the values of R, R_m and R_t for the t-th period; $y_t = R_t - R_{ft}$ represents the return on the asset in excess of the risk-free rate during the $t-$th period; $x_t = R_{mt} - R_{ft}$ is the excess return on the market portfolio of assets in the $t-$th period; and finally $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ represent random errors, $t = 1, 2, \ldots, n$. For the purpose of making inference, it is typically assumed that the share returns follow a normal distribution (Elton and Gruber 1995, Campbell et al. 1997). However, this method is known to be sensitive to atypical returns, which frequently arise in the real world, especially in Latin American markets (see for example, Duarte and Mendes 1997). Thus, outliers, leverage points, and gross errors can alter the beta estimate.

To deal with all the issues mentioned earlier, we adopt here a modeling approach that has two main components, which we describe next.

2.1 Main Components of The Model

2.1.1 Use of heavy-tailed distributions

Our approach to deal with atypical observations involves usage of symmetric distributions with heavier tails than the normal case, which is justified for this particular application by empirical evidence found by, e.g., Fama (1965), Blattberg and Gonedes (1974) and Zhou (1993). See also Hodgson et al. (2002) and Vorkink (2003). We consider here the subclass of elliptical distributions that are representable as scale mixture of normals, i.e., with density function $f(z|\mu, \sigma^2, h) = \sigma^{-1} h(\sigma^{-1}(z - \mu))$ and where $h(u) = \int_0^\infty (2\pi v)^{-1/2} e^{-u^2/(2v)}\, dG(v)$ for some distribution function G such that $G(0) = 0$. This class includes, as an important special case, the t distribution, which is useful due to its conceptual and computational simplicity (Taylor 1992, Geweke 1993).

2.1.2 Use of Partition Structures

A second important element in our modeling strategy is given by partition structures. Following the developments in Quintana and Iglesias (2003), the main idea is to allow α to change with t in (2) and group together those periods with identical α values. The groups are generally unknown so we need to define a probability model on partitions. Let $\rho = \{S_1, \ldots, S_k\}$ denote any partition of $S_0 = \{1, \ldots, n\}$. We adopt the class of *product distributions*, which is a convenient way to express uncertainty on the set $\mathcal{P}$ of all partitions ρ of S_0. The prior probability assigned to each partition $\rho \in \mathcal{P}$ is called product distribution if

$$P(\rho = \{S_1, \ldots, S_k\}) = \mathcal{K} \prod_{i=1}^{k} c(S_i), \tag{3}$$

where $c(s) \geq 0$ is a *cohesion* defined for each $S \subset S_0$, and $\mathcal{K}$ is a normalization constant.

2.2 The Model

We now combine the two aspects described in Section 2.1, considering the CAPM embedded in the structure provided by product distributions, with conditionally i.i.d. error terms distributed as scale mixture of normals. To do so, we assign the partition structure to the vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ of regression intercepts. Each partition $\rho = (S_1, \ldots, S_k)$ can be represented by means of group indicators $s_1, \ldots, s_n$, where $s_i = j$ if and only if $i \in S_j$. Denoting by $\alpha_1^*, \ldots, \alpha_k^*$ the set of unique values among coordinates of $\boldsymbol{\alpha}$ then $\alpha_t = \alpha_{s_t}^*$. Thus given ρ we assume $\boldsymbol{\alpha}$ adopts the form $\boldsymbol{\alpha}_\rho = (\alpha_{s_1}^*, \ldots, \alpha_{s_n}^*)$. An alternative representation of $\boldsymbol{\alpha}$ is $(\alpha_1^*, \ldots, \alpha_{|\rho|}^*, \rho)$, where $|\rho|$ is the number of elements in ρ.

Our modeling strategy considers that a priori, any point is a potential outlier. Identifying outliers is of practical interest, but we also view it as equivalent to selecting a partition ρ (or model M_ρ) to improve the estimation of β or to address another decision problem that is of specific interest. Given the above considerations, we postulate the following hierarchical model:

$$y_t|(\alpha_1^*,\ldots,\alpha_{|\rho|}^*,\rho),\beta,\sigma^2 \overset{ind}{\sim} N(\alpha_{s_t}^*+\beta x_t,\omega_t\sigma^2) \tag{4}$$

$$\omega_1,\ldots,\omega_n \overset{iid}{\sim} G \tag{5}$$

$$\alpha_1^*,\ldots,\alpha_{|\rho|}^*|\rho,\sigma^2 \overset{iid}{\sim} N(a,\tau_0^2\sigma^2) \tag{6}$$

$$\beta|\sigma^2\sim N(b,\gamma_0^2\sigma^2), \quad \rho \sim \text{product distribution}, \quad \sigma^2\sim IG(\nu_0,\lambda_0) \tag{7}$$

where $\boldsymbol{\omega}=(\omega_1,\ldots,\omega_n)$, assumed independent of $(\boldsymbol{\alpha},\rho,\sigma^2,\beta)$, are a auxiliary latent variables introduced to break the scale mixture of normals, and a, b, τ_0^2, γ_0^2, ν_0 and λ_0 are user-specified hyperparameters. Specification of the product distribution is done via defining the cohesions $c(S)$. A useful choice is $c(S)=c\times(|S|-1)!$ for some positive constant c. Indeed, Quintana and Iglesias (2003) argued that such cohesions tend to produce a reduced number of large clusters. This is desirable for outlier identification, where one typically does not wish to label too many observations as "outliers". Finally, we point out that (4) plus the specification of a product distribution as in (7) defines what is called a *product partition model* (PPM), a class of models that was introduced by Hartigan (1990) and Barry and Hartigan (1992).

2.3 Outlier detection using product partition models

To detect outliers under model (4) – (7) we use the algorithm presented in Quintana and Iglesias (2003). We aim at selecting a partition which ideally has the form of a large cluster of "normal" or typical observations and a reduced number of clusters with a small number of observations, representing the groups of outliers. The main idea is that each possible partition is a different model and the best model is chosen by minimizing a certain loss function.

Following Quintana and Iglesias (2003) we choose a loss function that combines the estimation and partition selection problems. Let $\boldsymbol{\eta}=(\beta,\sigma^2)$, and let $l(\rho,\boldsymbol{\alpha},\boldsymbol{\omega},\boldsymbol{\eta})$ be the loss originated by the successive choice of partition ρ and action a_ρ when $(\boldsymbol{\alpha},\boldsymbol{\omega},\boldsymbol{\eta})$

is the real "state of the world". Concretely, we choose the following loss function:

$$l(\rho, \boldsymbol{\alpha}_\rho, \boldsymbol{\omega}_\rho, \boldsymbol{\eta}_\rho, \boldsymbol{\alpha}, \boldsymbol{\omega}, \boldsymbol{\eta}) = \frac{\kappa_1}{n} \parallel \boldsymbol{\alpha}_\rho - \boldsymbol{\alpha} \parallel^2 + \frac{\kappa_2}{n} \parallel \boldsymbol{\omega}_\rho - \boldsymbol{\omega} \parallel^2 + \\ \kappa_3(\beta_\rho - \beta)^2 + \kappa_4(\sigma^2_\rho - \sigma^2)^2 + (1 - \kappa_1 - \kappa_2 - \kappa_3 - \kappa_4)|\rho|.$$

Here $\parallel \cdot \parallel$ is the Euclidean norm and $0 \leq \kappa_i$, $\sum_{i=1}^4 \kappa_i \leq 1$ are constants chosen by the user to represent the trade-off between accurate parameter estimation and model complexity. Following Quintana and Iglesias (2003) it can be shown that the expected loss minimization criterion leads to the choice ρ^* that minimizes

$$SC_{\kappa_1,\kappa_2,\kappa_3,\kappa_4}(\rho) = \frac{\kappa_1}{n} \parallel \hat{\boldsymbol{\alpha}}_B(\boldsymbol{y}) - \hat{\boldsymbol{\alpha}}_\rho(\boldsymbol{y}) \parallel^2 + \frac{\kappa_2}{n} \parallel \hat{\boldsymbol{\omega}}_B(\boldsymbol{y}) - \hat{\boldsymbol{\omega}}_\rho(\boldsymbol{y}) \parallel^2 + \\ \kappa_3(\hat{\beta}_B(\boldsymbol{y}) - \hat{\beta}_\rho(\boldsymbol{y}))^2 + \kappa_3(\hat{\sigma}^2_B(\boldsymbol{y}) - \hat{\sigma}^2_\rho(\boldsymbol{y}))^2 + (1 - \kappa_1 - \kappa_2 - \kappa_3 - \kappa_4)|\rho|, \quad (8)$$

where $\hat{\xi}_B(\boldsymbol{y}) = E(\xi|\boldsymbol{y})$ and $\hat{\xi}_\rho(\boldsymbol{y}) = E(\xi|\boldsymbol{y}, \rho)$ and ξ generically denotes any of the parameters in (8). Note that the Bayes estimate of $\boldsymbol{\alpha}$ can be theoretically expressed as an average over all possible partitions ρ: $\hat{\boldsymbol{\alpha}}_B(\boldsymbol{y}) = \sum_\rho \hat{\boldsymbol{\alpha}}_\rho(\boldsymbol{y}) P(\rho|\boldsymbol{y})$. In practice, however, the estimation is implemented via MCMC strategies. The basic idea exploited in Quintana and Iglesias (2003) relies on the connection between PPMs with cohesion given by $c(S) = c \times (|S| - 1)!$, and Bayesian nonparametric models (see, e.g., Dey et al. 1998) with a Dirichlet process prior (Ferguson 1973). This connection is a key element for adapting the Gibbs sampling algorithms proposed by Bush and MacEachern (1996) for nonparametric Bayesian inference to our specific model. See details in the Appendix.

The algorithm starts with the partition formed by S_0 only, and proceeds to detach one by one elements from S_0, placing these in new sets or joining previously detached elements. This process continues until no further improvement is achieved. See further details in Quintana and Iglesias (2003).

Table 1: *Posterior expectation for β and 95% credibility intervals under different distributional assumptions for the error term.*

Company Name	Error Distribution				
	Normal	$t(30)$	$t(4)$	$t(1)$	Normal (no cluster)
Cementos	0.659	0.664	0.669	0.655	0.906
Bío-Bío S.A.	(0.504,0.823)	(0.512,0.825)	(0.512,0.838)	(0.524,0.812)	(0.868,0.945)
Cía. Cervecerías	0.843	0.839	0.804	0.793	0.862
Unidas	(0.682,0.999)	(0.678,0.998)	(0.624,0.985)	(0.590,1.006)	(0.830,0.894)
CHILECTRA	0.884	0.885	0.941	0.966	0.882
	(0.758,1.010)	(0.754,1.015)	(0.796,1.066)	(0.825,1.084)	(0.855,0.907)
CHOLGUAN	0.777	0.777	0.787	0.762	0.963
	(0.587,0.956)	(0.590,0.955)	(0.602,0.966)	(0.548,0.990)	(0.912,1.014)
Viña Concha y	0.650	0.638	0.571	0.468	0.932
Toro S.A.	(0.505,0.790)	(0.495,0.780)	(0.417,0.732)	(0.329,0.619)	(0.889,0.975)
ENDESA	1.090	1.096	1.120	1.131	1.090
	(0.996,1.186)	(1.003,1.189)	(1.021,1.214)	(1.041,1.217)	(1.071,1.110)
ENTEL	1.107	1.097	1.071	1.047	1.108
	(0.903,1.303)	(0.901,1.292)	(0.884,1.252)	(0.818,1.260)	(1.069,1.148)
IANSA	0.734	0.739	0.702	0.645	0.821
	(0.559,0.912)	(0.559,0.917)	(0.507,0.886)	(0.381,0.836)	(0.785,0.856)
INFORSA	0.825	0.825	0.846	0.839	0.972
	(0.682,0.975)	(0.682,0.976)	(0.697,1.012)	(0.660,1.061)	(0.936,1.008)
MARINSA	0.486	0.486	0.488	0.509	0.746
	(0.354,0.624)	(0.357,0.625)	(0.361,0.615)	(0.372,0.653)	(0.712,0.780)
PASUR	0.335	0.333	0.338	0.314	0.474
	(0.226,0.465)	(0.226,0.457)	(0.231,0.450)	(0.194,0.429)	(0.445,0.503)
CCT	0.552	0.556	0.549	0.576	0.599
	(0.424,0.681)	(0.429,0.689)	(0.420,0.678)	(0.462,0.674)	(0.571,0.627)
CAP	0.865	0.857	0.834	0.770	1.105
	(0.671,1.055)	(0.669,1.040)	(0.627,1.043)	(0.611,0.960)	(1.059,1.151)
MINERA	0.282	0.277	0.256	0.214	0.412
	(0.181,0.388)	(0.174,0.385)	(0.154,0.361)	(0.114,0.324)	(0.386,0.438)
CTI	0.817	0.821	0.823	0.799	0.927
	(0.653,0.981)	(0.654,0.984)	(0.648,0.998)	(0.612,0.998)	(0.894,0.961)
COPEC	0.840	0.842	0.834	0.820	0.886
	(0.731,0.951)	(0.730,0.958)	(0.724,0.948)	(0.715,0.934)	(0.861,0.909)
CGE	0.601	0.604	0.616	0.568	0.751
	(0.462,0.751)	(0.463,0.755)	(0.457,0.772)	(0.407,0.759)	(0.720,0.783)
VAPORES	0.483	0.484	0.557	0.638	0.513
	(0.289,0.721)	(0.300,0.714)	(0.371,0.741)	(0.461,0.810)	(0.479,0.547)
CMPC	0.065	0.079	0.123	0.198	0.147
	(-0.088,0.224)	(-0.074,0.239)	(-0.030,0.279)	(0.061,0.330)	(0.114,0.180)

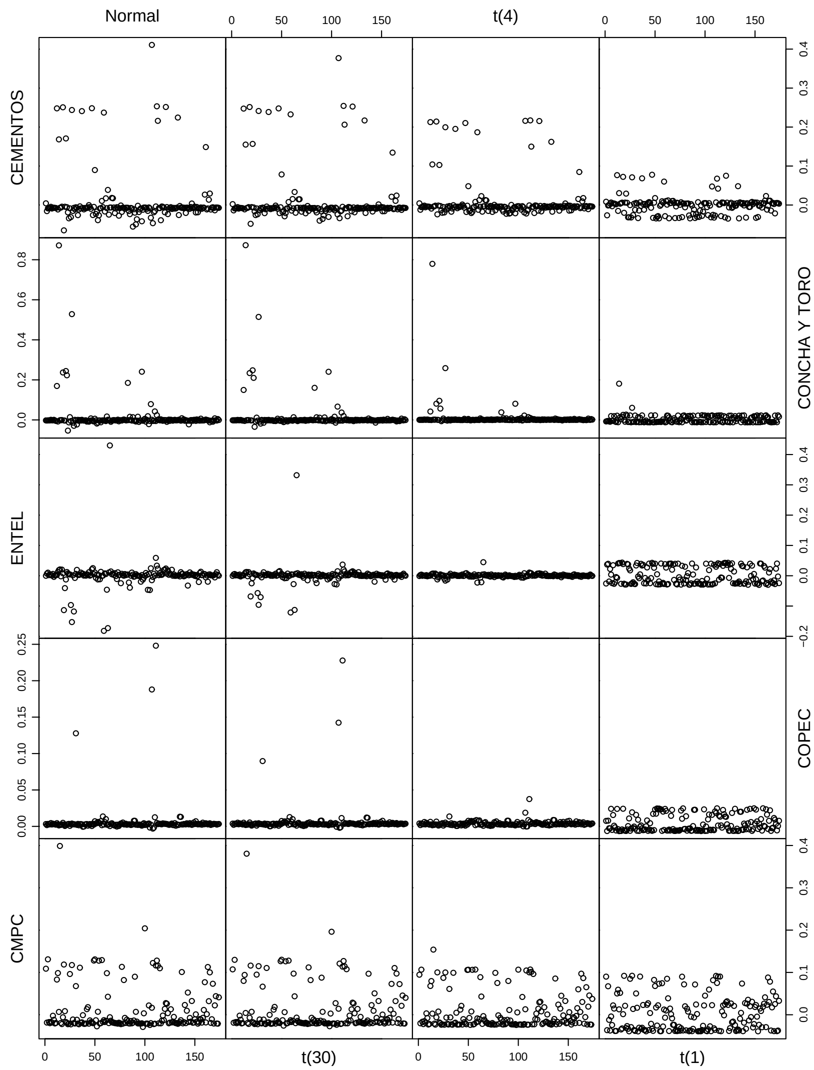
3 Robust estimation of the systematic risk in the Chilean stock market

We consider now data corresponding to monthly excess returns, adjusted by patrimonial variation, of 19 shares from the Chilean stock market, covering a wide range of economic activities, including small, mid-sized and large companies. We use the selective index of prices of shares (IPSA) as a measure of the market returns and the interest rate in sale of discount bonus of the Central Bank as the risk free rate. The data are available on a monthly basis, totaling 174 data points, from January 1990 through June 2004. A list of the 19 companies is included in Table 1.

In what follows we assume model (4) – (7) with G specified as $IG(\frac{d}{2}, \frac{d}{2})$, thus leading to $t(d)$-distributed errors, and the following hyperparameter values: $a = 0$, $b = 1$, $\tau_0^2 = \gamma_0^2 = 1000$, $\nu_0 = 2.0001$ and $\lambda_0 = 0.010001$. These choices imply prior expectation 0.01 and variance 1 for σ^2, reflecting what is known from past experience about the extremely small variability of stock returns, but trying to avoid being excessively informative on key parameters, namely, β and the α_i coefficients. Also, the choices for a and b are based on assessments made by market analysts, e.g., as discussed in Cademartori et al. (2003) and in Polson and Tew (1999).

Table 1 shows the posterior mean for β and the corresponding 95% credibility intervals under normal and $t(d)$ distributions for the error term. To facilitate comparison, the case of normal errors *without* cluster structure (i.e., a standard Bayesian linear regression model) is included.

Figure 1: *Estimated posterior means of intercept parameter α for five selected cases. Columns represent regression model with $t(d)$ errors and $d \in \{\infty, 30, 4, 1\}$; rows represent the indicated companies.*



There is a notorious variability among different companies. While some of them exhibit little changes in the β estimates (e.g. Entel or Copec), some others have moderate to large variability across choices of d (i.e. Concha y Toro or CMPC). The stability under error distributions seems to be correlated with the size of the company. Indeed, the figures suggest that big companies, which have substantial impact in the stock market, are much more stable under market fluctuations. At the same time, some (but not all) of the small companies have significant variations on the β estimates. We also note that the 95% credibility intervals generally tend to increase their lengths as d increases. This seems natural because the $t(d)$ distribution has heavier tails for smaller values of d, which implies more robustness in the estimation of β. In other words, the corresponding marginal posterior distribution of the systematic risk coefficients are more dispersed as d increases.

Figure 1 shows the estimated posterior means of the whole set of intercept coefficients $\{\alpha_i\}$ for five selected companies and $d \in \{1, 4, 30, \infty\}$. These coefficients are precisely the coordinates of the Bayes estimate $\hat{\alpha}_B(\boldsymbol{y})$ as discussed in Section 2.3. It is clear that the variability in the posterior means tends to increase with d, but this trend is reverted in the case of some companies when $d = 1$. In any case, these plots are important because the values of $\hat{\alpha}_B(\boldsymbol{y})$ are used to determine the order in which different partitions will be assessed, i.e. they play a key role in the clustering algorithm. Thus the effect of increasing d is the potential creation of additional clusters of data-points.

We consider next the identification of outliers via the explicit construction of clusters of months. To do so, we use the algorithm that minimizes (8) with $(\kappa_1, \kappa_2, \kappa_3, \kappa_4) = \frac{1}{2013}(1000, 1, 1000, 1)$. This choice is intended to give priority to estimation of β and $\boldsymbol{\alpha}$, setting almost no restriction on the formation of clusters, so that we can effectively detect outliers. It also considers little weight on the σ^2 parameter (which is far less important than β) and almost no weight on the latent $\{\omega_i\}$ variables. The results are given in Table 2 and they are in agreement with what is seen in Figure 1 and in Table 1.

The five companies are representative of the overall behavior. In all the cases reported, the number of detected outliers decreases as d decreases, going down to zero when $d = 1$, which reflects the robustness mentioned earlier. In some cases, the number

Table 2: *Best partitions determined by the clustering algorithm for six selected companies. We indicate here only the subsets detached from the main body of points. The symbol* $\emptyset$ *denotes that no outlier was identified*

Company	Distribution of Errors			
	Normal	$t(30)$	$t(4)$	$t(1)$
CONCHA Y TORO	{14},{21},{27}	{14},{21},{27}	{14},{21}, {27}	$\emptyset$
CEMENTOS	{12,18,27,37, 112,121},{21,113} {47,59,133},{107}	{12,27,37,47,112, 121,133},{18,107} {21},{59,113}	{18,27,37,59, 107,133},{113} {12,47,112},{121}	$\emptyset$
ENTEL	{65}	{65}	$\emptyset$	$\emptyset$
COPEC	{31,107},{111}	{111}	$\emptyset$	$\emptyset$
CMPC	{15}	{15}	{15}	$\emptyset$

of detected outliers is quite small and occurring only when the distribution of errors is normal or $t(30)$ (e.g. ENTEL and COPEC). This suggests that heavier tails in the errors distribution are capable to accommodate all the variability exhibited by the data, and that no partition structure is needed. In some other cases, the number of different groups is large (e.g. CEMENTOS), suggesting that these companies are much more sensitive to market variations. Interestingly, the outliers usually correspond to different months, revealing that companies have different responses to the market's behavior. The particular case of CEMENTOS is worth noting because the groups of outliers change with d, reflecting the relative variations in the estimated posterior means of all parameters, particularly the α_i intercepts.

Finally, it is important to point out that the choice of κ parameters could modify the final selection of groups. For instance, in the case of COPEC with $d = 30$ or $d = 4$, setting $\kappa_1 = 1$ detects more outliers than reported in Table 2. The reason for this is that such choice completely ignores estimation of all parameters but the α_i coefficients, and so the groups simply follow the patterns exhibited in Figure 1. Therefore, usage of the algorithm requires a prior assessment of the relative weight one is willing to assign to the different components of the decision problem.

4 Conclusions

We studied robust inference for the systematic risk β in CAPM. The basic idea is to consider models with the ability of detecting the presence of groups of observations y_i, including the case of outliers, via a partition structure. Using error distributions with heavy tails is useful because the financial market is particularly sensitive to external effects. The application in Section 3 suggests that outlying points can be accommodated either by a normal model with partition structure of the PPM type or by a simple regression model with student-$t(d)$ error distribution with small (or moderate) d.

The prior distributions for structural parameters were specified according to the opinion of experts in the field and the related literature. Although it is not our main goal, we carried out a small study of the sensitivity of inferences to changes in hyperparameter specifications. The analysis (data not shown) suggests that posterior inference for structural parameters is indeed sensitive to changes in hyperparameters, but the selected partition is quite stable. A possible explanation is the relatively high dimensionality of the parameter vector. Additional discussion about this issue in the context of change-point identification can be found in Loschi et al. (2003).

Appendix: Full Conditional Distributions

Under model (4) - (7) and denoting by g the density of G we have that

$$\beta|\sigma^2, \boldsymbol{\alpha}, \boldsymbol{y}, \boldsymbol{\omega} \sim N\left(\frac{\frac{b}{\gamma_0^2} + \sum \frac{(y_i-\alpha_i)x_i}{\omega_i}}{\frac{1}{\gamma_0^2} + \sum \frac{x_i^2}{\omega_i}}, \frac{\sigma^2}{\frac{1}{\gamma_0^2} + \sum \frac{x_i^2}{\omega_i}}\right), \tag{9}$$

$$\sigma^2|\boldsymbol{\alpha}, \beta, \boldsymbol{y}, \boldsymbol{\omega} \sim IG\Bigg(\nu_0 + \frac{n+k+1}{2}, \\ \lambda_0 + \frac{(\beta-b)^2}{2\gamma_0^2} + \frac{1}{2\tau_0^2}\sum_{j=1}^{|\rho|}(\alpha_j^* - a)^2 + \frac{1}{2}\sum_{i=1}^{n}\frac{(y_i - \alpha_i - \beta x_i)^2}{\omega_i}\Bigg), \tag{10}$$

$$p(\alpha_i|\boldsymbol{\alpha}_{-i}, \beta, \sigma^2, \boldsymbol{y}, \boldsymbol{\omega}) \propto \sum_{j \neq i} \exp\left\{-\frac{(y_i - \alpha_j - \beta x_i)^2}{2\sigma^2\omega_i}\right\}\delta_{\alpha_j}(\alpha_i) \\ + \frac{\exp\{-\frac{(y_i - \beta x_i - a)^2}{2\sigma^2(\omega_i + \tau_0^2)}\}}{\sqrt{1 + \frac{\tau_0^2}{\omega_i}}} \times N\left(\frac{\frac{y_i - \beta x_i}{\omega_i} + \frac{a}{\tau_0^2}}{\frac{1}{\omega_i} + \frac{1}{\tau_0^2}}, \frac{\sigma^2}{\frac{1}{\omega_i} + \frac{1}{\tau_0^2}}\right), \tag{11}$$

$$p(\omega_i|\boldsymbol{\omega}_{-i}, \boldsymbol{\alpha}, \beta, \sigma^2, \boldsymbol{y}) \propto \exp\left\{\frac{(y_i - \alpha_i - \beta x_i)^2}{2\sigma^2\omega_i}\right\} g(\omega_i). \tag{12}$$

Before proceeding to the next Gibbs iteration the locations are updated by independently drawing from

$$\alpha_j^* \sim N\left(\frac{\sum_{i \in S_j}\frac{(y_i - \beta x_i)}{\omega_i} + \frac{a}{\tau_0^2}}{\sum_{i \in S_j}\frac{1}{\omega_i} + \frac{1}{\tau_0^2}}, \frac{\sigma^2}{\sum_{i \in S_j}\frac{1}{\omega_i} + \frac{1}{\tau_0^2}},\right), \qquad j = 1, \dots, |\rho|.$$

When $\omega_i \equiv 1$ for $i = 1, \dots, n$ (normal model) we have that

$$\alpha_j^* \sim N\left(\frac{\sum_{i \in S_j}(y_i - \beta x_i) + \frac{a}{\tau_0^2}}{|S_j| + \frac{1}{\tau_0^2}}\ ,\ \frac{\sigma^2}{|S_j| + \frac{1}{\tau_0^2}}\right), \qquad j = 1, \dots, |\rho|$$

and when $\omega_1, \dots, \omega_n \overset{i.i.d.}{\sim} IG(\frac{d}{2}, \frac{d}{2})$ (t-model) we have that

$$\omega_i|\boldsymbol{\omega}_{-i}, \boldsymbol{\alpha}, \beta, \sigma^2, \boldsymbol{y} \sim IG\left(\frac{d}{2} + \frac{1}{2}, \frac{d}{2} + \frac{(y_i - \alpha_i - \beta x_i)^2}{2\sigma^2}\right).$$

References

Barry, D. and Hartigan, J. A. (1992). Product Partition Models for Change Point Problems, *The Annals of Statistics* **20**(1): 260–279.

Blattberg, R. C. and Gonedes, N. J. (1974). A comparison of stable and student distribution as statistical models for stock prices, *Journal of Business* **47**: 244–280.

Branco, M., Bolfarine, H. and Iglesias, P. (1998). Bayesian calibration under a Student-t model, *Computational Statistics* **13**: 319–338.

Brännäs, K. and Nordman, N. (2003). Conditional skewness modelling for stock returns, *Applied Economics Letters* **10**: 725–728.

Bush, C. A. and MacEachern, S. N. (1996). A semiparametric Bayesian model for randomised block designs, *Biometrika* **83**(2): 275–285.

Cademartori, D., Romo, C., Campos, R. and Galea, M. (2003). Robust estimation of systematic risk using the t distribution in the Chilean Stock Markets, *Applied Economics Letters* **10**: 447–453.

Campbell, J., Lo, A. and MacKinlay, A. (1997). *Econometrics of Financial Markets*, New Jersey: Princeton University Press.

Chatterjee, S. and Hadi, A. S. (1988). *Sensitivity analysis in linear regression*, Wiley.

Dey, D., Müller, P. and Sinha, D. (eds) (1998). *Practical Nonparametric and Semiparametric Bayesian Statistics*, New York: Springer-Verlag.

Duarte, A. M. and Mendes, B. (1997). Robust estimation of systematic risk in emerging stock markets, *Emerging Markets Quarterly* **1**: 85–95.

Elton, E. J. and Gruber, M. J. (1995). *Modern Portfolio Theory and Investment Analysis, 5th Edition*, New York: Wiley.

Fama, E. (1965). The behavior of stock market prices, *Journal of Business* **38**: 34–105.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems, *The Annals of Statistics* **1**: 209–230.

Geweke, J. (1993). Bayesian treatment of the independent Student-t linear model, *Journal of Applied Econometrics* **8S**: 19–40.

Hartigan, J. A. (1990). Partition Models, *Communications in Statistics – Theory and Methods* **19**(8): 2745–2756.

Harvey, C. R. and Zhou, G. (1990). Bayesian inference in asset pricing tests, *Journal of Financial Economics* **26**: 221–254.

Hodgson, D. J., Linton, O. and Vorkink, K. (2002). Testing the capital asset pricing model efficiently under elliptical symmetry: a semiparametric approach, *Journal of Applied Econometrics* **17**: 617–639.

Lange, K. L., Little, R. J. A. and Taylor, J. M. G. (1989). Robust statistical modeling using the t-distribution, *Journal of the American Statistical Association* **84**: 881–896.

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* **41**: 13–37.

Loschi, R., Cruz, F. R. B., Iglesias, P. L. and Arellano-Valle, R. B. (2003). A Gibbs sampling scheme to the product partition model: An application to change point problems, *Computers and Operations Research* **30**: 463–482.

Min, I. and Kim, I. (2004). A Monte Carlo comparison of parametric and nonparametric quantile regressions, *Applied Economics Letters* **11**: 71–74.

Mossin, J. (1966). Equilibrium in capital asset market, *Econometrica* **35**: 768–783.

Polson, N. and Tew, B. (1999). Bayesian Portfolio Selections: an Analysis of the SP500 index 1970-1996, *Journal of Business and Economic Statistics* **18**: 164–173.

Quintana, F. A. and Iglesias, P. L. (2003). Bayesian Clustering and Product Partition Models, *Journal of the Royal Statistical Society Series B* **65**: 557–574.

Sharpe, W. (1964). Capital asset prices: A theory of markets equilibrium under conditions of risk, *Journal of Finance* **19**: 425–442.

Taylor, J. (1992). Properties of modelling the error distribution with an extra shape parameter, *Computational Statistics and Data Analysis* **13**: 33–46.

Vorkink, K. (2003). Return Distributions and Improved Tests of Asset Pricing Models, *The Review of Financial Studies* **16**: 845–874.

Zhou, G. (1993). Asset-pricing tests under alternative distributions, *The Journal of Finance* **48**: 1927–1942.