Quaderni di Dipartimento

Optimal clustering in Bayesian Capital Asset Pricing Model

Maria Elena De Giuli
(Università di Pavia)

Claudia Tarantola
(Università di Pavia)

Pierpaolo Uberti
(Università di Pavia)

197 (04-07)

Dipartimento di economia politica
e metodi quantitativi
Università degli studi di Pavia
Via San Felice, 5
I-27100 Pavia

Aprile 2007

Optimal clustering in Bayesian Capital Asset Pricing Model

Maria Elena De Giuli
University of Pavia, Italy

Claudia Tarantola *
University of Pavia, Italy

Pierpaolo Uberti
University of Milano Bicocca, Italy

Abstract

We focus on robust Bayesian estimation of the systematic risk of an asset in presence of outlying points. We assume that the returns follow independent normal distributions with a product partition structure on the parameters of interest. A Bayesian decision theoretical approach is used to identify the partition that best separates standard and atypical data points. We apply a nonsmooth optimization algorithm to minimize the expected value of a given loss function. The methodology is illustrated with reference to the IPSA stock market index and the MIBTEL one.

Keywords: Capital Asset Pricing Model, Markov Chain Monte Carlo, outlier identification, product partition models, score function

1 Introduction

Following Quintana and Iglesias (2003) and Quintana *et all.* (2005a) we focus on Bayesian robust estimation of the systematic risk in Capital Asset Pricing Model (CAPM), see Sharpe (1964), Lintner (1965), Mossin (1966) and Black (1972).

The CAPM is a simple linear regression model relating the asset expected return to the market portfolio one. It takes into account the two different components of a portfolio risk: the systematic and the specific one. The systematic risk, also called

*Address for correspondence: Claudia Tarantola, Dipartimento di Economia Politica e Metodi Quantitativi, Via S. Felice 7, 27100 Pavia, Italy. `E-mail: claudia.tarantola@unipv.it`

covariance or market risk, is the risk of holding the market portfolio and cannot be eliminated in a diversified portfolio. It measures the sensitivity of an asset return to movement in the market and it corresponds to the slope of the regression model. Whereas, the specific risk, also called residual or non market risk, is unique to an individual asset and can be eliminated in a diversified portfolio. It represents the component of an asset return which is uncorrelated with general market moves and corresponds to the intercept of the model.

Almost all empirical analysis have been carried out in the classical framework. The systematic risk is usually estimated by the least square method which coincide with the maximum likelihood estimator under the assumption of normality. This approach has at least two disadvantages. Firstly, it is not possible to incorporate in the model prior beliefs about the data behaviour. Secondly, such estimation method is sensitive to the presence of atypical data, i.e. outliers (shift in the regression mean), leverage points and gross errors (see e.g. Chatterjee and Hadi, 1988).

A variety of methodologies have been proposed to take into account the presence of atypical returns, here we focus on robust estimation procedures. From a classical prospective the problem of robust linear estimation has been considered, among the others, by Huber, P.J. (1973), Rousseeuw and Leroy (1987), Lange *et all.* (1989), Staudta and Sheather (1990) and Cademartori *et all.* (2003). Whereas from a Bayesian point of view relevant works are e.g. Chaturvedi (1996), Fernández *et all.* (2001) Quintana and Iglesias (2003) and Quintana *et all.* (2005a, 2005b).

In particular, Quintana and Iglesias (2003) and Quintana *et all.* (2005a) show that outlying point can be accommodated either by a normal model with a product partition structure or by a simple regression model with t shape errors with small (or moderate) degrees of freedom d. In this paper we follow the former approach, hence we remain in a normal setting consistent with a Mean-Variance analysis even in presence of outlying points.

The use of a partition structure do not only allow to accommodate for anomalous points but also to identify them. Quintana and Iglesias (2003) apply a clustering algorithm to select the partition that best resemble, in terms of a quadratic score function, the Bayesian estimates of the parameters of interest. As they pointed out a weakness of their algorithm is that it could be trapped in local modes. Following their suggestion (see page 572 in Quintana and Iglesias, 2003) we apply an optimization algorithm to move efficiently in the space of all candidate solutions.

The methodology is illustrated with reference to the Chilean Stock Price Index (Índice de Precios Selectivo de Acciones, or IPSA) and the MIBTEL (Milano Indice Borsa Telematica) one. With reference to the IPSA data, we compare our results with the ones obtained by Quintana *et all.* (2005a). Our algorithm shows a better performance in terms of the optimal value of the score function. Both for the IPSA and the MIBTEL data, we provide a microeconomic analysis of the outliers.

The plan of the paper is the following. In Section 2 we briefly introduce the CAPM and present the product partition model (PPM) used for outliers accommodation and detection. In Section 3 we describe the optimization algorithm considered. In Section 4 we apply our methodology to the IPSA and MIBTEL data. Some final comments are stated in Section 5.

2 Background and preliminaries

The CAPM states that the asset expected return is a linear function of the market portfolio one. We use a time series regression to evaluate the return of a generic asset i for the t-th period in excess of the risk free-rate,

$$\begin{aligned} R_{i_t} - R_{f_t} &= \alpha_i + \beta_i (R_{m_t} - R_{f_t}) + \varepsilon_{i_t}, \\ y_{i_t} &= \alpha_i + \beta_i x_t + \varepsilon_{i_t} \quad i = 1, \dots, N \quad t = 1, \dots, T. \end{aligned} \tag{1}$$

In equation (1), R_{i_t} is the asset return, R_{f_t} is the risk-free rate return, R_{m_t} is the return on a market proxy, ε_{i_t} is an error term, α_i and β_i are parameters to be estimated. The coefficient β_i measures the systematic risk of the asset i, while the coefficient α_i defines whether the asset i outperforms the market index. Following Quintana and Iglesias (2003), we allow α_i to change with t, that is

$$y_{i_t} = \alpha_{i_t} + \beta_i x_t + \varepsilon_{i_t}, \tag{2}$$

and we group together the data with identical α_{i_t} values.

The vector $\boldsymbol{\alpha}_i = (\alpha_{i_1}, \ldots, \alpha_{i_T})'$ is identified by the corresponding time points, that is $S_0 = \{1, \ldots, T\}$. Let $(\alpha^*_{i_1}, \ldots, \alpha^*_{i_{|\rho|}})'$ be the vector of distinct values of α_i, and $\rho = \{S_1, \ldots, S_{|\rho|}\}$ be the corresponding partition of S_0 with $S_d = \left\{t : \alpha_{i_t} = \alpha^*_{i_d}\right\}$. Note that, given a finite set $\mathcal{S}$ we indicate with $|\mathcal{S}|$ the number of elements in $\mathcal{S}$, that is its cardinality. Since the groups of data are generally unknown we need to define a probability model on the set $\mathcal{P}$ of all partitions. To this end we rely on PPM (Hartigan, 1990), with special reference to its parametric version see Barry and Hartigan (1992). More precisely on each partition of the set S_0 we assign a prior probability given by

$$P(\rho = \{S_1, \ldots, S_{|\rho|}\}) = K \prod_{d=1}^{|\rho|} C(S_d), \tag{3}$$

where $C(S_d)$ is a cohesion function and K is the normalizing constant. Equation (3) is referred to as the product distribution for partitions. The cohesions represent prior weight on group formation and $C(S_d)$ can be thought of as formalizing our opinion on how tightly clustered the elements of S_d would be. The cohesions can be specified in different ways, an useful choice is

$$C(S_d) = c \times (|S_d| - 1)!, \tag{4}$$

for some positive constant c.

Indeed, Quintana and Iglesias (2003) argued that for moderate values of c, e.g. $c = 1$ or $c = 2$, these cohesions yield a prior distribution that favours the formation of partitions with a reduced number of large subsets. See e.g. Liu (1996) for the relation between the choice of c and the prior mean and variance of the number of clusters. For more details on the choice of c see also Quintana and Iglesias (2003), Quintana *et all.* (2005b) and Tarantola *et all.* (2007).

There is an interesting connection between parametric PPMs and the class of Bayesian nonparametric models based on mixture of Dirichlet Processes (Antoniak, 1974). Under the latter prior, the marginal distribution of the observables is a specific PPM with the cohesion functions specified by equation (4), see Quintana and Iglesias (2003). Efficient Markov Chain Monte Carlo (MCMC) algorithms have been developed for Bayesian nonparametric problems based on Mixtures of Dirichlet Processes, as the one by Bush and MacEachern (1996) that we apply here.

We consider the following Bayesian hierachical model

$$
\begin{aligned}
&y_{i_t}|\rho,(\alpha^*_{i_1},\dots,\alpha^*_{i_{|\rho|}}),\beta_i,\sigma^2_i \overset{ind}{\sim} N\left(\alpha_{i_t}+\beta_i x_t,\sigma^2_i\right)\\
&\alpha^*_{i_1},\dots,\alpha^*_{i_{|\rho|}}|\rho,\sigma^2_i \overset{IID}{\sim} N\left(a,\tau^2_0\sigma^2_i\right)\\
&\beta_i|\sigma^2_i \sim N\left(b,\gamma^2_0\sigma^2_i\right)\\
&\rho \sim \text{product distribution}\\
&\sigma^2_i \sim IG(v_0,\lambda_0)
\end{aligned}
$$

where a, b, τ^2_0, γ^2_0, v_0 and λ_0 are user-specified hyperparameters, the product distribution is defined in (3) and $IG(v_0,\lambda_0)$ is an inverted gamma distribution with $E(\sigma^2_i)=\lambda_0/(v_0-1)$. In the Appendix we describe the Gibbs algorithm applied to sample from the posterior distributions of the parameters of interest.

3 A nonsmooth optimization algorithm

As proposed in Quintana and Iglesias (2003) we work in a Bayesian decision theoretic framework, and we choose a loss function that combines estimation and the partition selection problem.

Given a generic asset i, let $\left(\boldsymbol{\alpha}_i,\beta_i,\sigma^2_i\right)$ be the vector of parameters of the model and $\left(\boldsymbol{\alpha}_{i\rho},\beta_{i_\rho},\sigma^2_{i\ \rho}\right)$ be the corresponding vector that results when fixing ρ. We consider the following loss function

$$
\begin{aligned}
L(\rho,\boldsymbol{\alpha}_{i\rho},\beta_{i_\rho},\sigma^2_{i\ \rho},\boldsymbol{\alpha}_i,\beta_i,\sigma^2_i) = \frac{k_1}{T}\ \parallel \boldsymbol{\alpha}_{i\rho}-\boldsymbol{\alpha}_i \parallel^2\ +k_2\ (\beta_{i_\rho}-\beta_i)^2\ +\\
+k_3\ (\sigma^2_{i_\rho}-{\sigma_i}^2)^2\ +\ (1-k_1-k_2-k_3)|\rho|, \qquad (5)
\end{aligned}
$$

where $\parallel\cdot\parallel$ is the Euclidean norm and k_i $(i=1,2,3)$ are positive cost-complexity parameters with $\sum_{i=1}^3 k_i \le 1$.

It follows (Quintana and Iglesias, 2003) that minimizing the expected value of (5) is equivalent to choosing the partition ρ^* that minimizes the following score function

$$
\begin{aligned}
SC(\rho) = \frac{k_1}{T}\ \parallel \hat{\boldsymbol{\alpha}}_{iB}(y)-\hat{\boldsymbol{\alpha}}_{i\rho}(y)\parallel^2\ +k_2\ (\hat{\beta}_{iB}(y)-\hat{\beta}_{i_\rho}(y))^2\ +\\
+k_3\ (\hat{\sigma}^2_{iB}(y)-\hat{\sigma}^2_{i\rho}(y))^2\ +\ (1-k_1-k_2-k_3)|\rho|. \qquad (6)
\end{aligned}
$$

In (6), a subscript "B" means that we consider the Bayesian estimates of the corresponding parameter whereas a subscript "ρ" indicates the estimate of the parameter (or vector of parameters) conditionally on a partition ρ. Formally, if we indicate with θ a generic parameter in (6), we get $\hat{\theta}_B(y) = E(\theta|y)$ and $\hat{\theta}_\rho(y) = E(\theta|y, \rho)$. The Bayesian estimate of θ is obtained via the MCMC method described in the Appendix. As pointed out by Quintana and Iglesias (2003) also the evaluation of any particular $\hat{\theta}_\rho(y)$ may itself require the use of MCMC methods. These are structurally simpler that the one described in the Appendix since the partition ρ is fixed.

Minimization of the score function (6) requires finding the partition ρ^* that attains the corresponding optimal value. In order to do this, we apply a nonsmooth optimization algorithm. The Literature on classical nonsmooth optimization is wide, see e.g. Rockafellar (1970), Clarke (1983) and Mäkelä and Neittaanmäki (1992). We consider a numerical procedure proposed by Uberti (2006) that allows to find the minimum of an univariate function with a finite number of jump discontinuities. This procedure belongs to the class of direct methods and falls under the category of sequential line search. Although, it was originally designed for nonsmooth functions it works also well for a class of non continuous ones, like the one considered here. See Maggi and Uberti (2007) for the multivariate version of the algorithm.

We now describe the algorithm with reference to the problem at hand. The algorithm consists of two nested cycles. In the following, we indicate with the superscript ℓ the iteration of the external cycle and with the subscripts h the iteration of the internal one. Given a generic asset i, $m = min(\boldsymbol{\alpha}_i)$ and $M = max(\boldsymbol{\alpha}_i)$ are, respectively, the minimum and the maximum element of the vector $\boldsymbol{\alpha}_i$. We fix a starting point $x_0 \in [m, M]$ and an initial step length $s^0 > 0$, such that $(x_0 + s^0) \in [m, M]$. We indicate with $\rho(x_0) = \{S_1, S_2\}$ the partition with elements $S_1 = \{t : \alpha_{i_t} < x_0\}$ and $S_2 = \{t : \alpha_{i_t} \geq x_0\}$. For any real number x the corresponding partition $\rho(x)$ can be obtained in a similar way.

The external cycle reduces progressively the step length and generates a sequence $\{s^\ell\}$ strictly decreasing, with $\{s^\ell\} \downarrow 0$. The sequence $\{s^\ell\}$ can be constructed in different ways, in the following we set $s^\ell = \alpha s^{\ell-1}$, with $\alpha < 1$.

For any step length s^ℓ, the internal cycle generates a new point

$$x_h^\ell = x_{h-1}^\ell + s^\ell \times sgn\left[-\Delta_h SC\left(x_h^\ell\right)\right]$$

where sgn is the signum function, $\Delta_h SC\left(x_h^\ell\right) = SC\left(\rho\left(x_h^\ell + s^\ell\right)\right) - SC\left(\rho\left(x_h^\ell\right)\right)$, and $SC(\cdot)$ is the score function defined in (6). The internal cycle stops when $\Delta_{h-1} SC\left(x_{h-1}^\ell\right) \times \Delta_h SC\left(x_h^\ell\right) \leq 0$.

Given ε a desired level of tolerance of the solution, the procedure stops when $s^\ell < \varepsilon$, and the partition $\rho(x_h^\ell)$ is proposed as a finite approximation of the local minimum point of the score function (6).

This method does not generally guarantee a monotone reduction of the absolute error between two successive external iterations. Hence it is not possible to qualify it as a method of order 1 in the sense of Definition 6.1 of Quarteroni *et all.* (2000).

4 Illustrative examples

The methodology described in Sections 2 and 3 is now illustrated on two real data sets: the IPSA and the MIBTEL ones.

For outliers detection we proceed as follow. Firstly, last trimmed square regression (LTSR) is applied to the full data in order to identify a reasonable set of potential outliers. We used LTSR regression because it has a very high breakdown point (close to 1/2) and tends to identify large numbers of observation as abnormal. All points such that the absolute value of the standardized residuals is greater than the threshold $\delta = 2.5$ are considered as potential outliers (Rousseeuw, 1984). Then, by applying the algorithm described in Section 3 we identify the subset of the points selected by LTSR corresponding to anomalous data; see Tables 1 and 3.

In the examples below we used the following values of the hyperparameters. We set $c = 1$ in the cohesion function in order to favour the creation of a small number of large clusters. We used a weakly informative prior setting $a = 0, b = 1, \tau_0^2 = \gamma_0^2 = 1000, v_0 = 3$ and $\lambda_0 = 2$. In particular, the chosen values for v_0 and λ_0 lead to a 95% prior credible interval $(0.27; 3.22)$ for σ_i^2. Since $E(\sigma_i^2) = 1$, this interval extends (approximately) from a quarter to three times the expected value, hence it can be reasonably taken as weakly informative. Finally, for the value of v_0, we chose the smallest integer value which admits a finite variance. Setting $(k_1, k_2, k_3) = \frac{1}{1002}(500, 500, 1)$ in (5) and (6), we gave priority to the estimation of $\boldsymbol{\alpha}_i$ and β_i imposing almost no restriction on the number of clusters.

We considered a run of 10 000 sweeps with a burn-in of 1 000 iterations. Convergence of the MCMC algorithm was assessed using standard convergence criteria, see e.g. Best et al.(1995) and Coweles and Carlin (1996). No specific indication of abnormal behaviour is obtained. For the two examples discussed in the Subsections 4.1 and 4.2 the MCMC algorithm required 16 and 10 minutes respectively per 10 000 iterations on a Pentium 4 3.4 GHz personal computer. The programs were written in MATLAB; it is expected that a lower level programming language could speed up the execution time by a factor of at least 5.

4.1 IPSA stock market data

The IPSA is the main index of the "Bolsa de Comercio de Santiago" (Santiago Stock Exchange). It corresponds to an indicator of returns of the 40 most heavily traded stocks, the list is revised quarterly.

For comparative purposes we focus our analysis only on the 5 shares listed in Table 1, for which Quintana *et all.* (2005a) provide a detailed analysis both of the estimates of the parameters of interest and of the selected partition. The data are relative to the period January 1990-June 2004. We use the IPSA index as a measure of the market return and the interest rate in sale of discount bonus of the Central Bank as the risk free rate.

In Table 1 we report the partition selected by our algorithm and the one proposed by Quintana *et all.* (2005a), respectively ρ^*_{DTU} and ρ^*_{QI}. The results are compared in terms of the observed values of the score function, for all 5 shares $SC(\rho^*_{DTU}) \leq SC(\rho^*_{QI})$.

TABLE 1 ABOUT HERE

We performed a microeconomic analysis of the societies under study, and we list below some events that could have produced the abnormal behaviours identified by the outliers. For a description of a number of extraordinary events in the Chilean history that could be related to the outliers in Table 1 see Loschi *et all.* (1999). All the information provided below are freely available on the World Wide Web.

1) CEMENTOS BÍO-BÍO S.A. The Cementos Bío Bío S.A. is a company involved in the production and sale of cement and lime products, wood and its by-products, premixed concrete and ceramics.
In 1998 (outlier 107) it expanded the cement plant in Antofagasta and started up a new cement plant in Curicó.
In 1999 (outliers 112, 113) Cementos de Mexico, the world's third-largest cement manufacturer, entered the Chilean market by acquiring the 12 % of the shares of the Cementos Bio Bio.

2) CMPC. The group's principal activity is manufacturing pulp and paper in Chile. It is an integrated company that undertakes its industrial work through five business affiliates (CMPC Celulosa, CMPC Papeles, CMPC Productos de Papel, CMPC Tissue, and Forestal Mininco), and owns industrial plants in Chile, Argentina, Peru and Uruguay.
The years from 1990 to 1992 (outlier 15) were characterised by an expansion in Latin America. In 1990 CMPC entered in Argentina by purchasing, in partnership with Procter & Gamble, Quimica Estrella San Luis S.A. (now Prodesa), a manufacturer of sanitary napkins and paper diapers. In 1992 CMPC formed a strategic alliance with Procter & Gamble to develop markets for the aforementioned products in Chile, Argentina, Bolivia, Paraguay, Peru, and Uruguay.

3) CONCHA Y TORO. Concha y Toro is one of the leading producers of wine in Chile. It produces and exports a wide range of wines. In 1994, Concha y Toro became the first Chilean winery to be listed on the New York Stock Exchange.
During the years 1991-1993 (outliers 14, 21, 22) important changes took place. Concha y Toro tripled the size of its vineyards to reduce dependence on outside grape growers and enrolled French and California oenologists. It modernized its production and transformed the original Concha y Toro mansion into the headquarter of the firm for its export operations.
In 1996 (outlier 83) Concha y Toro purchased a vineyard in the Mendoza region in Argentina.
In 1997 the company and the French firm Baron Philippe de Rothschild S.A. endorsed a joint venture with the aim of producing a wine to the standards of the French Grand Cru

Class. In 1998 (outlier 97) Concha y Toro lunched on the market the Vina Almaviva. In the same year the company ranked second among wine exporters to the United States.

4) COPEC S.A. Copec S.A. is a diversified Chilean financial holding company that participates through subsidiaries and related companies in different business sectors of the economy (energy, forestry, fishing, mining and power industries).
In 1999 (outlier 111) COPEC created in joint venture with BP Global Investmentsthe Air Bp Copec S.A. to commercialise fuels for national and international air lines.

5) ENTEL. Entel (Empresa Nacional de Telecomunicaciones) was created the 31 of August 1964 as a state company, and it was privatised in 1986. The group's principal activities are providing telecommunication services. It also operates in Central America and Peru aside from its centre of major operations which is located in Chile.
For this firm we did not identify any outliers.

4.2 MIBTEL stock market data

The MIBTEL is one of the main indices of the Borsa Italiana, it consists of 300 shares (mainly Italian and certain foreign) whose identification is based on liquidity criteria. The results that follow are related to the application of the methodology described in the previous sections to the returns of MIBTEL components in the period from January 1996 to April 2006. The data correspond to monthly excess returns of the 142 shares for which we have complete time series. We use the MIBTEL index as a measure of the market returns and 1-month Treasury Bill (BOT) as the risk free rate.

We performed a preliminary analysis, by using a standard Bayesian linear regression model, in order to identify shares with abnormal behaviour in terms of systematic risk and/or specific risk. In this way we selected the 8 shares listed in Tables 2 and 3. In Table 2 we provide the Bayesian estimates and 95% credible intervals for β_i; in Table 3 we report the partition selected by LTSR and the one selected by our algorithm, for which we also indicate the observed value of the score function.

TABLE 2 ABOUT HERE

TABLE 3 ABOUT HERE

In the following we discuss separately each of the examined shares, and we indicate some events that could have produced the abnormal behaviours identified by the outliers. All the information provided are freely available on the World Wide Web.

1) AUTOSTRADE S.p.A. The Autostrade S.p.A. operates in Italy, in the United States, in United Kingdom and in other European countries. Its principal activities are construction and management of toll motorways and tunnels under license, as well as the designing, implementing and financing of electronic fee collection and tolling systems for large scale networks. Other minor activities are advertising and telecommunication services.
Years 1995 and 1996 (outlier 4) were characterised by an expansion of the society both in Italy and abroad. At the beginning of 1995, the company acquired the Autostrade Finance S.A. and the Autostrade International S.p.A. in order to develop its presence on the international markets. Furthermore, at the end of the same year, Autostrade Telecomunicazioni S.r.l. was set up to further widen the activities of the Group. Finally, in 1996, the company Spea-Ingegneria Europea S.p.A. was acquired, as was control of Pavimental.
In 1997 (outlier 23) ANAS (the National Road Board) and Autostrade S.p.A. signed an agreement that committed Autostrade S.p.A. to carry out significant investments, including the construction of the Variante di Valico (this agreement will expire on 31 December 2038).
In 1999 (outlier 36) Autostrade S.p.A. was privatised.

2) AUTOSTRADA TO-MI S.p.A. The group Autostrada Torino-Milano S.p.A. (ASTM) is the second largest motorway operator in Italy. Its principal activities are maintenance and management of Turin-Milan motorway, including its access roads and intersections and the collection of traffic tolls, and other motorway's segments. ASTM is also involved in the construction of a high speed rail link between Milan, Turin and Genoa.
In 1999 (outliers 39, 41, 42, 48) ASTM achieved high performances thanks to a stronger alliance with Autostrade S.p.A..
In February 2002 (outlier 75) ASTM created as a spin-off SIAS (Società Iniziative Autostradali e Servizi).

3) BANCA IFIS S.p.A. IFIS was founded in 1983 as an industrial factor and it became a bank (BANCA IFIS) in 2000 (outliers 51, 53, 55). Nowadays, its principal activities are in the factoring and leasing sector and related services. It manages companies circulating capital and support their business credit policies. Its activities are located in Italy and East European countries.
Since 1999 Banca IFIS begun its expansion both in Italy and in the eastern Europe. In particular in 1999 (outlier 49), it started the acquisition of credits of enterprises located in Romania and Hungary.

4) BOERO BARTOLOMEO S.p.A. The principal activities of Boero Bartolomeo S.p.A. are manufacture and distribution of varnishes (paints, lacquers and enamel varnishes) for building, anti corrosions of ships and yachts. It operates mainly in the Italian market.
In 2000 the Boero Bartolomeo S.p.A. sold the controlled Apsa S.p.A.. In 2002 (outlier 76) the Boero Bartolomeo S.p.A. set aside a fund to face the risk deriving from the litigation with the purchasers of Apsa S.p.A..

5) GARBOLI CONICOS S.p.A. Garboli Conicos was founded in 1998 (outlier 25) through the merger of CON.I.COS and Garboli-REP S.p.A. as a consequence of the denationalization of the I.R.I. group. It operates in the field of constructions both in Italy and abroad.
In April 2005 (outlier 113) Astaldi did not succeeded to acquire the Garboli Conicos that was acquired unexpectedly from the Pizzarotti group.

6) LA GIOVANNI CRESPI S.p.A. Crespi Group is divided in several companies dealing with different business: synthetic materials, PU foam, advanced textile technology and health-care. It is an international reality with production units in Italy, Poland, China and Brazil.
In 1997 (outlier 17) it started the Chinese Joint Venture (Crespi Beijing).
In 1998 (outlier 27) it acquired the ITS Artea (Crespi/ITS Artea).
In 2000 (outlier 49) it started the production of synthetic leather in Crespi Do Brasil.
In 2002 (outlier 81) it opened a new plant in Pisticci for the production of non woven fabrics.

7) SAIPEM. Saipem was founded in 1957 through the merger of Snam Montaggi and SAIP. The name of the company derives from this tie-up: SAIP + E (Italian for “and”) + M(ontaggi). It is a world leader in the oil and gas contracting services sector, both onshore and offshore.

From 1996 to 1997 SAIPEM undertook a program of strong investment (outliers 4, 12, 23). The most relevant investments were related to the purchase of new marine vessels for offshore drilling and for a new floating production unit.

In February 2000 (outlier 50) Saipem acquired, in joint venture with other companies, a contract for offshore Construction in Malaysia.

In 2002 (outliers 74, 75) Saipem acquired from Bouygues Construction the majority stake in Bouygues Offshore S.A., the leading French provider of engineering services to the oil industry.

8) SCHIAPPARELLI 1824 S.p.A. Schiapparelli 1824 S.p.A. manufactures and distributes cosmetics and related items.

Since January 1997 (outlier 13) Schiapparelli 1824 S.p.A. is no more an operating holding company.

5 Concluding remarks

We considered robust Bayesian estimation of the systematic risk in CAPM. We assumed that the data follow a normal distribution and we imposed a partition structure on the specific risks. A clear advantage of this procedure is to remain in a Mean-Variance framework even in presence of abnormal points.

We worked in a Bayesian decision framework, and we selected as optimal partition the one minimizing the expected value of the quadratic loss function in (5). If the problem at hand requires the use of different metrics a suitable loss function should be used. However, by using the MCMC method is possible to derive the required posterior summaries.

To minimize the score function in (6) we applied the optimization procedure described in Section 3. Among the partition of cardinality 2 we identified the one that best separates “standard” observations from the “atypical” ones. We believe that, since at each iteration we move more that one data point from one group to the other, this procedure is superior to the one proposed by Quintana and Iglesias (2003).

Acknowledgements

The authors acknowledge Fernando Quintana for helpful discussion regarding the computational algorithm and Manuel Galea-Rojas for providing the IPSA data. The research of the second author was partially supported by MIUR, Rome (PRIN 2005132307), and the University of Pavia.

Appendix: Gibbs sampling distribution

Consider a generic asset i. Given the starting values α_0, β_0 and σ_0^2 we iteratively sample from the following distributions

$$\beta_i|\sigma_i^2, \boldsymbol{\alpha}_i, \mathbf{y}_i \sim N\left\{\frac{b/\gamma_0^2 + \sum_{t=1}^T (y_{i_t} - \alpha_{i_t})x_t}{1/\gamma_0^2 + \sum_{t=1}^T x_t^2}, \frac{\sigma_i^2}{1/\gamma_0^2 + \sum_{t=1}^T x_t^2}\right\}$$

$$\sigma_i^2|\boldsymbol{\alpha}_i, \beta_i, \boldsymbol{y}_i \sim IG\left\{v_0 + \frac{T + |\rho| + 1}{2}, \lambda_0 + \frac{(\beta_i - b)^2}{2\gamma_0^2} + \frac{1}{2\tau_0^2}\sum_{d=1}^{|\rho|}(\alpha_{i_d}^* - a)^2 \right.$$
$$\left. + \frac{1}{2}\sum_{t=1}^T (y_{i_t} - \alpha_{i_t} - \beta x_t)^2\right\}$$

$$\alpha_{i_t}|\boldsymbol{\alpha}_{i_{-t}}, \beta_i, \sigma_i^2, \boldsymbol{y} \propto \sum_{j\neq t} \exp\left\{-\frac{1}{2\sigma^2}(y_{i_t} - \alpha_{i_j} - \beta_i x_t)^2\right\}\delta_{\alpha_{i_j}}(\alpha_{i_t})$$
$$+ \frac{\exp\left\{-(y_{i_t} - \beta x_t - a)^2/2\sigma_i^2(1+\tau_0^2)\right\}}{\sqrt{1+\tau_0^2}} \; N\left(\frac{y_{i_t} - \beta_i x_t + a/\tau_0^2}{1 + 1/\tau_0^2}, \frac{\sigma_i^2}{1+1/\tau_0^2}\right)$$

where $\boldsymbol{\alpha}_{i_{-t}} = (\alpha_{i_1}, \ldots, \alpha_{i_{t-1}}, \alpha_{i_{t+1}}, \ldots, \alpha_{i_T})'$ and $\delta_{\alpha_j}(\cdot)$ is the delta function.

Note that β_i and σ_i^2 are sampled from the corresponding full conditional whereas each α_{i_t} is sampled from a mixture of point masses and a normal distribution. In this way we automatically update both the vector $\boldsymbol{\alpha}_i$ and the partition structure.

Before proceeding to the next Gibbs iteration we update the vector $\boldsymbol{\alpha}_i$ given the partition ρ sampling from

$$\alpha_{i_d}^* \sim N\left(\frac{\sum_{t\in S_d}(y_{i_t} - \beta_i x_t) + a/\tau_0^2}{|S_d| + 1/\tau_0^2}, \frac{\sigma_i^2}{|S_d| + 1/\tau_0^2}\right) \quad d = 1, \ldots, |\rho|$$

This last step was introduced in Bush and MacEachern (1996) to avoid being trapped in sticky patches in the Markov Space.

References

Antoniak, C.E. (1974) Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist.*, **2**, 1152-1174.

Barry, D. and Hartigan, J.A. (1992) Product partition models for change point problems. *Ann. Statist.*, **20**, 260-279.

Best, N.G., Cowles, M.K. and Vines, K. (1995) CODA: convergence diagnosis and output analysis software for Gibbs sampling output, versione 0.30. *Tecnical report*, Medical Research Council Biostatistics Unit, Cambridge.

Black, F. (1972) Capital market equilibrium with restricted borrowing. *J. Bus.*, **45**, 444-454.

Bush, C.A. and MacEachern, S.N. (1996) A semiparametric Bayesian model for randomised block designs. *Biometrika*, **83**, 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. *Appl. Econ. Lett.*, **10**, 447-453.

Chaturvedi, A. (1996) Robust Bayesian analysis of the linear regression models. *J. Statist. Plann. Inference*, **50**, 175-186.

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

Clarke F.H. (1983) *Optimization and Nonsmooth Analysis.* Wiley-Interscience, New York.

Coweles, M. and Carlin, B. (1996) Markov Chain Monte Carlo convergence diagnostics: a comparative review. *J. Am. Statist. Ass.*, **91**, 883-904.

Fernández, C. Osiewalski, J. and Steel, M. F. J. (2001) Robust Bayesian inference on scale parameters. *J. Multivariate Anal.*, **77**, 54-72.

Hartigan, J.A. (1990) Partition models. *Comm. Statist. Theory and Methods* , **19**, 2745-2756.

Huber,P. J. (1973) Robust Regression: Asymptotics, Conjectures and Monte Carlo. *Ann. Statist.*, **1**, 799-821.

Lange, K. L., Little, R. J. and Taylor, J. (1989) Robust statistical modelling using the t-distribution. *J. Am. Statist. Ass.*, **84**, 881-96.

Lintner, J. (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Rev. Econ. Statist.*, **47**, 13-37.

Liu, J.S. (1996) Nonparametric hierarchical Bayes via sequential imputations. *Ann. Statist.*, **24**, 911-930.

Loschi, R., Iglesias, P., and Arellano-Valle, R. (1999) Bayesian detection of change points in the Chilean Stock market. *Proc. Bayesian Statist. Sect. Am. Statist. Ass.*, 160-165.

Maggi, M.A. and Uberti, P. (2007) A general optimization method for non-continuous functions. *Manuscript*.

Mossin, J. (1966) Equilibrium in capital asset market. *Econometrica*, **25**, 768-783.

Mäkelä, M.M. and Neittanmäki, P. (1992) *Nonsmooth optimization: analysis and algorithms with applications to optimal control*. World Scientific, Singapore.

Quarteroni, A., Sacco, R. and Saleri, F. (2000). *Numerical Mathematics*. Springer, New York.

Quintana, F.A. and Iglesias, P.L. (2003) Bayesian clustering and product partition models. *J. R. Statist. Soc. B*, **65**, 557-574.

Quintana, F.A., Iglesias, P. L. and Galea-Rojas, M. (2005a) Bayesian robust estimation of systematic risk using product partition models. *Appl. Finan. Econ. Lett.*, **1**, 313-320.

Quintana, F.A., Iglesias, P. L. and Bolfarine, H.. (2005b) Bayesian Sensitivity Analysis for Measurement Error Models. *Adv. Complex Syst.*, **8**, 433-449.

Rockafeller R. T. (1970) *Convex Analysis*. Princeton University Press, Princeton, New Yersey.

Rousseeuw, P.J. and Leroy, A. (1987) *Robust Regression and Outlier Detection*. Wiley, New York.

Rousseeuw, P.J. (1984). Last median of square regression. *J. Am. Statis. Ass.*, **79**, 871-88.

Sharpe, W.F. (1964) Capital asset prices: a theory of market equilibrium under conditions of risk. *J. Finance*, **19**, 425-442.

Staudta, R.G. and Sheather, S.J. (1990) *Robust Estimation and Testing*. John Wiley, New York.

Tarantola, C., Consonni, G. and Dellaportas, P. Bayesian clustering for row effects models, submitted.

Uberti, P. (2006), *Modelli condizionali di selezione del portafoglio*. Degree thesis, University of Pavia.

Table 1: IPSA stock market data: comparison of the results obtained via LTSR, QI's algorithm and DTU's algorithm

Society	$LTSR(\delta = 2.5)$	ρ^*_{QI}	ρ^*_{DTU}	SC_{QI}	SC_{DTU}
CEMENTOS BÍO-BÍO S.A.	$\{12, 14, 18, 19, 21, 27, 37, 47, 59, 88, 91, 107, 112, 113, 121, 133, 161\}$	$\{12, 18, 27, 37, 112, 121\}, \{21, 113\}, \{47, 59, 133\}, \{107\}$	$\{107, 112, 113, 121\}$	0.0052	0.0038
CMPC	$\{3, 15, 49, 50, 54, 57, 100, 112\}$	$\{15\}$	$\{15\}$	0.0020	0.0020
CONCHA Y TORO	$\{12, 14, 18, 21, 22, 23, 27, 29, 32, 83, 97, 104, 106, 110, 144\}$	$\{14\}, \{21\}, \{27\}$	$\{14, 21, 22, 27, 83, 97\}$	0.0444	0.0058
COPEC S.A.	$\{31, 105, 107, 108, 109, 111\}$	$\{31, 107\}, \{111\}$	$\{107, 111\}$	0.0049	0.0040
ENTEL	$\{14, 27, 29, 48, 59, 63, 65, 106, 111, 112\}$	$\{65\}$	no partition	0.0010	0.0009

For partitions with $|\rho| > 1$ we give only the subsets that are formed by elements detached from S_0.
QI and DTU stand for Quintana and Iglesias and De Giuli, Tarantola and Uberti, respectively.

Table 2: MIBTEL stock market data: Bayesian estimates and 95% credible intervals for the systematic risk of the asset i

Society	$\widehat{\beta}_{iB}$	$(2.5\% - 97.5\%)$
Autostrade S.p.A.	0.7893	(0.5549, 1.0222)
Autostrada TO-MI S.p.A.	0.5222	(0.2891, 0.7524)
Banca IFIS S.p.A.	0.2972	(-0.1388, 0,7346)
Boero Bartolomeo S.p.A.	0.1440	(-0.0701, 0.3566)
Garboli Conicos S.p.A.	0.1395	(-0.2387, 0.5183)
La Giovanni Crespi S.p.A.	0.6340	(0.4430, 0.8220)
SAIPEM	1.2989	(1.1362, 1.4585)
Schiapparelli 1824 S.p.A.	0.2803	(-1.6326, 2.7226)

Table 3: MIBTEL stock market data: comparison of the results obatined via LTSR and DTU's algorithm

Society	$LTSR(\delta = 2.5)$	ρ^*_{DTU}	SC_{DTU}
Autostrade S.p.A.	$\{4, 20, 23, 24, 32, 36, 37, 38, 47, 123\}$	$\{4, 23, 36\}$	0.0037
Autostrada TO-MI S.p.A.	$\{24, 25, 39, 41, 42, 48, 62, 68, 75\}$	$\{24, 25, 39, 41, 42, 48, 75\}$	0.0036
Banca IFIS S.p.A.	$\{1, 5, 6, 9, 22, 24, 31, 49, 51, 53, 55\}$	$\{6, 9, 31, 49, 51, 53, 55\}$	0.0198
Boero Bartolomeo S.p.A.	$\{9, 21, 24, 30, 35, 39, 41, 43, 44, 76, 77, 85, 88, 110, 120\}$	$\{21, 24, 76\}$	0.0033
Garboli Conicos S.p.A.	$\{13, 25, 28, 30, 35, 38, 44, 50, 51, 52, 56, 70, 92, 104, 106, 107, 108, 113\}$	$\{13, 25, 113\}$	0.0071
La Giovanni Crespi S.p.A.	$\{6, 17, 26, 27, 49, 81, 87\}$	$\{17, 27, 49, 81\}$	0.0081
SAIPEM	$\{4, 12, 23, 50, 74, 75, 76\}$	$\{4, 12, 23, 50, 74, 75\}$	0.0093
Schiapparelli 1824 S.p.A.	$\{4, 13, 14, 24, 25, 26, 49, 87, 88, 95, 104, 115\}$	$\{13\}$	0.3508

For partitions with $|\rho| > 1$ we give only the subsets that are formed by elements detached from S_0.
DTU stands for De Giuli, Tarantola and Uberti.