UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 07-39
Business Economics Series 09
May 2007

Departamento de Economía de la Empresa
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 6249607

PRICING TRANCHED CREDIT PRODUCTS WITH GENERALIZED MULTIFACTOR MODELS*

M. Moreno, J.I. Peña and P. Serrano

Abstract

The market for tranched credit products (CDOs, Itraxx tranches) is one of the fastest growing segments in the credit derivatives industry. However, some assumptions underlying the standard Gaussian one-factor pricing model (homogeneity, single factor, Normality), which is the pricing standard widely used in the industry, are probably too restrictive. In this paper we generalize the standard model by means of a two by two model (two factors and two asset classes). We assume two driving factors (business cycle and industry) with independent t-Student distributions, respectively, and we allow the model to distinguish among portfolio assets classes. In order to illustrate the estimation of the parameters of the model, an empirical application with Moody's data is also included.

Keywords: Collateral Debt Obligations, Factor Models, Probit-Logit Models
JEL Classification: G13, C35, C51

* We want to thank to A. Novales for his helpful comments. Serrano acknowledges financial support from the Plan Nacional de I+D+I (project BEC2003-02084) and especially to Jose M. Usategui. Peña thanks financial support from MEC grant SEJ2005-05485. Manuel Moreno is from University of Castilla La-Mancha, Facultad CC. Jurídicas y Sociales, Cobertizo de San Pedro Martir, Dpto Análisis Económico y Finanzas, Toledo, Spain. E-mail: Manuel.Moreno@uclm.es. Juan Ignacio Peña is from University CarlosIII, Dpto. Economía de la Empresa, Madrid, Spain. E-mail: ypenya@eco.uc3m.es. Pedro Serrano is from University of Basque Country, Facultad CC. Económicas, Dpto. Fundamentos del Análisis Económico II, Avda. Lehendakari Aguirre, 83, 48015 Bilbao, Spain. E-mail: pedrojose.serrano@ehu.es.

Pricing Tranched Credit Products with Generalized Multifactor Models

M. Moreno J. I. Peña P. Serrano *

April 30, 2007

*We want to thank to A. Novales for his helpful comments. Serrano acknowledges financial support from the Plan Nacional de I+D+I (project BEC2003-02084) and especially to Jose M. Usategui. Peña thanks financial support from MEC grant SEJ2005-05485.
Manuel Moreno is from University of Castilla La-Mancha, Facultad CC. Juridicas y Sociales, Cobertizo de San Pedro Martir, Dpto Analisis Economico y Finanzas, Toledo, Spain. E-mail: Manuel.Moreno@uclm.es. Juan Ignacio Peña is from University Carlos III, Dpto. Economia de la Empresa, Madrid, Spain. E-mail: ypenya@eco.uc3m.es. Pedro Serrano is from University of Basque Country, Facultad CC. Economicas, Dpto. Fundamentos del Analisis Economico II, Avda. Lehendakari Aguirre, 83, 48015 Bilbao, Spain. E-mail: pedrojose.serrano@ehu.es.

Abstract

The market for tranched credit products (CDOs, Itraxx tranches) is one of the fastest growing segments in the credit derivatives industry. However, some assumptions underlying the standard Gaussian one-factor pricing model (homogeneity, single factor, Normality), which is the pricing standard widely used in the industry, are probably too restrictive. In this paper we generalize the standard model by means of a two by two model (two factors and two asset classes). We assume two driving factors (business cycle and industry) with independent t-Student distributions, respectively, and we allow the model to distinguish among portfolio assets classes. In order to illustrate the estimation of the parameters of the model, an empirical application with Moody's data is also included.

Keywords: Collateral Debt Obligations, Factor Models, Probit-Logit Models

Journal of Economic Literature classification: G13, C35, C51

1 Introduction

The market of credit tranched products is one of the fastest growing segments in the credit derivative industry. As an example, Tavakoli (2003) reports an increase in market size from almost \$19 billions in 1996, to \$200 billions in 2001. Recent reports estimate market size to be \$20 trillions in 2006.[1] As a result, an increasing attention of the financial sector audience has focused on the pricing of these new products.

The development of pricing models for multiname derivatives is relatively recent. As was pointed out by Hull and White (2004), the standard approach on the credit risk literature tends to subdivide the pricing models for multiname derivatives in two groups: structural models, which are those inspired in Merton's (1974) model or Black and Cox (1976), or the intensity based approach, like Duffie and Garlenau (2001). Roughly speaking, their differences remain on how the probability of default of a firm is obtained: using their fundamental variables - assets and liabilities - as in the case of the structural models, or using directly market spreads, as in the intensity models approach. Up to a point, the structural based approach has been extensively implemented by the financial sector, maybe due to the extended use of industrial models like Vasicek (1987) or Creditmetrics model of Gupton, Finger and Bathia (1997).[2] However, recent academic literature analyze the prices of CDO tranches using intensity models, as Longstaff and Rajan (2006). We refer to Bielecki and Rutkowski (2002) for a general presentation of structural and intensity based models.

This paper presents an extension of the standard gaussian model of Vasicek (1991), in line with the structural models literature. Basically, Vasicek's (1991) model assumes that the value of a firm is explained by the weighted average of one common factor for every asset plus an independent idiosyncratic factor. By means of linking the realization of one systematic factor to every firm's values, Vasicek (1991) provides a simple way to reduce the complexity of dealing with dependence relationships between firms. Gibson (2004) or Gregory and Laurent (2004, 2005) provide additional insights about risk features of this model.

[1] See BBA Credit Derivatives Report (2006).

[2] It is worth mentioning that the appearance of techniques within the Structural framework that diminishes the traditional high computing cost of multiname credit derivatives (see Andersen, Sidenius and Basu (2003) or Glasserman and Suchitabandid (2006), among others) has contributed to the widely usage of structural models.

Our approach relies on the connection between the changes of value of a firm and the sum of two factors: systematic and idiosyncratic. Our approach overcomes the limitations of the standard Gaussian model: the different areas or *regions* of correlation that could compose a credit portfolio (see Gregory and Laurent, 2004). This article presents a model that captures some of the facts found in real data. Motivated by this fact, this article proposes an extension to the two Gaussian asset classes as in Schönbucher (2003). Our paper extends the existing literature in three ways: firstly, the assumption of asset homogeneity is relaxed by introducing two asset classes. Secondly, we consider an additional source of systematic risk by including another common factor related with industry factors. Finally the normality assumption on common factors is relaxed.

This paper is divided as follows: Section 2 presents the model. Section 3 studies the sensitivity to correlation and to changes in credit spreads. Section 4 addresses the econometric modelling. Finally, some conclusions are presented on section 5.

2 The model

To motivate our model we discuss some empirical features found in CDO data that are not captured by the standard Gaussian models and propose an extension to the asset class models posited by Schönbucher (2003). Notation is taken from Mardia, Kent and Bibby (1979).

2.1 The Standard Gaussian model

The Gaussian model introduced by Vasicek (1991) has become a standard in the industry. Basically, it addresses in a simple and elegant way the key input in CDOs price: the correlation in default probabilities between firms affects the price of the CDO.

Usually a CDO is based on a large portfolio of firms bonds or CDS.[3] Let $\mathbf{V}_{n\times 1}$ (subscript denotes matrix dimension) be a random vector with mean zero and covariance matrix $\mathbf{\Sigma}$. As standard notation in multivariate analysis, we will define the p-factor model as

$$\mathbf{V}_{n\times 1} = \mathbf{\Lambda}_{n\times p}\mathbf{F}_{p\times 1} + \mathbf{u}_{n\times 1} \tag{1}$$

[3]CDO tranches of NYME are composed by 100 firms.

where $\mathbf{F}_{p\times 1}$ and $\mathbf{u}_{n\times 1}$ are random variables with different distributions.

The interpretation of the model (1) is the following:

- The vector $\mathbf{V}_{n\times 1}$ represents the value of the assets for each of the individual n-obligors.
- The vector $\mathbf{F}_{p\times 1}$ captures the effect of systematic factors - business cycle, industry, etc. - that affect to the whole economy.
- By contrast, $\mathbf{u}_{n\times 1}$ represents the idiosyncratic factors that affect each of the n-companies.
- Finally, $\mathbf{\Lambda}_{n\times p}$ is called the *loading* matrix, and determines the correlation between each of the n-firms.

By assumption, we have that

$$E(\mathbf{F}) = \mathbf{0}, \qquad Var(\mathbf{F}) = \mathbf{I}, \tag{2}$$

$$E(\mathbf{u}) = \mathbf{0}, \qquad cov(\mathbf{u_i}, \mathbf{u_j}) = 0, \qquad i \neq j, \tag{3}$$

and

$$cov(\mathbf{F}, \mathbf{u}) = \mathbf{0} \tag{4}$$

where $\mathbf{I}$ is the identity matrix.

We will also assume that vector $\mathbf{u}_{n\times 1}$ is standardized to have zero mean and unit variance.

Using a simplified form of equation (1), Vasicek (1991) assumes that firm's values in the asset pool backing the CDO are affected by the sum of two elements: on one hand, a common factor to every firm which represents the systematic component represented by the factor F; on the other hand, an idiosyncratic component modelled by a noise ε_i. Both are assumed to be standard N(0,1) random variables

$$V_i = \rho_{i1} F_1 + \sqrt{1 - \rho_{i1}{}^2}\varepsilon_i, \quad \text{with } i \leq n \tag{5}$$

By means of equation (5) it is possible to capture the correlation between different firms in a portfolio. As equation (5) reveals, Vasicek (1987) assumes that correlation coefficient is homogeneous for each pair of firms.[4] Additionally, the simplicity of their assumptions permits a fast computation of CDO prices under this framework.

As an immediate consequence of equation (5), a further step is given by generalizing the number of factors that affects firms values. Thus, Lucas, Klaassen, Spreij and Staetmans (2001) consider the following factor model

$$V_i = \sum_{j=1}^{p} \rho_{ij} F_j + \sqrt{1 - \sum_{j=1}^{p} \rho_{ij}{}^2 u_i}, \text{ with } i \leq n \tag{6}$$

where V_i represents the value of the i-company, F_j, $j = 1, ..., p$ capture the effect of p-systematic factors and u_i is the idiosyncratic factor associated to the i-firm. Needless to say, assumptions on the distribution of common and idiosyncratic factors (F and u, respectively) could be imposed. Hull and White (2004) also includes an extension to many factors, including the t-Student distributed case. Finally, Glasserman and Suchitabandid (2006) implements in a recent paper numerical approximations to deal with these multifactor structures within a Gaussian framework.

With the purpose of getting intuition about the behaviour of these different models, Figure 1 exhibits the loss distribution generated by three alternative models:

- The standard Gaussian model (Vasicek, 1991),

$$V_i^G = \rho_{i1} F_1 + \sqrt{1 - \rho_{i1}{}^2} \varepsilon_i, \quad \text{and } F_1, \varepsilon_i \sim N(0, 1)$$

- The one factor double-t model (Hull and White, 2004),

$$V_i^S = \rho_{i1} F_1 + \sqrt{1 - \rho_{i1}{}^2} \varepsilon_i,$$

  where F_1 and ε_i follows t-Student distributions, with 6 degrees of freedom.

- The two-factor Gaussian model (Hull and White (2004), or Glasserman and Suchitabandid, 2006),

$$V_i^{DG} = \rho_{i1} F_1 + \rho_{i2} F_2 + \sqrt{1 - \rho_{i1}{}^2 - \rho_{i2}{}^2} \varepsilon_i$$

  with $F_1, F_2, \varepsilon_i \sim N(0, 1)$ and $\rho_{i1}^2 + \rho_{i2}^2 < 1$.

[4]For a detailed exposition of assumptions in structural models see Bielecki and Rutkowski (2002).

As we will see later, the loss distribution plays a role crucial in the CDO pricing. Different distributions are computed for a portfolio of 100 firms, with constant default intensities of 1% per year. Correlation parameters are 0.3 for all models, except for two-Gaussian factor, with 0.1 and 0.3. The double-t model considers a common and idiosyncratic factors each distributed as t-Student with 6 degrees of freedom. The picture shows that, under the same correlation parameters, the Standard Gaussian model assigns lower probability to high losses (see for instance the case of 20 firms) than the double-t model. By considering an additional factor, as in the 2-Gaussian case, the probability of high losses is substantially higher than in previous cases.

[INSERT FIGURE 1 AROUND HERE]

Table 1 presents the spreads (in basis points) of a 5-years CDO portfolio composed by 100 firms. The individual default probabilities are constant and fixed at 1%. The recovery rate is 40%, a standard in the market. Finally, the different correlation parameters are displayed in the table. We remember that all spreads are obtained considering only one asset class. As Table 1 shows, results are consistent with the loss distribution lines presented in Figure 1: the double t-Student distribution gives prices systematically bigger than those obtained for the Standard Gaussian model, keeping constant the correlation. In the 2 Gaussian factor model prices, we observe a mixture of effects due to different combinations of correlations in the portfolio.

[INSERT TABLE 1 AROUND HERE]

2.2 A 2-by-2 model

Generally, as considered by Schonbücher (2003) or Lando (2004), a credit portfolio is composed by different asset classes or buckets, attending to criteria of investment grade, non-investment grade assets or industry, among others. The exposure of a credit portfolio to a set of common risk factors could be significant between groups, but should be homogeneous within them. In line with this, the idea of two groups of assets treated in different ways is a more realistic assumption.

As was pointed out by Gregory and Laurent (2004), the one-factor model imposes a limited correlation structure on the credit portfolio, which is not realistic. Initially, one can argue that increasing the number of factors could be enough to capture a richer structure of correlation within the portfolio. However, this is not yet consistent with the idea of heterogeneity correlation among groups, due to the fact that every asset is exposed to the same degree of correlation. By contrast, a richer correlation structure could be imposed in the portfolio if these two groups are treated in different ways.

To illustrate these ideas, Figure 2 shows the yearly percentage default rates for investment and non-investment grades. Rates for investment data have been multiplied by ten with the intention of clarifying the exposition. Figure 2 reveals that correlation between these two assets groups varies through time: periods with high degree of default in non-investment grade assets do not match with high default rates for investment grade. With reference to this idea, Figure 3 displays the correlation coefficients computed using a moving window of five years. This also provides some additional insights on the degree of correlation between asset classes: the picture shows how the correlation among different groups changes during the sample period, from negative correlation (1975, 1977 or 1990), to zero (1993, 1996) or highly positive (1982, 1995 or 2000). These differences in default rates through time could reveal the existence of an idiosyncratic component between asset classes. This fact could support the idea of modelling in a different way the behaviour of different assets.

[INSERT FIGURES 2 AND 3 AROUND HERE]

This article considers a family of models that take account the existence of these different asset classes or *regions*, in line with the suggestions of Gregory and Laurent (2004). We analyze a model in the line of the two assets-two Gaussian factor model of Schonbücher (2004), where no distinction is made between the obligors which belongs to the same class. Our model generalizes that posited by Schonbücher (2004) by considering a t-Student distribution, which assigns a higher probability to high default events. Empirical evidence seems to go in this direction.[5] Our work contributes to the existing literature in the analysis of these asset class models. To the best of our knowledge, no similar studies has been reported yet in this direction.

The model we propose is a two-by-two factor model as follows[6]

[5] See, for example, Mashal and Naldi (2002).

[6] For the sake of brevity, we omit the graph of th loss distribution implied by this model.

$$\begin{aligned} V_{i1} &= \alpha_{11}F_1 + \alpha_{12}F_2 + \beta_1 u_{i1} \\ V_{i2} &= \alpha_{21}F_1 + \alpha_{22}F_2 + \beta_2 u_{i2} \end{aligned} \tag{7}$$

where

- V_{i1}, V_{i2} $(i < n)$ represents the value of the i-company which belongs to the different asset class
- F_j, $j = 1, 2$ captures the effect of systematic factors - business cycle and industry - with independent t- distributions with n_j degrees of freedom, and
- u_{i1}, u_{i2} are idiosyncratic factors distributed also with n_{i1}, n_{i2} degrees of freedom, respectively.

Under the assumptions of factor models, F_j are scaled to have variance 1, then $\alpha_{11} = \rho_{11}\sqrt{\frac{n_1-2}{n_1}}$ and so on. Idiosyncratic errors are also scaled, for instance $\beta_1 = \sqrt{\frac{n_{i1}-2}{n_{i1}}}\sqrt{1-\alpha_{11}^2-\alpha_{12}^2}$. Finally, we assume the same default barriers K_{i1}, K_{i2} for the obligors of the same class.

Needless to say that the model could be easily generalized to the case of m-asset classes, as follows:

$$V_{i,m} = \sum_{h=1}^{p} \rho_{mh} F_h + u_{i,m}\sqrt{1 - \sum_{h=1}^{p} {\rho_{mh}}^2}$$

where $V_{i,m}$ represents the value of the i-company which belong to the m-asset class, F_j, $j = 1, ..., p$ capture the effect of systematic factors and $u_{i,m}$ is the idiosyncratic factor corresponded to i-firm of the $m-$asset class. Generally speaking, assumptions relying on distribution factors or more asset classes could also be proposed. However, a trade-off between accuracy, parsimony and computing efficiency must be considered.

2.3 Conditional Default Probabilities

Without loss of generality we omit the subscript that refers to the i-firm for the ease of exposition. We want to study the probability of default for the

i-firm which belongs to an asset class m, with $m = 1, 2$,

$$P\left[V_m \leq K|\mathbf{F} = \mathbf{f}\right] = P\left[u_m \leq \frac{K - \alpha\mathbf{F}}{\beta_m}|\mathbf{F} = \mathbf{f}\right] = T_m\left(\frac{K - \alpha\mathbf{f}}{\beta_m}\right)$$

where T_m denotes the distribution function of a t-student with n_m degrees of freedom for the i-firm, and $\mathbf{F}$, α are the common vector factors and their coefficients, respectively.

It is usual to calculate the probability of having k default events, conditional to realization of vector factor $\mathbf{F}$ as

$$P[X = k|\mathbf{F}] = \sum_{l=0}^{k} b\left(l, N_1, T_1\right) b\left(k - l, N_2, T_2\right)$$

where $b\left(l, N, T\right)$ denotes the binomial frequency function, which gives the probability of observing l successes with probability T, where N represents the number of firms which belong to each asset class. In the same manner, the unconditional probability of k default events is obtained considering all possible realizations of factors $\mathbf{F}$

$$P[X = k] = \sum_{l=0}^{k} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} b\left(l, N_1, T_1\right) b\left(k - l, N_2, T_2\right) \psi\left(\mathbf{f}\right) d\mathbf{f}$$

where $\psi\left(\mathbf{f}\right)$ denotes the probability density function of $\mathbf{F}$.

Finally, the total failure distribution is just obtained as the sum of all the defaults up to level k,

$$\begin{aligned} P[X &\leq k] = \sum_{r=0}^{k} P[X = r] \\ &= \sum_{r=0}^{k} \sum_{l=0}^{r} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} b\left(l, N_1, T_1\right) b\left(r - l, N_2, T_2\right) \psi\left(\mathbf{f}\right) d\mathbf{f} \qquad (8) \end{aligned}$$

Lando (2004) refers to this problem as a *different buckets* problem in sense that by means of multinomial distributions we compute the total loss distribution of the portfolio. Our approach to this point would be the computation of this set of bucket probabilities but we will adopt a different approach.

2.4 Loss distribution

As is pointed out in Lando (2004), calculation of multinomial expressions like (8) is burdensome. Instead of computing (8) by brute force, we will use a simple idea due to Andersen, Sidenius and Basu (2003) that could improve the efficiency in terms of computing cost.[7]

Andersen, Sidenius and Basu (2003) provides an efficient algorithm which has become widely used by the industry. Roughly speaking, the main idea behind is to observe what happens to the total loss distribution of a portfolio when we increase its size by one firm.

Consider a portfolio that includes n credit references and let $p_n(i|\mathbf{f})$ be the probability of default of i firms in this portfolio conditional on factors $\mathbf{f}$. Let $q_{n+1}(\mathbf{f})$ be the default probability of an individual firm that is added to this portfolio. These two probabilities are conditional on the realization of the common factor vector $\mathbf{f}$.

Consider the total Loss Distribution (LD) in this portfolio. Intuition says that the probability of i-defaults in this portfolio - conditional on $\mathbf{f}$ - can be written as

$$LD(i|\mathbf{f}) = p_n(i|\mathbf{f}) \times (1 - q_{n+1}(\mathbf{f})) + p_n(i-1|\mathbf{f}) \times q_{n+1}(\mathbf{f}), \quad 0 < i < n+1$$

where the first term reflects that default is due to i-firms included in the initial portfolio while the new reference (just included in the portfolio) survives. In a similar way, the second term reflects the new firm (just included in the portfolio) defaults while the other $i-1$ defaulted firms were previously included in the original portfolio.

Moreover, for the extreme cases of default firms, we have

$$\begin{aligned} p_{n+1}(0|\mathbf{f}) &= p_n(0|\mathbf{f}) \times (1 - q_{n+1}(\mathbf{f})) \\ p_{n+1}(n+1|\mathbf{f}) &= p_n(n|\mathbf{f}) \times q_{n+1}(\mathbf{f}) \end{aligned}$$

Then, taking into account the last firm just included in the portfolio, these equations reflect that no firm defaults or all of them default, respectively.

As an example, consider an initial portfolio including two firms. Adding

[7] Their contribution has been also explored and extended in Hull and White (2004).

a third firm, the total default probabilities are given by

$$\begin{aligned} p_3(0|\mathbf{f}) &= p_2(0|\mathbf{f})(1-q_3(\mathbf{f})) \\ p_3(1|\mathbf{f}) &= p_2(1|\mathbf{f})(1-q_3(\mathbf{f})) + p_2(0|\mathbf{f})\, q_3(\mathbf{f}) \\ p_3(2|\mathbf{f}) &= p_2(2|\mathbf{f})(1-q_3(\mathbf{f})) + p_2(1|\mathbf{f})\, q_3(\mathbf{f}) \\ p_3(3|\mathbf{f}) &= p_2(2|\mathbf{f})\, q_3(\mathbf{f}) \end{aligned}$$

Using this iterative procedure, we can compute the unconditional total Loss Distribution by considering all possible realizations of $\mathbf{f}$.[8]

3 Results for 2-by-2 model

In this section we present some results of the model (7). Firstly, we discuss the spreads obtained using the two-by-two approach. Secondly, we give an approach useful for cases of high degrees of freedom based on Cornish-Fisher expansions, which is useful in terms of computing cost.

3.1 Numerical results

To give some results of model (7), we analyze different cases for a two asset classes, 5-year CDO with 100 firms with quarterly payments. We also assume that the recovery rate is fixed and equal at 40%. Risk-neutral default individual default probabilities are fixed at 1% and 5% for assets which belong to class 1 and 2, respectively. For ease of explanation, the size of each asset class in the portfolio is the same (50% for each one). More results concerning the size of the portfolio will be provided in Section 4.

Table 2 displays the main results obtained. First row corresponds to the three simulated base cases: the standard Gaussian model[9] with one factor, two asset classes; the two assets-two Gaussian factor; finally, the two assets-two t-Student factors.[10] t-Student distributions have been fixed at 5 degrees of freedom for idiosyncratic and systematic factors. Second row displays the correlation parameters of *each* asset class with *both* factors. For example,

[8] See Andersen, Sidenius and Basu (2003) or Gibson (2004) for more details.

[9] Standard Gaussian model has been computed using a Gauss-Hermite quadrature with 8 nodes.

[10] t-Student simulations have been computed using a Simpson's quadrature with 25 points.

0.1/0.3 in the two Gaussian case refers to a correlation of 0.1 (0.3) for elements of class 1 (2) with *the two* systematic factors. In line with Gregory and Laurent (2004), our idea is to check the behaviour of the CDO to a portfolio exposed to two different degrees of correlation.

[INSERT TABLE 2 AROUND HERE]

All the simulations have been carried on for different tranches values. Looking at the riskier tranche (equity tranche), we observe that, for the same degree of correlation, its spread is systematically bigger for the two t-Student case than for the 2 Gaussian factors. One conclusion that arises from Table 2 is that the two assets-one Gaussian factor spreads for equity tranche are close to those values obtained for two assets-two t-Student factors.

As expected, the same does not apply for mezzanine tranches: the addition of more factors provides more weight to extreme default events, which results in an increase in spread of mezzanine tranches. The same conclusions apply to senior tranche.

3.2 Approximation for n infinite

The asymptotic relationship between a t-distributed random variable and a normal random variable by means of the Cornish-Fisher expansion could be interesting for cases of big degrees of freedom. Shaw (2006) provides the Cornish-Fisher expansion for a t-distributed random variable, with zero mean and unit variance, in terms of a standard normal random variable distribution. This reduces considerably the computational time as, in this case, it is possible to use a double Gauss-Hermite quadrature, instead of a Simpson quadrature. Basically, Shaw (2006) provides the relationship

$$s = z + \frac{1}{4n} z \left(z^2 - 3\right) + \frac{1}{96n^2} z \left(5z^4 - 8z^2 - 69\right) + ... \tag{9}$$

where s is the t-distributed random variable with n degrees of freedom, and z is a standard normal random variable. To check the accuracy of the approximation (9), Table 3 displays the spreads (in basis points) for different tranches in a two asset classes, 50-named CDO with quarterly payments under different correlation parameters for various degrees of freedom. As in the previous section, correlation parameters correspond to factors of each asset classes. As in previous examples, risk neutral default probabilities have been also fixed at 0.1 and 0.3 for each asset class. The recovery rate is fixed at 40%.

[INSERT TABLE 3 AROUND HERE]

Basically, two general models have been computed: two Gaussian factors and a two t-Student factor. The last column refers to the two t-Student factor approximation using the Cornish-Fisher expansion. Correlation coefficients and degrees of freedom for each one are displayed in the table.[11] The Gaussian case is presented to get intuition of how far we are from the asymptotic result. As expected, the larger the degree of freedom, the higher the accuracy of the results. The differences between the equity tranche range from 20% for 10 degrees of freedom to 12% for a 15 degrees of freedom case. In a similar way, considering the mezzanine cases, differences go from 21% to 11%. There are no substantial changes in the case of senior tranche. It is worth to remember that computations under the *exact* t-Student distribution have been done using a numerical quadrature and, so, they are exposed to numerical errors.

4 Sensitivity analysis

Now, we are interested in the prices of the CDO under two different scenarios: changes in portfolio size and correlation. Firstly, we present some standard measures in risk management as the Value at Risk and the Conditional Value at Risk ones. Secondly, we analyze the sensitivity of different tranches to changes in correlation.

4.1 VaR and CVaR

Value at Risk (VaR) and Conditional Value at Risk (CVaR) are usually taken as representative risk measures for a portfolio. VaR is defined as the percentile of the distribution of portfolio losses given a certain level of confidence.[12] Artzner, Delbaen, Eber and Heath (1999) enumerates some limitations of the VaR measure and discuss some interesting properties of a proper measure of risk. According to this, we also include the CVaR measure,[13] de-

[11]The two Gaussian factor model is equivalent to a two t-Student factor model with infinite degrees of freedom.

[12]See Duffie and Pan (1997) for details.

[13]This measure was posited by Acerbi, Nordio and Sirtori (2001).

fined as the expected loss in a portfolio conditional to a certain loss threshold u, that is,

$$CVaR_u = E\left[x|x > u\right]$$

where the sign of the inequality has been changed because we are working directly with the total loss distribution.

Table 4 includes the VaR and CVaR measures for a 100-named CDO composed by two different risky asset classes under different correlation parameters and different proportions in the portfolio. The (yearly) risk-neutral default probabilities for each obligor of asset class 1 have been fixed at 0.01, and 0.05 for those which belong to asset class 2. For the sake of simplicity, all the factors - systematic and idiosyncratic - in the model (7) have been fixed at 5 degrees of freedom. The first column includes the correlation coefficients corresponding to both factors of each class (i.e. 0.1/0.3 means a correlation coefficient of 0.1 (0.3) for *both* factors in asset class 1 (2)). Correlation coefficients equal to zero refers to independence case between obligors. With the purpose of analyze the response of the loss distribution generated by the model (7) with respect to different sizes of the portfolio, the remaining columns show different percentage sizes of the portfolio. The first term refers to the portfolio percentage of class 1, and so on.

[INSERT TABLE 4 AROUND HERE]

Needles to say that two main results arise form Table 4. Firstly, the higher the correlation the higher expected extreme loss as measured by VaR and CVaR, as expected. Secondly, an increase in the percentage of the risky asset (asset class 2) produces an increase in the losses of the portfolio.

4.2 Sensitivity to correlation

To analyze the sensitivity to correlation of the model (7) we have created a CDO based on a portfolio of 50 names. Individual default probabilities have been fixed at 1% for asset class 1 and 5% for asset class 2, respectively. To reduce the computational cost of the implementation, we have set the distribution of the two systematic factors as t-Student ones with 15 degrees of freedom. We have used the results of Shaw (2006) developed in Section 2.5, without loss of generality. All simulations have been performed using a double Hermite quadrature with 64 nodes. Idiosyncratic factors have been fixed at 5 for each asset class. To search differences in the portfolio composition, we have applied our study to two different sized portfolio: equally weighted portfolio (50% asset class 1, 50% asset class 2) and risky portfolio (25% asset class 1, 75% asset class 2).

Figures 4 and 5 display the spreads obtained under different sets of correlations for the equally weighted and risky portfolios. In general, the convexity pattern of the equity-senior curves remains constant in both cases, which is consistent with the preference (aversion) for risk on equity (senior) tranche investors, as expected.

Regarding changes in correlation, Figure 4 reveals that correlation with risky asset classes are, by large, responsible of changes in the value of equity spreads. When it comes to the risky portfolio (Figure 5), it is interesting to observe how changes produced by the correlation in asset class 1 or 2 (see Equity and Mezzanine tranches) produce almost the same effects.

[**INSERT FIGURE 4 AROUND HERE**]

As it is also expected, an increase in global correlation raises the spreads of senior tranche. A higher correlation increases the probability of big losses, which is reflected in the senior spreads. This feature could be also mentioned (in a different scale) to the case of the risky portfolio in Figure 5.

[**INSERT FIGURE 5 AROUND HERE**]

5 Econometric Framework

This section focuses on the parameters estimation of the model (7). As pointed out in Embretchs, Frey and McNeil (2005), the statistical estimation of parameters in many industrial models are *simply assigned by means of economic arguments or proxies variables.* We will develop an exercise of formal estimation using some well known econometric tools as logit-probit regressions.[14] Due to the features of our data, some cautions must be taken to understand our results. This is due to the shortage of relevant data (for instance, rates of default of high-rated companies) or the sample size, as was also noticed by Embretchs, Frey and McNeil (2005). These authors also provide a more general discussion on the statistical estimation of portfolio credit risk models.

[14]Standard references on this type of regressions using grouped data can be found in Novales (1993) or Greene (2003).

5.1 Estimation Techniques

As suggested by Schonbucher (2003) or Embretchs, Frey and McNeil (2005), the estimation of parameters in the expression (7) will be carried on using the models for discrete choice of proportions data. Basically, the idea consists in explaining the sample rates of default p_i (where i refers to the asset class or group) as an approximation to the population rates of default P_i plus an error term, ε_i. The idea behind is to link the population probability with some function $F(\cdot)$ over a set of explanatory factors $\mathbf{x}_i$ and their coefficients β, as follows:

$$p_i = P_i + \varepsilon_i = F(\mathbf{x}_i'\beta) + \varepsilon_i \tag{10}$$

To be interpreted as a probability, the function $F(\cdot)$ must be bounded and monotonically increasing in the interval $[0, 1]$. Some widely used functions for F are the standard Normal distribution, which corresponds to the probit model, or the uniform distribution, which results in the linear probability model.

As suggested by Greene (2003), we could use regression methods as well as maximum likelihood procedures to estimate the set of coefficients β of the expression (10). For example, in the case of the probit regression, the relationship between the sample rates of default p_i and their population counterparts are

$$p_i = P_i + \varepsilon_i \to \Phi^{-1}(p_i) = \Phi^{-1}(P_i + \varepsilon_i)$$

which could be aproximated by (Novales, 1993)

$$\Phi^{-1}(p_i) \simeq \mathbf{x}_i'\beta + \frac{\varepsilon_i}{f(\mathbf{x}_i'\beta)}$$

where $\Phi(\cdot)$ denotes the distribution function of a standard Normal variable. As mentioned in Novales (1993), the last expression suggests that we can estimate the parameter vector β by regressing the sample probits $\Phi^{-1}(p_i)$ on the variables x. Considerations about the heteroscedasticity of the residual can be found in the cited reference.

To check the model's goodness of fit, Novales (1993) also provides a comparison of different regressions (probit, logit or lineal) in terms of the mean square error (MSE) . The statistic s is defined as

$$s = \sum_{1}^{T} \frac{n_i\left(p_i - \widehat{P}_i\right)^2}{\widehat{P}_i\left(1 - \widehat{P}_i\right)} \sim \chi^2_{T-k} \tag{11}$$

where n_i represents the sample size of the data (subscript i refers to asset class or group) and p_i, $\widehat{P}_i$ are the observed and estimated frequencies, respectively. The statistic s follows a chi-square distribution with $T-k$ degrees of freedom, sample length T and k restrictions.

5.2 Variables and estimation

With the intention of illustrating the estimation of the model (7), we have chosen a set of six explanatory variables for the rates of default: the real Growth Domestic Product (GDP), the Consumers Price Index (CPI), the annual return on the S&P500 index (SP_ret), its annualized standard deviation (SP_std), the 10-year Treasury Constant Maturity Rate (10_rate) and the Industrial Production Index (IPI).[15] As dependent variables we have the annual rates of default for two investment grades: non investment grade (SG) and investment grade (IG), both collected from Hamilton, Varma, Ou and Cantor (2005). Due to the availability of default rate data, the sample period has been taken from 1970 to 2004, which results in 35 observations. A summary of the main statistics and the correlation coefficients is presented in Tables 5 and 6, respectively.

[INSERT TABLES 5 AND 6 AROUND HERE]

To visualize the influence of the proposed explanatory variables in the default rates, Figures 6 and 7 represent the scatter plots of different investment classes versus different explanatory variables. Figure 6 seems to corroborate what we could guess departing from the correlation parameters included in Table 6: the standard deviation of the S&P 500 returns, the GDP and the CPI can be good candidates for explaining the default rate in the case of the non-investment firms. Additionally, at a certain degree, the S&P 500 return can be added to this list as a possible explanatory variable in the case of the investment firms.[16]

[INSERT FIGURES 6 AND 7 AROUND HERE]

[15] All data are available from the Federal Reserve Bank of St. Louis webpage (www.stlouisfed.org) except the S&P 500 index level, which has been taken from Bloomberg.

[16] As Figure 7 reveals, due to the high number of null observations in the IG sample, conclusions about the factors affecting IG rates should be taken carefully.

One step ahead is to compute how much of the sample can be explained by the set of variables under study. To answer this question, we regress the explanatory variables on the non-investment and investment rates. Table 7 shows the results. The first row corresponds to the different independent variables under study. The first column contains the model under study - linear, probit, logit - and the different regressed variables (SG and IG default rates). Second to eighth columns display different betas obtained under different models. Finally, the last column shows the s statistic defined in (11), which will be used as a naïve benchmark: if the whole set of independent variables explains some quantity of the sample, two variables would explain *"less"*: the pair of variables whose s value are closest to the benchmark could be good candidates as common factors in the model (7).

[**INSERT TABLE 7 AROUND HERE**]

We start with regressions on SG rates. One main reason recommends this procedure: their data are more relevant to determine which factors may cause default. Up to a point, conclusions on the factors will be more robust. Previous regressions suggest choosing the variables GDP, CPI, IPI and SP_stdas common factors in the model (7). These variables minimize the statistic (11) with respect to other pairs of alternatives. Finally, we select GDP and CPI as common factors according to two main reasons:

1. Firstly, the Industrial Production Index could be seen as a proxy of the GDP and its information could result redundant. Moreover, regressions of probit-logit models using these two variables support the choice of GDP against the IPI.

2. Secondly, regressions on the parameter SP_std give a beta close to the precision imposed to our estimated parameters (10^{-4}).

Table 8 presents in columns the OLS[17] estimates for betas of independent term, GDP variable and CPI variable, respectively, using the SG rates. Confidence intervals at the 95% level are displayed into brackets. The rows in this table also display regression results for linear, probit and logit models. The last row contains the value of the statistic (11) obtained for each case.

[17]GLS estimates have not been computed due to the sample size.

Attending to the goodness-of-fit criteria using s, the OLS probit model regression provides the best fit to the sample[18]. Overall, all the beta estimates corresponding to OLS regressions are negative, except for the independent term of the linear model, which leads to higher s statistic. Results concerning to OLS regressions can be interpreted as follows: a negative beta implies an increasing on default probabilities. In line with this, as expected, a decrease in GDP rates may produce an increase on SG default rates. Surprisingly, an increase in the CPI rate could diminish the rates of default, which might result counter-intuitive.

Table 9 shows regression results for linear, probit and logit models using IG rates as dependent variable. It is worth to notice that results are not conclusive as 63% of the sample under study are zeros. GLS estimations do not make sense in this context. In order to avoid numerical problems in the estimation, we have added the quantity 0.00005 to the sample, as suggested by Greene (2003). The first row displays the independent term and explanatory variables. Each pair of the following rows contains firstly the different values of betas obtained using two variables (GDP and CPI); secondly, their values using *only* the GDP variable. Maximum likelihood estimates (available upon request) for the probit and logit models are close to those parameters obtained for respective OLS models. We have estimated GDP variable alone with the intention of analyzing the explanatory power of the GDP on IG rates. First to second rows show the model and procedure used. The last column displays the value for the statistic s. At a certain degree, results on Table 8 could support the inverse relationship between the explanatory variables and the IG rates of default, as previously noted for the SG case.

[INSERT TABLES 8 AND 9 AROUND HERE]

5.3 Interpretation of coefficients

As was pointed out by Elizalde (2005), due to the difficulty of interpreting what the correlation term represents, estimating the correlation term in factor models is not an evident task. Looking at equation (10), the estimate β describes the effect from the explanatory factor x through a non-linear transformation of the firm's asset value, which itself is unobserved, as it is

[18]Maximum Likelihood estimates (available upon request) for models probit and logit are close to those parameters obtained for respective OLS models.

also noticed by Elizalde (2005). As this fact complicates understanding the proper correlation term, the author enumerates some measures used ad hoc by practitioners, as equity return correlations, to conclude about the insufficiency (and scarcity) of papers that deals with this problem.

Our interpretation of coefficients in the model (7) goes in the direction of the econometric explanation for the coefficients of the linear, logit and probit models, that is, the influence of the exogenous variables on the endogenous one. In other words, the (relative) impact of the explanatory variables on the probability of default. Following Novales (1993), this interpretation of estimates for the linear model must differ to that for the logit and probit models.[19] This is the main reason why we split our results in two tables, Tables 10 and 11, that include - respectively - the estimation of the linear and logit-probit models.

[INSERT TABLES 10 AND 11 AROUND HERE]

Regarding the estimates of the linear probability model, Table 10 reflects the contribution of the two explanatory random variables to the probability of default. The main conclusions are obtained from the default probabilities of non-investment grade assets (SG), but can also be extended to the investment grade (IG) ones. Looking at Table 10, it is interesting to observe the sign of the coefficients, which is negative: the more we decrease the GDP or the CPI, the more we increase the rates of default. Given the value of the coefficients, the variables have the same contribution to the default probability. With reference to the fit of the model to the data, under the null hypothesis that the goodness-of-fit to the sample is good, we cannot reject that the linear probability model could explain the results obtained.

Table 11 includes the ratio between estimates for SG and IG series for probit and logit models, respectively. By and large, the conclusions are the same for all the series under study. According to Novales (1993), the ratio between the estimated betas measures the relative contribution of the explanatory variables on the default probability. Results are consistent to those obtained for the linear probability model: the negative sign of the explanatory variables, which reflects an opposite effect between default ratios and the macroeconomic variables. Moreover, the relative contribution of

[19]For example, the relationship between the explanatory and explained variables in the probit model is non-linear while the linear probability model implies linearity between dependent and independent variables.

the explanatory variables remains equal, as was also derived from Table 10. Finally, we do not reject the goodness-of-fit of the model using confidence levels of 95% and 99%.

6 Conclusions

The current success of the credit derivatives market for tranched products is one of the biggest ones seen within the financial industry. The standard pricing model, widely used by the practitioners, is the Gaussian one-factor model (Vasicek, 1991). However, some assumptions underlying this model are probably too restrictive. These features concern, among others, to those of homogeneity of asset classes involved, or the exposure to one sources of systematic risk.

In a more realistic setting, Schonbücher (2003) or Lando (2004) pointed out that a credit portfolio is composed by different asset classes or buckets, attending to criteria as, for example, investment grade, non-investment grade assets or industry. The exposure of a credit portfolio to a set of common risk factors could be significant between groups, but should be homogeneous within them. In line with this, the idea of two groups of assets treated in different ways could become a more realistic assumption than that used previously in the literature.

With the aim of contributing to the current literature, this article considers a family of models that takes into account the existence of different asset classes or regions of correlation. Thus, we analyze a model in the line of the two assets-two Gaussian factor model of Schonbücher (2004). In this paper we generalize the standard model by means of a two by two model (two factors and two asset classes). We assume two driving factors (business cycle and industry) with independent t-Student distributions, respectively, and allow the model to distinguish between portfolio assets classes. One of the main implications of considering a t-Student distribution is that we assign a higher probability to high default events.

Our work contributes to the existing literature in the analysis of these asset class models. To the best of our knowledge, no similar studies has been reported yet in this direction. Regarding to distributional assumptions, we extend the standard gaussian model by considering the t-Student distribution. In this way, we deal with a more general model with the additional advantage that includes the Gaussian model as a particular case. We also provide the econometric framework for assessing the parameters of the posited model. Finally, an empirical application with Moody's data has been presented as an illustration of the methodology proposed.

References

[1] Acerbi, C., Nordio, C. and Sirtori, C. (2001): Expected Shortfall as a Tool for Financial Risk Management, Working Paper, Italian Association for Financial Risk Management.

[2] Andersen, L., Sidenius, J. and Basu, S. (2003): All your hedges in one basket, *Risk*, November: 67–72.

[3] Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999) Coherent Measures of Risk, *Mathematical Finance*, 9: 203–208.

[4] BBA Credit Derivatives Report (2006).

[5] Bielecki,A. and F. Rutkowski (2002): Credit Risk, Springer.

[6] Black, F. and Cox, J.C. (1976): Valuing Corporate Securities: Some Effects of Bond Indenture Provisions, *Journal of Finance*, 31: 351–367.

[7] Duffie, D. and Garlenau, N. (2001): Risk and Valuation of Collateralized Debt Obligations, *Financial Analysts Journal*, January/February: 41–59.

[8] Duffie, D. and Pan, J. (1997): An overview of VaR, *Journal of Derivatives*, 4: 98–112.

[9] Elizalde, A. (2005): Credit Risk Models IV: Understanding and pricing CDOs, Working Paper, CEMFI, (available at www.abelelizalde.com).

[10] Gibson, M. S. (2004): Understanding the Risk of Synthetic CDOs, Working Paper 2004-36, Federal Reserve Board.

[11] Glasserman, P. and Suchitabandid, S. (2006): Correlation Expansions for CDO Pricing, to appear in *Journal of Banking and Finance.*

[12] Green, W. H. (2003): Econometric Analysis, 5th edition. Pearson Education, New Jersey.

[13] Gregory, J. and Laurent, J.P. (2004): In the Core of Correlation, *Risk*, October: 87–91.

[14] Gregory, J. and Laurent, J.P. (2005): Basket Default Swaps, CDO's and Factor Copulas, *Journal of Risk*, 7, 4, 103–122.

[15] Gupton, G., Finger, C. and Bhatia, M. (1997): CreditMetrics. Technical document, JPMorgan, April. www.creditmetrics.com

[16] Hamilton, D. T., Varma, P., Ou, S. and Cantor, R. (2005): Default and Recovery Rates of Corporate Bond Issuers, 1920-2004, Moody's Investor Service, January.

[17] Hull, J. and White, A. (2004): Valuation of a CDO and an n-th to Default CDS Without Monte Carlo Simulation, *Journal of Derivatives*, Winter: 8–23.

[18] Lucas, A., Klaassen, P., Spreij, P. and Staetmans, S. (2001): An analytic approach to credit risk of large corporate bond and loan portfolios, *Journal of Banking and Finance*, 25 (9): 1635-1664.

[19] Longstaff, F. A. and Rajan, A. (2006): An empirical analysis of the pricing of collateralized debt obligations, Working Paper, UCLA Andersson School.

[20] Mashal, R. and Naldi, M. (2002): Calculating portfolio loss, *Risk*, August: 82–86.

[21] Mardia, K. V., Kent, J. T. and Bibby, J. (1979): Multivariate Analysis, Ed. Academic Press, London.

[22] Merton, R. C. (1974): On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, *Journal of Finance*, 29: 449-470.

[23] Novales, A. (1993): Econometría, 2^a ed. McGraw-Hill, Madrid.

[24] Tavakoli, J. M. (2003): Collateralized Debt Obligations and Structured Finance. New Developments in Cash and Synthetic Securitization, John Wiley and Sons, Inc., Hoboken, New Jersey.

[25] Vasicek, O. (1991): Limiting Loan Loss Probability Distribution, KMV Corporation.

Table 1: Spreads of different tranches for a 100-names CDO under different correlation parameters

Model	Standard gaussian		2 Gaussian factors		t-Student	
Correlation	0.1	0.3	0.1/0.1	0.1/0.3	0.1	0.3
Attachment points (%)	Tranche Spread (basis points)					
0–3	2117	1320	2186	1656	2485	2135
3–10	86	110	508	653	131	233
10–100	0.02	1	7	29	1	9

Table 2: Spreads of different tranches for a 100-names CDO under different correlation parameters

Model	Standard gaussian		2 Gaussian factors		t-Student	
Correlation	0.1/0.1	0.1/0.3	0.1/0.1	0.1/0.3	0.1/0.1	0.1/0.3
Attachment points (%)	Tranche Spread (basis points)					
0–3	7166	5016	5421	3154	6899	4977
3–10	1040	929	1581	1401	1254	1302
10–100	5	11	43	122	27	65

Table 3: Spreads of different tranches for a 50-names CDO under different correlation parameters

Model	2 Gaussian factors	t-Student		CF t-Student	
Correlation	0.1/0.3	0.1/0.3	0.1/0.3	0.1/0.3	0.1/0.3
Degrees of freedom (n1/n2)	∞/∞	10/10	15/15	10/10	15/15
Attachment points (%)	Tranche Spread (basis points)				
0–3	3015	3924	3711	3270	3296
3–10	1380	1236	1198	1025	1083
10–100	289	160	160	166	160

Table 4: VaR and CVaR measures for a 100-named CDO composed by two different risky asset classes under different correlation parameters and different proportions in the portfolio. Degrees of freedom are fixed at 5 for all factors. The first (second) coefficient refers to the proportion (in percentage) of Asset Class 1 (2) in the portfolio. Yearly risk-neutral probabilities have been fixed to 0.01 for each obligor of Asset Class 1, and 0.05 for those of Asset Class 2.

	VaR 99%			CVaR 99%		
Correlation	25/75	50/50	75/25	25/75	50/50	75/25
0.0/0.0	10	9	7	18	9	9
0.1/0.1	25	19	13	35	27	19
0.1/0.3	57	40	23	67	48	30
0.3/0.3	61	47	33	72	61	50

Table 5: Descriptive Statistics

Descriptive Statistics	mean	median	std	skewness	kurtosis	max	min
IG	0.0006	0	0.0011	2.2481	8.2051	0.0049	0
SG	0.0389	0.0345	0.0288	1.0213	3.0976	0.1058	0.0042
SP_ret	0.0297	0.0439	0.0675	-0.6110	2.5129	0.1310	-0.1208
SP_std	89.4702	39.3947	106.3397	1.4176	3.7198	390.1884	7.5537
GDP	0.0130	0.0148	0.0087	-0.6077	2.8766	0.0301	-0.0085
CPI	0.0205	0.0159	0.0128	1.2338	3.7824	0.0564	0.0052
10_rate	7.7577	7.4176	2.4231	0.8413	3.2283	13.9214	4.0139
IPI	0.0111	0.0124	0.0185	-0.4614	3.2866	13.9214	-0.0409

Table 6: Correlation matrix

	IG	SG	SP_ret	SP_std	GDP	CPI	10_rate	IPI
IG	1.0000							
SG	0.4106	1.0000						
SP_ret	-0.2388	-0.2153	1.0000					
SP_std	0.1760	0.3472	-0.0424	1.0000				
GDP	-0.1431	-0.3541	0.4988	0.0029	1.0000			
CPI	-0.1847	-0.3209	-0.2806	-0.4959	-0.4937	1.0000		
10_rate	-0.1043	-0.2344	0.0611	-0.6113	-0.1070	0.6060	1.0000	
IPI	-0.1499	-0.2930	0.2687	-0.0385	0.6671	-0.3252	-0.1344	1.0000

Table 7: Regression of Non-investment (SG) and Investment (IG) default rates with respect to different explanatory variables. SP_ret and SP_std are the yearly returns and standard deviation of the S&P500 Index. GDP, 10_rate, CPI and IPI are the Growth Domestic Product, ten year constant maturity rate, Consumer Price Index and Industrial Production Index, respectively. Finally, s is the fit statistic defined in (11).

Model	β_0	SP_ret	SP_std	GDP	10_rate	CPI	IPI	s
linear								
SG	0.0792 [0.0344 0.1241]	−0.0511 [−0.1904 0.0881]	0.0000 [−0.0001 0.0001]	−1.9518 [−3.4391 −0.4646]	0.0025 [−0.0022 0.0074]	−1.6900 [−2.6699 −0.7102]	−0.1222 [−0.7063 0.4617]	11.5543
IG	0.0011 [−0.0010 0.0033]	−0.0045 [−0.0113 0.0022]	0.0000 [−0.0000 0.0000]	−0.0179 [−0.0898 0.0541]	0.0001 [−0.0001 0.0003]	−0.0391 [−0.0865 0.0082]	−0.0060 [−0.0343 0.0222]	1.7799
probit								
SG	−1.3794 [−1.8624 −0.8963]	−0.8812 [−2.3799 0.6176]	0.0003 [−0.0007 0.0014]	−23.7382 [−39.7437 −7.7326]	0.0489 [−0.0028 0.1005]	−26.7956 [−37.3405 −16.2507]	−1.3676 [−7.6526 4.9174]	13.9414
IG	−3.3965 [−4.3212 −2.4717]	−1.2129 [−4.0822 1.6565]	0.0003 [−0.0018 0.0023]	−17.1498 [−47.7920 13.4924]	0.0482 [−0.0506 0.1471]	−16.6621 [−36.850 13.5258]	3.6291 [−8.4033 15.6616]	6.0474
logit								
SG	−2.3439 [−3.4585 −1.2293]	−2.1634 [−5.6218 1.2951]	0.0007 [−0.0017 0.0032]	−54.8455 [−91.7787 −17.9123]	0.1191 [−0.0001 0.2383]	−64.7566 [−89.0892 −40.4240]	−3.2499 [−17.7527 11.2528]	15.0393
IG	−8.1171 [−11.4907 −4.7434]	−4.1999 [−14.6678 6.2680]	0.0011 [−0.0064 0.0086]	−64.2390 [−176.0283 47.5503]	0.1772 [−0.1835 0.5380]	−60.2912 [−133.9410 13.3587]	14.5665 [−29.3303 58.4634]	6.7882

Table 8: Results for regressions of Non-investment (SG) default rates on real Growth Domestic Product (GDP) and Consumer Price Index (CPI) as explanatory variables.

SG	Model	β_0	GDP	CPI	s
linear	OLS	0.0979 [0.0733 0.1226]	−2.2380 [−3.2486 −1.2274]	−1.4690 [−2.1548 −0.7832]	21.8030
probit	OLS	−1.0785 [−1.3568 −0.8002]	−26.7705 [−38.1822 −15.3587]	−21.6288 [−29.3730 −13.8846]	14.7653
logit	OLS	−1.6293 [−2.2769 −0.9816]	−61.8990 [−88.4563 −35.3416]	−51.7790 [−69.8014 −33.7567]	15.4018

Table 9: Results for regressions of Investment (IG) default rates on real Growth Domestic Product (GDP) and Consumer Price Index (CPI) as explanatory variables.

IG	Model	β_0	GDP	CPI	s
linear	OLS	0.0017 [0.0005 0.0029]	−0.0391 [−0.0884 0.0102]	−0.0290 [−0.0624 0.0045]	5.0921
	OLS	0.0009 [0.0002 0.0015]	−0.0181 [−0.0623 0.0261]	–	2.2557
probit	OLS	−3.1421 [−3.6496 −2.6347]	−14.1617 [−34.9687 6.6453]	−11.2345 [−25.3546 2.8856]	6.4347
	OLS	−3.4781 [−3.7663 −3.1899]	−5.9888 [−24.5193 12.5417]	–	8.6336
logit	OLS	−7.1801 [−9.0318 −5.3284]	−50.7973 [−126.7275 25.1328]	−40.7355 [−92.2634 10.7923]	7.1658
	OLS	−8.3983 [−9.4493 −7.3472]	−21.1631 [−88.7533 46.4271]	–	9.6372

Table 10: Estimated coefficients for the linear probability model. s refers to the fit statistic.

Linear model	β_0	GDP	CPI	s	$\chi^2_{95\%}\,(32)$	$\chi^2_{99\%}\,(32)$
SG	0.0979	-2.2380	-1.4690	21.8030	No reject	No reject
IG	0.0017	-0.0391	-0.0290	5.0921	No reject	No reject

Table 11: Estimated coefficients for probit and logit models. s refers to the fit statistic.

	β_0	$\frac{GDP}{CPI}$	s	$\chi^2_{95\%}(32)$	$\chi^2_{99\%}(32)$
Probit model					
SG	-1.0785	1.2377	14.7653	No reject	No reject
IG	-3.1421	1.2606	6.4347	No reject	No reject
Logit model					
SG	-1.6293	1.1954	15.4018	No reject	No reject
IG	-7.1801	1.2470	7.1658	No reject	No reject

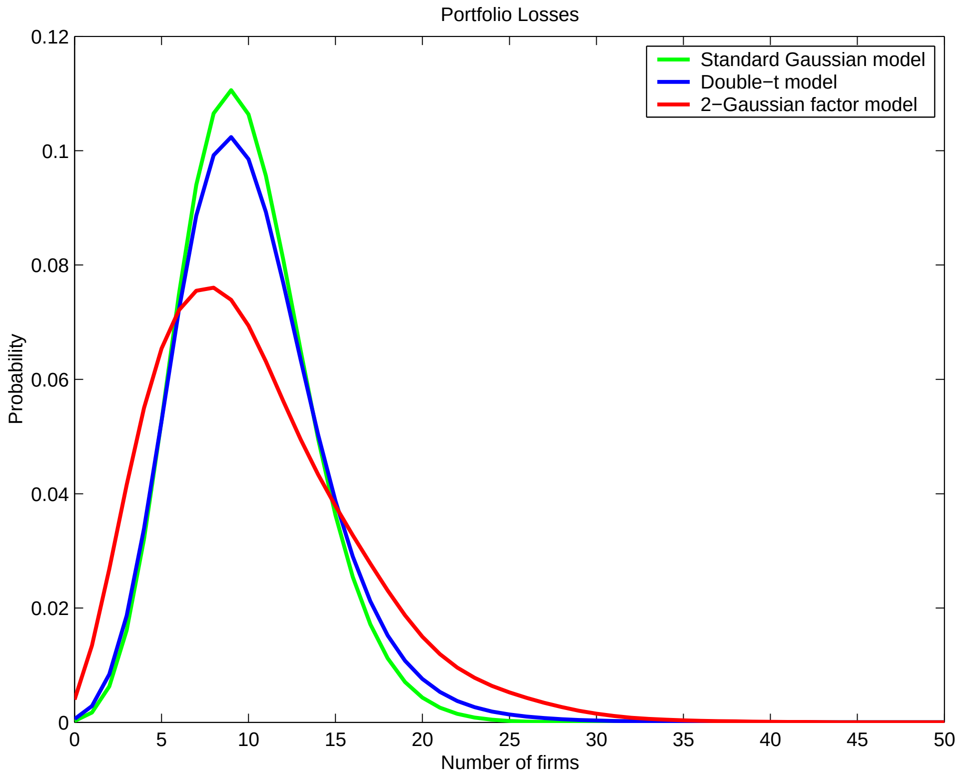

Figure 1: Total loss distribution of a 100-firms portfolio for different models. Individual default probabilities are fixed at 1%. Correlation for the Standard gaussian and t-Student distribution (with 6 degrees of freedom) is 0.3. Correlations for the Double Gaussian model are 0.1, 0.3.

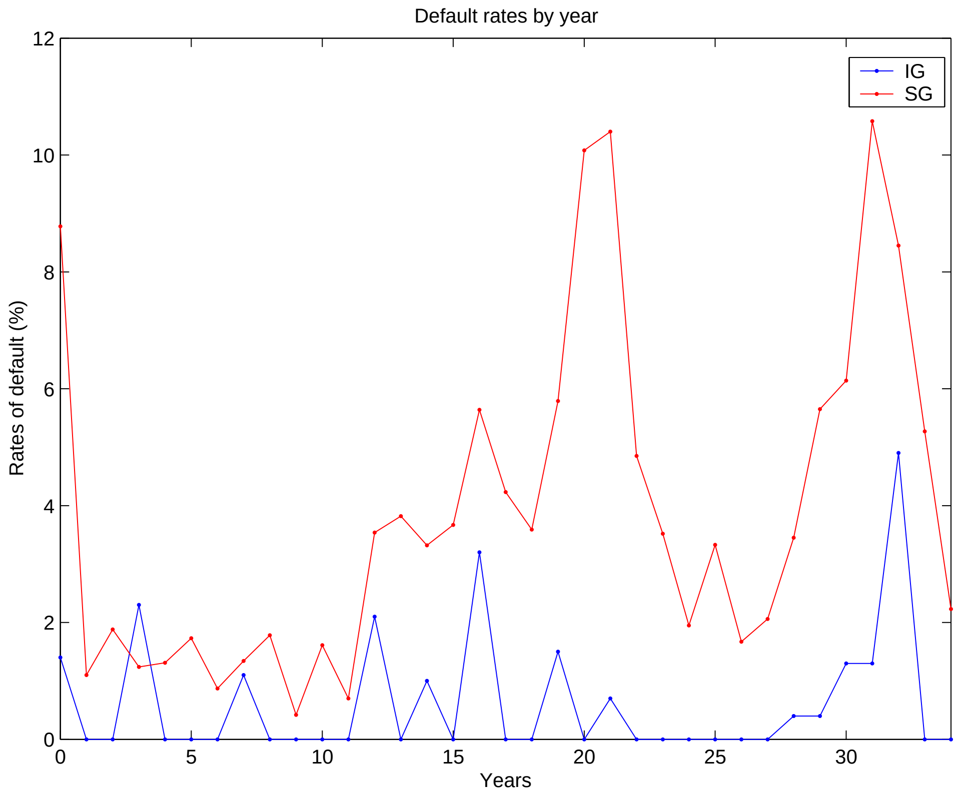

Figure 2: Default (yearly) rates for Investment and Non-investment grades. Rates for investment data have been multiplied by ten with the pourpose of comparing the data. Source: Moody's.

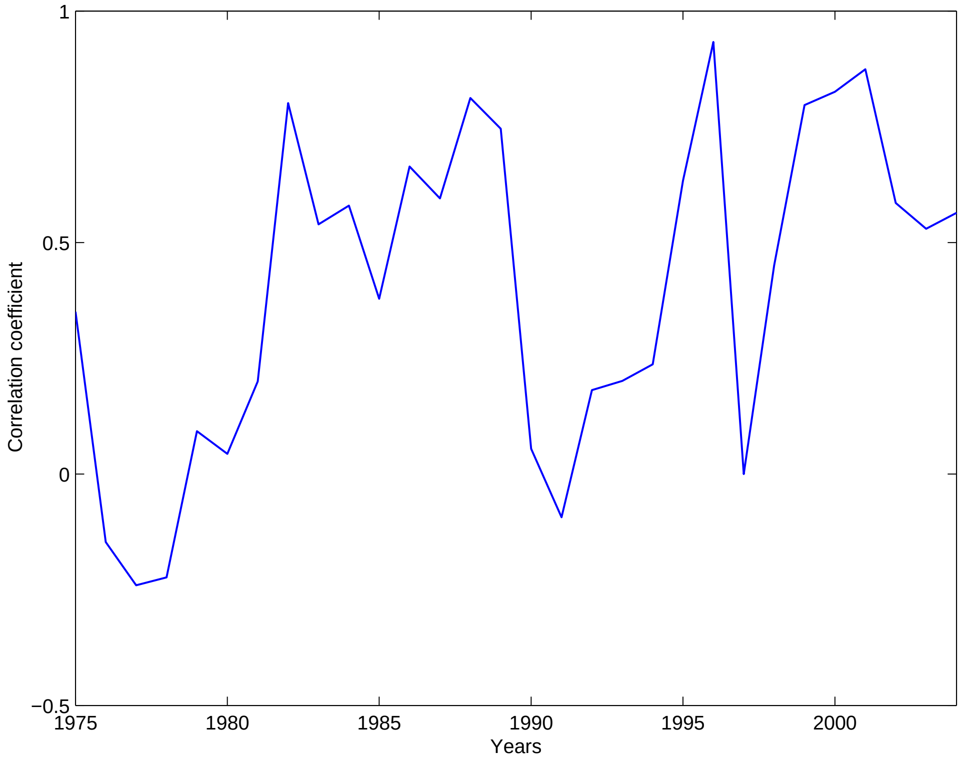

Figure 3: Correlation coefficients using a moving window of 5 years.

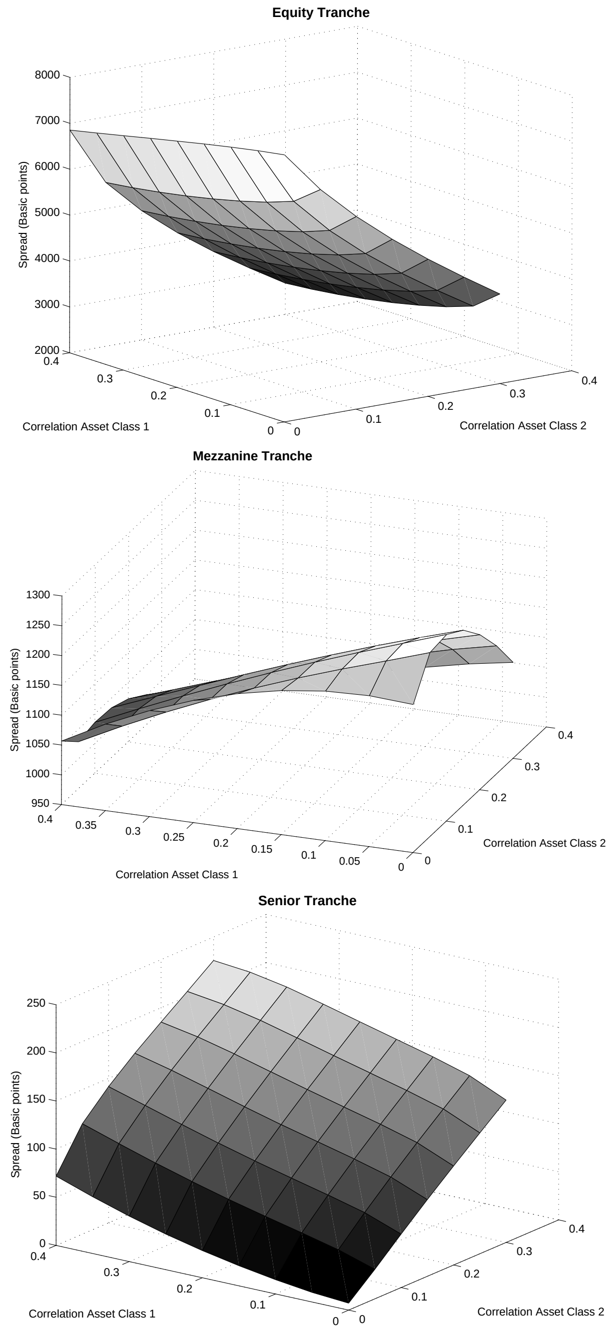

Figure 4: Spreads under different correlations for a equally weighted (50%-50%) portfolio.

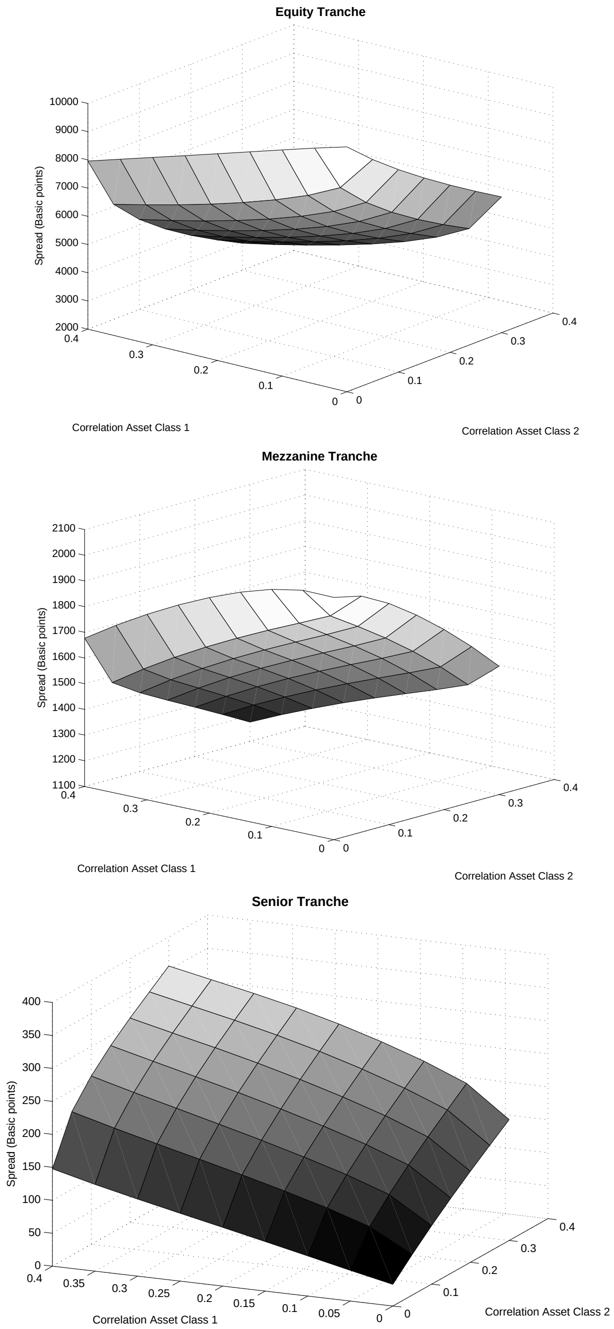

Figure 5: Spreads under different correlations for a 25%-75% portfolio.

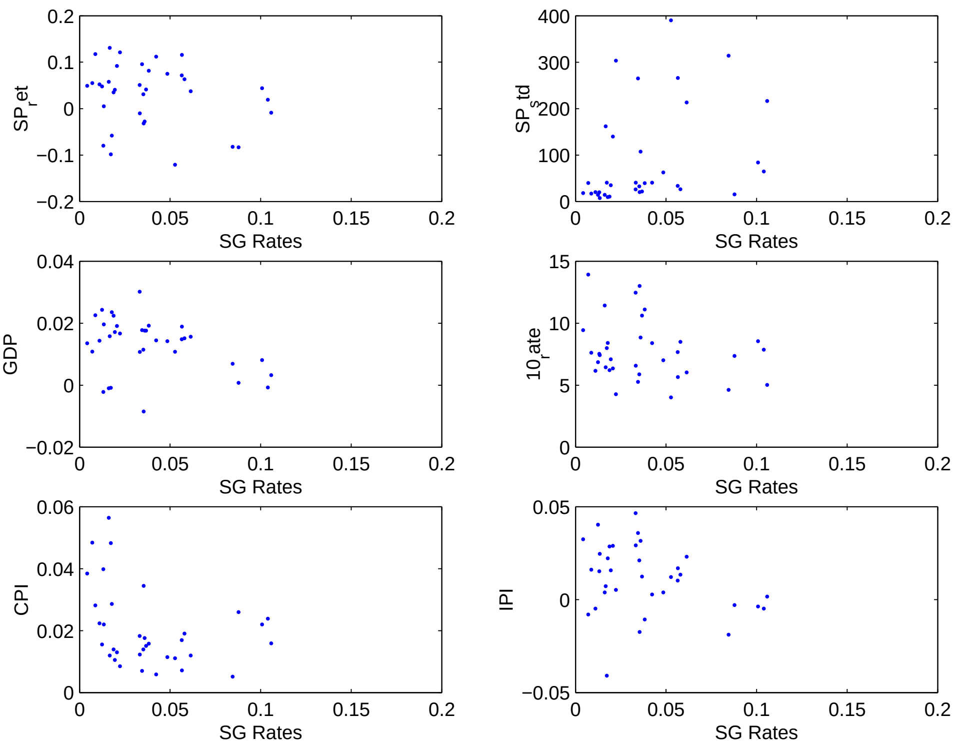

Figure 6: Non-investment grade rates of default (SG) versus different explanatory variables

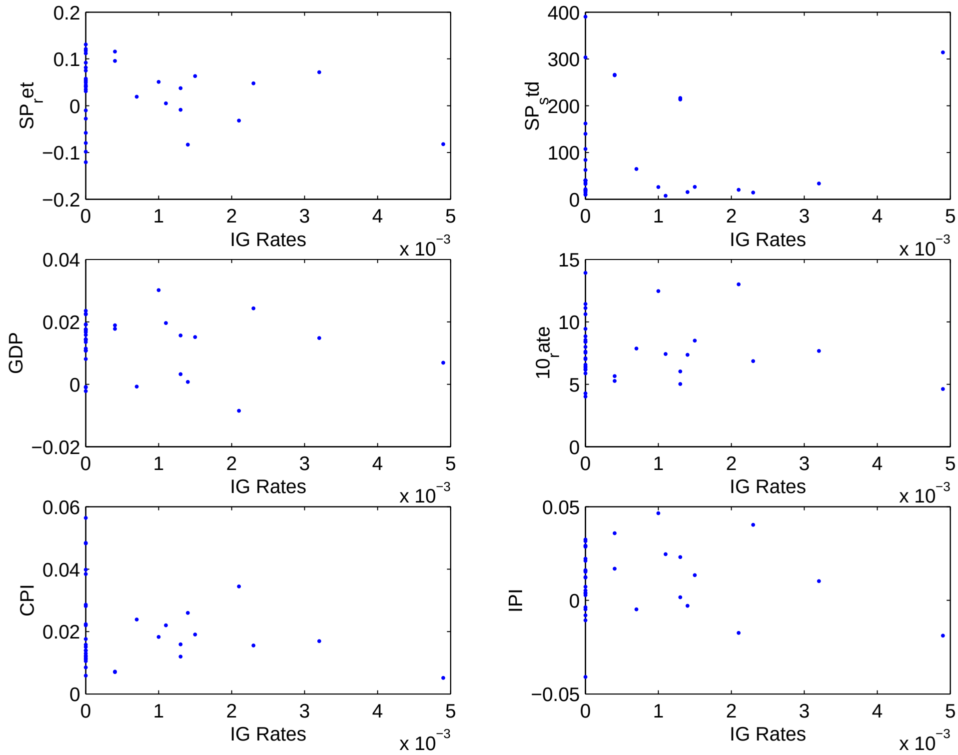

Figure 7: Investment grade rates of default (IG) versus different explanatory variables