Convex-Concave-Convex Distributions in Application to CDO Pricing

***Alex Veremyev*[1], *Peter Tsyurmasto*[2], *Stan Uryasev*[3], *R. Tyrrell Rockafellar*[4]**

RESEARCH REPORT # 2012-1

Risk Management and Financial Engineering Lab
Department of Industrial and Systems Engineering
University of Florida, Gainesville, FL 32611

Current Version: June 6, 2012

Correspondence should be addressed to: Stan Uryasev

Abstract

This paper proposed a new Convex-Concave-Convex (CCC) probability distribution function specified by a system of linear constraints. CCC distribution can be used for various applications involving calibration of probabilities. The suggested approach is used for calibrating of "implied copula" for Collateralized Debt Obligations (CDOs). The loss distribution is found by maximizing the entropy with no-arbitrage constraints based on bid and ask prices of CDO tranches. The case study shows that the approach has a stable performance. Codes used for conducting numerical experiments are provided for verification purposes.

Keywords: *Convex-Concave-Convex probability distribution, CCC distribution, implied copula, CDO pricing.*

[1]University of Florida, ISE Department, P.O. Box 116595, 303 Weil Hall, Gainesville, FL 32611-6595; e-mail: averemyev@ufl.edu.

[2]University of Florida, Risk Management and Financial Engineering Lab, ISE Department, P.O. Box 116595, 303 Weil Hall, Gainesville, FL 32611-6595; e-mail: tsyurmasto@ufl.edu.

[3]University of Florida, Risk Management and Financial Engineering Lab, ISE Department, P.O. Box 116595, 303 Weil Hall, Gainesville, FL 32611-6595; e-mail: uryasev@ufl.edu.

[4]Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195-4350 email: rtr@math.washington.edu

Introduction

The pricing of Collateralized Debt Obligations (CDOs) contracts is a difficult quantitative problem faced by credit risk markets. The main issue is uncertainty about obligors default risk. This paper considers a so-called "implied copula" CDO pricing model for calibrating default distribution. The idea of this model is that, conditional on different market states, the obligors have different hazard rates. For example, if the market goes up then the obligor may have a lower risk of default (low hazard rate), or if the market goes down then it is more likely for the obligor to default during the contract period (high hazard rate).

To find the probability distribution of hazard rates, Hull and White (2006) suggested the so-called "implied copula" model. This is not a specific copula like Gaussian, Student-t, or double-t. It is called implied because it can be deduced from market quotes. The CDO tranche quotes are used for calibration. The Hull and White (2006) model minimizes the sum of deviations from no-arbitrage equations and a smoothing term. The motivation in the deviation term comes from the equality between the mid-price of the CDO tranche and the expected payoff on this tranche (no-arbitrage constraint in risk-neutral setting). This equality may not be feasible for some CDO price quotes. The smoothing term is introduced to reduce the noise in the distribution. We observed, however, that optimal solution is quite sensitive to the smoothing term coefficient. Hull and White (2010) introduced a two parameter version of the implied copula model. One of the parameters was determined by default risk of the portfolio underlying the CDO, while the second parameter was related to the degree of default correlation among the names of the underlying portfolio.

Several others considered similar approaches, including the "Implied Factor Model" of Rosen and Saunders (2009) and Nedeljkovic et. al. (2010); the "Implied Archimedean Copula" of Vacca (2008); papers based on minimum entropy: Dempster et. al. (2007), Meyer-Dautrich and Wagner (2007), and Halperin (2009); and others (Walker (2006)).

This paper applies an "entropy approach" to the implied copula model. We found the distribution by maximizing the entropy with no-arbitrage constraints based on bid and ask prices of CDO tranches. In our numerical experiments, these constraints were feasible and we did not need to introduce penalties for deviation from no-arbitrage constraints. To reduce the noise in the data we introduced a new class of distributions, called Convex-Concave-Convex (CCC) distributions. This is a wide class of distribution containing the normal, gamma, and the F distributions. By definition, the PDF of CCC distribution is convex from the beginning to some point, then it is concave to some further point, and then it is again convex to the end. For discrete distributions, we described CCC distributions by a system of linear constraints. The class of CCC distributions is quite general and it can be used in various applications. This paper presents an application of CCC distribution for CDO calibration. The case study compares Hull and White (2006) and our approach based on CCC distributions.

We used December, 2006 iTraxx tranche quotes from Arnsdorf and Halperin (2007) containing the bid and ask quotes. We also considered more recent data where the market was in unstable condition. To do the case study we used the Portfolio Safeguard (PSG) package (MATLAB and Run-File Text Environments) by American Optimal Decisions (AORDa.com). The case study shows that the approach has a stable performance. We provided the MATLAB and Text codes used for conducting numerical experiments, see link[5].

Here we illustrate the application of the CCC class of distributions with a couple of graphs from the case study in this paper. We used the entropy approach to calibrate the loss distribution. Figure 1 and 2 represent distribution of hazard rates implied from 5 year iTraxx (100 hazard rate scenarios). Figure 1 shows the distribution without CCC constraint, while Figure 2 with CCC constraint. The graph in Figure 1 exhibits a noisy behavior (hump in the area $[-5, 0]$). Imposing additional CCC constraints removes the hump, as shown on Figure 2.

The paper proceeds as follows: Section 1 summarizes the implied copula model introduced by Hull and White (2006). Section 2 proposes the CCC distribution and describes how to calibrate it with the entropy approach. It provides the formal optimization problem statements and a heuristic algorithm for finding probability distribution. Section 3 discusses the case study.

1. Conventional Copula and the Implied Copula

This section summarizes the implied copula approach proposed by Hull and White (2006). For a full model description, a reader may refer to the Hull and White (2006) paper. A one-factor Gaussian copula model, first introduced by Li (2000), has became an industry standard. It models default intensities as a weighted sum of a market factor and an idiosyncratic term, a firm-dependent component. The model provides a correlation structure between default intensities of different obligors.

Define default intensities $X_i (1 \leq i \leq n)$ by:

$$X_i = \alpha_i V + \sqrt{1 - \alpha_i^2} W_i \ , \tag{1}$$

where V is a market factor and W_i is an idiosyncratic term (firm-dependent component). Let $Q_i(t)$ be the cumulative distribution of (unconditional) time to default o*f company i and let $F_i(t)$ be the cumulative distribution of X_i. Default intensity is then mapped to default time τ_i as $F_i(X_i) = Q_i(\tau_i)$.

A convenient way of defining Q_i is through a company hazard rate. The latter has an interpretation of default intensity if the default is modeled as the first event in a non-homogeneous Poisson process. The hazard rate $\lambda_i(\tau)$ is related to $Q_i(t)$ in the following way:

[5] http://www.ise.ufl.edu/uryasev/research/testproblems/financial_engineering/cs_calibration_copula/

Figure 1: Distribution obtained with the maximum entropy approach. The graph has a hump in the area $[-5, 0]$.

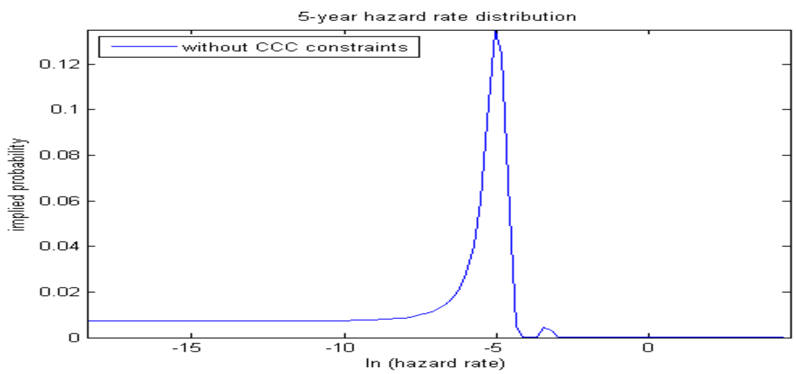
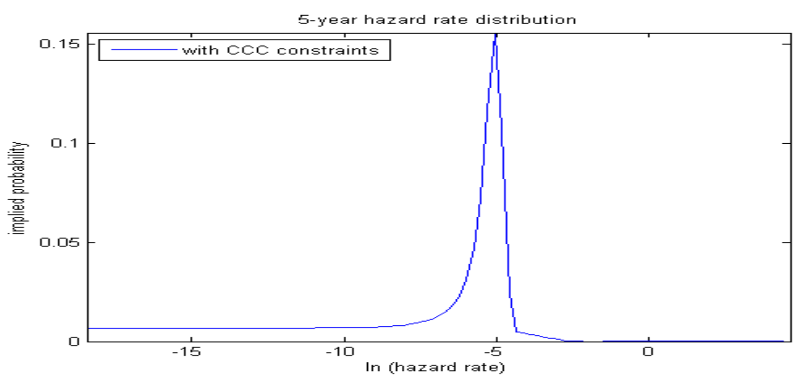


Figure 2: Distribution obtained with the maximum entropy approach and CCC constraints. The CCC constraint removed the hump in the area $[-5, 0]$.



$$\lambda_i(\tau) = \frac{1}{1 - Q_i(\tau)} \frac{dQ_i(\tau)}{d\tau} \ . \tag{2}$$

Hazard rates are popular in credit risk applications due to ease of implementation, convenient analytic expressions and clear physical interpretation.

We define a grid $\lambda_1, ..., \lambda_I$ of possible hazard rates[6]. In other words, we assume that in each scenario the hazard rate is constant and the same for all obligors. As in Hull and White (2010), we set the lowest hazard rate such that there is almost no chance to default ($\lambda_1 = 10^{-8}$), and the highest hazard rate such that almost all companies default immediately

[6]The hazard rate can be viewed as a severity of the credit environment over the life of the CDO.

($\lambda_I = 100$). The intermediate hazard rates are chosen so that the $ln\lambda_k$ are equally spaced. We present results for the number of hazard rates on the grid from 100 to 1,000. We tried to determine if the increasing number of scenarios of hazard rates leads to some stable limiting distribution. This property is expected from a "well defined" model where the increasing of precision improves the performance of the model. For a specific value of the market factor, defaults of each company or obligation are independent and described by their conditional hazard rates. These hazard rates are simultaneously higher or lower. Hull and White (2006) proposed a so-call "implied copula" model prescribing the same unconditional hazard rate to each company and then moved all hazard rates simultaneously (or, more precisely, proportionally) so that the collateral hazard takes on pre-defined values $\lambda_1, ..., \lambda_I$. The scenarios for hazard rate λ_i have probabilities p_i.

To fit a probability distribution for hazard rates to the market we consider CDO price data. We used the 5-year quotes for iTraxx index tranches on December, 2006[7]. By sampling default scenarios corresponding to each level of λ_i, the net payoff[8] of each tranche j can be determined, conditional on the hazard rate scenario λ_i. Denote this payoff by a_{ij}. Note that this net payoff is calculated with the mid-quotes for the spreads for every tranche. Later, we will describe how the bid and ask quotes can be used in no-arbitrage consideration. A probability p_i is assigned to $\{\lambda_i\}$ to form a probability distribution of hazard rates. No-arbitrage considerations in a risk-neutral setting assume that the expected net payoff of each CDO tranche is equal to zero[9]

$$\sum_{i=1}^{I} a_{ij} p_i = 0 \quad j = 1, \ldots, J \; . \tag{3}$$

The numerical experiments with the market data show that, in some cases, the constraints (3) are infeasible. In such cases, we need to find a distribution by solving equation (3) approximately. Some criterion has to be defined to choose a distribution the closest to a feasible one. Hull and White (2006) proposed solving the following optimization problem to find a suitable probability distribution:

[7]We used data from Arnsdorf and Halperin (2007).

[8]The difference between expected present value of premium leg payments and default leg payments.

[9]Tranche payoffs (with both payment legs included) have to be zero under no-arbitrage assumptions. The tranche spread has to be established at a level that the expected payoffs through the premium leg are precisely equal to the expected default losses, in other words, so that the premium leg has the same present value as the default leg.

Problem A

$$\min_p (D(p) + S(p))$$

subject to

probability distribution constraints

$$\sum_{i=1}^{I} p_i = 1 \ , \tag{4}$$

$$p_i \geq 0, \quad i = 1, \ldots, I \ . \tag{5}$$

where $D(p)$ is a deviation term

$$D(p) = \sum_{j=1}^{J} \left(\sum_{i=1}^{I} p_i a_{ij} \right)^2 \ , \tag{6}$$

and $S(p)$ is a smoothing term

$$S(p) = c \sum_{i=2}^{I-1} \left[\frac{p_{i+1} + p_{i-1} - 2p_i}{0.5(d_{i+1} - d_{i-1})} \right]^2 \ . \tag{7}$$

The deviation term penalizes deviations from zero of the net expected payoff of every tranche. The smoothing term enforces that every three consecutive points on the hazard rate distribution are approximately on the same line. The smoothing term is larger for larger differences from the straight line.

The smoothing term introduces distortion into resulting distribution, but a reasonable level of distortion may be better than a ragged distribution shape. The smoothing effect appears to decrease with the increase in the number of atoms in the distribution. Also, the coefficient c has to be chosen by trial and error.

Let us consider the case presented in the paper of Hull and White (2006). We used the data (see the Table 1) to simulate the expected cash flows on every tranche for different hazard rates. Then we solved the optimization Problem A. Figures 4 and 5 show graphs of optimal distributions obtained for different numbers of points and different values of the smoothing term coefficient c.

It seems that if we find a “good” smoothing coefficient c for a particular number of atoms in the distributions, it does not work the same way if we change the number of atoms. Therefore, the smoothing coefficient c should be chosen individually for every number of points on the hazard rate grid. Moreover, we can not find a “reasonable” justification of why one smoothing coefficient is better than any other one.

The next section discusses "the entropy approach" for finding the "best" probability distribution. With this approach we excluded from the model free parameters, such as the smoothing coefficient c.

2. Implied Copula: Entropy Approach with CCC constraints

Hull and White (2006) minimized the sum of squared deviations of tranche payoffs from "perfect fit" (6) and the smoothing term (7). We used an alternative maximum entropy principle and found the distribution in the class of CCC distributions.

The Maximum Entropy Principle (first introduced by Shannon, see also Golan (2002)) is popular in information theory. This principle is actively used in financial applications; see for instance Miller and Liu (2002), Chu and Satchell (2005), Mayer-Dautrich and Wagner (2007). The essence of the Maximum Entropy Principle is that, with some given information about the distribution (specified through equations and constraints), we maximize the entropy and select the most "unknown" distribution. Therefore, we find the most "unknown" distribution containing only available information about the distribution.

Compared to Hull and White (2006, 2010), instead of mid prices, we used bid and ask prices. Denote by $\underline{a}_{ij}$ and $\overline{a}_{ij}$ the expected net payoff of tranche j conditional on hazard rate i for ask and bid prices, respectively. Then, the no-arbitrage constraints are as follows: for ask prices the expected net payoff ($\sum_{i=1}^{I} \underline{a}_{ij} p_i$) of each CDO tranche is nonpositive and for bid prices ($\sum_{i=1}^{I} \overline{a}_{ij} p_i$) is nonnegative.

We maximized Shannon entropy $H(p) = -\sum_{i=1}^{I} p_i \ln p_i$ subject to the no-arbitrage constraints. We solved the following problem:

Problem B

$$\min_{p} \; -H(p)$$

subject to

no-arbitrage constraints

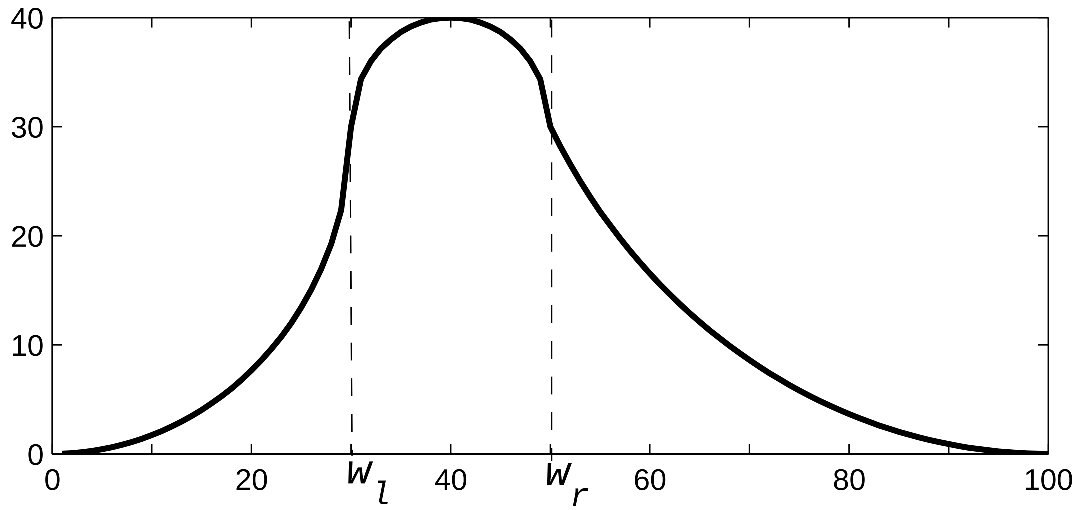
$$\sum_{i=1}^{I} \underline{a}_{ij} p_i \leq 0 \; , j = 1, \ldots, J \; , \tag{8}$$

$$\sum_{i=1}^{I} \overline{a}_{ij} p_i \geq 0 \; , j = 1, \ldots, J \; , \tag{9}$$

probability distribution constraints

$$\sum_{i=1}^{I} p_i = 1 \; , \tag{10}$$

Figure 3: Example of a CCC distribution, the first inflection point $w_l = 30$ and the second inflection point $w_r = 50$.



$$p_i \geq 0, \quad i = 1, \ldots, I \ . \tag{11}$$

The set of probability distributions satisfying constraints (8), (9) is bigger than for constraints (3). Therefore, constraints (8), (9) may have a feasible solution and we may not need to introduce the deviation term (6) to the objective.

The optimal solution of Problem A is very sensitive to the choice of c in (7) and to the number of points I on the grid (see Figures 4 and 5). Furthermore, it seems that with an increasing number of points I in the optimization Problem A, the optimal solution does not stabilize. Some kind of stabilization can be seen for $c = 10^{-5}$ at the right bottom graph in Figure 5). But, again, it is unclear why $c = 10^{-5}$ should be used.

We solved Problem B for different numbers of points also. We found that the shape of the optimal distribution stabilized for number of points ≥ 500. But we saw also that some "noise" was present in the optimal distribution. To cope with that we defined the CCC class of discrete probability distributions.

We say that a function belongs to CCC class if it is convex on the left up to some point, then concave up to a further point, and then again convex on the right. Figure 3 shows an example of CCC function.

By definition, a function $f : R \to R$ is convex if for any $x_1 \in R, x_2 \in R, \lambda \in [0, 1]$ the following inequality holds:

$$\lambda f(x_1) + (1 - \lambda) f(x_2) \geq f(\lambda x_1 + (1 - \lambda) x_2).$$

Let $x_3 = \lambda x_1 + (1-\lambda)x_2$, then $\lambda(x_2 - x_1) = x_2 - x_3$. Therefore, there is one to one correspondence between λ and x_3 and the convexity property can be rewritten as follows:

for any x_1, x_2, x_3 such that $x_1 \leq x_3 \leq x_2$ the following inequality holds,

$$(x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2) \geq (x_2 - x_1)f(x_3).$$

With this observation, we can generalize the concavity/convexity property to any set $X \subset R$, not necessarily convex, closed, etc. We say that $f : X \to R$ is convex on X, if for any $x_1, x_2, x_3 \in X$:

$$(x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2) \geq (x_2 - x_1)f(x_3).$$

Below is the formal definition of the CCC class of functions in general case.

Definition (general case). *Let $f : X \to R$. Then $f(x)$ belongs to CCC class if and only if there exist $w_l, w_r \in R$ such that the following inequalities hold:*

1. $w_l \leq w_r$,
2. $(x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2) \geq (x_2 - x_1)f(x_3)$, *for all* $x_1 \leq x_3 \leq x_2 \in (-\infty, w_l] \cap X$,
3. $(x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2) \leq (x_2 - x_1)f(x_3)$, *for all* $x_1 \leq x_3 \leq x_2 \in [w_l, w_r] \cap X$,
4. $(x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2) \geq (x_2 - x_1)f(x_3)$, *for all* $x_1 \leq x_3 \leq x_2 \in [w_r, +\infty) \cap X$.

First, we define the CCC class of continuous distributions.

Definition (continuous case). *Let $f : R \to R$ be a continuous density function of some continuous distribution. Then $f(x)$ belongs to CCC class of distributions if $f(x)$ is a CCC function.*

In our model we deal with discrete distributions. Let us define the CCC class of discrete distributions.

Definition (discrete case). *Let $f : \{d_1, \ldots, d_I\} \to [0,1]$ be a probability measure function on a sequence of points $d_1, \ldots, d_I : d_1 < d_2 < \cdots < d_I$, i.e $\sum_{i=1}^{I} f(d_i) = 1$. Then $f(x)$ belongs to CCC class of discrete distributions if the probability measure f belongs to the class of CCC functions.*

Clearly, $f(x)$ belongs to the CCC class, if and only if, the inequalities 2-4 in the Definition (general case) of the CCC class of functions hold for every three consecutive points d_{i-1}, d_i, d_{i+1}. In other words, the following proposition holds.

Proposition 1. *Let $f : \{d_1, \ldots, d_I\} \to [0,1]$ be a probability measure function on a sequence of points $d_1, \ldots, d_I : d_1 < d_2 < \cdots < d_I$, i.e $\sum_{i=1}^{I} f(d_i) = 1$. Then f belongs to the CCC class if and only if there exist indices w_l, w_r such that the following inequalities hold:*

1. $1 \leq w_l \leq w_r \leq I$,

2. $(d_{i+1}-d_i)f(d_{i-1})+(d_i-d_{i-1})f(d_{i+1}) \geq (d_{i-1}-d_{i+1})f(d_i)$*, for all i:* $1 < i < w_l$*,*

3. $(d_{i+1}-d_i)f(d_{i-1})+(d_i-d_{i-1})f(d_{i+1}) \leq (d_{i-1}-d_{i+1})f(d_i)$*, for all i:* $w_l < i < w_r$*,*

4. $(d_{i+1}-d_i)f(d_{i-1})+(d_i-d_{i-1})f(d_{i+1}) \geq (d_{i-1}-d_{i+1})f(d_i)$*, for all i:* $w_r < i < I$*.*

The proof is obvious and we leave it to the reader.

Further, we suppose that the distance between every two consecutive points d_i, d_{i+1} is the same. In this case, Proposition 1 simplifies to:

Proposition 2. *Let* $f : \{d_1, \dots, d_I\} \to [0,1]$ *be a probability measure function* $(\sum_{i=1}^{I} f(d_i) = 1)$ *on a sequence of points* $d_1, \dots, d_I : d_1 < d_2 < \cdots < d_I$*, such that the distance between every two consecutive points* d_i, d_{i+1} *is the same. Then,* $f(x)$ *belongs to CCC class if and only if there exist* d_{w_l}, d_{w_r} *such that the following inequalities hold:*

1. $1 \leq w_l \leq w_r \leq I$*,*

2. $f(d_{i-1})+f(d_{i+1}) \geq 2f(d_i)$*, for all i:* $1 < i < w_l$*,*

3. $f(d_{i-1})+f(d_{i+1}) \leq 2f(d_i)$*, for all i:* $w_l < i < w_r$*,*

4. $f(d_{i-1})+f(d_{i+1}) \geq 2f(d_i)$*, for all i:* $w_r < i < I$*.*

As it was mentioned earlier, our goal is to assign a probability p_i to every hazard rate λ_i to satisfy the market constraints. In this case, by a discrete distribution corresponding to a vector $(p_1, \dots, p_I)$ we mean a probability measure function $f : \{\lambda_1, \dots, \lambda_I\} \to [0,1]$ such that $f(\lambda_i) = p_i, i = 1, \dots, I$.

We solved Problem B with the additional constraints assuring that the distribution belongs to the CCC class. The CCC class of distributions can be specified in optimization problem by linear constraints. CCC constraints "regularize" the solution by reducing "noise" and they play the same role as the smoothing term $S(p)$ in Problem A. In this way we can avoid arbitrariness in the choice of the smoothing coefficient c. Also, the increasing number of points λ_i on hazard rate grid does not lead to additional noise in distribution.

To solve Problem B in the CCC class of discrete distributions, we used Proposition 2 to introduce CCC constraints. The CCC constraints include constraints on the left slope, right slope and hump:

Convexity of the left slope:

$$\frac{p_{i-1}+p_{i+1}}{2} \geq p_i,\ i = 2, ..., w_l - 1\ , \tag{12}$$

Concavity of the hump:

$$\frac{p_{i-1}+p_{i+1}}{2} \leq p_i,\ i = w_l + 1, ..., w_r - 1\ , \tag{13}$$

Convexity of the right slope:

$$\frac{p_{i-1} + p_{i+1}}{2} \geq p_i,\ i = w_r + 1, ..., I - 1\ . \tag{14}$$

The points w_1, w_2 may vary for different discrete distributions, therefore we incorporate them into the optimization problem as variables. By adding the CCC constraints to Problem B we have the following optimization problem:

Problem C

$$\min_{w_l, w_r, p} \; -H(p)$$

subject to

no-arbitrage constraints

$$\sum_{i=1}^{I} \underline{a}_{ij} p_i \leq 0\ , \tag{15}$$

$$\sum_{i=1}^{I} \overline{a}_{ij} p_i \geq 0\ , \tag{16}$$

CCC constraints:

constraint on inflection points

$$w_l \leq w_r,\ w_l \in \{1, \ldots, I\},\ w_r \in \{1, \ldots, I\}, \tag{17}$$

convexity of the left slope

$$\frac{p_{i-1} + p_{i+1}}{2} \geq p_i,\ i = 2, ..., w_l - 1\ , \tag{18}$$

concavity of the hump

$$\frac{p_{i-1} + p_{i+1}}{2} \leq p_i,\ i = w_l + 1, ..., w_r - 1\ , \tag{19}$$

convexity of the right slope

$$\frac{p_{i-1} + p_{i+1}}{2} \geq p_i,\ i = w_r + 1, ..., I - 1\ , \tag{20}$$

probability distribution constraints

$$\sum_{i=1}^{I} p_i = 1 \; , \tag{21}$$

$$p_i \geq 0, \quad i = 1, \ldots, I \; . \tag{22}$$

Let us consider a subproblem of this problem with fixed w_l, w_r.

Problem $C(w_l, w_r)$

$$\min_{p} \; -H(p) \tag{23}$$

subject to

no-arbitrage constraints

$$\sum_{i=1}^{I} \underline{a}_{ij} p_i \leq 0 \; , \tag{24}$$

$$\sum_{i=1}^{I} \overline{a}_{ij} p_i \geq 0 \; , \tag{25}$$

CCC constraints:
convexity of the left slope

$$\frac{p_{i-1} + p_{i+1}}{2} \geq p_i, \; i = 2, ..., w_l - 1 \; , \tag{26}$$

concavity of the hump

$$\frac{p_{i-1} + p_{i+1}}{2} \leq p_i, \; i = w_l + 1, ..., w_r - 1 \; , \tag{27}$$

convexity of the right slope

$$\frac{p_{i-1} + p_{i+1}}{2} \geq p_i, \; i = w_r + 1, ..., I - 1 \; , \tag{28}$$

probability distribution constraints

$$\sum_{i=1}^{I} p_i = 1 \; , \tag{29}$$

$$p_i \geq 0, \quad i = 1, \ldots, I \; . \tag{30}$$

To solve Problem C, we can solve Problem $C(w_l, w_r)$ for all possible pairs of integers w_l, w_r such that $1 \leq w_l \leq w_r \leq I$, and then choose the minimum among these solutions. The number of subproblems (Problem $C(w_l, w_r)$) is on the order of n^2. Originally we solved Problem B, but since its solutions had unexpected hump, we suggested to find the solution to Problem B in the CCC class of functions (Problem C). Further, we provide a heuristic algorithm for solving Problem C. We solved at first Problem B and then solved a sequence of Problem C(w_l, w_r) for different pairs of (w_l, w_r).

Here is the formal description of algorithm. Explanations are provided after the formal description.

Algorithm:

Step 0. Initial optimal solution.

- Solve Problem B and denote its solution obtained for optimization problem by p^*.
- Initialize $w_l = w_r = argmax\{p_i^* : i = 1, \ldots, I\}$[10]$, k = 0, H_0 = \infty$.

Step 1. Solve Problem C(w_l, w_r)

- Set $k = k + 1$, $exit\_flag = 0$.
- Solve Problem C(w_l, w_r) and obtain the optimal solution p_k^* and $H_k = H(p_k^*)$.

Step 2. Shifting w_r to the right

- If $w_r < I$ and $H_k \leq H_{k-1}$, then set $w_r = w_r + 1$, $exit\_flag = 1$, and go to Step 1.

Step 3. Initialization of shifting w_l to the left

- If $w_l > 1$, then set $w_l = w_l - 1$.
- If $w_l = 1$, then stop the algorithm, and p_{k-1}^* is an approximation of the optimal solution.

Step 4. Solve Problem C(w_l, w_r) (the same as Step 1)

- Set $k = k + 1$.
- Solve Problem C(w_l, w_r) and obtain the optimal solution p_k^* and $H_k = H(p_k^*)$.

Step 5. Shifting w_l to the left

- If $w_l > 1$ and $H_k \leq H_{k-1}$ then set $w_l = w_l - 1$, $exit\_flag = 1$, and go to Step 4.

[10] If the maximum is not unique, the algorithm should be performed for eash point in the set $argmax\{p_i^* : i = 1, \ldots, I\}$, and then the solution with the smallest objective value should be chosen.

- If $exit\_flag = 1$, then go to Step 1.
- If ($w_l = 1$ or $H_k > H_{k-1}$) and $exit\_flag = 0$, then stop the algorithm, and p^*_{k-1} is an approximation of the optimal point.

The idea of this algorithm is that we step-by-step change inflection points w_l, w_r and solve Problem C(w_l, w_r). In Step 0, we solve Problem B and obtain an optimal solution p^*. Then, we set $w_l = w_r = argmax\{p^*_i : i = 1, \ldots, I\}$. In other words, we find the maximum component of optimal vector p^* and make w_l, w_r equal to its index. In Step 1, we solve Problem C(w_l, w_r) with these w_l, w_r and obtain the optimal point and its objective value. Then, we shift w_r to the right, if it is possible, making $w_r = w_r + 1$. After that we go to Step 1 and again solve Problem C(w_l, w_r) to obtain the optimal point and its objective value. Then we compare this objective value with the previous one obtained in Step 1 (H_k and H_{k-1}). This procedure stops when the new objective value is greater then the previous one ($H_k > H_{k-1}$), or $w_r = I$. In Steps 3 to 5 we run the same procedure, but now we shift w_l to the left. The procedure also stops when the new objective value is larger then the previous one ($H_k > H_{k-1}$) , or $w_l = 1$. If during the steps 1 through 4, the smaller objective value is found by shifting w_r or w_l, then these steps should be performed again. In other words, we shift the points w_r and w_l to reach local optimality. Finally, the algorithm returns p^*_{k-1}, which is an approximation of the optimal point. We do not prove that this algorithm provides an optimal solution to Problem C. The case study in the following section shows that this algorithm provides reasonable solutions.

3. Case study

We used Portfolio Safeguard (2008) in MATLAB and Run-File Text Environment to do the case study (MATLAB and PSG Run-File text files are posted at the following link[11]). The provided files can be used for both simulating the expected cash flow matrices with the tranche quotes and solving the optimization problems. Appendix 1 contains information on running the case study with PSG.

For the case study, we have considered iTraxx index with different maturities. First, we used 5-year iTraxx tranche quotes to simulate the expected cash flow matrices. For bid prices, mid prices and ask prices we simulated different matrices $(\overline{a}_{ij})_{i=1,\ldots,I}^{j=1,\ldots,J}$, $(a_{ij})_{i=1,\ldots,I}^{j=1,\ldots,J}$, $(\underline{a}_{ij})_{i=1,\ldots,I}^{j=1,\ldots,J}$, for I=100, 200,..., 1,000. The number of tranches in the iTraxx index is six, so $J = 6$.

For particular i, j we simulated the times to default of 125 companies in the iTraxx index and the corresponding tranche cash flows 10,000 times and than took the average. As we mentioned earlier, the time to default of each company is exponentially distributed with parameter λ_j. For simulation, we used the minimum hazard rate $\lambda_1 = 10^{-8}$, the maximum hazard rate $\lambda_I = 100$, and the distances between $ln(\lambda_i)$ are equal. We assumed that the tranche payments are made quarterly, the recovery rate, in case of default, equals to 40% and the annual risk free rate is 4%. The reader may refer to the Hull and White (2006) to find more details on the simulation procedure.

We solved Problem A (Hull and White (2006)) for I=100, 300, 500 and 1,000 points. We used six different smoothing term coefficients in Problem A. The graphs are presented in Figures 4 and 5. The distribution functions in the graphs are not the actual solution vectors. We scaled functions so that the areas under the graphs are equal (the horizontal axis represents $ln(\lambda)$). These graphs are implied densities of hazard rate distributions. We found that the results are quite sensitive to the parameters c and I.

We compared our entropy approach with CCC constraints to the approach by Hull and White (2006). We ran the proposed heuristic algorithm, described at the end of Section 2, for I =100, 200, 300, 500, 800 and 1,000 points. The entropy maximization problem was solved with PSG in MATLAB environment by calling 'riskprog' optimization subroutine. To solve the problem, we need to put the matrix of constraints and 'entropyr' as parameters to the riskprog subroutine. We conducted the case study on the laptop with processor Intel Core 2 CPU @2GHz. The optimization time for Problem B varied from 0.01 sec. for $I = 100$ to 0.06 sec. for $I = 1,000$; for Problem C time varied from 0.36 sec. for $I = 100$ to 600 sec. for $I = 1,000$. Figure 6 shows six hazard rate distribution graphs for the six different values of I. Every graph shows initial distribution $\tilde{p}_0$, which is a solution of problem B and final optimal distribution $\tilde{p}_1$. We found that imposing CCC constraints has not changed significantly the shape of implied density functions. However, irregularities (humps, which

[11] `http://www.ise.ufl.edu/uryasev/research/testproblems/financial_engineering/cs_calibration_copula/`

we call "noise") were streamlined.

Figure 7 compares the final distributions $\tilde{p}_1$ for different numbers of hazard rates ($I = 100, 200, 300, 500, 800$, and $1,000$) on the grid. The last three graphs are almost identical, which seems quite natural. We did not observe the similar stability in the Hull and White (2006) even with a fixed smoothing coefficient c.

We applied our entropy approach with CCC constraint to the iTraxx thanches with different contract periods: 5 years, 7 years and 10 years. These calculations show whether the implied copula model can be used with the homogeneity assumption, i.e when hazard rate of a company stays the same during the whole contract period. If this approach is reasonable, we should obtain a similar distribution of hazard rates for different contract periods. We simulated the matrices of expected cash flows using the prices from Table 1 for I=100 for 5, 7 and 10-year contracts. Then, we used these matrices to solve Problem B and Problem C with proposed heuristics. Figure 8 represents scaled solutions with $ln(\lambda)$ on the horizontal axis. The graphs are quite similar and show little dependence of the length of the contract period.

We want to point out that the analyzed data were the market quotes for 5, 7, 10-year iTraxx on December 20, 2006. At that time the credit derivatives market was flourishing and expanding very fast. We also tested this model for the data taken for the later times when the market was very unstable.

First, we used the data for the market quotes for the 5-year iTraxx on four different dates: 10/31/07, 12/31/07, 6/30/08 and 9/30/08. The data contains only the closing prices. To get the bid and ask prices we used typical bid-ask spreads for that times varying from 2% to 7% depending on the tranche. Then, using the simulation technique described in the beginning of this section, we simulated expected cash flow matrices for the bid and ask prices with the number of hazard rate grid points $I = 100$. The implied density functions were obtained by solving Problem B. Figure 9 shows corresponding graphs. The graphs show the evolution of the hazard rate distribution function over the time. We want to mention that the Problem C is infeasible with the assumed bid-ask spreads.

Second, we picked the two latest dates for which we have the price information for the market quotes for 5, 7, 10-year iTraxx. The expected cash flow matrices were simulated the same way. Figure 10 shows the hazard rate distributions for this case.

References

Andersen, L., Sidenius, J., 2004. Extensions to the Gaussian Copula: Random Recovery and Random Factor Loadings. Journal of Credit Risk, Vol. 1, No. 1, 29–70.

Arnsdorf, M., Halperin, I., 2007. BSLP: Markovian Bivariate Spread-Loss Model for Portfolio Credit Derivatives. Quantitative Research, JP Morgan.

Bielecki, T., Jeanblanc, M., Rutkowski, M., 2004. Modeling and Valuation of Credit Risk, Springer.

Burtschell, X., Gregory, J., Laurent, J.P., 2005. A Comparative Analysis of CDO Pricing Models. BNP-Paribas.

Chu, Ba., Satchell, S., 2005. Computing the Most Entropic Copulas. Working paper.

Dempster, M., Medova, E., Yang, S. Empirical Copulas for CDO Tranche Pricing Using Relative Entropy, 2007. Final draft.

Duffie, D., 1999. Credit Swap Valuation. Financial Analysts Journal, 73–87.

Golan, A., 2002. Information and Entropy Econometrics: Editor's view. Journal of Econometrics 107, 1–15.

Halperin, I., 2009. Implied Multi-Factor Model for Bespoke CDO Tranches and other Portfolio Credit Derivatives. Working paper.

Houweling, P., Vorst, T., 2005. Pricing default swaps: Empirical evidence. Journal of International Money and Finance 24, 1200-1225.

Hull, J., White, A., 2006. Valuing Credit Derivatives Using an Implied Copula Approach. Journal of Derivatives, Vol.18, No.2, 8-28.

Hull, J., White, A., 2010. An Improved Implied Copula Model and Its Application to the Valuation of Bespoke CDO tranches. Journal of Investment Management.

Laurent, J.P., Gregory, J., 2003. Basket Default Swaps, CDO's and Factor Copulas. ISFA Actuarial School, University of Lyon, BNP-Paribas.

Li, D.X., 2000. On Default Correlation: A Copula Function Approach. Journal of Fixed Income 9, 43-54.

Marsh, I.W., 2001. What Central Banks Can Learn about Default Risk from Credit Markets. Centre for Central Banking Studies, Bank of England.

Meyer-Dautrich, S., Wagner, C., 2007. Minimum-Entropy Calibration of CDO Tranches.

Miller, D., Liu, W.H., 2002. On the Recovery of Joint Distributions from Limited Information. Journal of Econometrics 107, 259–274.

Nedeljkovic, J., Rosen, D., Saunders, D. Pricing and Hedging Collateralized Loan Obligations with Implied Factor Models. Journal of Credit Risk, 2010. Vol. 6, No. 3, 53-97.

Portfolio Safeguard, 2008. Version 2.1, `www.aorda.com`.

Rosen, D., Saunders, D., 2009. Valuing CDOs of Bespoke Portfolios with Implied Multi-Factor Models. Journal of Credit Risk, Vol. 5, No. 3, 3-36.

Schonbucher, P., Schubert, D., 2001. Copula-Dependent Default Risk in Intensity Models. Working paper, Bonn University.

Schonbucher, P., 2003. Credit Derivatives Pricing Models. Wiley.

Vacca,L., 2008. Market Implied Archimedian Copulas. Risk, January.

Walker, M., 2006. CDO Models - Towards the Next Generation: Incomplete Markets and Term Structure. Working Paper.

Yu, F., 2005. Default Correlation in Reduced-Form Models. Journal of Investment Management, Vol. 3, No. 4, 33-42.

Appendix 1: Running Case Study with Portfolio Safeguard (PSG)

PSG has several syntax formats for running optimization problems in MATLAB environment:

- Optimization subroutines for optimizing nonlinear functions. Subroutines (e.g., "riskprog") use as a parameter the name of a nonlinear function (e.g. "entropyr"), which is optimized.
- General PSG format.

With PSG optimization language in general format, the problem solving typically involves three main stages:

1. *Mathematical formulation of a problem with a meta-code using PSG nonlinear functions.* Typically, a problem formulation involves 5-10 operators of a meta-code. See in the end of the Appendix 1 the PSG meta-code for Problem C(w_l, w_r).
2. *Preparation of data for the PSG functions in an appropriate format.* For instance, the *meansquare* error function is defined by the matrix of loss scenarios. One of those matrices should be prepared if we use this function in the problem statement.
3. *Solving the optimization problem with PSG using the predefined problem statement and data for PSG functions.* The problem can be solved in several PSG environments, such as MATLAB environment and Run-File (Text) environment.

Further we present the PSG meta-code for solving Optimization Problem C(w_l,w_r), see formulas (23)-(30). Meta-code, data and solutions can be downloaded from the link[12].

Meta-Code for Optimization Problem C(w_l,w_r)

Problem: problem_CCC, type = **minimize**
Objective: objective_h, linearize = 1
entropyr_h(matrix_h)
Constraint: constraint_a, lower_bound = vector_bl, upper_bound = vector_b
linearmulti_a (matrix_a)
Constraint: constraint_aeq, lower_bound = 1, upper_bound = 1
linearmulti_aeq (matrix_aeq)
Box_of_Variables: lowerbounds = 0
Solver: VAN, precision = 5

Here we give a brief description of the presented meta-code. We boldfaced the important parts of the code. The keyword **minimize** tells a solver that the Problem C(w_l,w_r) is a

[12] `http://www.ise.ufl.edu/uryasev/research/testproblems/financial_engineering/cs_calibration_copula/`

minimization problem. The keyword **Objective** is used to define an objective function. The objective function (23), that is the Shannon entropy function of the distribution, is defined in lines 2,3 with the keyword **entropyr** and the data matrix, located in the file **matrix_h.txt**. Each constraint starts from the keyword **Constraint**. The constraints (24)-(28) are system of linear inequalities, defined in lines 4,5 with the keyword **linearmulti**. The coefficients for the linear inequalities are given in the file **matrix_a.txt**. The probability distribution constraint (29), is defined in lines 6,7 with keyword **linearmulti** and the matrix of unit coefficients, located in the file **matrix_aeq.txt**. The **Box_of_Variables** in line 8 sets the non-negativity of the variables (probabilities).

Table 1: The data are taken from Arnsdorf and Halperin (2007). Market quotes for 5, 7, 10-year iTraxx on December 20, 2006. Quotes for the 0 to 3% tranche are the percent of the principal that must be paid up front in addition to 500 basis points per year. Quotes for other tranches and the index are in basis points.

Maturity	Low stike	High Strike	Bid	Ask
20-Dec-11	0%	3%	11.75%	12.00%
20-Dec-11	3%	6%	53.75	55.25
20-Dec-11	6%	9%	14.00	15.50
20-Dec-11	9%	12%	5.75	6.75
20-Dec-11	12%	22%	2.13	2.88
20-Dec-11	22%	100%	0.80	1.30
20-Dec-11	0%	100%	24.75	25.25
20-Dec-13	0%	3%	26.88%	27.13%
20-Dec-13	3%	6%	130.00	132.00
20-Dec-13	6%	9%	36.75	38.25
20-Dec-13	9%	12%	16.25	18.00
20-Dec-13	12%	22%	5.50	6.50
20-Dec-13	22%	100%	2.40	2.90
20-Dec-13	0%	100%	33.50	34.50
20-Dec-16	0%	3%	41.88%	42.13%
20-Dec-16	3%	6%	348.00	353.00
20-Dec-16	6%	9%	93.00	95.00
20-Dec-16	9%	12%	40.00	42.00
20-Dec-16	12%	22%	13.25	14.25
20-Dec-16	22%	100%	4.35	4.85
20-Dec-16	0%	100%	44.50	45.50

Figure 4: Distributions of the collateral hazard rate implied by 5-year iTraxx tranche spreads. The distributions were found by solving Problem A for numbers of variables 100 and 300, and different smoothing term coefficients c.

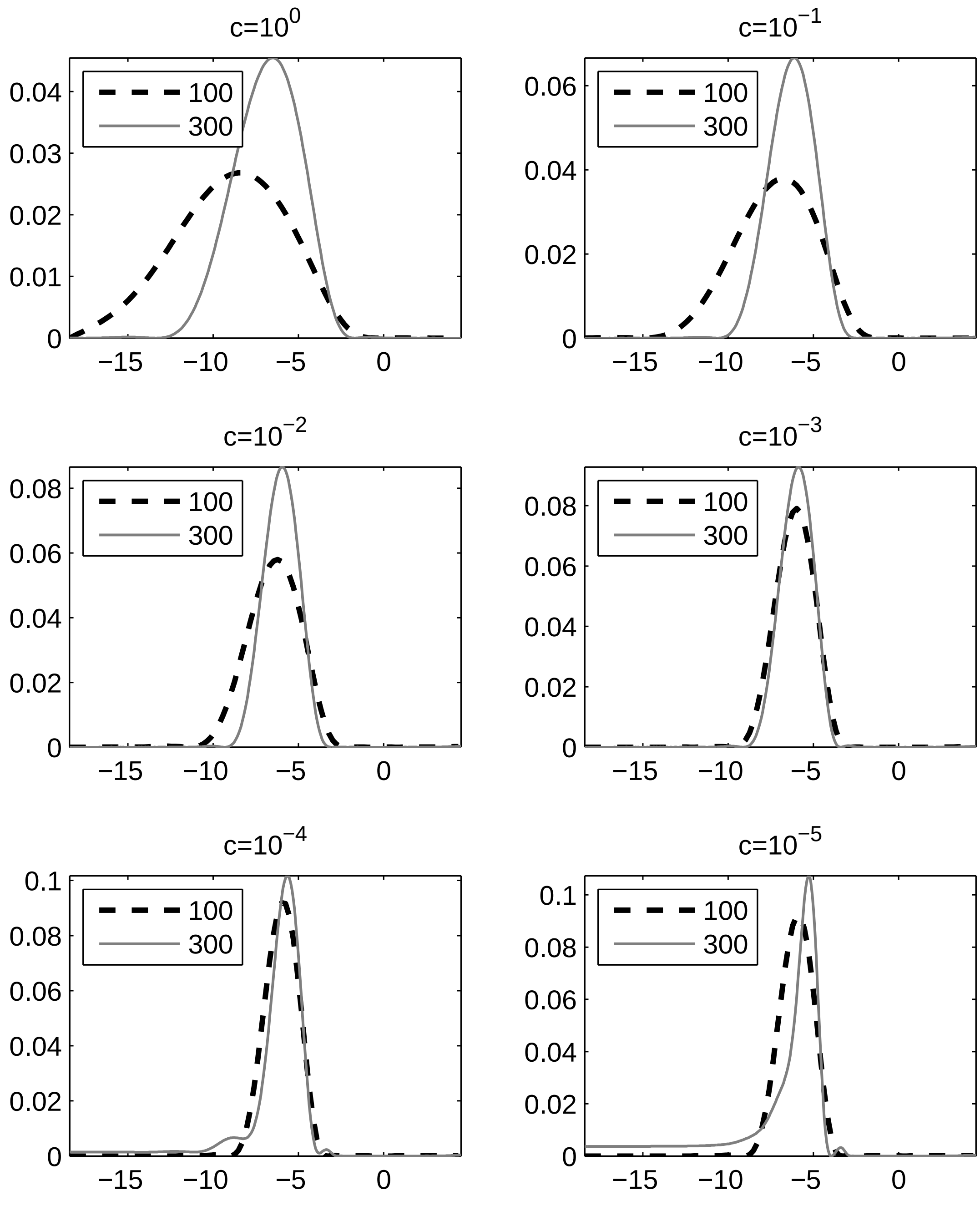


Figure 5: Distributions of the collateral hazard rate implied by 5-year iTraxx tranche spreads. The distributions were found by solving Problem A for numbers of variables 500 and 1,000, and different smoothing coefficients c.

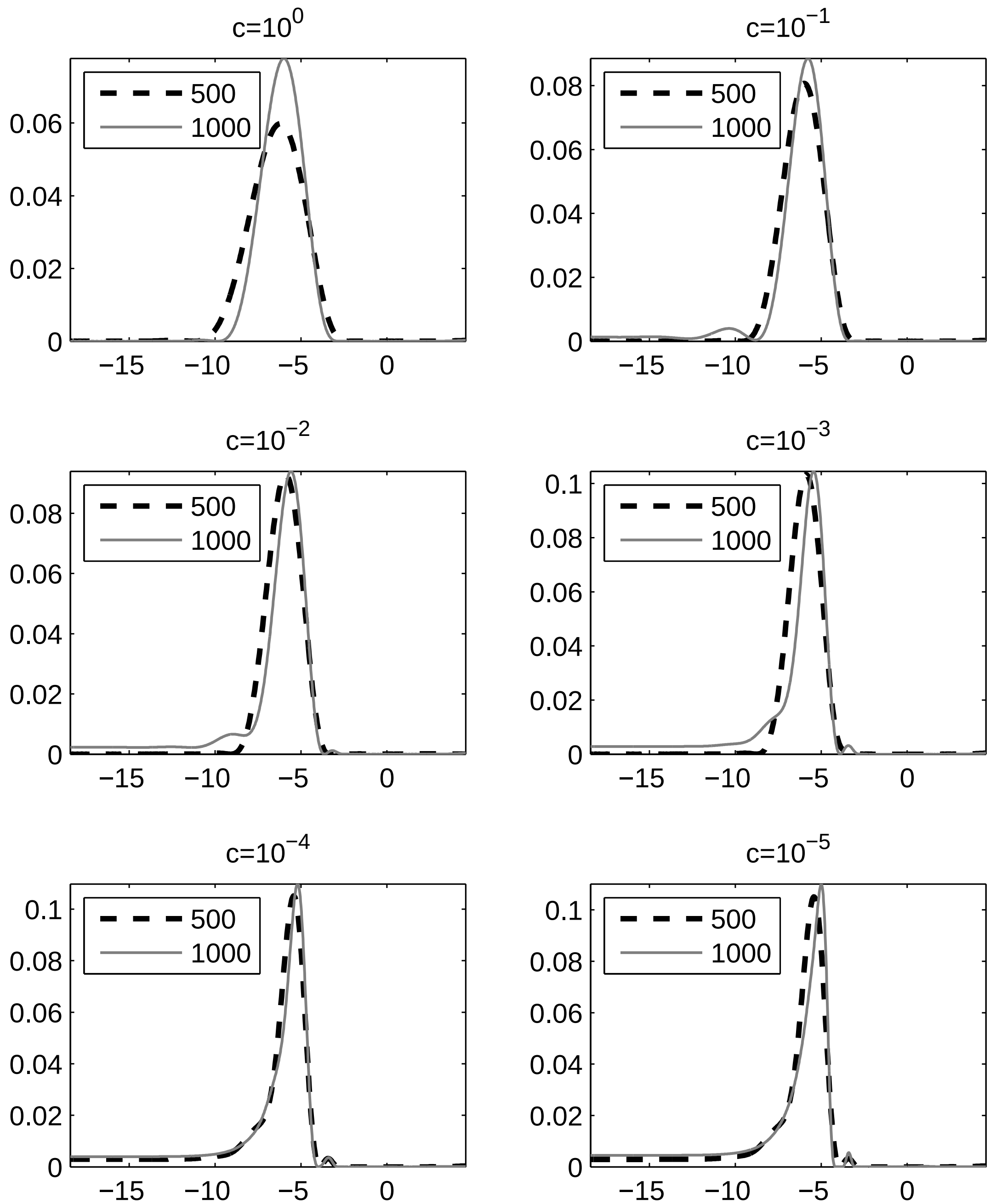


Figure 6: Distributions of the collateral hazard rate implied by 5-year iTraxx tranche spreads. These graphs compare solutions of Problem B and Problem C obtained with the heuristic algorithm for 6 cases with 100, 200, 300, 500, 800, 1,000 decision variables. Hump in the area [-5,0] was removed by imposing CCC constraint

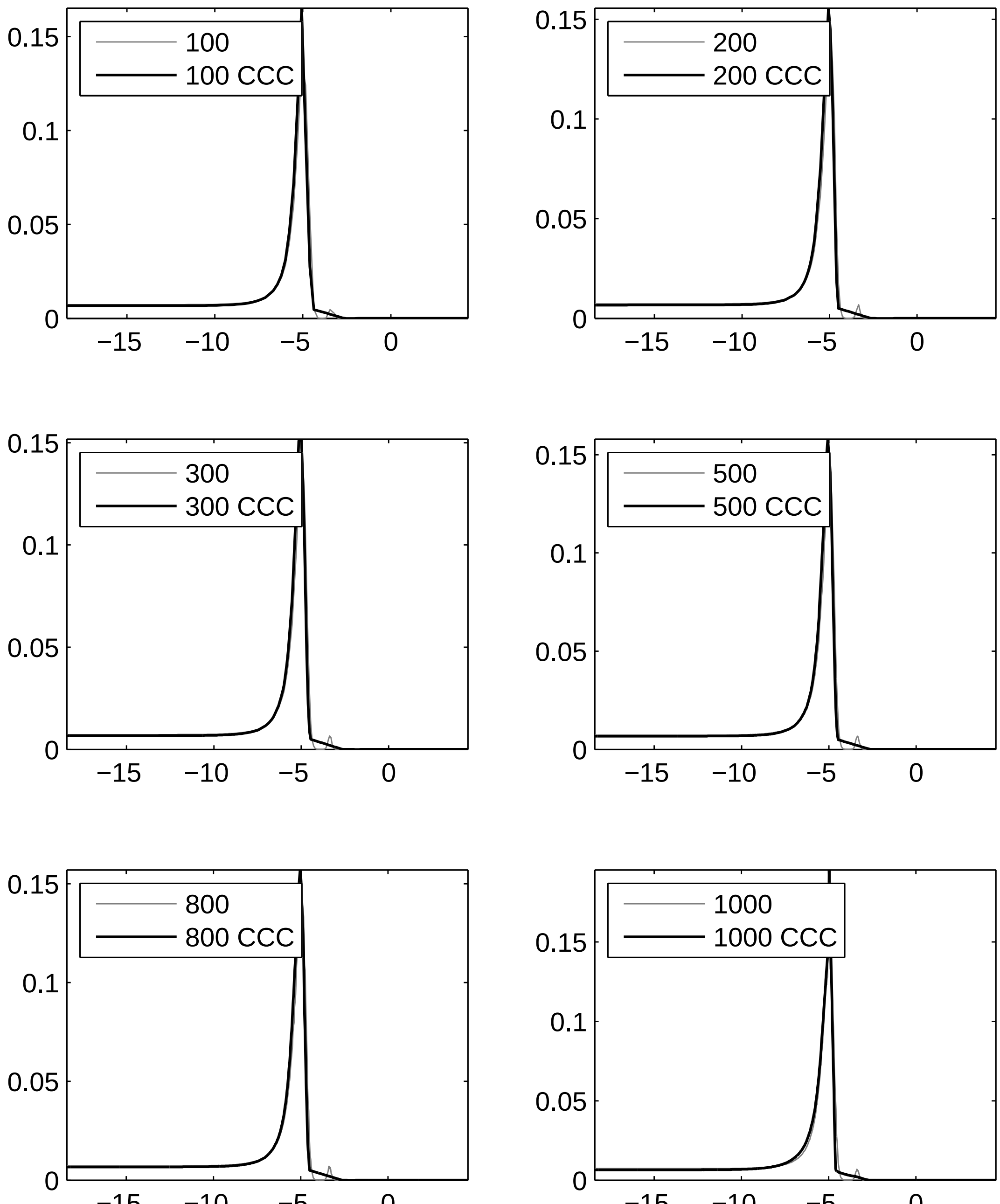


Figure 7: Distributions of the collateral hazard rate implied by 5-year iTraxx tranche spreads. The distributions in the CCC class were found by heuristic algorithm for 100, 200, 300, 500, 800, 1,000 decision variables.

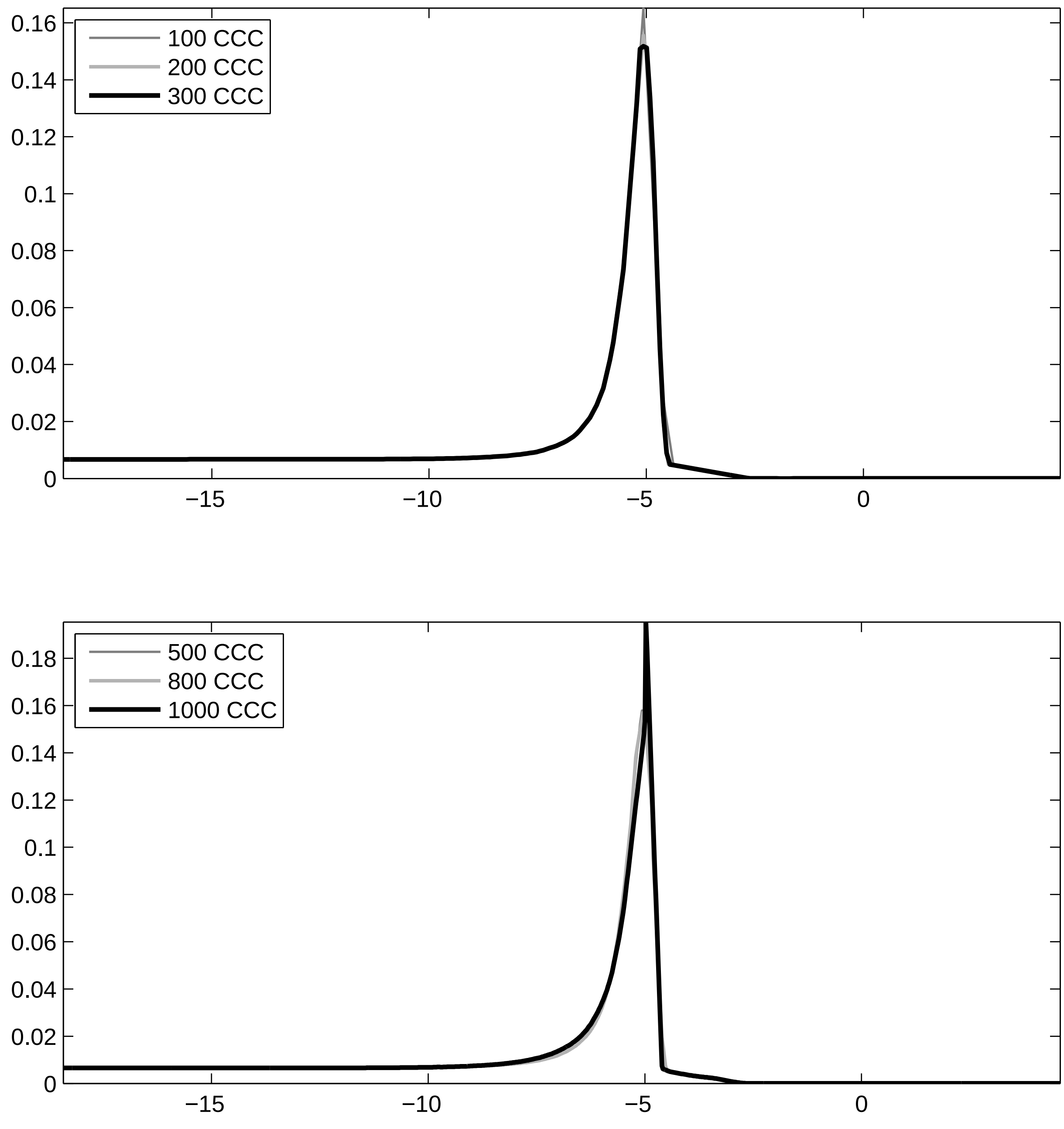
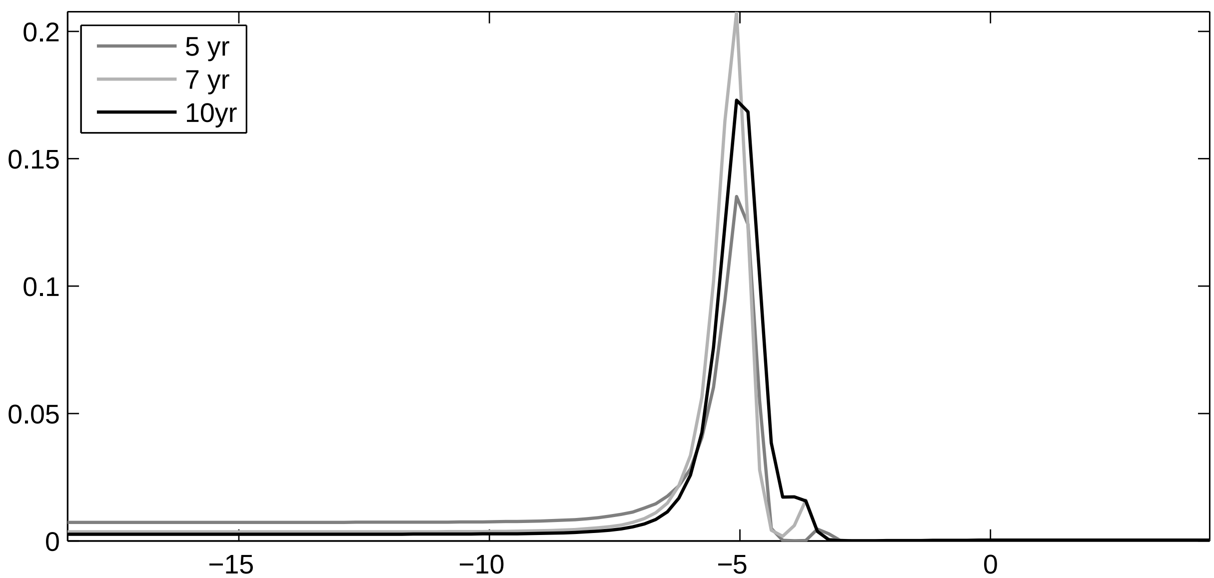


Figure 8: Distributions of the collateral hazard rate implied by 5, 7 and 10-year iTraxx tranche spreads. The distributions were found by solving Problem B (upper chart) and Problem C (lower chart) using heuristic algorithm for 100 decision variables.



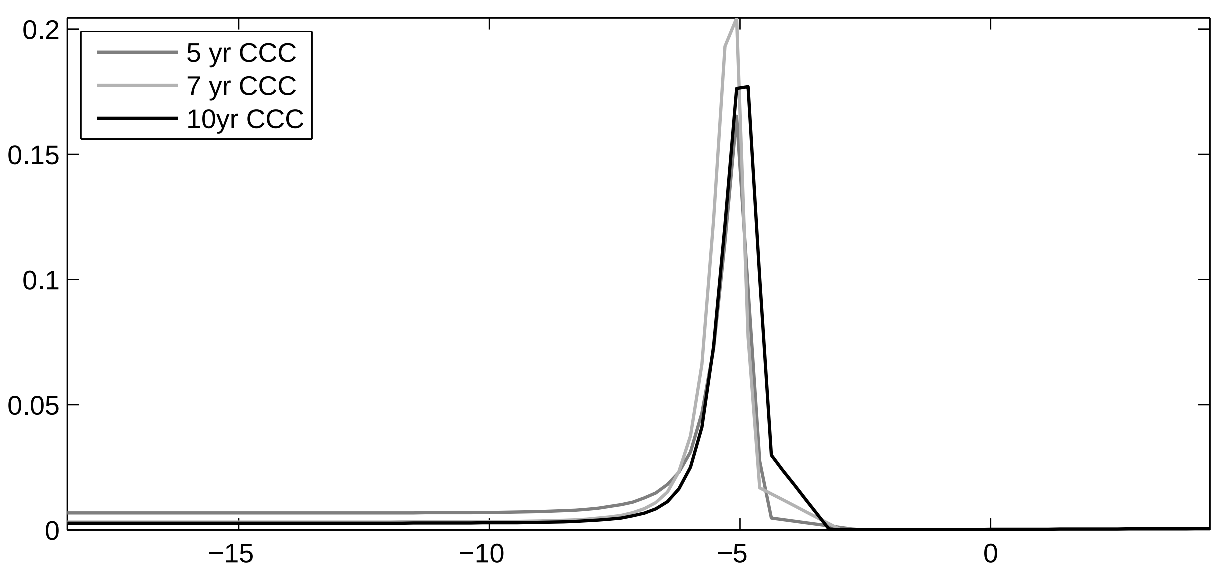


Figure 9: Distributions of the collateral hazard rate implied in 5-year iTraxx tranche spreads for different dates. The distributions were found by solving Problem B with 100 decision variables.

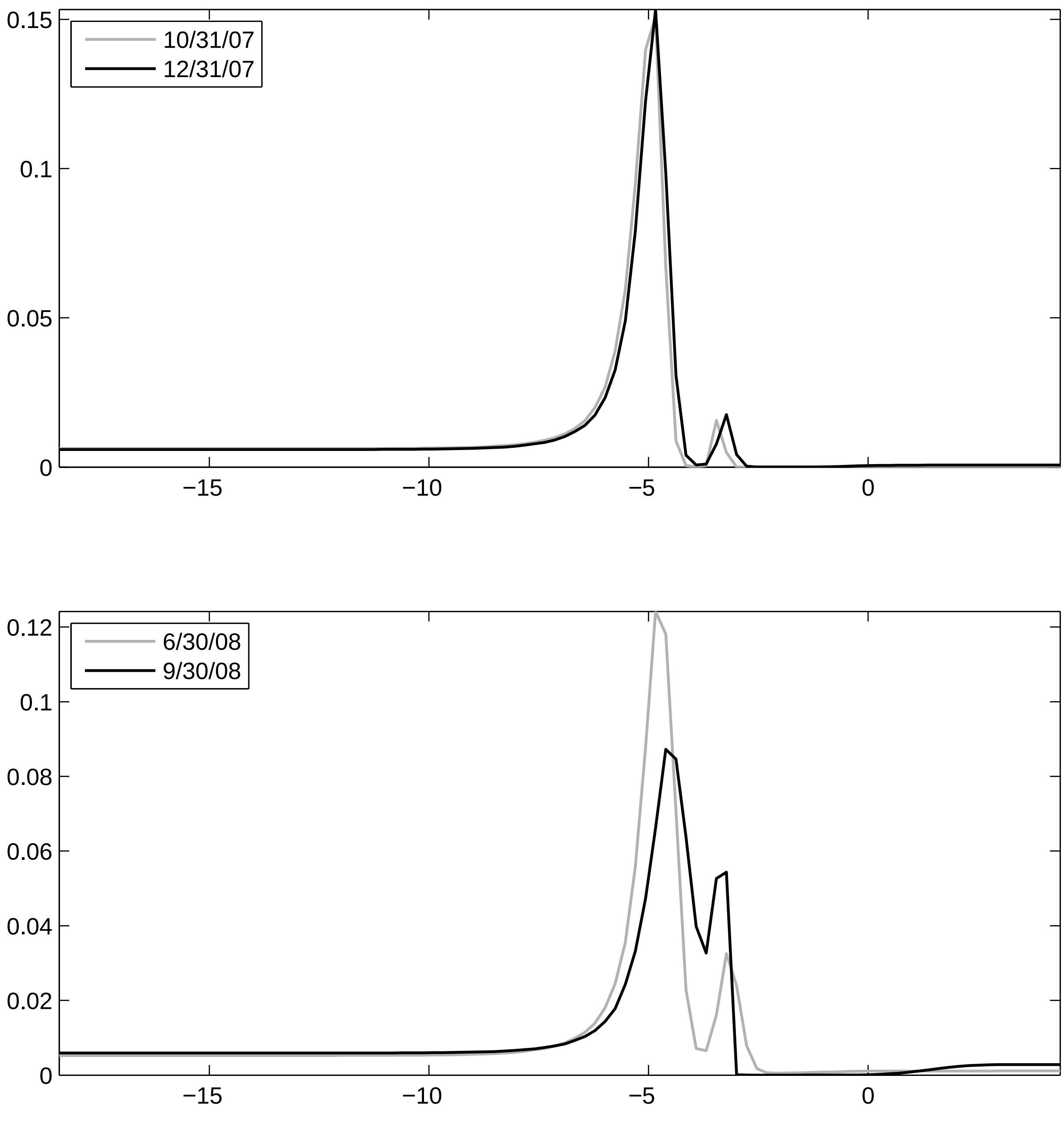


Figure 10: Distributions of the collateral hazard rate implied by 5, 7 and 10-year iTraxx tranche spreads at two different dates. The distributions were found by solving Problem B with 100 decision variables.