Journal of Empirical Finance 18 (2011) 728–742

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

Checking for asymmetric default dependence in a credit card portfolio: A copula approach

Jonathan Crook [1], Fernando Moreira *

Credit Research Centre, University of Edinburgh, 29 Buccleuch Place, Edinburgh, EH8 9JS, UK

ARTICLE INFO

Article history:
Received 12 July 2010
Received in revised form 5 May 2011
Accepted 12 May 2011
Available online 18 May 2011

JEL classification:
G20
G21
C14
C46

Keywords:
Credit risk
Asymmetric dependence
Consumer loans
Copulas

ABSTRACT

Traditional credit risk models adopt the linear correlation as a measure of dependence and assume that credit losses are normally-distributed. However some studies have shown that credit losses are seldom normal and the linear correlation does not give accurate assessment for asymmetric data. Therefore it is possible that many credit models tend to misestimate the probability of joint extreme defaults.

This paper employs Copula Theory to model the dependence across default rates in a credit card portfolio of a large UK bank and to estimate the likelihood of joint high default rates. Ten copula families are used as candidates to represent the dependence structure. The empirical analysis shows that, when compared to traditional models, estimations based on asymmetric copulas usually yield results closer to the ratio of simultaneous extreme losses observed in the credit card portfolio.

Copulas have been applied to evaluate the dependence among corporate debts but this research is the first paper to give evidence of the outperformance of copula estimations in portfolios of consumer loans. Moreover we test some families of copulas that are not typically considered in credit risk studies and find out that three of them are suitable for representing dependence across credit card defaults.

1. Introduction

Many credit risk models assume that returns from loans are normally distributed, not only individually but also at the portfolio level. This implies relatively fewer occurrences of simultaneous extreme values than if more appropriated distributions were used and therefore may lead to biased estimations if returns do not follow that particular distribution.

Since the 1960s there is abundant evidence in the literature showing that asset returns in general are not normally distributed (see Fama, 1965; Mandelbrot, 1963). Since then many empirical studies have confirmed this behavior for several classes of investments, including loan portfolios (Rosenberg and Schuermann, 2006). Moreover, it has also been found that returns are more correlated in the left tail (i.e. when investments result in losses or lower returns) than in the right tail. See, for instance, Ang and Bekaert (2002), Ning (2010) (who cites many other studies that reach the same conclusion) and Patton (2006). According to Das and Geng (2006) and Di Clemente and Romano (2004), returns of credit assets also present asymmetric (tail) dependence.

* Corresponding author. Tel.: +44 0755 123 7270; fax: +44 0131 668 3053.
E-mail addresses: J.Crook@ed.ac.uk (J. Crook), F.F.Moreira@sms.ed.ac.uk (F. Moreira).
[1] Tel.: +44 0755 123 7270; fax: +44 0131 668 3053.

0927-5398/$ – see front matter
doi:10.1016/j.jempfin.2011.05.005

Copulas are an effective way of capturing diverse dependence structures regardless of the individual distributions and symmetry. They have been used in finance since the end of the 90s and started being applied to credit risk a couple of years later. However in this latter field the application of copulas has been concentrated on corporate debts and derivatives.

The first contribution of this paper is the empirical estimation of best-fit copulas for consumer loans by using a credit card dataset provided by a large UK bank. Then, estimations of joint extreme default rates based on copulas are compared to estimations conditional on the assumption of normality. The second contribution is the test of five copulas that are not usually included in research pertaining to credit risk.

A third innovation is the use of goodness-of-fit tests (GoF) based only on the *right tail* of the variables' distributions (instead of the usual procedures that consider whole distributions). This strategy was implemented because the principal objective of finding the best-fit copulas here is to employ them to estimate the probability of simultaneous high defaults.

A sample of credit card loans was split into five segments according to a score provided by the Bank. Then the association between each of the ten pairs of segments was modeled by the best-fit copula. Most of the pairs of segments present right-tail dependence which suggests the existence of flaws in estimations of joint high defaults derived from traditional models. In other words, such structure means that higher default rates are more associated and the Bank is subject to larger losses in downturns than would be calculated with traditional techniques. We also find that some of the pairs have dependence appropriately represented by three of the five less popular copulas inserted in this study.

After finding the best representation for the dependence across the credit card loans, we compare estimations of conjunct high default rates following conventional assumptions of multivariate normality and Copula Theory. In most cases, the latter method generates values closer to the observed default rates in the dataset.[2] Considering each pair of segments separately and six risk levels (loss percentiles), the copula approach gave overall better results for all pairs.

The remainder of this paper is organized as follows. Section 2 contains a brief review of copulas, their application in credit risk and techniques to estimate copula parameters and decide which copula is the best one among many candidates. Next, we describe the data used in the empirical analysis. The ten copula families taken as candidates to represent the dependence structure among credit card loans are introduced in Section 4. Then, we estimate the dependence structure (copulas) between pairs of segments in a credit card portfolio of a large UK bank. Section 6 compares estimations of joint high defaults in the portfolio studied according to two approaches: by assuming normality and by using the best-fit copula. Final comments are in the last section.

2. Copulas, tail dependence and credit risk

2.1. Copulas and tail dependence

Copulas are functions that link univariate distributions to the multivariate distribution of the related variables. Let x and y be random variables. Then a copula may be represented as:

$$H(x,y) = C(F_X(x), G_Y(y))$$

where $F_X(x)$ and $G_Y(y)$ are the cumulative distribution functions of X and Y evaluated at x and y respectively and $C(.)$ is the copula that links those distributions in order to form the joint distribution $H(x, y)$. So, the copula C gives the probability that X and Y are simultaneously below x and y, i.e. $\Pr(X<x, Y<y)$, regardless of the shape of the distributions F_X and G_Y.

The probability that the variables are *above* specific points may be found by the survival copula, represented by $\hat{C}$:

$$\Pr(X>x, Y>y) = 1-F_X(x)-G_Y(y) + \Pr(X<x, Y<y) = \hat{C}\left(\overline{F}_X(x), \overline{G}_Y(y)\right)$$

where $\overline{F}_X(x) = 1-F_X(x)$ and $\overline{G}_Y(y) = 1-G_Y(y)$. Introductory explanations on Copula Theory may be found, for instance, in Joe (1997) and Nelsen (2006).

Copulas allow us to identify different levels of dependence across the distribution (i.e. when different levels of the variables present diverse association). In this paper, we are interested in the upper (right) tails of the default rate distributions and in the dependence present in this particular region (i.e. association among "high" default rates). This relationship can be measured by means of the upper tail dependence parameter, λ_U, given by (see, e.g., Joe, 1997; Trivedi and Zimmer, 2007):

$$\lambda_U = \lim_{p\to 1^-} \Pr\left[X > F_X^{-1}(p) \,|\, Y > G_Y^{-1}(p)\right] = \lim_{p\to 1^-} \Pr\left[Y > G_Y^{-1}(p) \,|\, X > F_X^{-1}(p)\right]$$

where p is the extreme percentile considered and F_X^{-1} and $G_Y^{-1}$are the inverse distributions of X and Y respectively. So, λ_U is the probability of one of the variables (X, for example) being greater than a specific percentile in its marginal distribution (F_X) given that the other variable (Y) is greater than that same percentile in its individual distribution (G_Y). Whenever $\lambda_U \in (0, 1]$, the variables present upper tail dependence and there is no upper tail dependence if $\lambda_U = 0$. The lower tail dependence parameter, λ_L, can be similarly calculated for variables smaller than specific cutoffs when the percentile p approaches zero. When $\lambda_L > 0$, the data is lower-tail dependent. Both tail dependence parameters are directly associated to the parameters of some copula families (see, for instance, Nelsen, 2006; Nikoloulopoulos et al., in press).

[2] Albeit the difference between the estimations via the two approaches was usually not statistically significant (see Section 6).