
ELSEVIER
Economic Modelling
Volume 27, Issue 6
November 2010
Co-Editors:
STEPHEN HALL
PETER PAULY
Associate Editors:
R. BARRELL
R.C. BRYANT
W.W. CHAREMZA
J. FEDDERKE
M. FUNKE
B.S. GANGNES
S. HOLLY
M. JUILLARD
L.R. KLEIN
L.J. LAU
D. LAXTON
J.-L. LIN
P. McADAM
M. MOORE
F. MOSCONE
P. NIJKAMP
D. OSBORN
G. TAVLAS
A. WELFE
S. WREN-LEWIS
ISSN 0264-9993
CONTENTS
Special Issue: P.A.V.B Swamy
Edited by Lawrence Klein, Stephen Hall, George Tavlas and Arnold Zellner
Arnold Zellner: January 2nd 1927–August 11th 2010 1337
S.G. Hall, L.R. Klein, G.S. Tavlas
Introduction: P.A.V.B. Swamy's contribution to Econometrics 1338
S.G. Hall, L.R. Klein, G.S. Tavlas, A. Zellner
Existence of singularity bifurcation in an Euler-equations model of the United States economy: Grandmont was right 1345
W.A. Barnett, S. He
Tests of hypotheses arising in the correlated random coefficient model 1355
J.J. Heckman, D. Schmierer
Panel data inference under spatial dependence 1368
B.H. Baltagi, A. Pirotte
Tests for structural change, aggregation, and homogeneity 1382
E. Maasoumi, A. Zaman, M. Ahmed
Bayesian shrinkage estimates and forecasts of individual and total or aggregate outcomes 1392
A. Zellner
Conditional volatility and correlations of weekly returns and the VaR analysis of 2008 stock market crash 1398
B. Pesaran, M.H. Pesaran
The 'Puzzles' methodology: En route to Indirect Inference? 1417
V.P.M. Le, P. Minford, M. Wickens
(contents continued on outside back cover)

Economic Modelling 27 (2010) 1529–1536

Contents lists available at ScienceDirect

Economic Modelling

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

Tail return analysis of Bear Stearns' credit default swaps

Liuling Li [a], Bruce Mizrach [b,*]

[a] Nankai University, P.R. China
[b] Department of Economics, Rutgers University, United States

ARTICLE INFO

JEL classification:
C11
G13

Keywords:
Bear Stearns
Credit default swaps
Bayesian analysis
Exponential power distribution

ABSTRACT

We compare several models for Bear Stearns' credit default swap spreads estimated via a Markov chain Monte Carlo algorithm. The Bayes Factor selects a CKLS model with GARCH–EPD errors as the best model. This model best captures the volatility clustering and extreme tail returns of the swaps during the crisis. Prior to November 2007, only four months ahead of Bear Stearns' collapse though, the swap spreads were indistinguishable statistically from the risk-free rate.

1. Introduction

The first major investment bank to fail during the global financial crisis was Bear Stearns. Bear Stearns was an 85-year old institution that until the summer of 2007 had never had a losing quarter. Nine months later, it was gone, absorbed into J.P. Morgan Chase in a shotgun marriage. This paper looks at credit default swap (CDS) prices to see if they shed any light on this story.

The academic literature on credit risk suggests that the bulk of price discovery is going on in the CDS market. Blanco et al. (2005) study 27 single name CDS, including 15 financial firms, and find that, on average, the CDS market contributes 80% of price discovery. They attribute this to more informed trading in the CDS market. A second factor is the deep liquidity in the CDS market. In the first half of 2008, the Bank for International Settlements reported that notional volume of single name CDS outstanding peaked at $33.4 trillion. There was another $24 trillion in multi-name instruments.

Prior to the financial crisis, the CDS spreads on investment grade debt narrowed substantially. The benchmark 5-year investment grade CDX index from Markit fell from 60 to 20 basis points between 2004 and early 2007.[1] For this reason, we first consider the swap spread as a risk-free rate, using the model of Chan et al. (1992).

As the crisis unfolded though, spreads gradually widened. We find that the volatility in the series requires adding GARCH errors. As Bear Stearns flirts with bankruptcy, an exponential power distribution is needed to capture these statistically improbable events. Wu (2006) has proposed using a similar model for stock returns and argues that they better approximate the tail behavior of financial assets.

We begin our analysis in Section 2 by describing the cash flows in a credit default swap, and then proceed to analyze the Bear Stearns series. In Section 3, we consider progressively more general models for the swap spread. We start with the CKLS model, then add GARCH errors, and finally, add an exponential power distribution to describe the fat tails that remain after GARCH modeling. Estimation via Markov chain Monte Carlo is described in Section 4. In Section 5, we report results and find that the Bayes Factor selects the general CKLS–GARCH–EPD model as the best. We test for structural breaks in Section 6, and confirm that the swap spread is statistically similar to a risk-free process prior to November 2007. Section 7 concludes.

2. Data

2.1. Credit default swaps

Credit default swaps are derivative securities that pay off in the case of a *credit event* by the *reference entity*, typically a default. The protection seller agrees to provide any missing cash flows from the *reference obligation* to the buyer, including interest and principal. The protection buyer generally pays an up-front fixed fee and a swap spread that varies with the market's assessment of the credit risk of the firm.

After signing the CDS contract, the buyer makes periodic payments, generally quarterly, to the seller until the maturity of the CDS or until a credit event occurs. The payment is calculated using the swap *spread*. Quoted in basis points (bp), or 0.01%, a spread of 180 bp, for example, implies that a protection buyer will pay $18,000 per year

* Corresponding author. Tel.: +1 732 932 7363; fax: +1 732 9327416.
E-mail address: mizrach@econ.rutgers.edu (B. Mizrach).
URL: http://snde.rutgers.edu (B. Mizrach).

[1] Gearson Lehman Group, "The Great CDO Unwind," October 31, 2008.

0264-9993/$ – see front matter
doi:10.1016/j.econmod.2010.07.023

to insure $1 million worth of par value. A higher spread, holding other factors constant, indicates a greater likelihood of default.

In our setting, Bear Stearns is the reference entity. The reference obligations are 5-year bonds which represent, according to Jorion and Zhang (2007), 85% of the CDS market. Although Bear Stearns came under severe stress, their takeover by J.P. Morgan prevented any credit events for bond holders. In contrast, the Lehman Brothers CDS wound up paying $91.375 for every $100 of par value insured following their October 2008 bankruptcy.

2.2. The story behind Bear Stearns' swap spreads

We use a daily time series of Bear Stearns' 5-year CDS spreads. It spans from April 2006 to March 2008 which is the month of the takeover by JP Morgan. There are 501 observations. This data was purchased from GFI, a major international broker dealer with a strong presence in the over-the-counter derivative markets.

The CDS spreads are plotted in Fig. 1.

It is straightforward to map the changes in CDS spreads into the time line of news events in Table 1. These headlines were collected from the *Wall Street Journal*, the *Financial Times*, and *Bloomberg*.

The swap spreads traded at 30 bp or less until February 2007, just days before Bear Stearns reported its first ever loss on the High Grade Structured Credit Strategies Fund (SCSF). This was the less heavily leveraged of two Bear Stearns' hedge funds with exposure to the subprime mortgage market. Mizrach (2010) notes that SCSF had gone 40 straight months without a loss, producing a cumulative 50% return.

Surprisingly though, spreads narrowed and fell back below 30 bp until June 2007 when Bear Stearns had to engineer a $3.2 billion bailout of its own funds. Spreads crossed 100 bp just before the two funds filed for bankruptcy in August 2007. Bear Stearns' credit risk further deteriorated when Warren Spector, head of the fixed income division and co-president (with Alan Schwartz) resigned on August 6, 2007.

Problems began to spread beyond Bear Stearns at that point. BNP Paribas suspended redemptions in several of its funds, and this was soon followed by liquidity injections from both the European Central Bank (ECB) and the Federal Reserve.

Table 1
News about Bear Stearns companies (BSC).

Date	Events
01-Mar-2007	Bear Stearns (BSC) reports first ever loss on the High Grade Structured Credit Fund
14-Jun-2007	BSC reports a 10% decline in quarterly earnings.
18-Jun-2007	Merrill Lynch seizes collateral from BSC hedge funds.
22-Jun-2007	BSC commits $3.2 billion to High Grade Structured Credit Fund.
17-Jul-2007	BSC tells clients that assets in Enhanced Leverage Fund are essentially worthless.
01-Aug-2007	BSC hedge funds file for bankruptcy.
06-Aug-2007	Warren Spector, Co-President resigns.
20-Sep-2007	BSC reports 68% drop in quarterly income.
26-Sep-2007	Rumor that Warren Buffet may buy 20% stake in BSC.
22-Oct-2007	BSC announces deal with Citic.
01-Nov-2007	*Wall Street Journal* article about CEO Cayne accuses him of smoking marijuana.
14-Nov-2007	CFO Molinaro says BSC will write down $1.62 billion and take a 4th quarter loss.
28-Nov-2007	BSC lays off another 4% of its staff.
20-Dec-2007	BSC takes $1.9 billion write-down. Cayne says he will skip his 2007 bonus.
07-Jan-2008	CEO Cayne retires under pressure. Schwartz takes over.
22-Jan-2008	Fed cuts rates 75 bps.
14-Feb-2008	UBS writes down $2 bn in Alt-A which BSC was long (paired with subprime short).
03-Mar-2008	Thornburg Mortgage fails to meet margin calls.
10-Mar-2008	Rumors of BSC liquidity problems begin to surface.
11-Mar-2008	Goldman Sachs e-mails clients that it will not do derivative deals with BSC.
14-Mar-2008	BSC announces $30 billion in funding from JP Morgan (JPM), via the Federal Reserve.
17-Mar-2008	JPM announces acquisition of BSC for $2 a share.
24-Mar-2008	JPM raises bid for BSC to $10 a share.

In retrospect, October of 2007 looks like the eye of a hurricane. The stock market rallied, and the Dow Jones Index reached an all-time high of 14,164 on October 9, 2007. Bear Stearns' swap spreads fell back to 70 bp.

Rumors about Bear Stearns' subprime exposure and liquidity needs persisted though and concerns about the firm were raised again in November 2007 when a *Wall Street Journal* article portrayed James

Fig. 1. Bear Stearns' 5-year credit default swap (CDS) spreads.

Cayne as a distracted, drug-using CEO. Swap spreads did not fall back below 100 bp from that point on.

The market's fears about Bear Stearns proved justified. The firm announced its first quarterly loss ever on November 14, 2007 and by the end of the month, swap spreads crossed 200 bp. Cayne resigned early in 2008, and despite further monetary injections by the Fed in February 2008, the upward trend in swap spreads remained intact.

To appreciate the velocity of Bear Stearns' collapse though, you have to recall that spreads did not reach 300 bp until March 3 when the second largest independent mortgage originator, Thornburg, could not meet margin calls. The run on Bear Stearns' remaining cash then began in earnest. A rumor that Goldman Sachs would no longer trade with Bear Stearns drove spreads above 600 bp on Tuesday March 11. By Friday March 14, Bear Stearns needed an emergency loan from the Fed with J.P. Morgan as conduit. While this announcement did not stabilize Bear Stearns' stock price, it did finally start to bring down the CDS spreads.

The takeover did not restore Bear Stearns to completely normal credit conditions. There were concerns about whether Bear Stearns would fight the takeover, and eventually, J.P. Morgan agreed to raise the acquisition price to $10 a share. With that offer in place on March 24, 2008, swap spreads fell back below 150 bp.

2.3. Descriptive statistics

We begin to move to a more formal analysis by providing some descriptive statistics in Table 2.

The skewness is significantly positive at 2.41, and the kurtosis is 10.77, which is greater than 3, indicating fat tails in the data. The probability density function of the first difference of CDS spreads is plotted in Fig. 2.

Compared to the normal distribution, both the left tail and the right tail of CDS spreads are longer.

3. Models

The short-term interest rate process is a fundamental input into a variety of asset pricing models. Econometric modeling of the short-rate remains an active area of research. While Bear Stearns' debt was hardly risk free, they were rated A+ and had been upgraded by Standard and Poor's in October 2006. They maintained that rating until August 2007.

The models we propose will also help understand how the Bear Stearns' CDS first gradually and then suddenly became much riskier.

3.1. CKLS

The paper by Chan et al. (1992) is the beginning of our search for an appropriate model. They proposed a differential equation, which we have discretized, that nests a number of popular models in the literature,

$$r_t = a + b_1 r_{t-1} + r_{t-1}^c e_t, \quad e_t = \sigma \varepsilon_t \tag{1}$$

where ε_t is distributed as Normal.

Table 2
Descriptive statistics for Bear Stearns' 5-year CDS spreads.

Obs.	Mean	Std. dev.	Minimum	Maximum	Skewness	Kurtosis
501	82.57	96.20	19	660	2.41	10.77

Notes: There are 501 observations spanning April 3, 2006 to March 31, 2008.

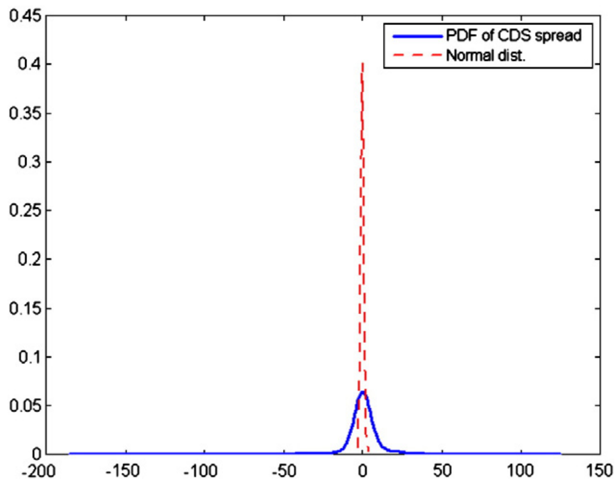

Fig. 2. Probability density function of CDS spreads.

Setting $c = 1/2$, you obtain the important Cox et al. (1985) model. Setting $c = 0$, which removes the effect from the level of the interest rate on volatility, you have the Ornstein–Unlenbeck process used by Vasicek (1977).

3.2. CKLS–GARCH model

In many financial time series, we observe unconditional fat tails and conditional volatility. Volatility clustering refers to the phenomenon that there are periods of high and low variances. That means large changes of variance tend to be followed by large changes, and small changes by small changes. Engle (1982) shows conditional heteroskedasticity may cause the fat tails in the unconditional distribution and suggests an ARCH model to capture the conditional heteroskedasticity. Bollerslev (1986) extends the ARCH model to a GARCH model, which is now the benchmark model for volatility clustering.

We append the GARCH model to Eq. (1). The CKLS–GARCH model is given by

$$\begin{aligned} r_t &= a + b_1 r_{t-1} + r_{t-1}^c e_t \\ e_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{j=1}^{r} \alpha_j e_{t-j}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \\ \alpha_0 &> 0, \quad \alpha_j \geq 0, \; j = 1, \ldots, r, \quad \beta_j \geq 0, \; j = 1, \ldots, s. \\ 1 &\geq \sum_{j=1}^{\max(r,s)} \left(\alpha_j + \beta_j\right) \end{aligned} \tag{2}$$

where ε_t follows the normal distribution. Research on CKLS–GARCH model can be found in Brenner et al. (1996), Koedijk et al. (1997), and Demirtas (2006). Brenner et al. (1996) show that the CKLS–GARCH model outperforms the CKLS model in maximum likelihood estimation of the 3-month Treasury bill.

3.3. CKLS–GARCH–EPD model

Many papers examining the time varying volatility in financial time series have questioned the conditional normality of the GARCH model. Typically, the standardized GARCH residuals ε_t/σ_t are not normally distributed. Bollerslev (1987), for example, suggests a conditional t-distribution for ε_t. Haas et al. (2006) have also utilized the stable Paretian density.

In this paper, we propose to capture the outlying volatility shocks ε_t using an exponential power distribution (EPD) with the probability density function (PDF)

$$f(\varepsilon_t) = \frac{1}{\lambda 2^{1+1/\alpha}\Gamma(1+1/\alpha)}\exp\left\{-\frac{1}{2}\left|\frac{\varepsilon_t}{\lambda}\right|^{\alpha}\right\} \tag{3}$$

where λ is a normalizing constant to make the variance of ε_t equal to unity:

$$\lambda = \sqrt{\frac{2^{-2/\alpha}\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}.$$

In Fig. 3, we plot a class of PDFs for the exponential power distribution. If parameter $\alpha=2$, the exponential power distribution will be the Normal distribution. If parameter $\alpha=1$, it will be the Laplace distribution, which has fatter tails than Normal distribution. As α increases from 2 to 3, the PDF will become more platykurtic. As α decreases from 2 to 1, the PDF will become more leptokurtic. The exponential power distribution in our model can be used to capture the jump component in the swap spread. For more information of the EPD distribution, one can refer to Nelson (1991) and Bali and Wu (2006).

The kurtosis of ε_t is defined as follows (see Nadarajah, 2005):

$$Kurtosis(\varepsilon_t) = \frac{\Gamma(1/\alpha)\Gamma(5/\alpha)}{\Gamma^2(3/\alpha)}.$$

The CKLS–GARCH–EPD model is given by

$$\begin{aligned} r_t &= a + b_1 r_{t-1} + r_{t-1}^c e_t \\ e_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{j=1}^{r} \alpha_j e_{t-j}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \\ \alpha_0 &> 0, \quad \alpha_j \geq 0,\ j=1,\dots,r, \quad \beta_j \geq 0,\ j=1,\dots,s. \\ 1 &\geq \sum_{j=1}^{\max(r,s)} \left(\alpha_j + \beta_j\right) \end{aligned} \tag{4}$$

where ε_t is drawn from Eq. (3).

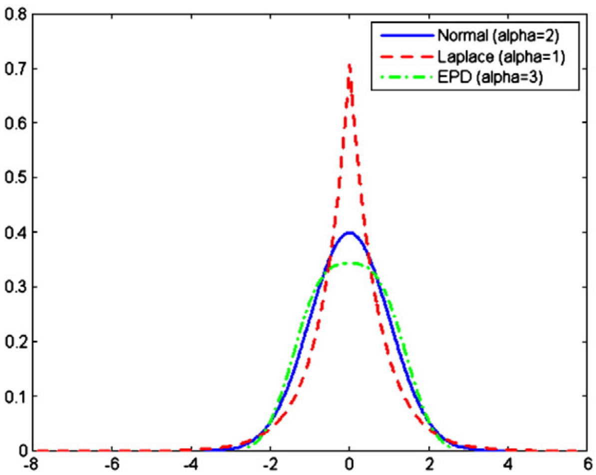

Fig. 3. Class of PDFs for the exponential power distribution (EPD).

4. Methodology

We estimate these models by the Bayesian method. Markov chain Monte Carlo algorithms are used. For example, the posterior PDF of the CKLS–GARCH–EPD model is given by

$$p(\Delta|data) \propto p(\Delta)\prod_{t=1}^{n}\frac{r_{t-1}^{-c}\sigma_t^{-1}}{\lambda 2^{1+1/\alpha}\Gamma(1+1/\alpha)}\exp\left\{-\frac{1}{2}\left|\frac{\hat{\varepsilon}_t}{\lambda}\right|^{\alpha}\right\} \tag{5}$$

where

$$\begin{aligned} \Delta &= \{a, b_1, c, \alpha_0, \Gamma', \Lambda', \alpha\} \\ \Gamma &= (\alpha_1, \dots, \alpha_r)' \\ \Lambda &= (\beta_1, \dots, \beta_s)' \end{aligned}$$

and

$$\hat{\varepsilon}_t = \frac{r_t - a - b_1 r_{t-1}}{\sigma_t r_{t-1}^c}, \quad t = 1, \cdots, n, \quad \hat{\varepsilon}_0 = \cdots = \hat{\varepsilon}_{-q} = 0.$$

As the prior, we set

$$\begin{aligned} p(\Delta) = {} & N\left(\bar{a}, \bar{\sigma}_a^2\right) \times N\left(\bar{b}_1, \bar{\sigma}_{b_1}^2\right) \times N\left(\bar{\alpha}_0, \bar{\sigma}_{\alpha_0}^2\right) I(\alpha_0 > 0) \\ & \times N\left(\bar{\Gamma}, \bar{\sigma}_\Gamma^2\right) I(\Gamma > 0) \times N\left(\bar{\Lambda}, \bar{\sigma}_\Lambda^2\right) I(\Lambda > 0) \times p(c) \times p(\alpha) \end{aligned} \tag{6}$$

where "$\bar{\bullet}$" denotes the prior parameters, $p(c)=1$, $p(\alpha)=1$, and $I(\cdot)$ is an indicator function. We set all $\bar{\sigma}^2$ sufficiently large and all the prior mean parameters zero. Similarly, we can derive the posterior densities for other models.

The MCMC algorithms for the CKLS–GARCH–EPD model include the following blocks (see Li et al. (2009)):

1. *draw*$[a, b_1]$;	2. *draw* α_0;	3. *draw* Γ;
4. *draw* Λ;	5. *draw* c;	6. *draw* α.

The convergence of the MCMC algorithm is judged by the Kolmogorov–Smirnov tests (KST) that are explained in Goldman et al. (2008). The robustness of our MCMC algorithm is checked by changing both priors and initial values, respectively.

5. Results

5.1. CKLS

The CKLS model is estimated in the first column of Table 3.

We calculate the estimated spreads from the CKLS model as follows:

$$\hat{r}_t = \hat{a} + \hat{b}_1 r_{t-1} \tag{7}$$

and the standardized residuals are:

$$\hat{\varepsilon}_t = \frac{r_t - \hat{r}_t}{r_{t-1}^{\hat{c}}\ \hat{\sigma}}. \tag{8}$$

The Lagrange multiplier (LM) test with null hypothesis of no ARCH effects is applied to the standardized residuals. The LM test statistic is 15.88 with a p-value of 0, which means we can reject the null hypothesis at the 5% significance level. Hence, we conclude that ARCH effects exist in the residuals of the CKLS model.

Table 3
Posterior means and t-statistics for the CDS spreads.

Models	M_1	M_2	M_3
a	2.7048	1.9928	0.1096
	(0.97)	(0.68)	(0.91)
b_1	0.8702	0.9169	0.9953
	(4.82)	(6.09)	(622.06)
c	0.8469	0.7879	0.4970
	(6.44)	(8.21)	(236.67)
α_0	0.4852	0.5448	0.5178
	(1.84)	(2.14)	(1.84)
α_1	–	0.7222	0.6778
		(4.54)	(4.19)
β_1	–	0.1673	0.1744
		(1.12)	(1.46)
α	2	2	0.2826
	(–)	(–)	(8.70)
B_{3*}	3110	975	–

Notes:B_3 means the value of Bayes Factor for M_3 over model *. M_1 is the CKLS model in Eq. (1). M_2 is the CKLS–GARCH model in Eq. (2). M_3 is the CKLS–GARCH–EPD model in Eq. (4).

5.2. CKLS–GARCH

We add the GARCH terms in Eq. (2) to mitigate the fat tails. The estimated model is in the second column of Table 3. We get the estimated spread from the CKLS–GARCH model as follows:

$$\hat{r}_t = \hat{a} + \hat{b}_1 r_{t-1} \tag{9}$$

and the standardized residuals are:

$$\hat{\varepsilon}_t = \frac{r_t - \hat{r}_t}{r_{t-1}^{\hat{c}} \hat{\sigma}_t},$$

where

$$\hat{\sigma}_t^2 = \hat{\alpha}_0 + \sum_{j=1}^{r} \hat{\alpha}_j \hat{e}_{t-j}^2 + \sum_{j=1}^{s} \hat{\beta}_j \hat{\sigma}_{t-j}^2 \tag{10}$$

$$\hat{e}_{t-j} = \hat{\sigma}_{t-j} \varepsilon_{t-j}. \tag{11}$$

The LM test statistic calculated for the standardized residuals is 0.31 with a p-value of 58%. That means under 5% significance level we accept the null hypothesis and conclude that there are no ARCH effects in the residuals of the CKLS–GARCH model.

The estimates for parameter α_1 and parameter β_1 are 0.7222 and 0.1673. The sum, $\alpha_1 + \beta_1 = 0.8895$, is well inside the integrated GARCH boundary.

We perform a Jarque–Bera test of the null hypothesis that the residuals come from a normal distribution. The p-value for the Jarque–Bera test is 0, so we conclude that the standardized residuals are not normally distributed at the 5% significance level.

5.3. CKLS–GARCH–EPD

We allow the GARCH error to follow an EPD and report estimation results in the third column of Table 3. We get the estimated spread from the CKLS–GARCH model as follows:

$$\hat{r}_t = \hat{a} + \hat{b}_1 r_{t-1} \tag{12}$$

and the standardized residuals are:

$$\hat{\varepsilon}_t = \frac{r_t - \hat{r}_t}{r_{t-1}^{\hat{c}} \hat{\sigma}_t} \tag{13}$$

$$\hat{\sigma}_t^2 = \hat{\alpha}_0 + \sum_{j=1}^{r} \hat{\alpha}_j \hat{e}_{t-j}^2 + \sum_{j=1}^{s} \hat{\beta}_j \hat{\sigma}_{t-j}^2 \tag{14}$$

$$\hat{e}_{t-j} = \hat{\sigma}_{t-j} \varepsilon_{t-j}. \tag{15}$$

We calculate the LM test for the standardized residuals. The LM test statistic is 1.97 with a p-value of 16%, so we can accept the null hypothesis at the 5% significance level. That is, there are no ARCH effects in the residuals of the CKLS–GARCH–EPD model.

We perform a Jarque–Bera test of the null hypothesis that the standardized residuals come from a normal distribution. The p-value for the Jarque–Bera test is 0, so we conclude the residuals are not normally distributed at the 5% significance level.

Next, we generate a sample from the EPD with $\hat{\alpha} = 0.2826$ and plot the empirical CDF in Fig. 4.

Compared to the empirical CDF of the standardized residuals, we can see they are very close to each other. For comparison, we also plot the CDF of standard normal. We conclude that the standardized residuals are closer to the EPD than to the normal.

The p-value of sign test for the standardized residuals and the sample of EPD with $\hat{\alpha} = 0.2826$ is 7.39%, higher than 5% significance level. So we conclude that there is no difference between the standardized residuals and the sample of EPD with $\hat{\alpha} = 0.2826$.

Following Bollerslev (1987), we compare the conditional kurtosis of EPD with that of the standardized residuals. The conditional kurtosis of EPD is calculated using the formula in Nadarajah (2005),

$$Kurtosis(\varepsilon_t) = \frac{\Gamma\left(1/\hat{\alpha}\right)\Gamma\left(5/\hat{\alpha}\right)}{\Gamma^2\left(3/\hat{\alpha}\right)}.$$

The estimated conditional kurtosis of EPD is 234, which is very similar to the sample analogue of $\left(\hat{\varepsilon}_t/\hat{\sigma}_t\right)^2$, $K = 279$. This result is similar to that of Bollerslev (1987).

The estimates for α_1 and β_1 are 0.6778 and 0.1744. $\alpha_1 + \beta_1$ is 0.8522, which again lies comfortably inside the integrated GARCH frontier.

5.4. Model selection by Bayes Factor

The Bayes Factor is used as model selection criterion (see Appendix A). The result shows the CKLS model with GARCH–EPD error terms provides a better fit. The critical values of log Bayes Factor are reported in Table 4. For example, for the CKLS–GARCH–EPD model (denoted as M_3), the Bayes Factor of M_3 over M_1 is 3110, which is greater than 2. That means, the evidence supporting M_3 is decisive.

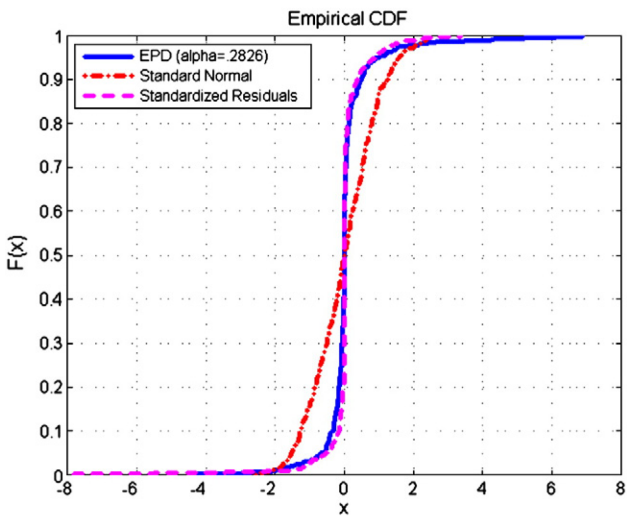

Fig. 4. Empirical CDF of EPD distribution.

Table 4
Critical values for log Bayes Factor.

$\log_{10}BF_{12}$	Evidence support M_1	$\log_{10}BF_{21}$
0 to 0.5	Weak	−0.5 to 0
0.5 to 1	Substantial	−1 to −0.5
1 to 2	Strong	−2 to −1
>2	Decisive	<−2

Notes: The values of log Bayes Factor can be negative. For example, if $\log_{10}BF_{12}=0.5$, then $\log_{10}BF_{21}=-0.5$ since $\log_{10}BF_{21}=\log_{10}(1/BF_{12})$.

We quantify the evidence of the Bayes Factor by the "energy" of changing the prior probability of the null hypothesis to a posterior probability of the null hypothesis. If one is not highly convinced of model 1 (3% prior probability of the null hypothesis) before analyzing the data, a Bayes Factor of 3110 surely will. It shifts the posterior probability to 99% in favor of the null hypothesis. In other words, a Bayes Factor of 3110 is strong enough to move one from being only 3% sure of the null hypothesis to being 99% sure.

The same analysis can be applied to compare M_3 with M_2. The Bayes Factor of M_3 over M_2 is 975, which is greater than 2. That means, the evidence supporting M_3 is decisive. From Table 5 we know Bayes Factor of 975 is strong enough to move one from being only 3% sure of the null hypothesis to being 98% sure. As a result, we conclude that model 3 is better.

For M_3, the CKLS model with a GARCH–EPD error process, we document a significant level effect. The t-statistic of parameter c is 237, which is significant at the 5% level. Also, the t-statistic of parameter α_1 is 4.19, which supports significant volatility clustering in Bear Stearns' CDS spreads. Last, the t-statistic of parameter α in EPD is 8.7, which provides significant evidence of jumps in the CDS spreads.[2] And the value of parameter α is 0.2826, far from 1 (Laplace distribution) and 2 (Normal distribution). That means, a conditional normal error distribution assumption is not adequate for CDS spreads.

In summary, the CKLS model with a GARCH–EPD error process best fits the Bear Stearns' CDS spreads. We document significant level effects, volatility clustering and the jumps in the crisis episode.

6. Structural breaks

To understand better the importance of structural breaks, if any, in the swaps, we utilize the Zivot and Andrews (1992) test. The null in Zivot and Andrews is a unit root, and the alternative is an $I(0)$ process with one structural break. Their procedure searches through the sample for the minimum t-stat on θ_1, the coefficient on the lagged level in the regression,

$$\Delta r_t = \alpha_0 + \alpha_1 t + \alpha_2 D_t^B + \theta_1 r_{t-1} + \sum_{j=1}^{J} \phi_j \Delta r_{t-j} + u_{i,t}. \tag{16}$$

t is a time trend, and D_t^B is a dummy variable that is zero prior to date B and equal to 1 thereafter. The lags J are included from a maximum of 12 lags using a t-test criterion at the 10% level. Zivot and Andrews provide asymptotic critical values of −4.80 for a 5% test and −5.34 for a 1%.

For the swaps, the t-test sets $J=12$, and the strongest evidence for a structural break comes on November 1, 2007 when $t(\theta_1)=-3.73$. This is when the *Wall Street Journal* reported on the erratic behavior of Bear Stearns' CEO Cayne.

We then separate the sample of CDS spreads into two parts using this date. The estimation results for the subsample (April 3, 2006–October 31, 2007) are reported in Table 6. The Bayes Factor of M_3 over M_2 is −611.

The data favor the CKLS–GARCH model over CKLS–GARCH–EPD. The t-statistic of the EPD shape parameter α is 1.23, which is less than 2. That means the EPD shape parameter is not significant at 5% significance level.

For comparison, we also estimate the model for the 3-month Treasury bill rate obtained from the Federal Reserve Board. The estimation results are listed in Table 7.

The Bayes Factor of M_3 over M_2 is −221, favoring the CKLS–GARCH model. The t-statistic of EPD shape parameter α is 1.89, which is again not significant at the 5% level. These results are similar to those of CDS spreads.

Hence, we conclude that during period of April 3, 2006–October 31, 2007, investors may have been correct to conclude that Bear Stearns' swaps were not that much riskier than safe government debt. Time was to prove them wrong.

7. Conclusion

We compare several models for credit default swap spreads using a Markov chain Monte Carlo algorithm. The Bayes Factor is used as a

[2] Potentially discontinuous price movements (jumps) can be accounted for by EPD distribution assumption. Bali and Wu (2006) use exponential power distribution(EPD) to capture the jumps in the interest rate. For more researches on jumps, one can refer to Das (2002), Piazzesi (2005) and Mancini and Reno (2008).

Table 5
Posterior probability of the null hypothesis.

Evidence for model 3	$\log BF_{3^*}$	Change in probability of the null hypothesis(model 3) From (%)	To (%)
Decisive	2	75	86
		50	67
		3	5
Decisive	3110	75	100*
		50	100
		3	99
Decisive	975	75	100
		50	100
		3	98

Notes: The table reports the posterior probability of the null hypothesis after observing the Bayes Factors with given prior probability. Posterior odds = Bayes Factor × Prior odds. Posterior probability = odds/(1 + odds). For example, the number marked by an asterisk in Table 5 is calculated as following example. Posterior odds = Bayes Factor × Prior odds = 3110 × 75%/(1 − 75%) = 9330. Posterior probability = odds/(1 + odds) = 9330/(1 + 9330) = 100%.

Table 6
Posterior means and t-statistics for the subsample of CDS spreads.

Models	M_1	M_2	M_3
a	−0.4115	6.8433	−7.2984
	(−0.14)	(1.21)	(−0.51)
b_1	−1.1459	0.7196	−0.6377
	(−0.77)	(3.01)	(−0.39)
c	1.2782	0.6369	0.5654
	(8.60)	(2.29)	(1.15)
α_0	0.4811	0.4798	0.4574
	(1.77)	(1.83)	(1.58)
α_1	–	0.7234	0.6975
		(4.06)	(4.18)
β_1	–	0.1743	0.1729
		(1.27)	(1.30)
α	2	2	0.7414
	(–)	(–)	(1.23)
B_{3^*}	2690	−611	–

Notes: We analyze the spread models over the period April 3, 2006 to October 31, 2007, just prior to the structural break identified by the Zivot–Andrews test. B_3 means the value of Bayes Factor for M_3 over model *. M_1 is the CKLS model in Eq. (1). M_2 is the CKLS–GARCH model in Eq. (2). M_3 is the CKLS–GARCH–EPD model in Eq. (4).

Table 7
Posterior means and *t*-statistics for the 3-month T-bill rate.

Models	M_1	M_2	M_3
a	0.2274	2.2591	4.6642
	(0.08)	(2.33)	(0.90)
b_1	−1.7254	0.5371	−1.1549
	(−0.91)	(2.78)	(−0.46)
c	1.2604	0.0101	0.2918
	(5.47)	(0.58)	(.6670)
α_0	0.4494	0.5191	0.5344
	(1.60)	(1.85)	(1.89)
α_1	–	0.7315	0.7376
		(4.34)	(4.80)
β_1	–	0.1516	0.1388
		(1.27)	(1.19)
α	2	2	0.7541
	(–)	(–)	(1.89)
B_{3*}	17,000	−221	–

Notes: We analyze the 3-month T-bill rate models over the period April 3, 2006 to October 31, 2007, just prior to the structural break in the CDS spreads identified by the Zivot–Andrews test. B_{3*} means the value of Bayes Factor for M_3 over model *. M_1 is the CKLS model in Eq. (1). M_2 is the CKLS–GARCH model in Eq. (2). M_3 is the CKLS–GARCH–EPD model in Eq. (4).

model selection criterion. We find that the CKLS model with GARCH volatility and exponential power distribution errors provides the best fit. This establishes that level effects, volatility clustering and jumps are statistically significant components of CDS spreads.

Our analysis also documents the calm before the storm. As late as October 2007, when policy makers and industry participants were assuring us that the subprime crisis was contained, Bear Stearns' CDS spreads were evolving smoothly like most other investment grade debt. While Bear Stearns' March 2008 collapse was ultimately confined through a toxic asset ring-fence and a merger with a larger, less leveraged institution, events would eventually spin out of control just six months later.

Acknowledgment

We would like to thank Shiliang Li for assistance with the MCMC algorithms.

Appendix A. Bayes Factor

Appendix A.1. Bayes Factor, posterior odds ratio and likelihood ratio

Given data Y, we want to compare models M_1 and M_2, which are with parameter sets Θ_1 and Θ_2, respectively. The *likelihood function* for model M_i is $p(Y_i|\Theta_i,M_i)$.

The *marginal likelihood function for model* M_i, or the *integrated likelihood* or the *evidence for model* M_i is:

$$p(Y|M_i) = \int p(Y,\Theta_i|M_i)d\Theta_i = \int p(Y|\Theta_i,M_i)p(\Theta_i|M_i)d\Theta_i.$$

The last equation is obtained by applying Bayes' theorem.

The *posterior probability* of model i given data Y, $p(M_i|Y)$, can be derived by Bayes' theorem,

$$p(M_i|Y) = \frac{p(M_i,Y)}{p(Y)} = \frac{p(Y|M_i)p(M_i)}{p(Y)}.$$

Hence, the *posterior odds ratio* for model M_1 against model M_2 is:

$$\frac{p(M_1|Y)}{p(M_2|Y)} = \frac{p(Y|M_1)}{p(Y|M_2)} \times \frac{p(M_1)}{p(M_2)}.$$

i.e.,

$$posterior\ \ odds\ \ ratio = Bayes\ Factor \times prior\ \ odds\ \ ratio.$$

Solving for the *Bayes Factor*, you find

$$\frac{p(Y|M_1)}{p(Y|M_2)} = \frac{p(M_1|Y)}{p(M_2|Y)} / \frac{p(M_1)}{p(M_2)}.$$

The Bayes Factor differs from the posterior odds ratio by eliminating the effect of priors.

In general, prior odds ratio is set to be 1. i.e., $p(M_1)=p(M_2)=0.5$. In this case, Bayes Factor equals the posterior odds ratio. That's the reason why sometimes these two terminologies can be inter-exchangeable. Further, if these two models are assumed with no parameters, then there will be no integration with respect to parameters. In this case, the Bayes Factor is just the *likelihood ratio*.[3]

Appendix A.2. Model selection via Bayes Factor

In a model selection problem, we have to choose between M_1 and M_2 on the basis of the data Y. In theory, *Bayes Factor* BF_{12} is given as follows:

$$\begin{aligned} BF_{12} &= \frac{p(Y|M_1)}{p(Y|M_2)}, \\ &= \frac{\int p(Y,\Theta_1|M_1)d\Theta_1}{\int p(Y,\Theta_2|M_2)d\Theta_2}, \\ &= \frac{\int p(Y|\Theta_1,M_1)p(\Theta_1|M_1)d\Theta_1}{\int p(Y|\Theta_2,M_2)p(\Theta_2|M_2)d\Theta_2}. \end{aligned}$$

where $p(Y|M_i)$ is the marginal likelihood for model i. $p(Y|\Theta_i,M_i)$ is the likelihood for model i. And Θ_i is the parameter set in model i.

To calculate the Bayes Factor from the MCMC algorithms, we use following method (Kass and Raftery (1995), p.779),

$$BF_{12} = \frac{\int p(Y|\Theta_1,M_1)p(\Theta_1|M_1)d\Theta_1}{\int p(Y|\Theta_2,M_2)p(\Theta_2|M_2)d\Theta_2} \approx \frac{\frac{1}{n}\sum_{j=1}^{n} p\left(Y|\Theta_1^{(j)},M_1\right)}{\frac{1}{n}\sum_{j=1}^{n} p\left(Y|\Theta_2^{(j)},M_2\right)}$$

where j is the jth draw of parameter Θ from the MCMC algorithms. n is the number of the accepted draws in the MCMC algorithms.

The critical regions in Table 4 are from page 777 of Kass and Raftery (1995). For example, if the value of $\log_{10}BF_{12}$ falls into the interval of [0.5,1], we conclude that there is substantial evidence supporting model 1.

Followed Goodman (1999), we quantify the evidence of Bayes Factor by the "energy" of changing the prior probability of the null hypothesis to a posterior probability of the null hypothesis.

For instance, if one is highly convinced of model 1 (75% prior probability of the null hypothesis) before analyzing the data, a Bayes Factor of 0.001 will convince that the null hypothesis is not true (3% posterior probability of the null hypothesis). In other words, to achieve 5% posterior probability of the null hypothesis with a Bayes Factor of 0.001, one needs to have an 85% prior probability of the null hypothesis.

A Bayes Factor of 2 is strong enough to move one from being 75% sure of the null hypothesis to being 85% sure (see Table 5).

References

Bali, Turan G., Wu, Liuren, 2006. A comprehensive analysis of the short term interest-rate dynamics. Journal of Banking and Finance 30, 1269–1290.

[3] See Kass and Raftery (1995).

Blanco, Roberto, Brennan, Simon, Marsh, Ian, 2005. An empirical analysis of the dynamic relationship between investment-grade bonds and credit default swaps. Journal of Finance 60, 2255–2281.
Bollerslev, Tim, 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.
Bollerslev, Tim, 1987. A conditional heteroscedastic time series model for speculative prices and rates of return. Review of Economics and Statistics 69, 542–547.
Brenner, Robin J., Harjes, Richard H., Kroner, Kenneth F., 1996. Another look at models of the short-term interest rate. The Journal of Financial and Quantitative Analyst 31, 85–107.
Chan, K.C., Andrew Karolyi, G., Longstaff, Francis A., Sanders, Anthony B., 1992. An empirical comparison of alternative models of the short-term interest rate. Journal of Finance 47, 1209–1228.
Cox, John C., Ingersoll Jr., Jonathan E., Ross, Stephen A., 1985. A theory of the term structure of interest rates. Econometrica 53, 385–408.
Das, Sanjiv R., 2002. The surprise element: jumps in interest rates. Journal of Econometrics 106, 27–65.
Demirtas, K. Ozgur, 2006. Nonlinear asymmetric models of the short-term interest rate. The Journal of Futures Markets 26, 869–894.
Engle, Robert F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. Econometrica 50, 987–1008.
Goldman, Elena, Valiyeva, Elmira, Tsurumi, Hiroki, 2008. Kolmogorov–Smirnov, fluctuation and Z_g tests for convergence of Markov chain Monte Carlo draws. Communications in Statistics, Simulations and Computations 37, 368–379.
Goodman, Steven N., 1999. Toward evidence-based medical statistics 2: the Bayes Factor. Annals of Internal Medicine 130, 1005–1013.
Haas, Markus, Mittnik, Stefan, Mizrach, Bruce, 2006. Assessing central bank credibility during the EMS crises: comparing option and spot market-based forecasts. Journal of Financial Stability 2 (2006), 28–54.
Jorion, Philippe, Zhang, Gaiyan, 2007. Good and bad credit contagion: evidence from credit default swaps. Journal of Financial Economics 84, 860–883.
Kass, Robert E., Raftery, Adrian E., 1995. Bayes factors. Journal of the American Statistical Association 90, 773–795.
Koedijk, Kees G., Nissen, Francois G.J.A., Schotman, Peter C., Wolff, Christian C.P., 1997. The dynamics of short-term interest rate volatility reconsidered. European Finance Review 1, 105–130.
Li, Liuling, Li, Shiliang, Tsurumi, Hiroki, 2009. Empirical test for the non-linear mean of the short term interest rate. Working Paper. Rutgers University.
Mancini, Cecilia, Reno, Roberto, 2008. Threshold Estimation of Jump Diffusion Models and Interest Rate Modeling. (Available at SSRN: http://ssrn.com/abstract=1158439).
Mizrach, Bruce, 2010. The role of leverage, VaR and balance sheets: a discussion of Danielsson, Shin and Zigrand. NBER Conference on Systematic Risk.
Nadarajah, Saralees, 2005. A generalized normal distribution. Journal of Applied Statistics 32, 685–694.
Nelson, Daniel B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59, 347–370.
Piazzesi, Monika, 2005. Bond yields and the Federal Reserve. Journal of Political Economy 113, 311–344.
Vasicek, Oldrich, 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177–188.
Wu, Liuren, 2006. Dampened power law: reconciling the tail behavior of financial security returns. Journal of Business 79, 1445–1473.
Zivot, Eric, Andrews, Donald W.K., 1992. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. Journal of Business & Economic Statistics 10, 251–270.