Insurance: Mathematics and Economics 64 (2015) 106–125

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

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

State price densities implied from weather derivatives

Wolfgang Karl Härdle [a,b], Brenda López-Cabrera [a], Huei-Wen Teng [c,*]

[a] C.A.S.E Center for Applied Statistics and Economics and Ladislaus von Bortkiewicz Chair of Statistics, Humboldt-Universität zu Berlin, Germany
[b] Sim Kee Boon Institute for Financial Economics, Singapore Management University, Singapore
[c] Graduate Institute of Statistics, National Central University, Taoyuan City, Taiwan

ARTICLE INFO

Article history:
Received July 2014
Received in revised form
April 2015
Accepted 4 May 2015
Available online 14 May 2015

Keywords:
Weather derivatives
Temperature derivatives
HDD
CDD
State Price Density
Quadrature
Bayesian
Data sparsity

ABSTRACT

A State Price Density (SPD) is the density function of a risk neutral equivalent martingale measure for option pricing, and is indispensable for exotic option pricing and portfolio risk management. Many approaches have been proposed in the last two decades to calibrate a SPD using financial options from the bond and equity markets. Among these, non and semiparametric methods were preferred because they can avoid model mis-specification of the underlying. However, these methods usually require a large data set to achieve desired convergence properties. One faces the problem in estimation by e.g., kernel techniques that there are not enough observations locally available. For this situation, we employ a Bayesian quadrature method because it allows us to incorporate prior assumptions on the model parameters and hence avoids problems with data sparsity. It is able to compute the SPD of both call and put options simultaneously, and is particularly robust when the market faces the data sparsity issue. As illustration, we calibrate the SPD for weather derivatives, a classical example of incomplete markets with financial contracts payoffs linked to non-tradable assets, namely, weather indices. Finally, we study related weather derivatives data and the dynamics of the implied SPDs.

1. Introduction

A State Price Density (SPD) is the density function of a Risk Neutral (RN) equivalent martingale measure for option pricing, and it is a measure more tied to uncertainty than to volatility and it is indispensable for (exotic) option pricing and portfolio risk management. It does not only reflect a risk-adaptive behavior of investors based on historical assessment of the futures market, but it also gives insights about the preferences and risk aversion of a representative agent, see for example Aït-Sahalia and Lo (2000), Jackwerth and Rubinstein (1996) and Rosenberg and Engle (2002).

Consider a European call option with maturity date T and strike price K. Under the non-arbitrage principle, its price at t can be given as:

$$C(K) = \mathrm{e}^{-r\tau} \int \max(x - K, 0) f(x) dx \tag{1}$$

where r is the risk-free interest rate, τ time to maturity and $f(x)$ is the defined SPD. The advantage of extracting the SPD directly from market prices is that volatility and other moments can easily be calculated using this SPD independent of any particular pricing model.

There are many approaches to calibrate the SPD using financial options from the bond and equity markets. Assuming a Black and Scholes (B&S) model implies that the RN measure is a lognormal distribution which may result in severe bias of the SPD estimation since certain volatility properties are not correctly reflected. As observed by Breeden and Litzenberger (1978), the SPD of any risky asset can be derived as the second derivative with respect to the strike price of an estimate of the pricing function C. A number of econometric techniques have been developed to address this calibration issue. The most notable examples include the stochastic volatility models and the GARCH models. Derman and Kani (1994), Dupire (1994) and Rubinstein (1994) implied SPDs using binomial trees, hence avoiding too strong stochasticity assumption like e.g., Geometric Brownian motion. Others like Abadir and Rockinger (2003) use hypergeometric distributions. Although useful in a variety of contexts, these (parametric) models are still susceptible to model specification.

Various non-parametric models have been employed to overcome this problem. Aït-Sahalia and Lo (1998) introduce a semiparametric alternative where the volatility of the B&S formulation is modeled non-parametrically. From a statistical point of view, estimating the SPD becomes estimating the second derivative of a

* Corresponding author.
E-mail address: venteng@gmail.com (H.-W. Teng).

http://dx.doi.org/10.1016/j.insmatheco.2015.05.001
0167-6687/

regression function, but the SPD needs to be a proper density function (non negative and integrates to one). This dictates that the price is decreasing and convex in terms of the strike price. How to impose these constraints presents the main difficulties of direct applications of nonparametric regression. Aït-Sahalia and Duarte (2003), Yatchew and Härdle (2006) and Härdle and Hlávka (2009) stress the importance of enforcing such shape constraints. Fan and Mancini (2009) use a non-parametric technique to estimate the state price distribution but not the density because the former is easier to estimate. Giacomini et al. (2008) use mixtures of scales and shifted t-distributions, while Yuan (2009) uses a mixture of lognormals. Curve fitting method have been presented in Rubinstein (1994) and Jackwerth and Rubinstein (1996). Liechty and Teng (2009) introduce the Bayesian quadrature model, where both the locations and weights of the support points for approximating the SPD are random variables. Most nonparametric methods require a rich body of data to achieve desired convergence properties. The main goal of this paper is to infer the SPD from markets, where trading activities are less frequently occurred.

For this purpose, we employ a Bayesian quadrature method as a calibration method for the SPD from option prices, because it allows us to incorporate prior assumptions on the model parameters and hence avoids problems with data sparsity. This approach takes a prior distribution which can be parametric (e.g. lognormal) or a uniform density. The posterior distribution of the SPD is calibrated to market data. This method is a special case of a mixture model, where the component densities are point measures.

The novelty of the Bayesian quadrature approach relies on the fact that it uses unequal weights and is in a Bayesian framework. Approximating the state price density with weighted sum of δ-functions produces good model fitting by using a parsimonious model. Bayesian inference gives a straightforward probabilistic framework and provides reasonable credible regions for the implied state price density, which can be further used for various purposes such as hedging and pricing.

We show that the proposed method has some advantages over other nonparametric methods: (1) it considers the locations and weights of the support points in the finite representation of the SPD as random variables, (2) it is parsimonious and allows for statistical inference, it enables us to construct credible regions for the current value of the SPD (3) it is computationally efficient in the sense that a Markov chain Monte Carlo algorithm with Gibbs sampler can be adopted, so that no additional tuning procedures are required for exploring the posterior distribution and (4) it is robust even if the market faces data sparsity issues. (5) These classes of Risk Neutral probabilities do not stem from market-risk-price assumptions.

We conduct our empirical analysis based on weather derivative (WD) data traded at the Chicago Mercantile Exchange (CME). WDs are newly developed financial instruments. Key features of weather derivatives are that the underlying process, i.e., temperature or rainfall index is not tradable and cannot be replicated by other risk factors (Benth et al., 2007; Härdle and López-Cabrera, 2012; López-Cabrera et al., 2013). Consequently, the Black–Scholes formula is unsuitable since an essential element of it is the tradability of the underlying. In addition, the temperature index shows apparent seasonality and it is determined by physical phenomena. An interesting feature is that weather futures and options are rarely traded and traded only at a few strike prices compared with other more frequently traded equity markets. The CME (the official WD platform) provides closing prices, which are however not the real trading prices negotiated by the market participants. The SPD enables to price options with complicated payoff functions simply by numerical integration of the payoff with respect to this density. However, data sparsity makes the SPD estimation a statistical challenge. In addition, we study the dynamics of the SPD which provides useful insight into the economic behavior of agents sensitive to weather conditions and the time inhomogeneity of the market.

This paper is structured as follows. Section 2 describes the quadrature approach and its comparison to other popular SPD density estimation methods. Section 3 conducts the empirical analysis of SPDs from CME weather option data, studies the dynamics of the SPD weather type, and gives economic interpretations from the implied SPD. In Section 4, we address the data sparsity issue by addressing why other nonparametric methods fail particularly when options with only a few strike prices are traded. Section 5 concludes the paper. All quotations of currency in this paper will be in USD and therefore we will omit the explicit notion of the currency. All the SPDs computations were carried out in Matlab version 7.6. The option data on temperature indices were obtained from CME and are also available from the research data center of the CRC 649 "Economic Risk".

2. The Bayesian quadrature method

Options are contingent claims on an underlying asset. Plain vanilla option is of either put or call type with a fixed maturity, i.e., the value of the underlying is compared to a strike price K at maturity T. Let x denote the underlying asset's price at maturity (in our application this will be equivalent to futures prices on weather indexes). For a call option, one has the payoff $\max(x-K,0)$ and for a put $\max(K-x,0)$. If we denote a put as $i=1$ and a call with $i=2$, and observed strike prices E_{ij} for $i=1,2$ and $j=1,\ldots,N_i$ indexing all possible strike prices on any given day t, then the payoff function at maturity, denoted by $\wp_{ij}(x)$, can be represented by one formula,

$$\wp_{ij}(x) = (-1)^i(x-E_{ij})\mathbf{I}\left\{(-1)^i(x-E_{ij})>0\right\}(x),$$

where $\mathbf{I}\{A\}$ is an indicator function for a set A. Let t be the current time. The fair option price is given as (1) as the discounted value of the expected payoff function:

$$C_{ij} = \exp(-r\tau)\,\mathsf{E}^{\mathsf{Q}}[\wp_{ij}(x)],$$

where $\tau = T-t$ is the time to maturity and $\mathsf{E}^{\mathsf{Q}}[\cdot]$ is the expectation operator taken under the risk-neutral measure. The density $f(x)$ under this risk-neutral measure is the defined SPD. When the SPD $f(x)$ exists, this equals:

$$C_{ij} = \exp(-r\tau)\int \wp_{ij}(x)f(x)dx. \tag{2}$$

The left hand side of (2) is observed on the market for different payoff types depending on put/call ($i=1,2$), strike price E_{ij}, and time to maturity τ. The interest of statistical calibration is to infer the SPD $f(x)$ from a set of observed option prices.

2.1. The quadrature method

The word "quadrature" means a numerical method to approximate an integral either analytically or numerically, see Ueberhuber (1997) for example. In this research, we work the adverse way, since the interest is to infer the unknown density from the observed integrals (option prices). Define the δ-function $\delta_{\varpi}(\cdot)$ as a unit point measure at the location s by

$$\delta_s(x) = \mathbf{I}\{s=x\}.$$

The basic idea of the quadrature method is to approximate the SPD $f(x)$ by $f_N(x|w,\theta)$, a weighted sum of δ-functions:

$$f_N(x|w,\theta) = w_1\delta_{\theta_1}(x)+\cdots+w_N\delta_{\theta_N}(x), \tag{3}$$

with unknown locations $\theta = (\theta_1,\ldots,\theta_N)^\top$ and weights $w = (w_1,\ldots,w_N)^\top$. Here, N is a non-negative integer (smoothing)

parameter. To produce a legitimate probability density, the locations θ are constrained to be non-negative quantities, and the weights w are constrained to be nonnegative quantities and sum up to one. From a modeling perspective, the quadrature method (3) can be seen as a finite mixture distribution with the point measure as the component density. Fig. 1 illustrates (3) for $N = 5$.

The option price (2) under $f_N(x|w, \theta)$ is:

$$C_{ij}^N(w, \theta) = \exp(-r\tau) \sum_{n=1}^{N} w_n \wp_{ij}(\theta_n). \tag{4}$$

Note that (4) is an approximation to (2) and the aim of calibration is to extract (w, θ) by matching $C_{ij}^N(w, \theta)$ to the observed option prices. More specifically, a call option price calculated with (3) is:

$$C_{2j}^N(w, \theta) = \exp\{-r\tau\} \sum_{n=1}^{N} w_n \max(\theta_n - E_{ij}, 0), \tag{5}$$

whereas a put option price under the quadrature method is:

$$C_{1j}^N(w, \theta) = \exp\{-r\tau\} \sum_{n=1}^{N} w_n \max(E_{ij} - \theta_n, 0). \tag{6}$$

2.2. Bayesian modeling and computation

Empirical observations show that options having higher prices usually have higher price variation, see Ghysels et al. (1995) and Ghysels et al. (1997). Hence for the calibration task as a variance stabilizing transformation, we consider the logarithm of option prices. The observations y_{ijk} are perturbations of the model option price $C_{ij}^N(w, \theta)$:

$$\log y_{ijk} = \log C_{ij}^N(w, \theta) + \varepsilon_{ijk} \tag{7}$$

for $i = 1, 2$, $j = 1, \ldots, N_i$, $k = 1, \ldots, N_{ij}$, where the error $\varepsilon_{ijk} \sim \mathrm{N}(0, \sigma^2)$. ε_{ijk} is attributed to market friction and the approximation discrepancy (Garcia et al., 2010; Renault, 1997). In Section 3, residual analysis of our empirical studies will support this error assumption.

These parameters, w, θ, and σ^2, are estimated in a Bayesian framework instead of a maximum likelihood method. Following (7), the likelihood is

$$L(y|w, \theta, \sigma^2) = \prod_{i=1}^{2} \prod_{j=1}^{N_i} \prod_{k=1}^{N_{ij}} (2\pi\sigma^2)^{-\frac{1}{2}} \times \exp\left[-\frac{\left\{\log y_{ijk} - \log C_{ij}^N(w, \theta)\right\}^2}{2\sigma^2} \right]. \tag{8}$$

A natural prior distribution for the weights w is the Dirichlet distribution, which ensures w being positive and summing up to one. The Dirichlet distribution with parameter $\gamma = (\gamma_1, \ldots, \gamma_N)^\top$ has the density function,

$$f(w|\gamma) = \frac{1}{B(\gamma)} \prod_{n=1}^{N} w_n^{\gamma_n - 1} \tag{9}$$

for $w_n > 0, n = 1, \ldots, N$, and $w_1 + \cdots + w_N = 1$. The normalizing constant $B(\gamma)$ is defined as

$$B(\gamma) = \frac{\prod_{n=1}^{N} \Gamma(\gamma_n)}{\Gamma\left(\sum_{n=1}^{N} \gamma_n\right)}$$

where $\Gamma(\cdot)$ is the gamma function (Chen and Shao, 1997).

Let $K_{\min}$ and $K_{\max}$ denote the minimum and maximum of the observed strike prices E_{ij}, respectively. To avoid label switching

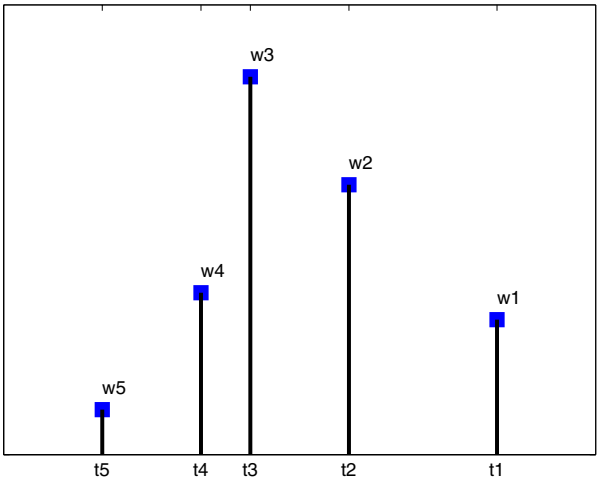

Fig. 1. The SPD $f_N(x|w, \theta)$ from (3) for $N = 5$.

problems for θ, we assume that the locations are ordered, i.e., $\theta_1 \le \cdots \le \theta_N$. Moreover, to avoid model option prices in (4) being zeros, assume a priori that the smallest location, θ_1, is less than the minimum of the observed strike price, and that the largest location, θ_N, is larger than the maximum of the observed strike prices. Therefore, we assume that the distribution of the locations θ is uniformly distributed over the set $\{\theta_1 \le \theta_2 \le \cdots \le \theta_N, \theta_1 < K_{\min},\ \theta_N > K_{\max}\}$:

$$f(\theta|K_{\min}, K_{\max}) \propto \mathbf{I}\{\theta_1 \le \cdots \le \theta_N,\ \theta_1 < K_{\min},\ \theta_N > K_{\max}\}(\theta). \tag{10}$$

For simplicity, we consider an inverse-gamma distribution with shape parameter α and scale parameter β as a prior distribution for σ^2, denoted by $\sigma^2 \sim IG(\alpha, \beta)$. The prior density of σ^2 is

$$f(\sigma^2|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} \exp\left(-\frac{\beta}{\sigma^2}\right). \tag{11}$$

Putting things together allows a conjugate prior for σ^2, as described in Casella and Berger (2001).

Note that (9)–(11) can be changed in cases where appropriate information is available. Bayesian inference for the parameters of interest is based on the posterior distribution of w, θ, and σ^2:

$$f(w, \theta, \sigma^2|y, \alpha, \beta, \gamma, K_{\min}, K_{\max}) \propto L(y|w, \theta, \sigma^2) f(w|\gamma) f(\theta|K_{\min}, K_{\max}) f(\sigma^2|\alpha, \beta). \tag{12}$$

Because of the complexity of (12), it is difficult to derive a closed-form formula for the posterior distribution (Liechty and Teng, 2009). The Markov chain Monte Carlo (MCMC) simulation is therefore used to sample w, θ, and σ^2. Because of the monotonicity of parameters w and θ in (4), an MCMC algorithm with slice samplers can be used to avoid manual tuning procedures in the MCMC simulation. In the following, we summarize major steps to run the MCMC algorithm. Let $U(A)$ denote the uniform distribution on the set A.

1. Start w, θ, and σ^2 randomly.
2. At each iteration, repeat the following steps until the samples appear to converge.
   (a) Sample $w_n \sim U(T_n)$ for $n = 1, \ldots, N-1$, where T_n is a properly derived open interval. Set $w_N = 1 - w_1 - \cdots - w_{N-1}$.
   (b) Sample $\theta_n \sim U(S_n)$ for $n = 1, \ldots, N$, where S_n is a properly derived open interval.

Table 1
The volume for HDD–CDD monthly, seasonal strips and average temperature products in each US city.

Index	City	Future					Option						
		HDD monthly	CDD monthly	HDD strips	CDD strips	Avg	HDD monthly	CDD monthly	HDD strips	CDD strips	Avg	MS (%)	Rank
1	Atlanta	49 621	35 567	14 400	2150	50	56 431	11 647	117 165	71 950	0	9.44	3
2	Baltimore	6 633	3 545	600	700	0	2 600	100	12 500	1 100	0	0.73	16
3	Boston	24 178	19 066	2 200	1150	0	11 029	550	42 174	19 450	0	3.15	13
4	Chicago	90 585	54 950	3 975	2800	0	39 676	19 300	107 616	67 725	0	10.17	2
5	Cincinnati	50 155	38 035	2 967	1700	455	29 280	28 910	73 255	74 975	0	7.89	4
6	Colorado Springs	1 936	1 450	0	0	0	15 025	8 750	0	0	0	0.71	17
7	Dallas	27 206	55 540	3 700	1961	200	13 085	39 775	47 450	94 850	0	7.47	5
8	Des Moines	40 929	30 510	3 190	1450	50	38 631	4 460	64 790	60 900	0	6.44	6
9	Detroit	2 185	351	50	50	0	0	0	0	0	0	0.07	23
10	Houston	16 901	18 229	1 400	1700	0	3 700	5 000	52 950	33 950	0	3.52	12
11	Jacksonville	100	1 600	0	0	0	0	16 575	0	0	0	0.48	19
12	Kansas City	36 513	23 145	1 325	1350	1100	11 025	7 200	45 050	33 750	0	4.22	10
13	Las Vegas	12 680	26 635	325	1650	0	3 100	14 200	34 600	76 650	0	4.47	9
14	Little Rock	120	105	0	0	0	0	12 250	0	0	0	0.33	20
15	Los Angeles	100	400	0	0	0	0	50	0	0	0	0.01	24
16	Minneapolis	50 085	27 955	2 150	1500	0	18 206	3 850	63 350	34 000	0	5.29	8
17	New York	187 264	154 605	6 700	4860	0	90 620	35 175	141 850	136 350	0	19.93	1
18	Philadelphia	16 441	34 449	2 300	2250	150	6 408	18 210	56 000	76 150	0	5.59	7
19	Portland	10 329	10 855	725	450	0	1 720	450	48 200	76 450	0	3.92	11
20	Raleigh Durham	550	1 500	0	0	0	23 700	0	0	0	0	0.68	18
21	Sacramento	6 383	23 401	550	750	0	2 850	1 675	16 200	48 000	0	2.63	14
22	Salt Lake City	739	504	150	0	0	0	0	4 500	3 500	0	0.25	22
23	Tucson	7 283	16 965	350	750	0	2 700	3 010	28 750	27 800	0	2.30	15
24	Washington	550	25	250	0	0	650	1 500	6 650	2 000	0	0.31	21

(c) Sample

$$\sigma^2 \sim IG\left(\sum_{i=1}^{2}\sum_{j=1}^{N_i}\alpha + M/2, \beta + \sum_{i=1}^{2}\sum_{j=1}^{N_i}\sum_{k=1}^{N_{ij}}(\log y_{ijk} - \log C_{ij}^{N}(w,\theta))^2\right)\Big/2,$$

where $M = \sum_{i=1}^{2}\sum_{j=1}^{N_i}\sum_{k=1}^{N_{ij}} 1$ is the number of observed options.

The derivations of open intervals T_n and S_n are rather lengthy and are hence omitted here for brevity. Please refer to Liechty and Teng (2009) for details.

2.3. Kernel smoothing density estimate of the quadrature method

The density $\hat{f}_N(x|w,\theta)$ from (3) is a weighted sum of δ functions and hence is not a continuous density. However, in many cases, it is interesting to visualize the SPD as a smoothed density. The kernel density for a set of M observed points $\vartheta = (\vartheta_1, \ldots, \vartheta_M)^\top$ is:

$$\hat{f}(x|\vartheta) = \int \hat{g}(u)K_h(x-u)du = \frac{1}{M}\sum_{m=1}^{M}K_h(x-\vartheta_m) \tag{13}$$

where $K_h(\cdot) = h^{-1}K(\cdot/h)$ is a kernel function with a bandwidth h and $\hat{g}(u)$ the sum δ-functions

$$\hat{g}(u) = M^{-1}\sum_{m=1}^{M}\delta_{\vartheta_m}(u)$$

with locations ϑ. Obviously, different values of h will change the appearance of $\hat{f}(x|\vartheta)$. Silverman's rule of thumb suggests a bandwidth

$$h_G = 1.06\hat{\sigma}M^{-1/5} \tag{14}$$

where $\hat{\sigma}$ is the sample standard deviation of ϑ and a normal kernel $K = \varphi$ the pdf of $\mathsf{N}(0,1)$ (Silverman, 1986).

Note that each ϑ_m for $m = 1, \ldots, M$ appears with equal probability $1/M$. However, in the Bayesian quadrature method, θ_n appears with probability w_n, for $n = 1, \ldots, N$. Therefore, we need to adjust the sample size and use $\hat{g}(u) = \sum_{n=1}^{N} w_n\delta_{\theta_n}(u)$ instead. The smoothed density version of (3) becomes

$$f_N^s(x|w,\theta) = \sum_{n=1}^{N} w_nK_h(x-\theta_n). \tag{15}$$

To apply Silverman's rule in the case of unequal weights in (15), we round off each w_n to the second decimal and adjust the sample size to be 100. The smoothed SPD appears to be reasonable. Indeed, it is possible to consider a more precise approximation: Round off each w_n to the q-th decimal, and set the sample size M to be 10^q. In the ith swipe of the MCMC algorithm, we obtain $w^{(i)}$ and $\theta^{(i)}$ and the smoothed SPD $f_N^s(x|w^{(i)},\theta^{(i)})$. We then report the posterior mean and 90% credible regions of the smoothed SPD based on $\hat{f}(x|w^{(i)},\theta^{(i)})$ point-wisely.

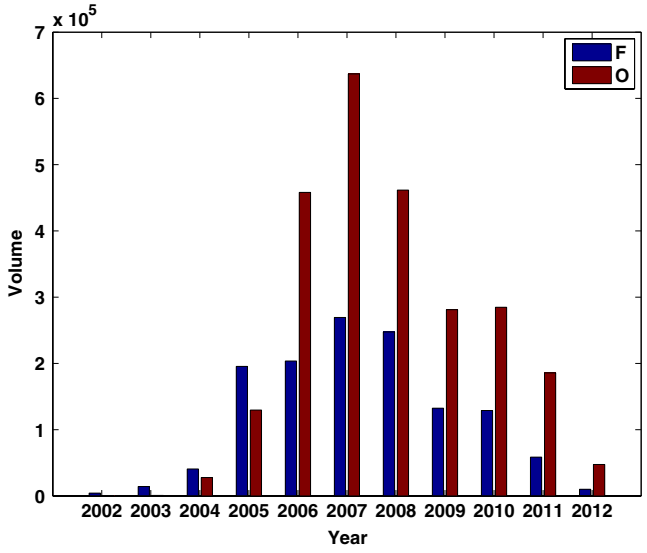

Fig. 2. The volume for US temperature futures (F) and options (O).

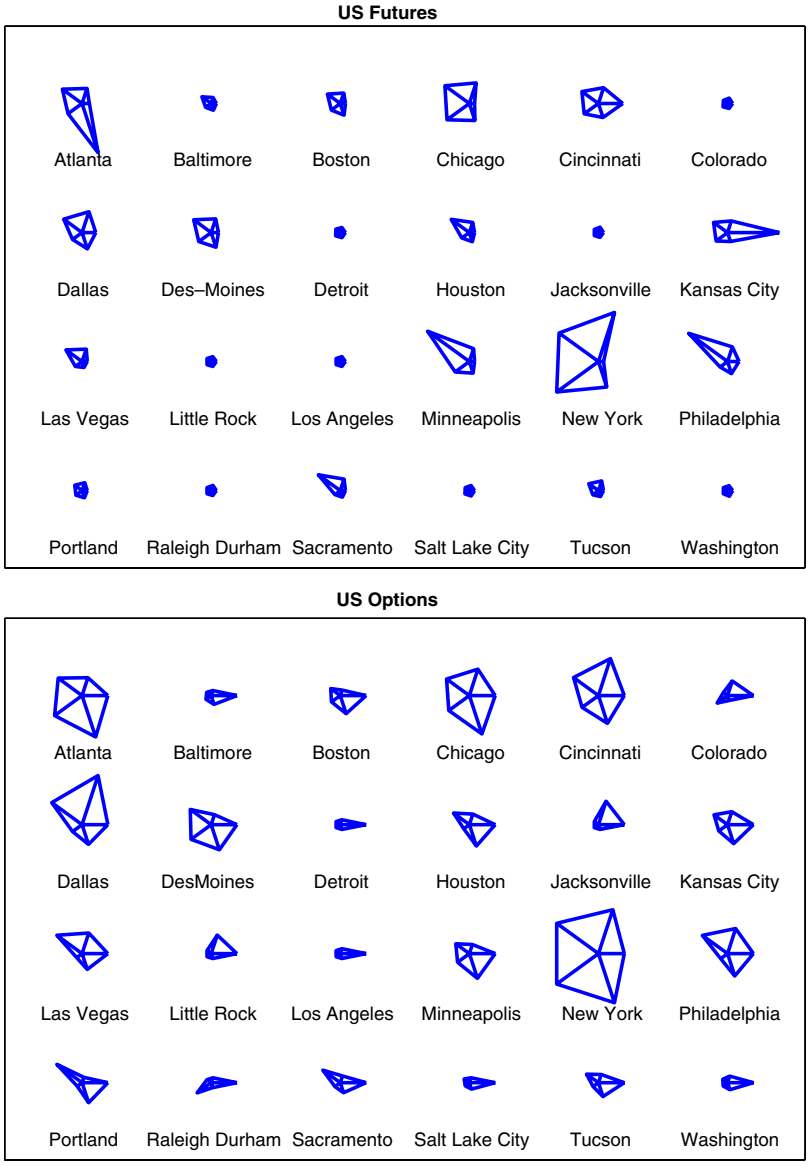

Fig. 3. Star plots representing the volume for US temperature contracts (HDD–CDD monthly, HDD–CDD seasonal strips, and weekly average) futures (upper panel) and options (lower panel) for each city. Each city is represented as a star whose ith spoke is proportional in length to the volume size of ith product (HDD Monthly, CDD Monthly, HDD Strips, CDD strips, Average) of the observed city.

As a remark, the bandwidth can be adjusted to other kernels by a canonical bandwidth, Härdle et al. (2004). For example for the quartic kernel:

$$K(u) = \frac{15}{16}(1-u^2)^2 \mathbf{I}\{|u| \leq 1\}, \tag{16}$$

Silverman's rule of thumb h_G transforms into:

$$h_{QUA} = 2.62 \cdot h_G \tag{17}$$

3. Empirical analysis

This section introduces the weather derivatives (WD) market and presents an overview on WD data. One major feature of the WD market is data sparsity, which makes most existing methods for estimating the SPD challenging and difficult. We then apply the described Bayesian quadrature technique to estimate the implied SPD on WD data, conduct an out-of-sample analysis, and study its dynamics.

3.1. Weather derivatives

WDs are financial contracts designed to hedge weather risk. The most common contracts traded at CME are based on temperature indices linked to the temperature at time t, denoted by T_t. These are the Heating Degree Days (HDD), the Cooling Degree Days (CDD), and the cumulative average temperature (CAT):

$$HDD(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} \max(c - T_t, 0)$$

$$CDD(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} \max(T_t - c, 0)$$

$$CAT(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} T_t \tag{18}$$

where c is a threshold (usually 65°F or 18 °C) and $[\tau_1, \tau_2]$ with $\tau_1 < \tau_2$ is the temperature measurement period. The standard is

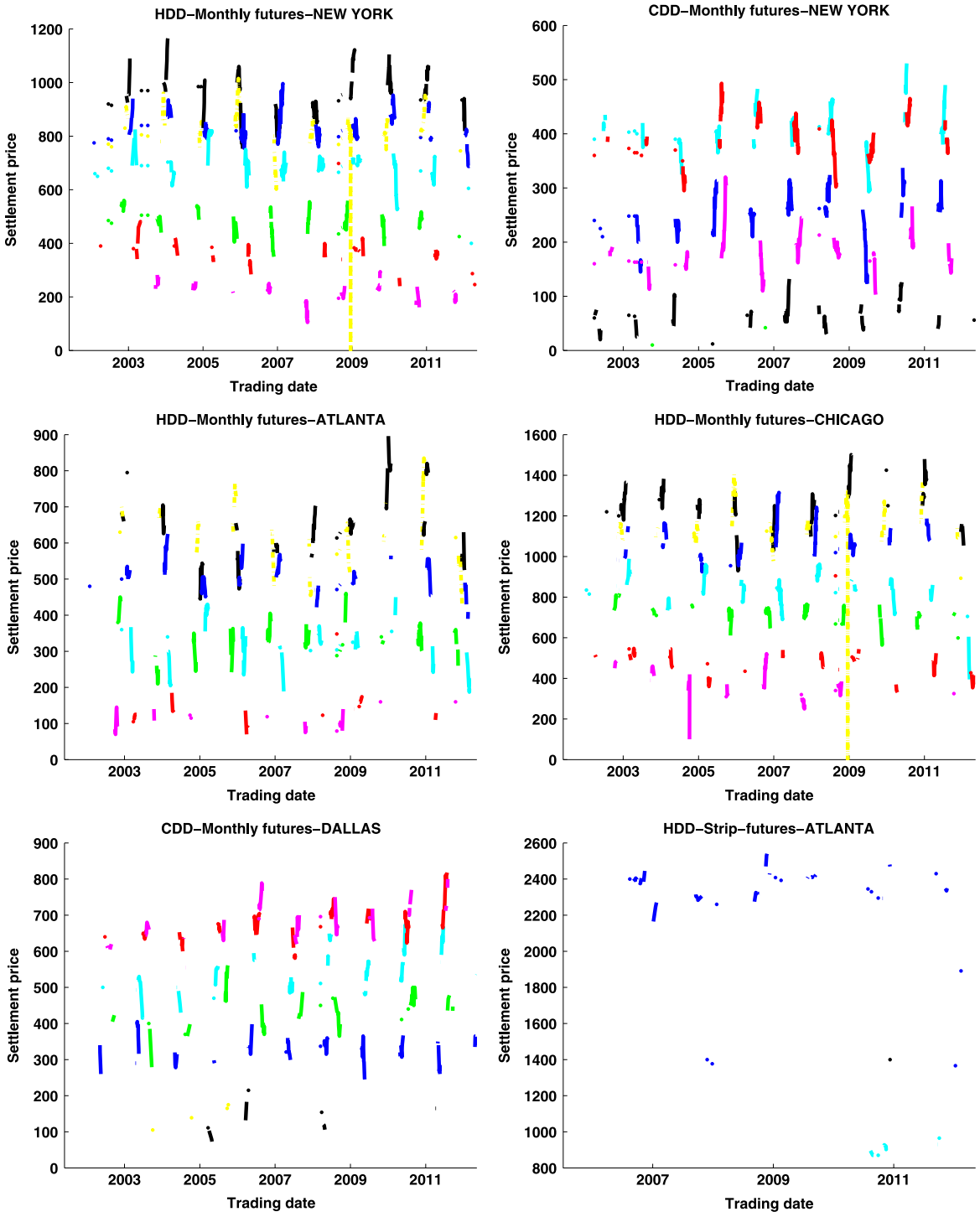

Fig. 4. Time series plots of New York, Atlanta, Chicago, Dallas HDD/CDD monthly and seasonal strips futures prices. HDD monthly futures with the measurement period of January (Black), February (Blue), March (Cyan), September (Red), October (Magenta), November (Yellow) and December (Green). CDD monthly futures with the measurement period of May (Black), June (Blue), July (Cyan), August (Red), and September (Magenta). HDD seasonal strip with the measurement period in January (Blue), March(Cyan) and December (Cyan). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that $[\tau_1, \tau_2]$ denotes a month of the year or as seasonal strips. The futures in question are delivering over a period $[\tau_1, \tau_2]$, and not at a fixed delivery time τ. The HDD index measures the demand for heating, while the CDD index measures the demand for cooling. Consequently, temperature indices are the underlying and not the temperature by itself.

Financial mathematical tools given in Benth et al. (2007, 2011) and Härdle and López-Cabrera (2012) allow the pricing of the non-tradable underlying by risk adjusted conditional expectation. Hereby, the futures temperature contract price on the sum of temperature $I(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} T_t$ with accumulation period $[\tau_1, \tau_2]$ is given by:

$$F(t, \tau_1, \tau_2) = \mathrm{E}^{\mathrm{Q}}\left[I(T_{\tau_1,\tau_2})|\mathcal{F}_t\right] \tag{19}$$

where $\mathrm{E}^{\mathrm{Q}}[\cdot]$ is any equivalent martingale measure and $\mathcal{F}_t$ is a filtration information set.

Table 2
The number of transactions—trading days (TD) and volume (vol) of New York/Atlanta/Chicago/Dallas HDD and CDD monthly and HDD seasonal strip options with respect to time to maturity (τ) in month and the number of strike prices.

		HDD—New York					HDD—Atlanta						HDD—Chicago					
Number of strike prices																		
τ		1	2	3	4	Total	1	2	3	4	5	Total	1	2	3	4	Total
≤1	TD	71	23	7	1	102	56	12	4	1	1	74	50	10	–	–	60
	vol	17 495	12 650	9900	1400	41 445	12 861	4700	2950	700	1250	22 461	10 961	4975	–	–	15 936
(1, 2]	TD	54	26	3	4	87	39	26	2	1	–	68	32	13	2	2	49
	vol	12 450	21 700	1075	5400	40 625	10 245	19 825	2800	1000	–	33 870	50	2000	–	–	2 050
(2, 3]	TD	3	1	–	–	4	1	–	–	–	–	1	2	–	–	–	2
	vol	1 000	1 000	–	–	2 000	100	–	–	–	–	100	2 000	–	–	–	2 000
(3, 4]	TD	2	1	–	–	3	–	–	–	–	–	–	1	–	–	–	1
	vol	300	2 000	–	–	2 300	–	–	–	–	–	–	2 000	–	–	–	2 000
(4, 5]	TD	1	1	–	–	2	–	–	–		–	–	1	–	–	–	1
	vol	250	2 000	–	–	2 250	–	–	–	–	–	–	2 000	–	–	–	2 000
(5, 6]	TD	–	1	–	–	1	–	–	–		–	–	1	–	–	–	1
	vol	–	2 000	–	–	2 000	–	–	–	–	–	–	2 000	–	–	–	2 000

		CDD—New York					HDD strips—Atlanta								CDD—Dallas				
Number of strike prices																			
τ		1	2	3	4	Total	1	2	3	4	5	6	7	8	Total	1	2	3	4	Total
≤1	TD	43	3	1	–	47	1	6	–	6	–	–	–	–	13	40	–	90	1	131
	vol	17 425	2000	600	–	20 025	200	3 100	–	2700	–	–	–	–	6 000	12 400	9150	1250	–	
(1, 2]	TD	34	13	2	–	49	6	6	1	–	–	–	–	–	13	30	–	10	1	41
	vol	8 200	5750	1200	–	15 150	4700	6 250	1875	–	–	–	–	–	–	9 125	5300	750	–	
(2, 3]	TD	–	–	–	–	–	3	9	–	3	–	–	–	–	15	–	1	–	–	1
	vol	–	–	–	–	–	2240	7 500	–	5500	–	–	–	–	–	–	450	–	–	450
(3, 4]	TD	–	–	–	–	–	–	9	–		–	–	–	–	9	–	1	–	–	1
	vol	–	–	–	–	–	–	10 400	–		–	–	–	–	10 400	–	450	–	–	450
(4, 5]	TD	–	–	–	–	–	1	11	1	4	–	1	1	–	19	–	1	–	–	1
	vol	–	–	–	–	–	1750	10 500	1000	9500	–	6500	6500	–		–	450	–	–	450
(5, 6]	TD	–	–	–	–	–	1	11	–	3	–	–	–	1–	16	–	–	–	–	–
	vol	–	–	–	–	–	250	11 700	–	6750	–	–	–	4250		–	–	–	–	–
(6, 7]	TD	–	–	–	–	–	10	–	–	–	–	–	–	–	1	–	–	–	–	–
	vol	–	–	–	–	–	9000	–	–	–	–	–	–	–	9 000	–	–	–	–	–
(7, 8]	TD	–	–	–	–	–	3	–	–	–	–	–	–	–	3	–	–	–	–	–
	vol	–	–	–	–	–	6000	–	–	–	–	–	–	–	6 000	–	–	–	–	–

Consequently, the European temperature call option price written on the futures price is defined as:

$$C(K) = \exp\{-r\tau\} \int \max\{F(t, \tau_1, \tau_2) - K, 0\} f(x)dx. \tag{20}$$

In order to compute (19), (20), it is necessary to know the stochastic properties of the temperature process T_t under the "physical measure" P and then adjust the risk measure Q, see Härdle and López-Cabrera (2012). In other words, the temperature derivative price is given by finding a model for the daily weather process consisting of a trend, a seasonality, an autoregressive part, seasonal variance and normally distributed residuals. Then one could specify a class of probability measures using the Radon–Nikodym derivative determined by the Esscher transform, see López-Cabrera et al. (2013). Another way is to model the index directly, see Dorfleitner and Wimmer (2010).

Here we estimate the SPD, different to the afore mentioned approach, directly under the risk neutral measure Q from real option data. Note that (20) is exactly (5) for $f = f_N$.

The options at CME are cash settled, i.e., the owner of a future receives 20 times the Degree Day Index at the end of the measurement period, in return for a fixed price. At time t, CME trades contracts with different measurement periods $t \leq \tau_1 < \tau_2$ or different maturities $\tau = \tau_2 - t$. The measurement period for CAT/HDD futures is typically during April–November, while CDD futures are measured during November–April.

3.2. Overview on the WD data

The WD data was purchased from CME for the study period from 2002/01/02 to 2012/05/11. The reported price is the settlement price for the future or option contract, and the volume is the number of contracts traded.

Depending on the measurement period, temperature products in the US market are further categorized into monthly, seasonal strips, and average products. HDD monthly products have seven contract months: October, November, December, January, February, March, April, and CDD monthly products have seven contract months: April, May, June, July, August, September, and October. For HDD seasonal strips, the contract period covers from October to April, and for CDD seasonal strips, the contract period covers from April to October. Contract for weekly average products covers all five weeks. Table 1 gives an overview of the volume of the temperature market.

Fig. 2 illustrates the volume for US temperature futures and options in the study period. The trading activity increased dramatically since 2002 but declined after the 2008 financial crisis. This is surprising since one could expect that these markets are uncorrelated with financial markets. However we believe that the decline is because the temperature market is not yet well known as a intermediary for diversification of weather risk. Star plots in Fig. 3 divide the volume into HDD–CDD monthly, HDD–CDD seasonal strips, and weekly average for futures and options for each US city. A star plot represents each city as a star whose ith spoke is proportional in length to the volume size of ith product (HDD Monthly, CDD Monthly, HDD Strips, CDD strips, Average) of the observed city. Clearly, monthly products are the most popular traded products, followed by seasonal strips. Nevertheless, no weekly average products are really traded in the US temperature market.

The volumes of HDD, CDD, Average monthly and seasonal strips futures and options for all US cities are reported in Table 1. New

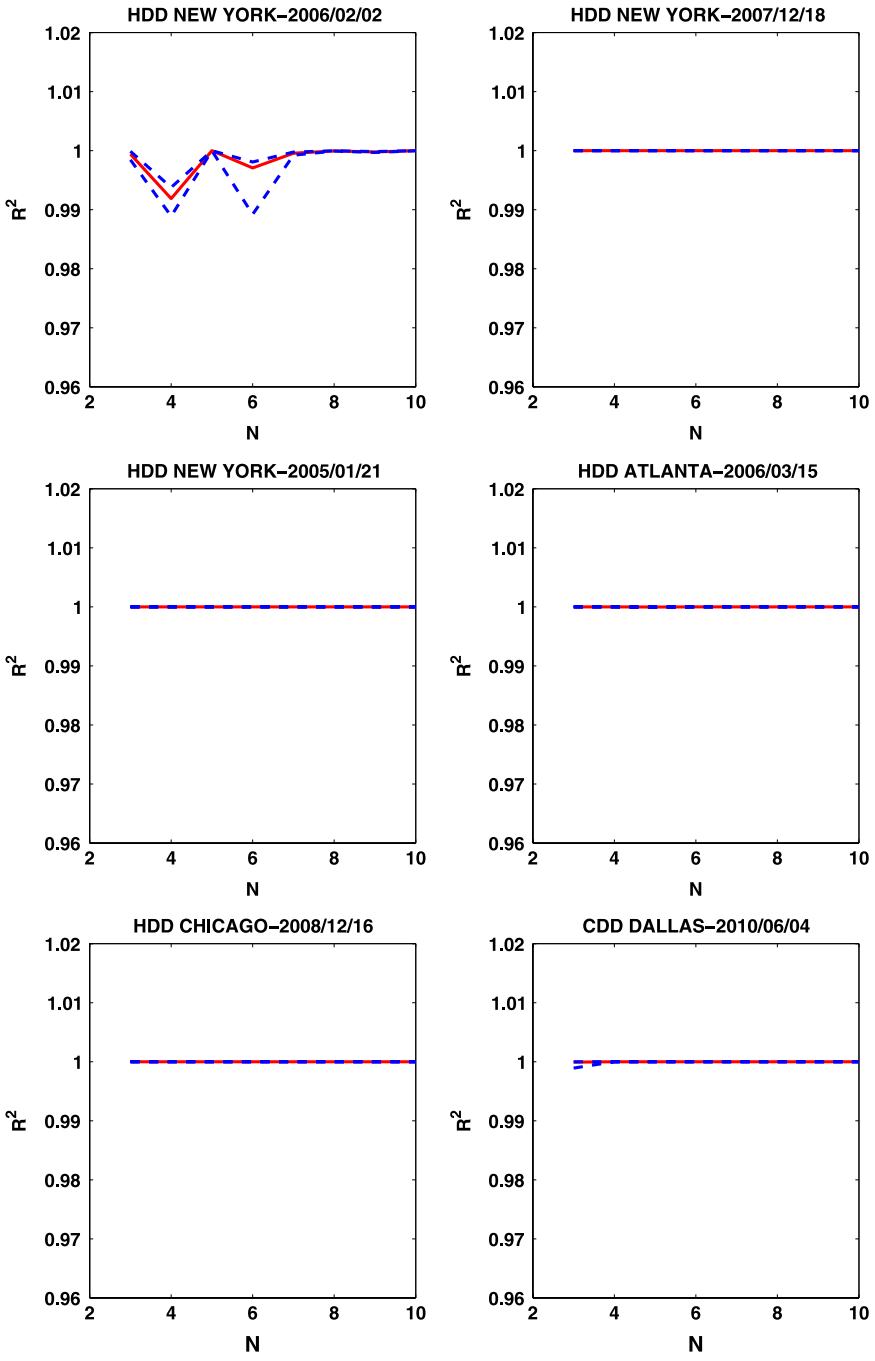

Fig. 5. The 90% credible regions (in blue dashed lines) and posterior means (in red lines) of R^2 when fitting New York/Atlanta/Chicago/Dallas HDD–CDD monthly options using the Bayesian quadrature approach against different number of support points N. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

York is to be the biggest temperature market and takes about 20% of the market volume, followed by Chicago (10%), Atlanta (9%), Cincinnati (8%), and Dallas (7%). The market share of these five cities exceeds 50% of the US temperature market. Following this, we took these cities as the most representative cities. Fig. 4 gives time series plots for New York, Atlanta, Chicago and Dallas monthly HDD, CDD monthly and seasonal strip future prices. The futures market is more liquid but also more volatile than option prices. In addition, most HDD and CDD futures are traded only with time to maturities less than a year. These features of future prices make the pricing mechanism for weather derivatives unique and challenging.

We further divide the volume of HDD and CDD monthly and HDD seasonal strip options with respect to strike prices and time to maturities, as summarized in Table 2. It is shown that most options are traded with only a few number of strike prices and of a short time to maturity (within one month and less than a year). Because of the fact that options are only traded with a few number of strike prices, this data sparsity problem makes most existing nonparametric methods (such as mixture of lognormal models or kernel methods) very difficult.

3.3. Implementation of the technique

As depicted in Fig. 8, we calibrate the SPD for HDD–CDD monthly and Seasonal strip options. These four plots present a typical pattern of option prices of weather options: options were

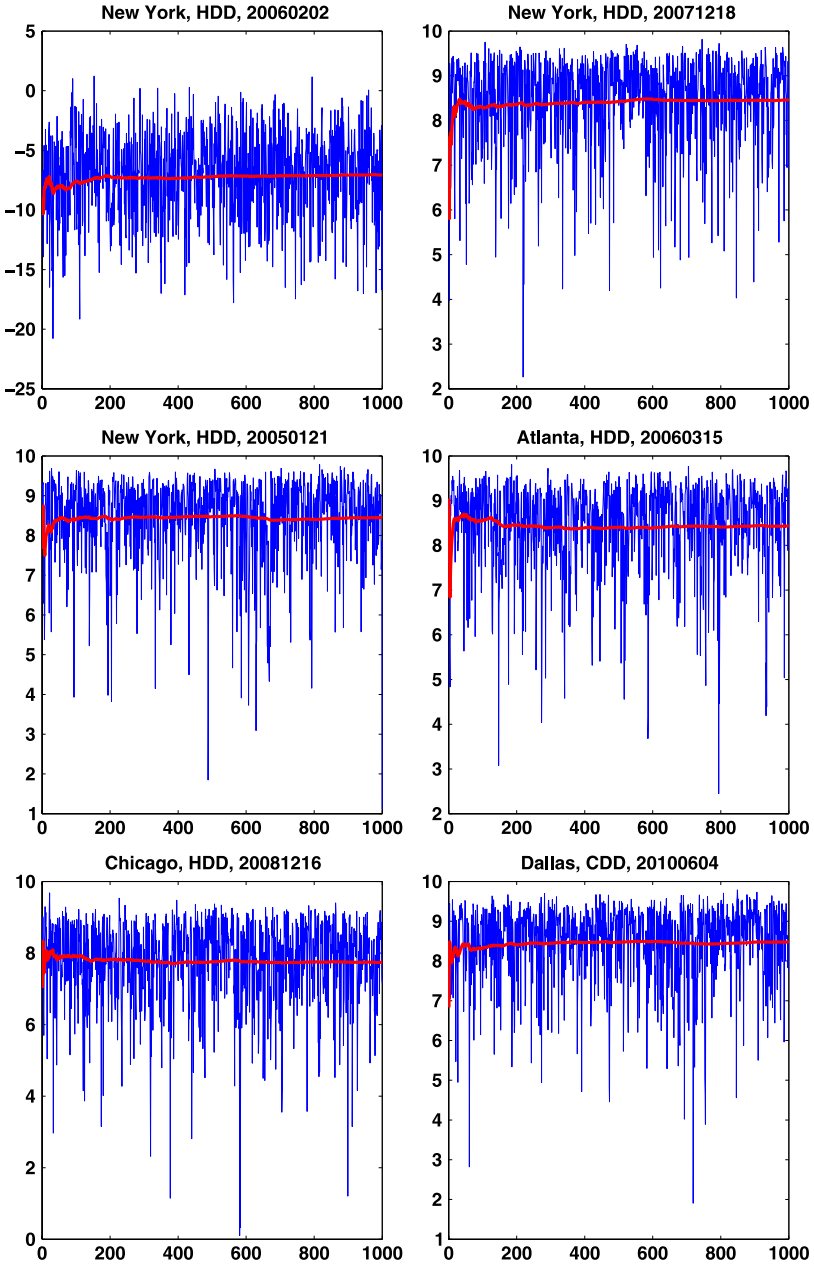

Fig. 6. Trace plots (in blue lines) and cumulative averages (in red lines) of the log-likelihood in the MCMC algorithm of different HDD/CDD monthly products. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

traded only with a very few number of strike prices, sometimes only call options or put options were traded or the both of them. The Bayesian quadrature method allows us to incorporate prior assumptions on the model parameters and hence avoid problems with data sparsity. It is able to compute the SPD of both call and put options simultaneously, and is particularly robust.

There is a trade-off in the selection of N. When N is larger, one produces better fit because there exist more free parameters in the model, but drawbacks of model complexity and computational demanding come along with. We provide more information on the sensitivity of the Bayesian quadrature approach with respect to N: Fig. 5 depicts posterior means and 90% credible regions of R^2 in fitting HDD/CDD options using the Bayesian quadrature method versus different number of support points N from three to ten. As shown in Fig. 5 that all R^2's are close to one, we conclude that the Bayesian quadrature approach provides good model fit with small N. Based on Fig. 5, we select $N = 5$ in our following analysis because it gives a simple model yet producing good model fit.

To calibrate (5) and (6), we implement an MCMC algorithm to explore the posterior distribution in (12). Because of the employment of unequal weights and the adoption of a Bayesian framework, inferring these parameters is computationally challenging. For this reason, we use an MCMC algorithm with slice samplers for making statistical inference. In our analysis, we discard the first 500 iterates in the MCMC algorithm (the burn-in period), and use the following 1000 iterates. Trace plots of the loglikelihood in Fig. 6 show that the MCMC algorithm appears to converge very fast, and autocorrelation plots of the loglikelihood in Fig. 7 indicate that samples in the MCMC algorithm are efficient.

Fig. 8 imposes model prices of the last swipe of the MCMC algorithm and demonstrates the fit, because model prices are close to market prices.

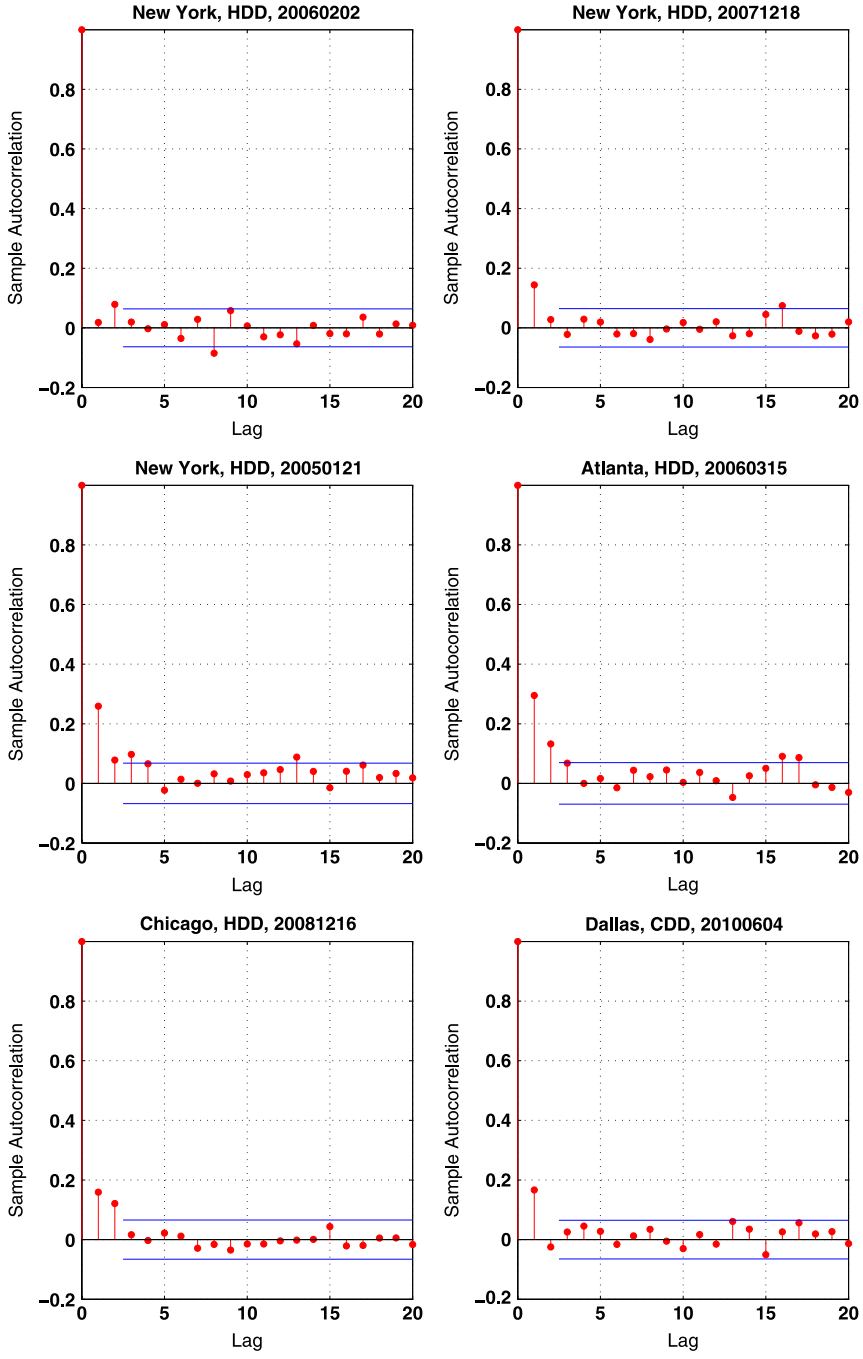

Fig. 7. Autocorrelation plots of the log-likelihood in the MCMC algorithm of different HDD/CDD monthly products.

For residual analysis, we calculate the residual at each swipe of the MCMC algorithm by

$$r_{ijk} = \log y_{ijk} - \log C_{ij}^{N}(w, \theta) \tag{21}$$

and provide kernel smooth density plots of the posterior distribution of these residuals in Fig. 9. All these four panel plots demonstrate that the residuals have mean zero, and are symmetric about zero when comparing with the normal KDE. This visual presentation supports our error assumption as a normal distribution in (8).

The density (3) approximates the SPD by weighted sum of δ functions and is discontinuous by its nature. As described earlier, to produce a smoothed SPD for visualization, we round off each w_n to the second decimal, and set the adjusted sample size as 100. Then we employ the kernel density estimation with a Gaussian kernel $K(\cdot) = \varphi(\cdot)$ and a bandwidth selected using the rule of thumb in (14) to calculate a smoothed SPD at each swipe in the MCMC algorithm.

Thus, it is clear that the smoothed density version (15) becomes:

$$\begin{aligned} f_N^s(x|w, \theta) &= \sum_{n=1}^{N} w_n K_h(x - \theta_n) \\ &= \sum_{n=1}^{N} w_n \frac{1}{h} \varphi\left(\frac{x - \theta_n}{h}\right) \\ &= \sum_{n=1}^{N} w_n \varphi(x; \theta_n, h) \end{aligned}$$

where $\varphi(x; \theta_n, h)$ is the pdf of $\mathsf{N}(\theta_n, h^2)$ distribution.

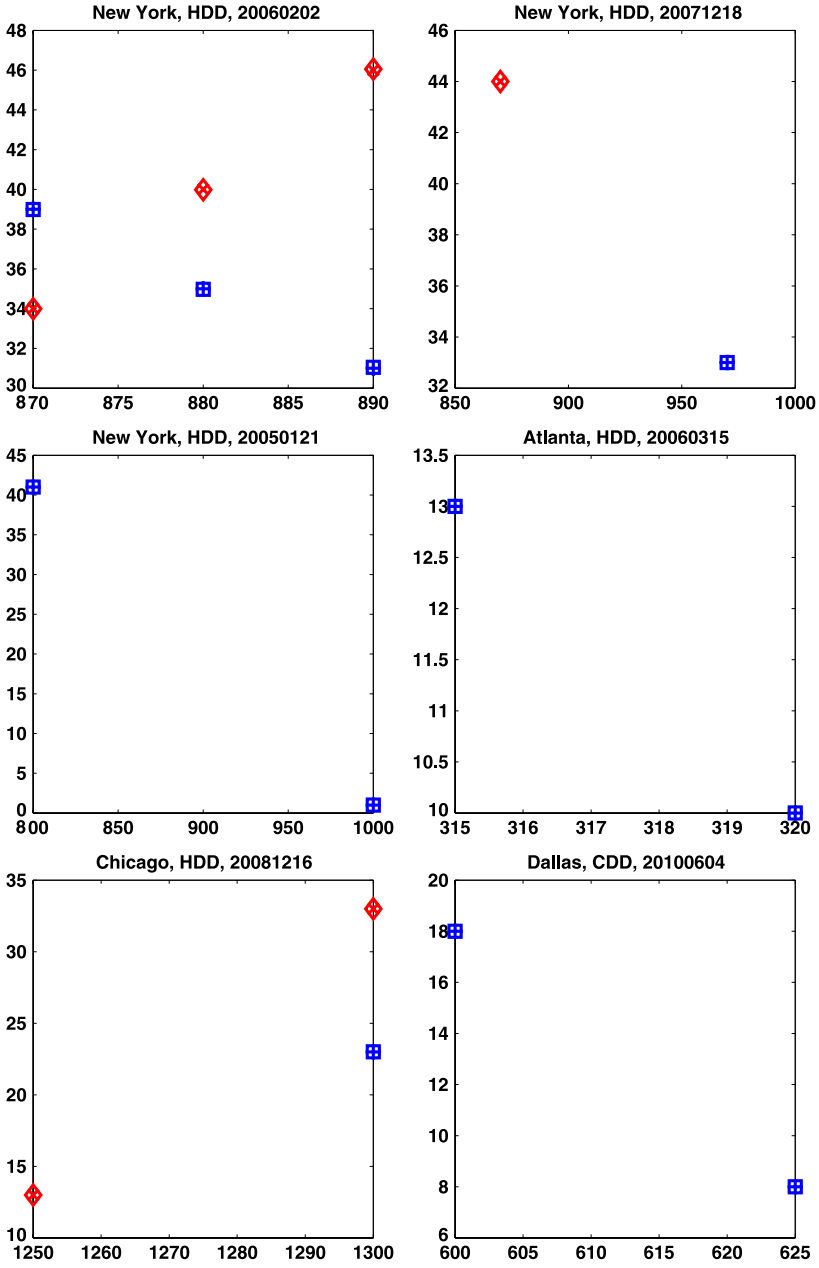

Fig. 8. Plots of market prices of New York/Atlanta/Chicago/Dallas HDD–CDD monthly options and model prices of the last swipe in the MCMC algorithm using the Bayesian quadrature method with $N = 5$. For market prices, a call option is indicated with a blue plus and a put is indicated with a red cross. For model prices, a call option is indicated with a blue diamond, and a put option is indicated with a red square. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Collecting these smoothed SPD, Fig. 10 gives the posterior mean (red line) and 90% credible regions (blue dotted lines) of the implied SPD. The right-upper and left-lower pictures show that the 90% credible regions are tight to the posterior mean of the smoothed SPD, whereas the other pictures depict that the 90% credible regions are wide. The cluster of star points in the horizontal axis denotes the future prices.

In Bayesian analysis, the 90% credible region for the smoothed SPD provides a region where 90% of the posterior distribution of the smoothed SPD will fall into. In the case of HDD New York options with maturity in 2 months traded at 20050121 and the case of CDD Dallas Option one month to maturity traded at 20100604, the left tail of 90% credible region appears to be extremely wide. This feature is not surprising though, because the data set for calibration consists of call options with only two strike prices, namely, 800 and 1000 and 600 and 625 respectively. Indeed, a call option price is simply the expected future price larger than the strike price under the SPD. Therefore, an option price only provides information for the right tail of the SPD. Once a few quadrature points in the right tail have achieved a high likelihood, points of the quadrature in the left tail (in this case, smaller than 800) do not affect the likelihood. As a result, these points are influenced only by its prior distributions. The prior assumptions in (9) and (10) put simply vague information for the weights and locations in the quadrature method. Such an assumption allows points in the left tail of the quadrature method moving freely, and causes a wider credible region in the left tail, as demonstrated in the left-upper panel in Fig. 10.

Similarly, for the right-lower panel, the 90% credible region is wider around strike 1000 but is tight in two side tails. This is because the data set consists of one call with strike 970 and one put with strike 870. As a result, the call option price gives

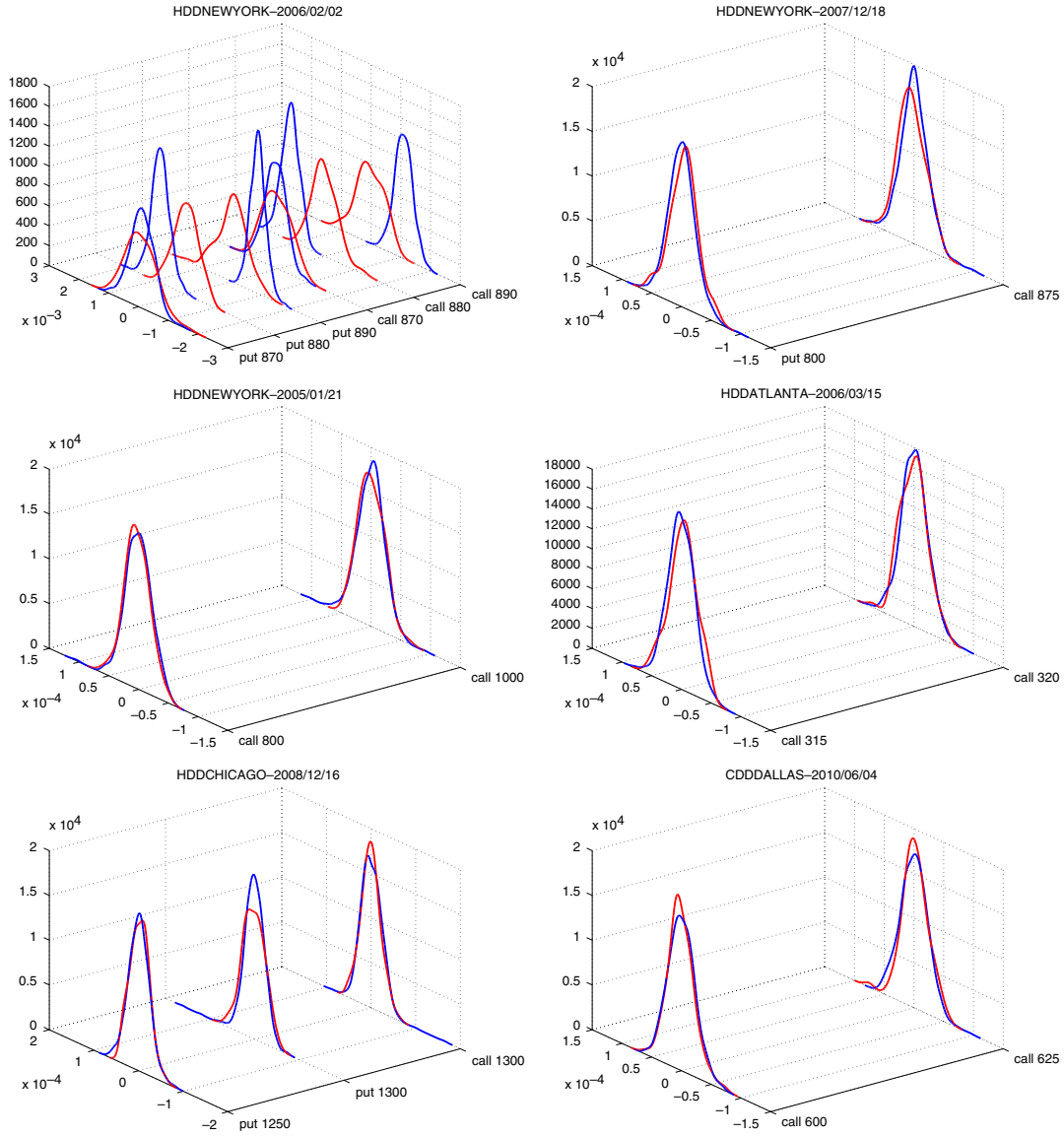

Fig. 9. Kernel density estimate (KDE) plots for the posterior distribution of residuals for HDDs and CDDs products (in blue lines) versus the normal KDE plots (in red lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

information of the right tail of the SPD, whereas the put option price gives information of the left tail of the SPD. When some points of both right and left tails in the quadrature method have achieved a high likelihood, points of the quadrature around 920 would not affect the likelihood. These points are determined by their prior assumptions again, and provide a wider credible region around 920.

Selecting prior distributions for the quadrature method is critical. In this research, we choose vague prior assumptions for the parameters and the analysis successfully reveals the fact that the width of the 90% credible region depends highly on the information provided by option prices and the prior assumptions on the parameters in the quadrature method. One may adopt more sophisticated prior distributions based on experience and knowledge. This flexibility may be considered as a technical advantage of the Bayesian quadrature method.

3.4. Out-of-sample analysis

In incomplete markets, no unique martingale measure exists. As a consequence, this may have a negative effect that the parameters estimated fit well the in-sample data, but they are inaccurate with out-of-sample data. In the following, we modify the quadrature model to forecast out-of-sample data and provide an empirical analysis confirming that our quadrature model preforms well for the out-of-sample data.

Recall that t is the current time, τ is the time to maturity, and $F(t, \tau_1, \tau_2)$ is the underlying temperature index future price at time t with accumulation period $[\tau_1, \tau_2]$. Let $\tilde{t}$ be the forecast time and $\tilde{\tau}$ be the time to maturity from time $\tilde{t}$ to maturity. Similar to Dumas et al. (1998) and Fan and Mancini (2009), we consider a one-week-ahead forecast horizon. Both $F(t, \tau_1, \tau_2)$ and $F(\tilde{t}, \tau_1, \tau_2)$ are known in our out-of-sample analysis.

To begin with, the option prices at time t are used to calibrate the Bayesian quadrature model, where w and θ are parameters. To provide a plausible yet simple quadrature model for forecast at time $\tilde{t}$ with parameters $\tilde{w}$ and $\tilde{\theta}$, we first set $\tilde{w} = w$. Recall that r is the risk-free interest rate. To adjust the current underlying temperature index future price and the discounted factor, define $\bar{\theta} = (\bar{\theta}_1, \ldots, \bar{\theta}_N)'$ as the normalized location parameters extracted from the quadrature model by

$$\theta_i = F(t, \tau_1, \tau_2)\mathrm{e}^{r\tau}\bar{\theta}_i,$$

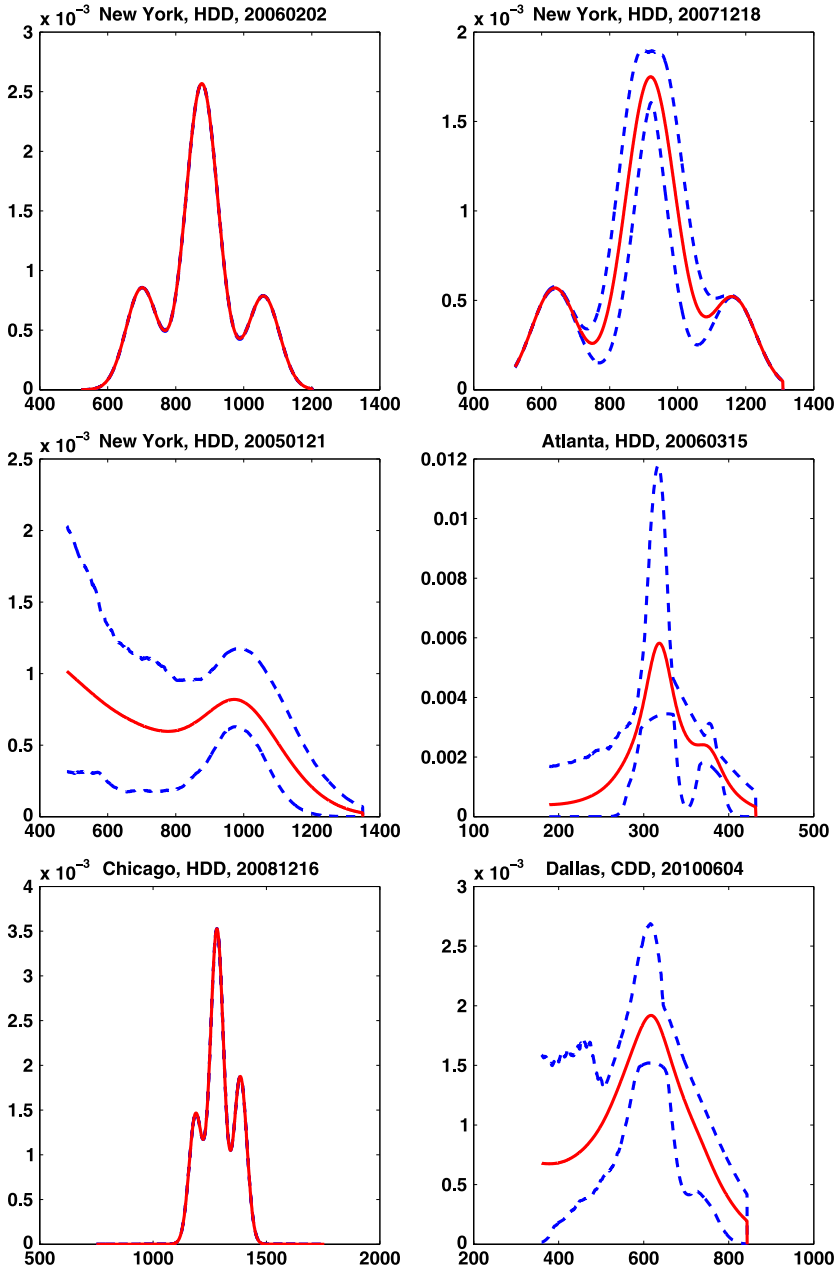

Fig. 10. Posterior means (in red solid lines) and the 90% credible regions (in blue dashed lines) of the smoothed SPD implied from New York HDD monthly options with respect to trading dates and time to maturity. The cluster of star points in the horizontal axis denotes the future prices. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

for $i = 1, \ldots, N$. Likewise, $\tilde{\theta}$ is linked to θ via the normalized location parameters by

$$\tilde{\theta}_i = F(\tilde{t}, \tau_1, \tau_2)\mathrm{e}^{r\tilde{\tau}}\bar{\theta}_i = \frac{F(\tilde{t}, \tau_1, \tau_2)}{F(t, \tau_1, \tau_2)}\mathrm{e}^{r(\tilde{\tau}-\tau)}\theta_i, \tag{22}$$

for $i = 1, \ldots, N$. As a result, the forecast quadrature model at time t_1 is

$$w_1\delta_{\tilde{\theta}_1}(x) + \cdots + w_N\delta_{\tilde{\theta}_N}(x),$$

where $\tilde{\theta}$ is given in (22), and can be used to calculate option prices forecast at time $\tilde{t}$ directly.

In the Bayesian out-of-sample analysis, we report the 90% Bayesian prediction intervals for option prices at time $\tilde{t}$ and averaged R^2 in the MCMC algorithm. Table 3 provides detailed information on the data used for the out-of-sample analysis: For each case of weather derivatives, Table 3 lists its current time t, forecast time $\tilde{t}$, maturity, underlying temperature index future prices at time t and $\tilde{t}$, and averaged R^2 in the MCMC algorithm. We remark that because the weather derivatives markets are less frequently traded, given t, if there is no settlement prices in the one-week-ahead horizon, we use options traded nearest to the one-week-ahead horizon as the out-of-sample data.

Fig. 11 depicts the 90% Bayesian prediction intervals for the forecast and realized market option prices traded at time $\tilde{t}$. It is shown that realized market option prices are within or close to the 90% Bayesian prediction intervals. Together with the numerical results that the averaged R^2 in Table 3 ranges from 0.82 to 0.99, we conclude that our quadrature method empirically performs well in this out-of-sample analysis.

Table 3
Data and averaged R^2 in the out-of-sample analysis. This table lists the type of weather derivatives, current time t, forecast time $\tilde{t}$, maturity, underlying temperature index future prices at times t and $\tilde{t}$, and averaged R^2 in the MCMC algorithm.

Product	t	$\tilde{t}$	Maturity	$F(t, \tau_1, \tau_2)$	$F(\tilde{t}, \tau_1, \tau_2)$	Averaged R^2
New York, HDD	20060202	20060207	Feb. 2006	875	778	0.91
New York, HDD	20071218	20071224	Jan. 2008	905	910	0.87
New York, HDD	20050121	20050126	Feb. 2005	805	845	0.89
Atlanta, HDD	20060315	20060322	Mar. 2006	320	342	0.82
Chicago, HDD	20081216	20081223	Dec. 2008	1290	1315	0.99
Dallas, CDD	20100604	20100611	Jun. 2010	600	617	0.94

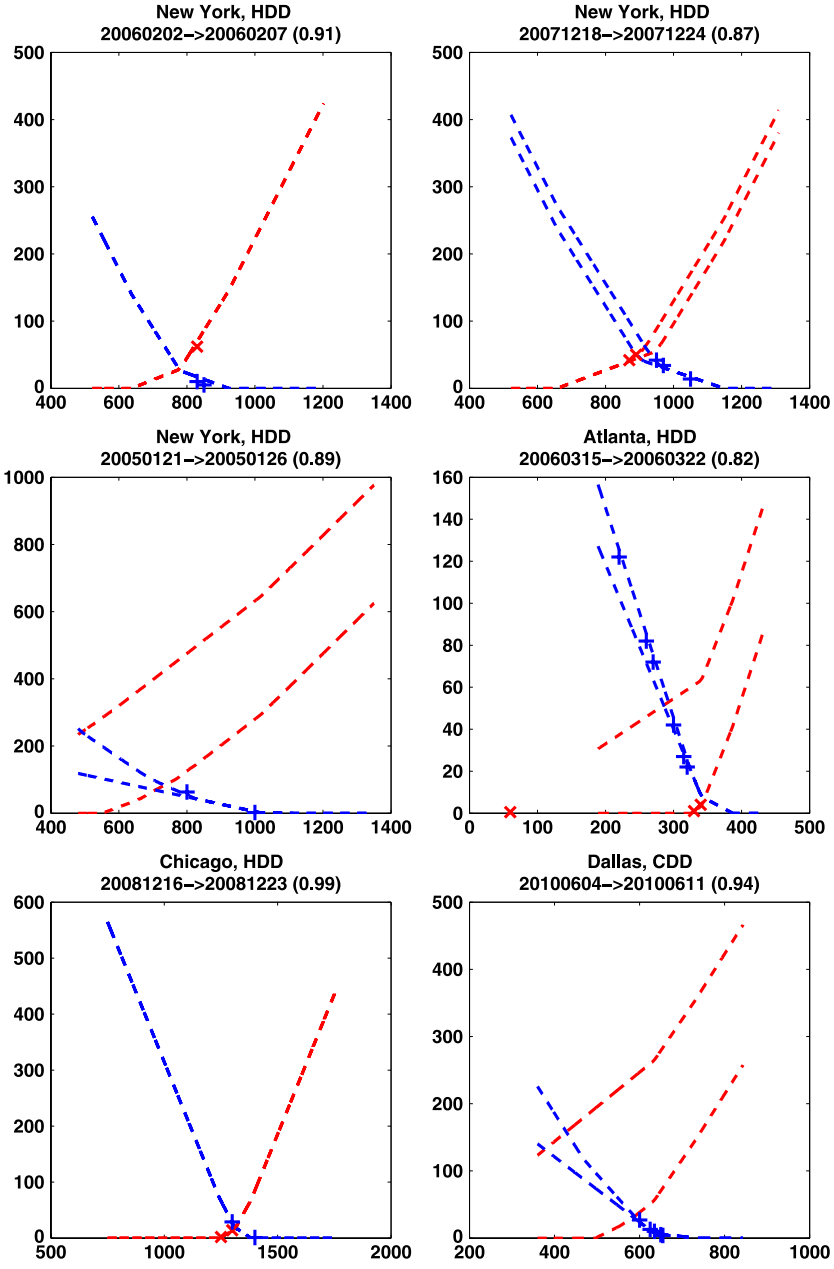

Fig. 11. Out-of-sample analysis for different contracts. For realized market prices, a call option is indicated with a blue plus and a put is indicated with a red cross. The 90% Bayesian prediction intervals of call and put option are in blue and red dashed lines, respectively. Averaged R^2 is recorded in parentheses in the title of each panel plot. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.5. Dynamics of SPD

Table 4 records the number of trading days with respect to the number of strike prices and trading months for the New York HDD/CDD monthly options with time to maturity τ less than one month and Atlanta HDD Seasonal Strips with $\tau = 6$. Again, the data sparsity remains an issue in the biggest weather market, the New York market. For illustration, we implement (4) for every trading day in March 2006. For other months, most of the number of strike in each trading day is simply one, and such case makes the SPD estimation very difficult in the sense that option price only provides information for one side of the SPD.

Table 4
Number of trading days with respect to the trading month and the number of strike prices for HDD/CDD monthly options with time to maturity less τ than one month and Atlanta HDD Seasonal Strips with $\tau = 6$ (time of measurement period of 5 months).

Year	Type	City	Month	Nr. of strike prices							Year	Type	City	Month	Nr. of strike prices							
				τ	1	2	3	4	5	Total					τ	1	2	3	4	5	6	Total
2002	HDD	NY	11	1	1	–	–	–	–	1	2006	HDD-Strip	Atlanta	10	6	–	–	1				1
2004	HDD	NY	1	1	2	–	–	–	–	2	2007	HDD-Strip	Atlanta	10	6	–	–	3	1	2	1	1	8
2004	HDD	NY	2	1	1	–	–	–	–	1	2008	HDD-Strip	Atlanta	10	6	–	–	4		1	1	1	5
2004	HDD	NY	3	1	1	–	–	–	–	1	2009	HDD-Strip	Atlanta	10	6	–	–	1		1			3
2005	HDD	NY	1	1	1	–	–	–	–	1	2010	HDD-Strip	Atlanta	10	6	–	1 –	2					2
2005	HDD	NY	2	1	2	–	–	–	–	2	2002	HDD	Chicago	11	1	1	–						1
2005	HDD	NY	3	1	1	–	–	–	–	1	2002	HDD	Chicago	12	1	2	–						2
2005	HDD	NY	12	1	1	1	–	–	–	2	2004	HDD	Chicago	2	1	1	–						1
2006	HDD	NY	1	1	2		–	–	–	2	2005	HDD	Chicago	2	1	2	–						2
2006	HDD	NY	2	1	5	3	1	–	–	9	2005	HDD	Chicago	12	1	4	–	–					4
2006	HDD	NY	3	1	3	5	4	1	–	13	2006	HDD	Chicago	3	1	5	2						7
2006	HDD	NY	10	1	1	1	–	–	–	2	2006	HDD	Chicago	10	1	2	1						3
2006	HDD	NY	11	1	1	2	–	–	–	3	2006	HDD	Chicago	11	1	2							2
2006	HDD	NY	12	1		1	–	–	–	1	2007	HDD	Chicago	1	1		3						3
2007	HDD	NY	1	1	2	–	–	–	–	2	2007	HDD	Chicago	2	1		1						1
2007	HDD	NY	2	1	4	–	–	–	–	4	2007	HDD	Chicago	3	1	1							1
2007	HDD	NY	3	1	1	–	–	–	–	1	2007	HDD	Chicago	12	1		1						1
2007	HDD	NY	11	1	1	–	–	–	–	1	2008	HDD	Chicago	1	1		1						1
2007	HDD	NY	12	1	1	2	–	–	–	3	2008	HDD	Chicago	2	1	3							3
2008	HDD	NY	1	1	6	1	2	–	–	9	2008	HDD	Chicago	3	1	2	1						3
2008	HDD	NY	2	1	3	–	–	–	–	3	2008	HDD	Chicago	12	1	2	2						4
2008	HDD	NY	12	1	2	1	–	–	–	3	2009	HDD	Chicago	1	1	2							2
2009	HDD	NY	1	1	4	–	–	–	–	4	2009	HDD	Chicago	12	1	1							1
2009	HDD	NY	2	1		2	–	–	–	2	2010	HDD	Chicago	3	1	1							1
2009	HDD	NY	3	1		1	–	–	–	1	2010	HDD	Chicago	11	1	1							1
2009	HDD	NY	11	1	3	–	–	–	–	3	2010	HDD	Chicago	12	1	2							2
2009	HDD	NY	12	1		1	–	–	–	1	2011	HDD	Chicago	1	1	3							3
2010	HDD	NY	3	1	1	–	–	–	–	1	2011	HDD	Chicago	2	1	2							2
2010	HDD	NY	11	1	5	–	–	–	–	5	2011	HDD	Chicago	11	1	2							2
2010	HDD	NY	12	1	2	–	–	–	–	2	2011	HDD	Chicago	12	1	2							2
2011	HDD	NY	1	1	4	1	–	–	–	5	2004	CDD	Dallas	9	1	1							1
2011	HDD	NY	2	1	5	–	–	–	–	5	2005	CDD	Dallas	8	1	1							1
2011	HDD	NY	3	1	1	1	–	–	–	2	2006	CDD	Dallas	6	1	1							1
2011	HDD	NY	11	1	1	–	–	–	–	1	2006	CDD	Dallas	9	1	1							1
2011	HDD	NY	12	1	2	–	–	–	–	2	2007	CDD	Dallas	7	1	1							1
2012	HDD	NY	2	1	1	–	–	–	–	1	2008	CDD	Dallas	5	1	6							6
2004	CDD	NY	9	1	–	1	–		–	1	2008	CDD	Dallas	6	1	1							1
2005	CDD	NY	5	1		1	–	–	–	1	2008	CDD	Dallas	7	1	2							2
2005	CDD	NY	6	1		8	1	–	–	9	2008	CDD	Dallas	8	1	1							1
2005	CDD	NY	7	1			–	–	−1	1	2009	CDD	Dallas	5	1	4	1						5
2005	CDD	NY	8	1		4	–	–	–	4	2009	CDD	Dallas	6	1	1		1					2
2006	CDD	NY	6	1		4	–	–	–	4	2009	CDD	Dallas	8	1	1							1
2006	CDD	NY	7	1		1	–	–	–	1	2009	CDD	Dallas	9	1	1							1
2006	CDD	NY	8	1		3	–	–	–	3	2010	CDD	Dallas	5	1	1							1
2006	CDD	NY	9	1		3	–	–	–	2	2010	CDD	Dallas	6	1	2	3						5
2007	CDD	NY	7	1		1	–	–	–	1	2010	CDD	Dallas	7	1		2						2
2007	CDD	NY	8	1		1	–	–	–	1	2010	CDD	Dallas	8	1		2						2
2007	CDD	NY	9	1		1	–	–	–	1	2010	CDD	Dallas	9	1		3	1					4
2008	CDD	NY	6	1		5	–	–	–	5	2011	CDD	Dallas	5	1			1					1
2008	CDD	NY	7	1		1	–	–	–	1	2011	CDD	Dallas	6	1	3	1						4
2008	CDD	NY	8	1		1	–	–	–	1	2011	CDD	Dallas	7	1	3							3
2009	CDD	NY	8	1		2	–	–	–	2	2011	CDD	Dallas	8	1	2							2
2010	CDD	NY	5	1		2	–	–	–	2	2012	CDD	Dallas	2	1	2							2
2010	CDD	NY	5	1		2	–	–	–	2													
2010	CDD	NY	6	1		1	–	–	–	1													
2010	CDD	NY	7	1		1	–	–	–	1													
2010	CDD	NY	9	1		1	–	–	–	1													
2011	CDD	NY	9	1		1	–	–	–	1													
2011	CDD	NY	6	1		3	–	–	–	3													

Fig. 12 plots give the evolution of New York-HDD, Atlanta HDD-Seasonal and option prices with time to maturities in one and six months respectively, against strike prices and trading days in March 2006 and October 2007. It is clear that options were traded with very few strike prices (from one to four strike prices) during this month. We used five support points ($N = 5$) in the quadrature method, and calculate R^2 in a logarithmic scale:

$$R^2 = 1 - \frac{\sum_{i=1}^{N_i}\sum_{j=1}^{N_i}\sum_{k=1}^{N_{ij}}\left\{\log y_{ijk} - \log C_{ij}^N(w,\theta)\right\}^2}{\sum_{i=1}^{N_i}\sum_{j=1}^{N_i}\sum_{k=1}^{N_{ij}}\log y_{ijk}^2}.$$

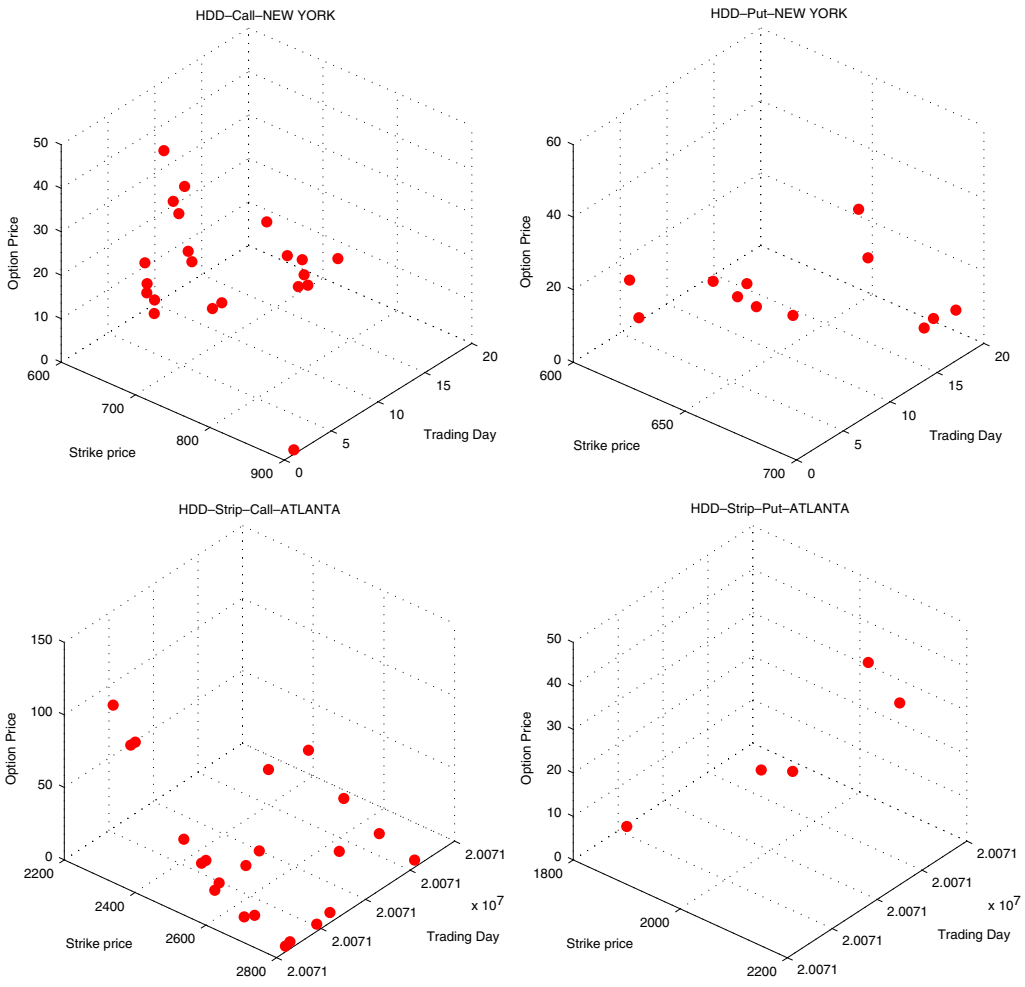

Fig. 12. New York monthly HDD option prices with time to maturity in one month against strike prices and trading days in March 2006, HDD Atlanta Seasonal option prices with time to maturity in 6 months and trading days in October 2007.

Recall that N_1 and N_2 are numbers of different strike prices for call and put options, respectively. N_i relies on the given data set. N_{ij} is the number of repeated options given a strike price and a type of option. In our empirical analysis, because we just take one daily closing price, we set $N_{ij} = 1$.

When R^2 is close to one, model prices are close to market prices and the model produces nice fit. In the Bayesian quadrature method, we calculate R^2 at each swipe of the Markov chain Monte Carlo algorithm, and summarize its posterior mean and quantiles for inference. Because all the mean, median, and the 2.5% and 97.5% quantiles of the R^2 calculated in the MCMC algorithm are close to one, we conclude that the Bayesian quadrature method produces an almost perfect fit for all these trading days. Fig. 13 presents dynamics of the smoothed implied SPD, all of which deviate from lognormality.

A simple way to investigate the dynamics of the implied SPD at each trading day is to calculate moments based on the quadrature method. In each swipe of the Markov chain Monte Carlo algorithm, we calculate the mean (μ), volatility (υ), skewness (s), and kurtosis (κ) of the quadrature method, by the following formulas,

$$\mu = \sum_{n=1}^{N} w_n \theta_n$$

$$\upsilon = \sqrt{\sum_{n=1}^{N} w_n (\theta_n - \mu)^2}$$

$$s = \sum_{n=1}^{N} w_n (\theta_n - \mu)^3 / \upsilon^3$$

$$\kappa = \sum_{n=1}^{N} w_n (\theta_n - \mu)^4 / \upsilon^4.$$

Fig. 14 gives the dynamics of the posterior means of the SPDs. Table 5 shows the posterior means of these four quantities of the quadrature method. However, skewness and kurtosis of weather options can be either positively or negatively skewed depending on futures maturity.

WD-SPD's tend to be positively skewed for short maturity contracts indicating that the tail on the right side is longer or fatter than the left side as the call-options only provide information in the right tail of the SPD. Conversely, negative skew indicates that the tail on the left side of the SPD, provided by put-options, is longer or fatter than the right side.

The heterogeneity beliefs on investors (hedgers versus speculators), reflected by the weather sensitivity preferences among agents, lead the SPD spread to the tails and even becomes bimodal. As shown in Jackwerth and Rubinstein (1996) and Rubinstein (1994), it is common to get in incomplete markets, like the stock index options, multimodal risk neutral densities. The presence of severe modes might be caused due to nonlinear relationship between the seasonal variance of the temperature process T_t and the underlying temperature index futures in (19), see Härdle and López-Cabrera (2012), Benth et al. (2007). Temperature

Table 5
Posterior mean of the mean (μ), volatility (υ), skewness (s), and kurtosis (κ) of the quadrature method at each trading day (TD) calibrated from New York HDD monthly options in March 2006, Atlanta HDD seasonal strip options in October 2007 and Dallas CDD monthly options in June 2010.

	TD	2	3	6	7	8	9	10	14	15	16	17	20
HDD	μ	511.16	676.00	618.80	676.31	660.00	645.00	660.00	422.78	554.10	708.58	698.31	380.26
NY	υ	248.98	81.57	57.02	73.88	61.38	59.89	45.07	240.72	217.04	41.54	42.11	205.93
	s	0.09	−0.92	0.25	4.30	0.48	0.21	0.38	0.43	−0.42	1.47	1.65	0.60
	κ	1.54	2.70	2.79	56.28	2.26	3.28	2.98	2.16	1.70	5.66	7.59	2.90
	TD	2	3	4	9	12	24	31					
HDD	μ	2225.54	1741.96	1865.99	1630.19	1470.13	2335.12	2340.70					
Strips	υ	324.45	818.35	757.39	751.52	835.64	287.44	258.52					
Atlanta	s	0.58	−0.05	−0.33	0.21	0.59	0.64	0.07					
	κ	3.44	1.42	1.73	2.15	2.44	7.59	3.00					
	TD	4	10	14									
CDD	μ	489.96	390.42	596.63									
Dallas	υ	175.79	195.06	155.40									
	s	−0.83	0.18	−0.57									
	κ	2.72	1.98	2.76									

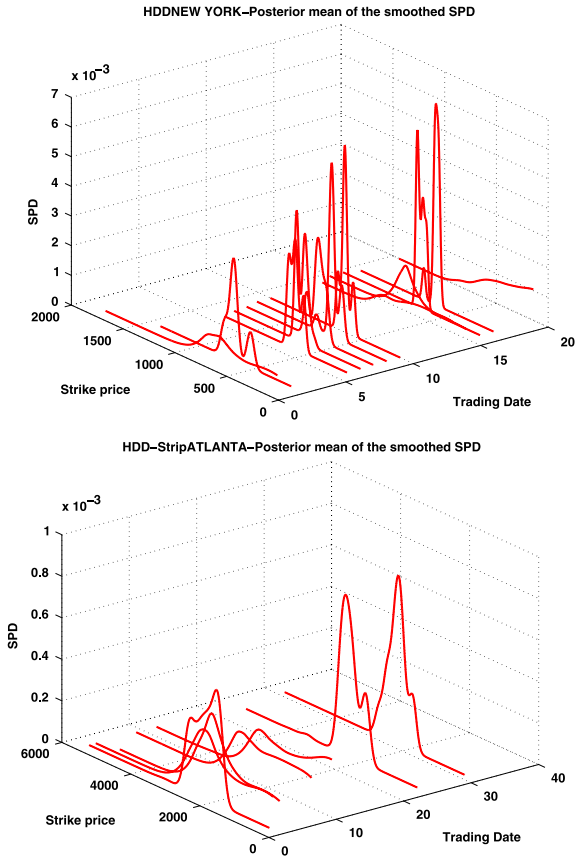

Fig. 13. Quadrature method and smoothed SPDs implied from New York monthly HDD option prices with time to maturity in one month against strike prices and trading days in March 2006, HDD Atlanta Seasonal option prices with time to maturity in 6 months and trading days in October 2007, Dallas monthly CDD options with time to maturity in one month traded in June 2010.

tends to stay stable during periods with low seasonal variance. Thus, SPDs are depending on the conditional volatility: the SPD is wider when the conditional volatility is high. This is different with what documented on index options market (unimodal densities) (Bakshi et al., 2010), interest rate derivatives market (lognormal densities) (Li and Zhao, 2007) and temperature markets (unimodal normal densities) (Benth et al., 2007), but similar to rainfall markets (skewed densities) (López-Cabrera et al., 2013). This is also explained by the economic behavior of agents sensitive to weather conditions. Investors expect that temperature variations, that affect their cash flows, will occur with high probability in winter times than in summer times (conversely for WDs in Australia). Hence some investors will use these option contracts for hedging purposes in presence of negative expected payoffs to eliminate their risk, while others will act as speculators from bearing hedgers' risk. The results show, as expected, that the option temperature market offers a much greater premium than the futures temperature market (Härdle and López-Cabrera, 2012).

4. The infeasibility of other nonparametric methods

Here, we compare the feasibility of our method with other known nonparametric approaches, which are popular tools avoiding risk of misspecification. Some of these methods estimate the SPD by differentiating an interpolation of smoothing of option prices. In this context, data sparsity makes the estimation of the SPD a statistical challenge. Let us now explain why the kernel regression method and mixtures of lognormals do not work well in the context of WD implied SPDs.

The cross-section of the call and put prices is given in Fig. 12. Different grids correspond to different contracts with different times to measurement periods and consequently one can argue that option prices can be extrapolated as a smoothed function of the strike.

4.1. Kernel regression

The kernel regression method (KRM) takes advantage of differentiating twice (1):

$$f(K) = \mathrm{e}^{r\tau}\frac{\partial^2}{\partial K^2}C(K). \tag{23}$$

In order to employ (23), one needs more observations as the option price function is treated as a continuous function of strikes and therefore relies on the put–call parity to transfer put option prices to call option prices. When the market is rarely traded, it is not promising to employ the put–call parity though. In our empirical data analysis most options are traded with only a few strike prices. Very often, an option was traded only with one or two strike prices. When the kernel method is applied to a data set of such a case, it is even difficult to find an option function $C(K)$, not to mention to find its second derivatives and interpolated version, may not yield a density estimate that guarantee to be positive and integrate to one. Consequently, KRM is sensitive to data sparsity.

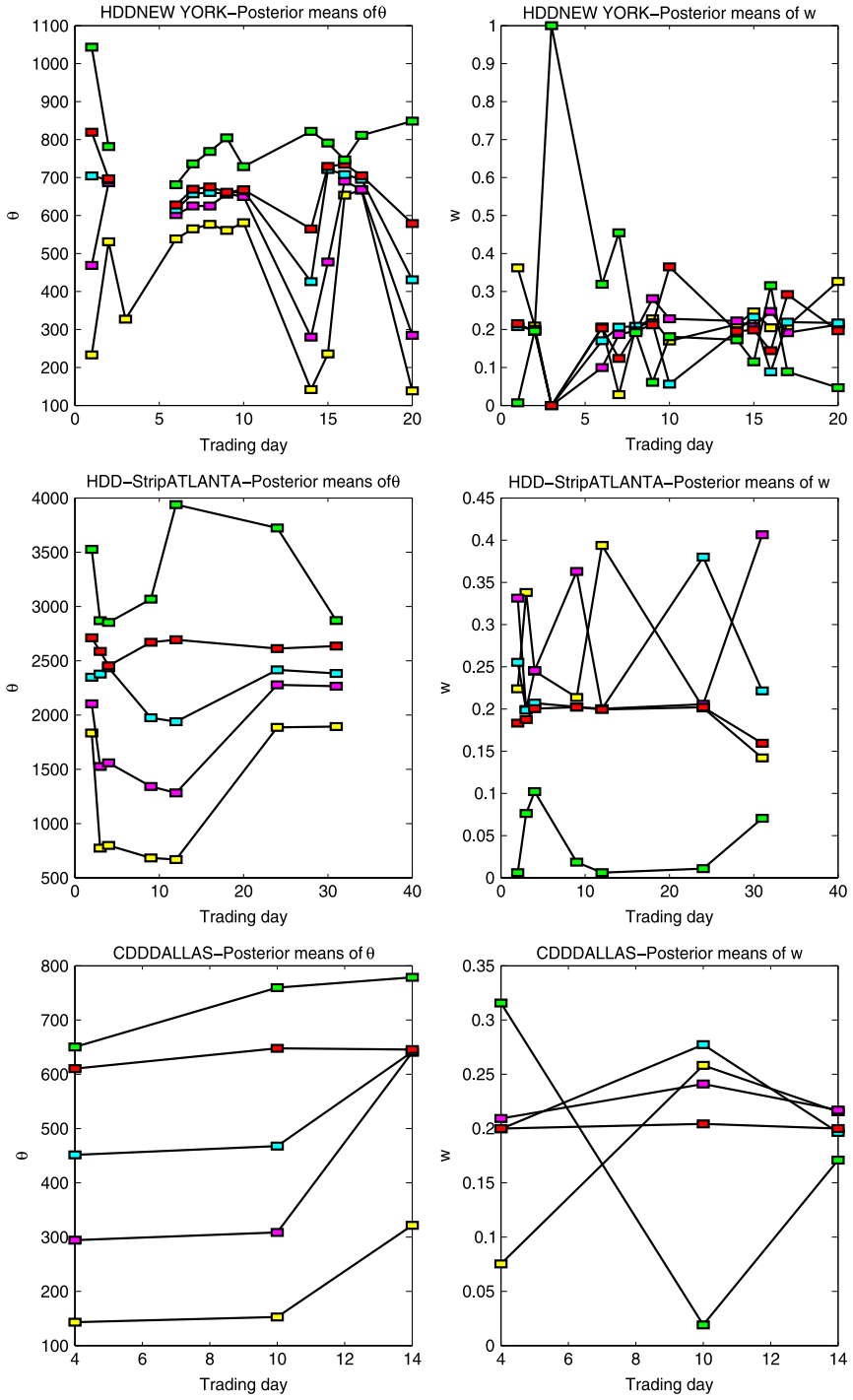

Fig. 14. Dynamics of the parameter of the quadrature method implied from New York monthly HDD option prices with time to maturity in one month against strike prices and trading days in March 2006, HDD Atlanta Seasonal option prices with time to maturity in 6 months and trading days in October 2007, Dallas monthly CDD options with time to maturity in one month traded in June 2010.

4.2. Mixture of lognormals

When applying the mixture of lognormal methods, it is necessary to specify the range of the variances of the lognormal density. The selection is objective and influences the estimated SPD dramatically. When the data set consists of a few data point, it is possible to produce two totally different densities (particularly in terms of variances) which produce the same quality of model fit.

We estimate the SPD using a mixture of lognormal for New York HDD monthly options, given in Fig. 8. For mixture of lognormals, we will show that two different SPD using mixture of lognormal produce the same model fit, but they have quite different higher moments.

Yuan (2009) proposed a function class:

$$\mathcal{F} = \left\{ f(\cdot) : f(x) = \int f(x|\mu, \sigma^2) dG(\mu, \sigma), \right.$$
$$\left. \text{supp}(g) \subset [-M, M] \times [\underline{\sigma}, \bar{\sigma}] \right\}$$

where $M < \infty$ and $0 < \underline{\sigma} \leq \bar{\sigma} < \infty$, $f(x|\mu, \sigma^2)$ is the pdf of the lognormal distribution with location μ and scale σ and

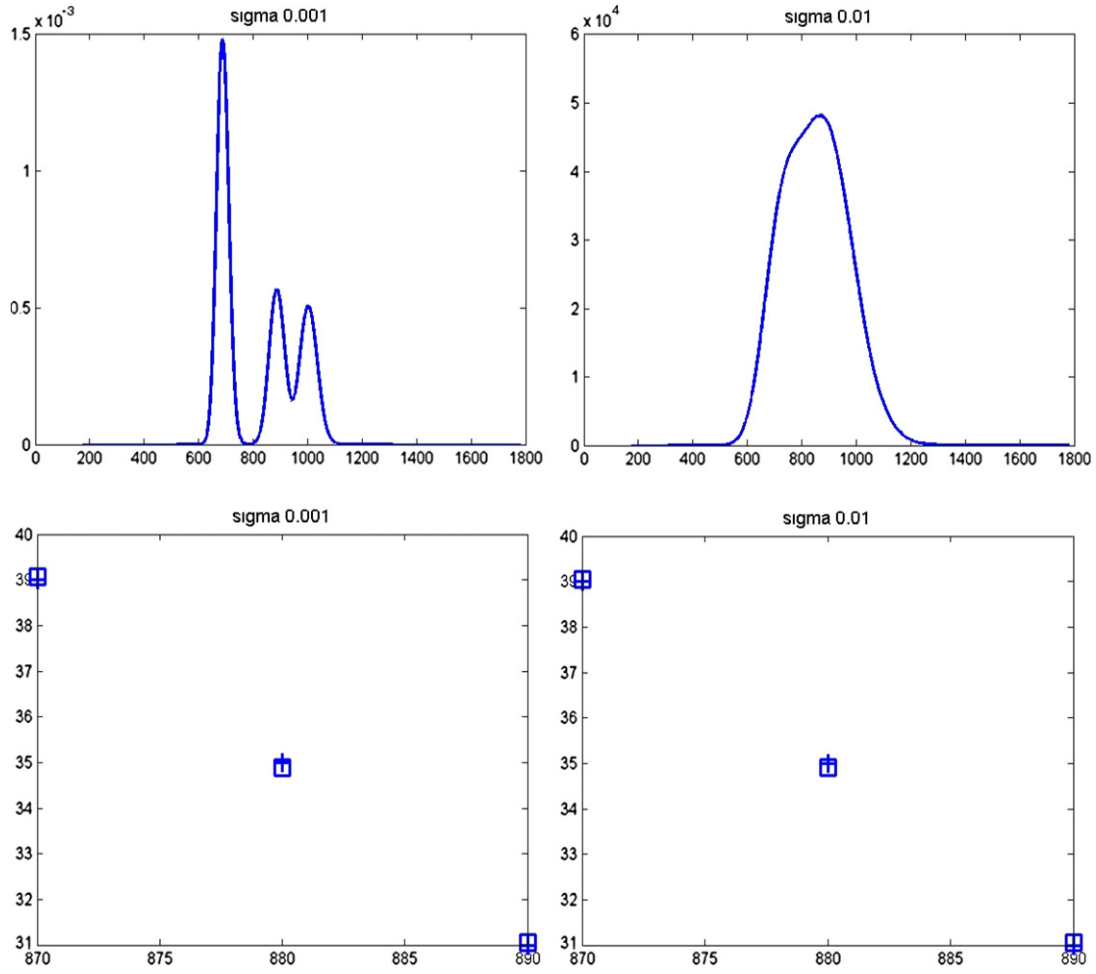

Fig. 15. Estimated SPD using mixtures of lognormal. The left-upper and right-upper panels give estimated SPD using $\underline{\sigma}$ equal to 0.01 and 0.001, respectively. The left-lower and right-lower panels give market prices indicated by a blue cross and model prices indicated by a blue square. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

G determines the mixing distribution. The corresponding pricing function in this case is similar to (2):

$$C(X;\mu,\sigma^2) = \exp(-r\tau)\int_0^\infty \wp_{ij}(x)f(x)dx$$

$$C(X;G) = \int C(X;\mu,\sigma^2)dG(\mu,\sigma) \tag{24}$$

as the SPD $f(x)$ is defined in the previous family $\mathcal{F}$. The least squares estimate of the pricing function can be written as:

$$\hat{G}(\cdot) = \underset{G\in\mathcal{G}}{\arg\min}\, n^{-1}\sum_{i=1}^{n}\left\{y_{ijk} - C(X;G)\right\}^2 \tag{25}$$

where $\mathcal{G}$ is the collection of all probability measures on μ and σ^2. Note that the minimization is taken over a function space of infinite dimensions, however the solution can be represented in a finite dimensional space. In particular, all solutions can be expressed as a convex combination of at most $n+1$ Black–Scholes type of pricing functions.

This model has several nice theoretical properties. For example, as the sample size n increases, the pricing functions can be recovered with squared error converging to zero at the rate of $\log^2 n/n$, which is close to the parametric rate of convergence $1/n$. However, practical difficulties arise when fitting mixtures of lognormal distributions (or other mixtures models) to real data. The feature that weather options are traded with a few number of strike prices make mixture models inapplicable, because mixture models need to select corresponding scale parameters and the number of components. For example, when options are traded with n different strike prices, maximum likelihood suggests to use $n/2$ support points. When n is large, this leads a very complicated model and possible over fitting problem. When n is small, the resulting model may be inappropriate. In addition, numerical procedure for searching the maximum likelihood estimate is particularly difficult for large n.

We apply mixture of lognormals by Yuan (2009) to the New York monthly HDD call options traded on 2006/02/02, with two different manually selected variances. Fig. 15 shows that these two estimated SPD are quite different in shapes, although they produce similar quality of model fit. Therefore, this illustration shows that the estimated SPD is very sensitive to the selection of $\underline{\sigma}$. In practical implementation, Yuan (2009) suggests to determine $\underline{\sigma}$ by cross-validation. This however is very computationally demanding.

5. Conclusions

We estimate SPDs for WDs using the Bayesian quadrature method. The WD market is characterized by its incompleteness and less frequently traded activities. This makes the estimation of the SPD a statistical challenge. However, the quadrature method, in advantage to the parametric and other non-semiparametric techniques, avoids model miss-specification and allows the SPD estimation by a parsimonious model. The technique is

computationally fast and robust. The obtained SPD do not stem from market-risk-price assumptions. We present empirical results on real CME temperature derivatives data, which help us to understand the dynamics of SPD. The results suggest that the SPD of weather derivatives exhibits a non-normal behavior type.

Acknowledgments

The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 "Ökonomisches Risiko", Humboldt-Universität zu Berlin is gratefully acknowledged. Huei-Wen Teng's research is supported by the Ministry of Science and Technology, Taiwan (ROC), under grant 103-2633-M-008-001.

References

Abadir, K., Rockinger, M., 2003. Density functionals, with an option-pricing application. Econometric Theory 19 (5), 778–811.
Aït-Sahalia, Y., Duarte, J., 2003. Nonparametric option pricing under shape restrictions. J. Econometrics 116, 9–47.
Aït-Sahalia, Y., Lo, A.W., 1998. Nonparametric estimation of state-price densities implicit in financial asset prices. J. Finance 53, 499–547.
Aït-Sahalia, Y., Lo, A.W., 2000. Nonparametric risk management and implied risk aversion. J. Econometrics 94, 9–51.
Bakshi, G., Madan, D., Panayotov, G., 2010. Returns of claims on the upside and the viability of u-shaped pricing kernels. J. Financ. Econ. 97 (1), 130–154.
Benth, F., Benth, S., Koekebakker, S., 2007. Putting a price on temperature. Scand. J. Statist. 34, 746–767.
Benth, F., Härdle, W.K., López-Cabrera, B., 2011. Pricing Asian temperature risk. In: Cizek, , Härdle, , Weron, (Eds.), Statistical Tools for Finance and Insurance, second ed.. Springer Verlag, Heidelberg.
Breeden, D., Litzenberger, R., 1978. Price of state-contingent claims implicit in option prices. J. Bus. 51, 621–651.
Casella, G., Berger, R.L., 2001. Statistical Inference. Duxbury Press.
Chen, M., Shao, Q., 1997. On Monte Carlo methods for estimating ratios of normalizing constants. Ann. Statist. 25 (4), 1563–1594.
Derman, E., Kani, I., 1994. Riding on the smile. Risk 7, 32–39.
Dorfleitner, G., Wimmer, M., 2010. The pricing of temperature futures at the Chicago Mercantile exchange. J. Banking Finance 34 (6), 1360–1370.
Dumas, B., Fleming, J., Whaley, R.E., 1998. Implied volatility functions: Empirical tests. J. Finance 53 (6), 2059–2106.
Dupire, B., 1994. Pricing with a smile. Risk 7, 18–20.
Fan, J., Mancini, L., 2009. Option pricing with model-guided nonparametric methods. J. Amer. Statist. Assoc. 104 (488), 1351–1372.
Garcia, R., Ghysels, E., Renault, E., 2010. Econometrics of option pricing models. In: Aït-Sahalia, Y., Hansen, L.P. (Eds.), Handbook of Financial Econometrics, Vol. 1. North Holland, Amsterdam.
Ghysels, E., Harvey, A., Renault, R., 1995. Stochastic volatility. In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics 14, Statistical Methods in Finance. North Holland, Amsterdam.
Ghysels, E., Patilea, V., Renault, E., Torres, O., 1997. Nonparametric methods and option pricing. In: Hand, D., Jacka, S. (Eds.), Statistics in Finance. Edward Arnold, London.
Giacomini, R., Gottschling, A., Haefke, C., White, H., 2008. Mixtures of t distributions for finance and forecasting. J. Econometrics 144, 175–192.
Härdle, W.K., Hlávka, Z., 2009. Dynamics of state price densities. J. Econometrics 150, 1–15.
Härdle, W.K., López-Cabrera, B., 2012. Infering the market price of weather risk. Appl. Math. Finance 19 (1), 59–95.
Härdle, W., Müller, M., Sperlich, S., Werwatz, A., 2004. Nonparametric and Semiparametric Models. Springer Verlag, Heidelberg.
Jackwerth, J., Rubinstein, M., 1996. Recovering probabilities distributions from option prices. J. Finance 51, 1611–1631.
Li, H., Zhao, F., 2007. Nonparametric estimation of state-price densities implicit in interest rate cap prices. Rev. Financ. Stud. 22 (11), 4335–4376.
Liechty, J., Teng, H.-W., 2009. Bayesian quadrature approaches to state price density estimation. working paper.
Available at: http://www.personal.psu.edu/jcl12/Bayesian%20Quadrature.pdf.
López-Cabrera, B., Odening, M., Ritter, M., 2013. Pricing rainfall futures at the CME. J. Banking Finance 37 (11), 4286–4298.
Renault, E., 1997. Econometric models of options pricing errors. In: Kreps, D., Wallis, K. (Eds.), Advances in Economics and Econometrics: Theory and Applications * Seventh World Congress, Vol. III. Cambridge University Press, Cambridge.
Rosenberg, J., Engle, R., 2002. Empirical pricing kernels. J. Financ. Econ. 64, 341–372.
Rubinstein, M., 1994. Implied binomial trees. J. Finance 49, 771–818.
Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman and Hall.
Ueberhuber, C.W., 1997. Numerical Computation 2: Methods, Software, and Analysis. Springer-Verlag.
Yatchew, A., Härdle, W., 2006. Nonparametric state price density estimation using constrained least squares and the bootstrap. J. Econometrics 133 (2), 579–599.
Yuan, M., 2009. State price density estimation via nonparametric mixtures. Ann. Appl. Stat. 3 (3), 963–984.