The pricing of temperature futures at the Chicago Mercantile Exchange*

Gregor Dorfleitner† Maximilian Wimmer‡

This version: December 3, 2009

Abstract

This paper analyzes observed prices of U.S. temperature futures at the Chicago Mercantile Exchange (CME). Results show that an index modeling approach without detrending captures the prices exceptionally well. Moreover, weather forecasts significantly influence prices up to 11 days ahead. It is shown that valuations of temperature futures relying on a model without detrending yield biased valuations by overpricing winter contracts and underpricing summer contracts. Several trading strategies are devised to exploit the mispricing observed at the CME and to demonstrate that speculating on temperature futures can not only generate high overall returns, but also perform well on a risk-adjusted basis.

Keywords: Weather derivatives, Index modeling, Weather forecast, Futures pricing, Trading strategy

JEL Classification: C53, G13, G29, Q54

*This is a preprint of an article whose final and definitive form is in press in the Journal of Banking & Finance, doi:10.1016/j.jbankfin.2009.12.004. We gratefully acknowledge the support of part of this research by the German Research Foundation via the Collaborative Research Center 649 *Economic Risk*. We would like to thank Günter Bamberg, Robert Ferstl, Josef Hayden, Michaela Leidl and Bernhard Zwergel for fruitful comments and discussion.

†Corresponding author; Department of Finance, University of Regensburg, 93040 Regensburg, Germany, email: gregor.dorfleitner@wiwi.uni-regensburg.de.

‡Department of Finance, University of Regensburg, 93040 Regensburg, Germany, email: maximilian.wimmer@wiwi.uni-regensburg.de.

1. Introduction

Weather derivatives, i.e., financial instruments written on meteorological data such as temperature or precipitation, have existed since 1996. While, as is common for new derivatives, the first contracts were arranged over-the-counter (OTC), the Chicago Mercantile Exchange (CME) commenced trading weather contracts in 1999. In accordance with Jewson and Brix (2005), we define weather derivatives by the following components: the period of time (the *measurement period*) from the starting date to the ending date in which the *weather variable* is measured at a specific *weather station*. Additionally a *weather index* aggregating the weather variable during the measurement period is required. The weather index is the actual underlying whose value is fed into the *payoff-function* of the derivative to calculate the resulting cash flow shortly after the end of the measurement period. In option-like contracts the buyer will likely have to pay a premium to the seller. It is obvious that weather derivatives can only be cash settled, since physical delivery is impossible.

Whereas the CME began trading with monthly heating and cooling degree days (HDD and CDD) futures and options for several U.S. cities only, the range of products has broadened noticeably by beginning of 2009, comprising seasonal and monthly futures and options written on CDD, HDD, CAT (cumulated average temperature, for some European and Asian-Pacific cities), frost days, snowfall, or hurricane indices, to name the most important underlyings. According to PricewaterhouseCoopers (2006), a large portion of weather derivatives is still written on temperature indices like HDD and CDD and is traded on the CME, while a large part constituting the market beyond these indices takes place OTC. Besides the common separation of options and futures market participants into hedgers, speculators, and arbitrageurs, it is noteworthy that many of the CME temperature market participants are insurance companies, which are traditionally weather risk takers, or utility companies, which use weather derivatives as a means for hedging volume risks (PricewaterhouseCoopers, 2006). Several studies show the dependence of various industries on weather today (see e.g. Dutton, 2002; Harrod et al., 2007). In the light of climate change, causing not only higher temperatures but also a higher temperature volatility (Intergovernmental Panel on Climate Change, 2007), the use of weather derivatives can be expected to grow further.

Since the underlying of weather derivatives is not tradable, the pricing of these instruments is less straightforward than it is for other derivatives and leads to several possible valuation techniques that do not cause direct arbitrage opportunities against the underlying. Thus, it is not surprising that there is no general consensus regarding the question concerning how to price weather derivatives. The pricing mechanisms suggested in the literature range from simple burn analysis and index modeling approaches to models for the daily weather variable process, the so-called daily simulation approach, with a more or less sophisticated pricing argument on top. Burn analysis valuates a derivative simply by averaging all payoffs that would have been realized in the past. Index modeling—as the name suggests—goes one step further and models the weather index with a probability distribution whose parameters are usually estimated based upon historical data. The actual pricing then proceeds by taking the expected value of the derivatives' payoff plus a possible risk premium. Such pricing approaches are rather actuarial than financial, which can be explained by the fact that weather derivatives often substitute traditional weather insurance. In the context of this paper, we focus on the most commonly traded type of contract, namely futures contracts written on temperature indices. For these contracts, the frontier between burn analysis and index modeling becomes blurred because we are mainly interested in the expected index value.

A lot of research on weather derivatives in the last decade has focused on developing adequate pricing models. Davis (2001) develops a pricing formula using marginal utility techniques; Pirrong and Jermakyan (2008) compare weather derivatives to power derivatives and suggest using finite difference methods for valuation. The theory of index modeling is treated thoroughly in Jewson and Brix (2005).

Another class of pricing models is comprised of daily simulation approaches. These models can be driven by a continuous process (see e.g. Dischel, 1998; Dornier and Queruel, 2000; Alaton et al., 2002; Brody et al., 2002; Benth and Šaltytė-Benth, 2005; Zapranis and Alexandridis, 2008), or by a discrete autoregressive-type process (see e.g. Cao and Wei, 2000; Caballero et al., 2002; Jewson and Caballero, 2003a; Campbell and Diebold, 2005; Benth and Šaltytė-Benth, 2007). Using continuous-time autoregressive processes, Benth et al. (2007) and Härdle and López Cabrera (2009) combine both daily simulation branches. One of the advantages of daily simulation models illustrated by academics is the possibility to derive, after estimating the market price of risk, no-arbitrage prices based on a continuous-time hedging strategy. While most of these papers simply assume a market price of risk of zero, Cao and Wei (2004) demonstrate in a Lucas (1978) equilibrium model that the market price of risk associated with temperature futures is indeed insignificant in many cases. Following this approach, Hamisultane (2009) computes the market price of risk from temperature futures quotations of New York, but finds that the estimate is not stable.

Several studies (see Oetomo and Stevenson, 2005; Schiller et al., 2008) compare different valuation approaches using backtesting analysis and confirm the superiority of daily simulation models compared to index modeling approaches. However, these studies also emphasize the large spread of valuations, which the different daily simulation models yield.

It is surprising that up to now there has not been a single study which has attempted to identify a pricing model capable of describing real weather derivatives' prices. Of course it is very difficult to obtain over-the-counter market prices. However, the prices of the CME contracts are publicly available. Consequently, in this paper we fill this gap for U.S. temperature futures contracts.

In this work we restrict our analysis to index modeling approaches for several reasons. According to Jewson and Brix (2005), most practitioners rely on the well-understood index modeling. Furthermore, since daily simulation models include a high number of parameters, they inhibit the chance of misspecification and need to be calibrated for each individual weather station (cf.

Schiller et al., 2008). Moreover, in the case of futures contracts the weather index is aggregated over the measurement period, and therefore it is not necessary to detail the daily movement of the weather variable. When assuming a market price of risk of zero, as most of the studies above do, the no-arbitrage price of daily simulation models is reduced to the expected value. Hence, the advantage of daily simulation models of deriving no-arbitrage prices diminishes in this case. Yet the question remains concerning how well a parsimonious index modeling approach can explain real temperature futures prices.

The remainder of this paper is structured as follows: In section 2, we first verify the adequacy of index modeling with and without detrending using a variable number of preceding years for HDD and CDD data corresponding to real futures traded on the CME. In section 3 we then analyze how well index modeling can capture real futures prices. Since analysis of the weather data shows that the market model can principally be outperformed by a detrending model using 30 years of data, we then explore in section 4 the economic significance of real prices' deviation from this model through trading strategies involving various levels of sophistication. Section 5 concludes the paper with a short summary of our findings and consequences for participants in weather futures markets.

2. Index modeling with and without detrending

The temperature futures traded on the CME for U.S. weather stations are written on HDD and CDD indices. Therefore, these notions first require clarification. Let $[\tau_1, \tau_2]$ be the measurement period and T_t^{avg} be the average of the minimal and maximal temperature measured at a certain weather station on a certain day $t \in [\tau_1, \tau_2]$. Then HDD and CDD are defined by

$$\text{HDD}(t) := \max(65\,°\text{F} - T_t^{\text{avg}}, 0), \quad \text{CDD}(t) := \max(T_t^{\text{avg}} - 65\,°\text{F}, 0)\,. \tag{1}$$

As mentioned above, we view the index modeling approach as the most apt way to valuate weather derivatives in practice. The main idea of this approach is calculating the expected future payoff of a derivative directly by considering the payoffs that the same derivative would have yielded in the corresponding measurement periods in the past years. If, for example, a HDD derivative for a measurement period $[\tau_1, \tau_2]$ is to be priced for the year $n+1$, one would calculate the fictitious indices the same derivative had in the year n, $n-1$, $n-2$, etc. This yields a series $Y_1, Y_2, \ldots, Y_n$ with $Y_i = \sum_{t=\tau_1}^{\tau_2} T_{t,i}$ where $T_{t,i}$ denotes the HDD of day t in year i. The pricing of the instrument is based on this series $Y_i, i = 1, \ldots n$. So far this has been the idea of burn analysis. If additionally there are assumptions about a distribution of the series of past indices the procedure is named index modeling. The actual valuation is carried out by taking the expected payoff given the distribution of the index Y_{n+1} under the real probability measure.

As usual the value of a futures contract is zero at initiation. Therefore, with Y_{n+1} denoting the index corresponding to the contract under consideration, the futures price F_t is determined by

$$0 = \mathbb{E}_t^P(Y_{n+1} - F_t) = \mathbb{E}_t^P(Y_{n+1}) - F_t, \tag{2}$$

where P symbolizes the real-world measure corresponding to the chosen index model.

2.1. Linear models

In its simplest form, the index modeling approach averages the past n indices to estimate the estimated index value of (2). However, it is a well known fact that historical temperature time series usually exhibit a trend, and thus we can expect the index time series to display one, too. The reason for the trend in the temperature time series may not only be attributed to the effects of global warming, but also to urbanization effects that have lead to local warming (Cotton and Pielke, 2007). It is clear that the average temperature in high-density areas is above the temperature in sparsely populated areas due to waste heat from the buildings and the reduced circulation of air. Hence, increasing building density around a weather station leads to a warming trend in the historical temperature and index data.

We follow the approach of Jewson and Penzer (2004) to estimate the bias and uncertainty in both no-detrending and linear detrending models. In order to derive a rigorous mathematical argument, we commence using the linear model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{3}$$

with the design matrix $\boldsymbol{X}$ capturing the fixed effect and possibly time dependency as well. As usual, we make four assumptions:

1. The expected error $\mathbb{E}(\varepsilon_i) = 0$ for all years $i = 1, \dots, n+1$.
2. The variance of the errors $\mathrm{Var}(\varepsilon_i) = \sigma^2$ is constant for all years $i = 1, \dots, n+1$.
3. The covariance of the errors $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ for all years $i \neq j$.
4. The errors ε_i, $i = 1, \dots, n+1$, are independent and identically normally distributed.

The latter assumption motivates the term *index modeling*, since $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$ implies the distribution of the indices $Y_i \sim \mathrm{N}(\boldsymbol{X}_i\boldsymbol{\beta}, \sigma^2)$. We briefly review the main properties of the linear model. With the unbiased estimator $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$ of $\boldsymbol{\beta}$, the variance of the error of a prediction $\hat{y}_0 = \boldsymbol{x}_0'\hat{\boldsymbol{\beta}}$ of $y_0 = \boldsymbol{x}_0'\boldsymbol{\beta}$ becomes

$$\mathrm{Var}(\hat{y}_0 - y_0) = \boldsymbol{x}_0'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_0\sigma^2.$$

In the case of linear detrending, we let $i = 1, \dots, n$ and estimate the expected value of the variable for the next period $n+1$. The model (3) is thus specified by

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \quad \boldsymbol{x}_0 = \begin{bmatrix} 1 \\ n+1 \end{bmatrix}, \tag{4}$$

yielding a variance of

$$\mathrm{Var}(\hat{y}_0 - y_0) = \frac{4n+2}{n(n-1)}\sigma^2.$$

Since the prediction $\hat{y}_0$ is unbiased, the mean square error of the prediction with respect to the expected value y_0 equals the variance of the error, i.e.,

$$\mathrm{MSE}(\hat{y}_0) = \frac{4n+2}{n(n-1)}\sigma^2. \tag{5}$$

Next we wish to compute the size of the error if there is no detrending applied. To distinguish the notation from the detrending case, we put bars above all variables in the no-detrending setup. Here, we specify the model (3) by

$$\bar{\boldsymbol{Y}} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}, \quad \bar{\boldsymbol{X}} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \bar{\boldsymbol{\beta}} = \beta_1, \quad \bar{\boldsymbol{x}}_0 = 1, \tag{6}$$

and calculate the variance as

$$\mathrm{Var}(\hat{\bar{y}}_0 - y_0) = \frac{1}{n}\sigma^2. \tag{7}$$

Assuming the actual temperature data exhibits a linear trend, the no-detrending model yields a bias of

$$\mathbb{E}(\hat{\bar{y}}_0 - y_0) = -\frac{n+1}{2}\beta_2, \tag{8}$$

where β_2 denotes the actual trend. Putting equations (7) and (8) together, the mean square error of the no-detrending model becomes

$$\mathrm{MSE}(\hat{\bar{y}}_0) = \left(\frac{n+1}{2}\right)^2 \beta_2^2 + \frac{1}{n}\sigma^2. \tag{9}$$

If the temperatures exhibit only a minor trend β_2, the no-detrending MSE (9) clearly outperforms the linear detrending MSE (5). However, with increasing trend to variance level $|\beta_2|/\sigma$ and increasing number of years n of data used, the bias of the no-detrending model increases. Technically, the MSE (9) is smaller than the MSE (5) if

$$\frac{|\beta_2|}{\sigma} < \sqrt{\frac{12}{(n+1)n(n-1)}}. \tag{10}$$

Note that if we refer to a prediction of the realized value $y_0 + \varepsilon_{n+1}$ instead of the expected (time $n+1$) index value y_0 the relation (10) remains the same since the bias (8) is unaltered. The MSE (5) is then replaced by

$$\frac{(n+2)(n+1)}{n(n-1)}\sigma^2, \tag{11}$$

and the MSE (9) is replaced by

$$\left(\frac{n+1}{2}\right)^2 \beta_2^2 + \frac{n+1}{n}\sigma^2. \tag{12}$$

It is exactly this interpretation of the prediction and the MSE which is used when backtesting the models in the next subsection.

2.2. Analysis of past U.S. weather data

As inequation (10) shows, the problem of whether the linear detrending model outperforms the no-detrending model from a MSE perspective depends on the strength of the trend. Since this paper is dedicated to analyzing real weather futures prices corresponding to major U.S. cities, we empirically investigate the general theoretical result of the previous subsection with the weather data relevant for the futures under consideration.

We use historical temperature data for 18 weather stations across the United States.[1] The data originally stems from the U.S. National Weather Service and consists of daily minimum and maximum temperatures. Since the data contains gaps due to failures in measurement equipment or data transmission, and jumps due to changes in measurement equipment, it was pre-processed by Earth Satellite Corporation to fill in such gaps and remove such jumps (Boissonnade et al., 2002). Most temperature series begin in 1950, the only exceptions being Chicago (1958), Houston (1969), Kansas City (1972), and Detroit (1959), and end in 2006.

In order to test whether or not the linear detrending model (4) and the non-detrending model (6) describe the relevant weather index data well and therefore justifies the use of the estimators presented above, we perform a backtesting analysis. This implies that we examine how well the models would have performed in the past in predicting (expected) HDD and CDD indices corresponding to virtual weather derivatives. To do so, we chose the ten most common monthly contracts and compute the error each model makes for each contract at each station using the past $n = 1, \dots, 40$ years of data. The contracts are for the five summer months (May to September) which we use for CDD indices, and for the five winter months (November to March) which we use for HDD indices. For each weather station and month in a specific year we define a virtual futures contract. We then calculate the forecast for the expected index value of the next year with and without detrending and using the data of the n preceding years, provided that there is enough historical temperature data available. Let $I^{n,j}$ and $\bar{I}^{n,j}$ denote the index modeling value for virtual contract j using n preceding years of data with and without

[1] The actual stations are listed in Table 1 below.

detrending, respectively, i.e., $\hat{y}_0$ and $\hat{\hat{y}}_0$ from above, respectively.

The mean error is then computed by aggregation over all of these virtual contracts

$$\mathrm{ME}(n) = \frac{1}{\#j} \sum_j \left(I^{n,j} - Y^j \right), \tag{13}$$

where Y^j is the corresponding ex-post realized index value for virtual futures contract j. We define the ME analogously for index modeling without detrending. We also define the root mean squared error for the procedure with detrending

$$\mathrm{RMSE}(n) = \sqrt{\frac{1}{\#j} \sum_j \left(I^{n,j} - Y^j \right)^2}, \tag{14}$$

and analogously without detrending.

Figure 1: Mean error of index modeling of fictitious monthly HDD and CDD contracts. The regular dashed lines show the expected mean errors of the no-detrending model according to equation (8) for different trends β_2. The solid lines show the empirical mean error of the fictitious contracts according to equation (13).

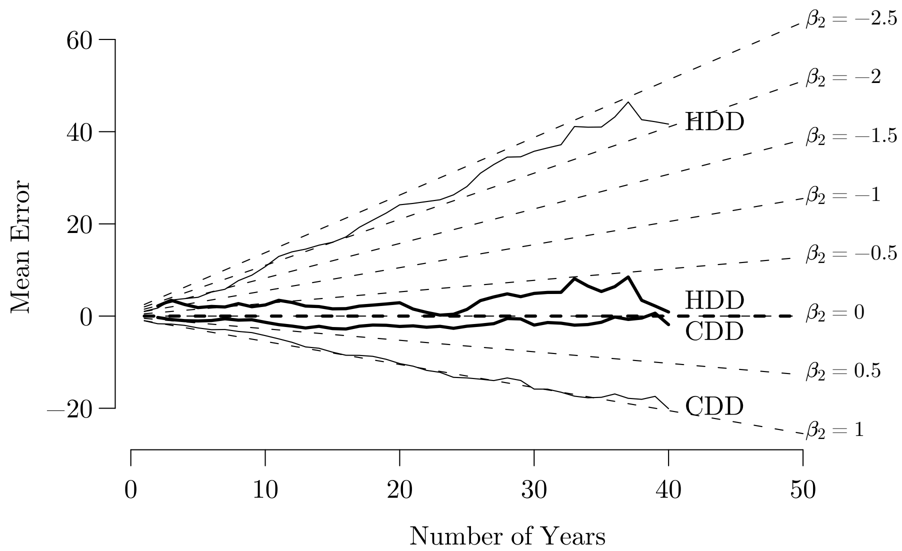
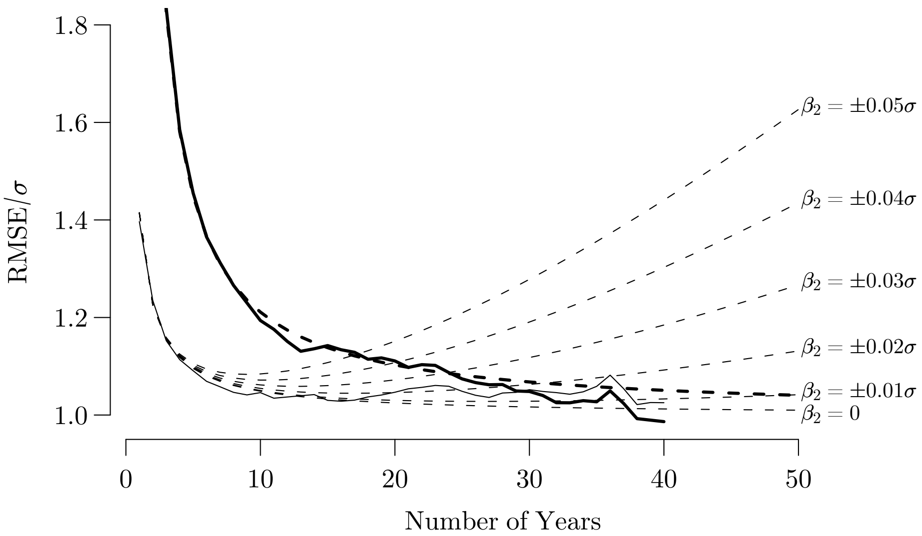


Figure 1 shows the MEs of all HDD and CDD contracts, respectively, depending on the number of years n of data used, and the theoretical curves according to equation (8). Even if for aggregation purposes virtual contracts of different weather stations with possibly differently strong trends and variances are mixed, the linear model appear to be generally apt for the situation. While the errors of the linear detrending models remain relatively low no matter how many years of data have been used, the errors of the no-detrending model show an significantly increasing bias, the more years are used. The data also suggests that trends tend to be stronger in the winter than in the summer, a result that is generally confirmed by Intergovernmental

Panel on Climate Change (2007). In order to check the significance of the difference of the mean error with and without detrending, we calculate the one-sided confidence intervals using the standard bootstrap percentile method. We find that in the CDD case, the mean errors differ significantly on a 5% level for $n = 5$ and on a 1% level for $n \geq 6$. In the HDD case, the mean errors differ significantly on a 5% level for $n = 6$ and on a 1% level for $n \geq 7$.

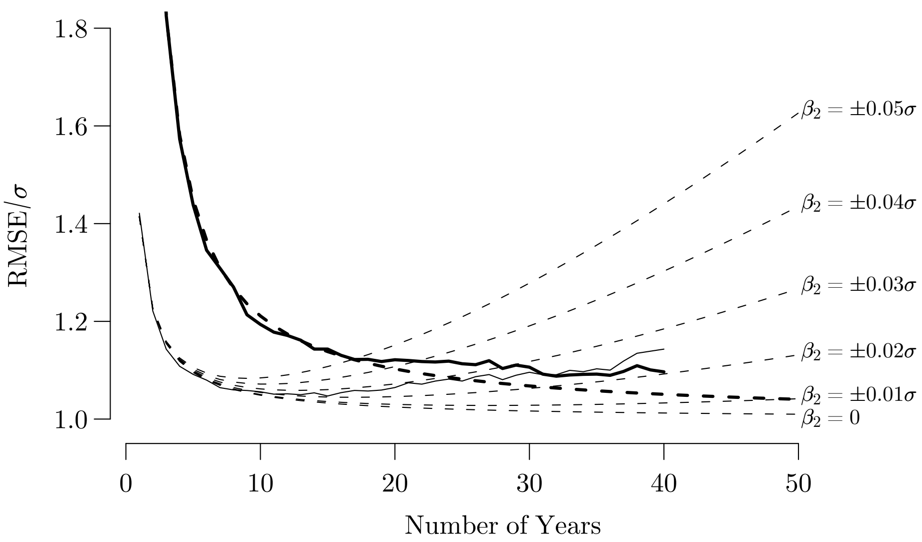
Figure 2: RMSE of index modeling for fictitious monthly HDD contracts. The bold dashed line shows the RMSE of the linear detrending model according to equation (11). The regular dashed lines show the RMSE of the no-detrending model according to equation (12) for different trends β_2. Both lines are normalized by dividing by σ. The solid lines show the empirical RMSEs according to equation (14) of the fictitious contracts, both normalized by a implied σ, ensuring that the detrending RMSE starts at same point as the theoretical RMSE line.



However, the picture changes when looking at the RMSEs of the same analysis (Figures 2 and 3). We can assert that in this analysis, the no-detrending model using approximately 10 years of data performs very well whereas for HDD the detrending model using approximately 30 years of data is generally better. For the CDD data the detrending model does not perform better on the aggregated level from an RMSE perspective. Yet for individual weather contracts detrending with 30 years of data can still be superior to no-detrending with 10 years if the trend is strong enough. In the HDD data one can identify an average $|\beta_2|/\sigma$ ratio of about 0.01 to 0.02, whereas the same figure for CDD data may lie between 0.02 and 0.03.

In summary, we can assert that detrending may have a higher variance but leads to an unbiased predictor of next year's index value. From an MSE point of view, its performance is not superior to no-detrending in general. However, in situations where there is a strong trend it should be preferred to no-detrending. Moreover, from an ME point of view, the detrending model outperforms the no-detrending model significantly on a 1% level if more than six years of data are used. We will make use of this insight below when developing a trading strategy for

Figure 3: RMSE of index modeling for fictitious monthly CDD contracts. The bold dashed line shows the RMSE of the linear detrending model according to equation (11). The regular dashed lines show the RMSE of the no-detrending model according to equation (12) for different trends β_2. Both lines are normalized by dividing by σ. The solid lines show the empirical RMSEs according to equation (14) of the fictitious contracts, both normalized by a implied σ, ensuring that the detrending RMSE starts at same point as the theoretical RMSE line.



real CME futures.

3. Valuation of temperature futures at CME

3.1. Data

To empirically analyze the price formation of weather futures, we use daily settlement prices of weather futures traded at CME from May 2000 to September 2006. Measurement periods are either seasonal, i.e., they correspond to the summer season from May until September and the winter season from November until March, or monthly. April and October contracts are traded as HDD and CDD, while all other winter and summer contracts are traded as HDD and CDD, respectively. The weather stations with available data are summarized in Table 1.

For each day and each contract under consideration we calculate the theoretical value of the contract using index modeling with and without linear detrending using the past one to thirty years of data. The solid line in Figure 4 exemplifies the evolution of the prices of one specific contract.

The index modeling approach using linear models described in the previous section considers only the cases in which a contract is priced before the beginning of the measurement period. A straightforward extension to valuing contracts within the measurement period can be seen in the following: Let t be the day when a contract with measurement period $[\tau_1, \tau_2]$ is priced, with

Table 1: First delivery dates of the CME weather contracts used in this study. "–" indicates not available contracts, "*" indicates not traded contracts.

	Monthly		Seasonal	
City	**HDD**	**CDD**	**HDD**	**CDD**
Atlanta	10/2002	5/2002	3/2005	9/2004
Chicago	10/2002	6/2002	3/2004	9/2004
Cincinnati	10/2002	5/2002	3/2004	9/2004
New York	11/2002	5/2002	3/2004	9/2004
Dallas	10/2002	6/2000	3/2006	9/2004
Philadelphia	10/2002	5/2002	3/2004	9/2005
Portland	10/2002	6/2002	*	9/2005
Tuscon	11/2002	5/2000	*	9/2005
Des Moines	10/2002	7/2000	3/2004	9/2004
Las Vegas	11/2002	6/2000	–	9/2005
Boston	10/2003	5/2004	–	–
Houston	10/2003	9/2003	–	–
Kansas City	10/2003	10/2003	–	–
Minneapolis	10/2003	10/2003	–	–
Sacramento	10/2003	10/2003	–	–
Salt Lake City	10/2005	7/2005	*	*
Detroit	10/2005	9/2005	3/2006	*
Baltimore	11/2005	5/2006	3/2005	*

$\tau_1 \le t \le \tau_2$. We split the measurement period into two parts, $[\tau_1, t-1]$ and $[t, \tau_2]$. Note that by definition $Y_i = \sum_{k=\tau_1}^{t-1} T_{k,i} + \sum_{k=t}^{\tau_2} T_{k,i}$, where $T_{k,i}$ is the temperature index (HDD or CDD) at day k in the year i. Since at day t, which lies within the measurement period, the first sum is already known, we only need to use index modeling for the second sum, whose measurement period lies in the future. Therefore, we define the index modeling price at day t as

$$I_t^n(\tau_1, \tau_2) = \sum_{k=\tau_1}^{t-1} T_{k,n+1} + I_t^n(t, \tau_2), \tag{15}$$

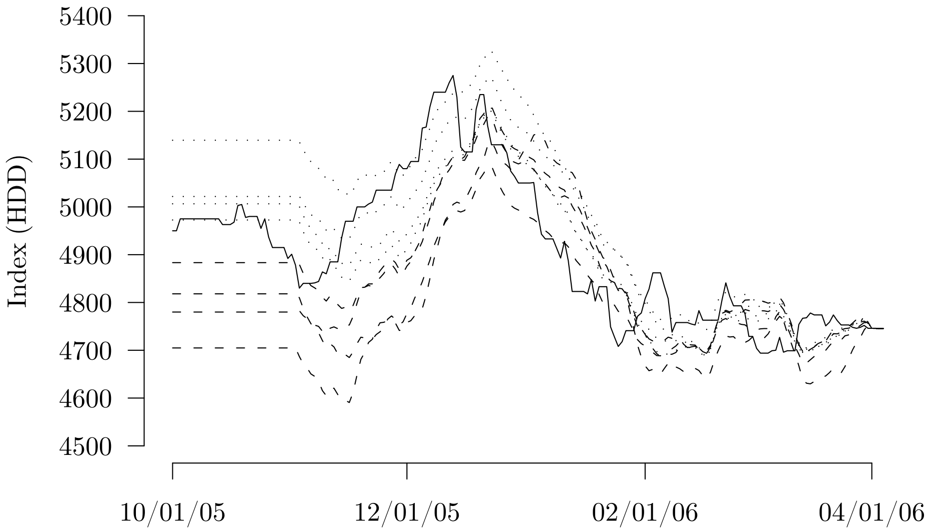
with year $n+1$ representing the year of the actual measurement period. Note that in the case of index modeling without detrending, equation (15) is reduced to

$$\bar{I}_t^n(\tau_1, \tau_2) = \sum_{k=\tau_1}^{t-1} T_{k,n+1} + \frac{1}{n} \sum_{i=1}^{n} \sum_{k=t}^{\tau_2} T_{k,i}. \tag{16}$$

3.2. Weather forecasts

Figure 4 shows the evolution of the market prices of one specific HDD futures contract accompanied by different index modeling prices. By looking at Figure 4 it seems apparent that

Figure 4: Seasonal winter contract for Chicago 2005–2006. The solid line indicates the daily settlement price at CME, the dotted and dashed lines represent index modeling prices for 5, 10, 20, and 30 years without detrending and with detrending, respectively.



the CME prices are ahead of the index modeling prices by a few days. While market participants have short-term and mid-term weather forecasts available at the time of trading, our index modeling approach so far uses only the temperature data available until the day of trading. Thus, if we wish to study which index modeling approach with how many years of historical data can best explain market prices, we have to include historical weather forecasts to our analysis.

The National Weather Service (NWS) issues forecasts for the daily minimum and maximum temperatures for each day up to seven days in advance and updates its forecast each hour on the hour. We obtained the historical forecasts from the National Digital Forecast Database; unfortunately, the data is only available after 06/06/2004. When pricing a contract on a certain day, we use the forecasts issued at 3:00 p.m. Central Time[2], which is the last forecast available for public before the CME halts trading from 3:15 p.m. to 5:00 p.m. Since the NWS issues point forecasts for the daily minimum and maximum temperatures, we redefine the index modeling price (15) as

$$I_t^n(\tau_1, \tau_2) = \sum_{k=\tau_1}^{t-1} T_{k,n+1} + \sum_{k=t}^{t+7} \hat{T}_{k,n+1}(t) + I_t^n(t+8, \tau_2), \tag{17}$$

given the temperature forecasts $\hat{T}_{k,n+1}(t) = \frac{1}{2}(\hat{T}^{\max}_{k,n+1}(t) + \hat{T}^{\min}_{k,n+1}(t))$ for forecast days $k = t, t+1, \ldots, t+7$ in the current year $n+1$ issued at day t.

[2]In the cases, where there are no forecasts available at that time due to failures at the NWS, we use the last available forecasts before 3:00 p.m.

3.3. Deviation of theoretical and market prices

In a first descriptive analysis, we consider the mean squared differences of the index modeling prices $I_t^{n,j}$ and the market prices M_t^j for each traded contract j on day t:

$$\text{MSE}(n) = \frac{1}{\#j,t} \sum_{j,t} \left(M_t^j - I_t^{n,j} \right)^2 ,$$

and analogously without detrending. Since we do not have historical temperature forecasts before 06/06/2004, we restrict our further analysis in this section to temperature futures traded after this date.

Table 2: Mean squared deviation of the prices observed at CME from index modeling prices with different numbers of years used. We analyze index modeling without detrending (wd), and index modeling with linear detrending (ld). The minimum is obtained without detrending using 10 years of data.

n	**MSE(wd)**	**MSE(ld)**	n	**MSE(wd)**	**MSE(ld)**	n	**MSE(wd)**	**MSE(ld)**
1	6091.9	—	11	1212.8	2245.8	21	1395.7	1463.1
2	4991.1	—	12	1144.6	2497.0	22	1346.0	1902.6
3	3084.8	25606.8	13	1106.5	2372.5	23	1432.5	1579.3
4	2140.5	15880.2	14	1121.7	1582.0	24	1505.2	1966.9
5	1235.8	11568.2	15	1190.9	1543.7	25	1540.3	2079.8
6	1152.4	7481.1	16	1264.3	1563.8	26	1683.1	2308.1
7	1330.4	7383.5	17	1281.1	1469.0	27	2040.5	3163.7
8	1589.4	5867.7	18	1242.7	1408.3	28	2448.7	4375.1
9	1472.7	3527.3	19	1298.3	1266.9	29	2553.9	4364.5
10	1076.3	3048.6	20	1311.4	1353.5	30	2425.9	3160.1

The results of this analysis are displayed in Table 2 which shows that the minimal squared difference of 1076.3 between the market prices and the index modeling prices is obtained using exactly 10 years of historical data with no detrending. With a mean value of the market prices of 1131.21 and a mean absolute error of 20.37, we calculate an average error of approximately 1.80% for the parsimonious index modeling pricing using 10 years of data with no detrending.

3.4. Time series approach

Until now, we relied on the fact that market participants only have weather forecasts for the next seven days available. In reality however, temperature forecasts are available for a longer time horizon. Although the NWS does not issue point forecasts for more than the next seven days, it does indicate whether the temperatures of days 8–14 ahead are expected to be above or below the climatological average. While including such forecasts into the pricing is not as trivial as with point forecasts, we can still expect market participants to gain an advantage from these forecasts and include the information into the prices. A general overview of the use of weather forecasts in the pricing of weather derivatives is given in Jewson et al. (2002). While embedding weather forecasts into daily simulation models is straightforward, Jewson and Caballero (2003b) show how probabilistic weather forecasts can be incorporated into an index modeling valuation approach.

By refining the deviation analysis from above, we restrict ourselves to index modeling using 10 years without detrending. Using inductive statistics, we investigate if the market prices are lagged to our index modeling prices. If the market prices are lagged by several days, this indicates that the market participants have further temperature forecasts available. To analyze this question, we consider the differentiated time series of market prices $\nabla M_t = M_t - M_{t-1}$ and index modeling prices $\nabla \bar{I}^{10}_{t+l} = \bar{I}^{10}_{t+l} - \bar{I}^{10}_{t+l-1}$ using 10 years without detrending with lag l days. Note that from equation (17) it follows that

$$\begin{aligned}\nabla \bar{I}^n_t &= T_{t-1,n+1} - \hat{T}_{t-1,n+1}(t-1) \\ &+ \sum_{k=t}^{t+6} \left(\hat{T}_{k,n+1}(t) - \hat{T}_{k,n+1}(t-1) \right) \\ &\quad + \hat{T}_{t+7,n+1}(t) - \frac{1}{n} \sum_{i=1}^{n} T_{t+7,i},\end{aligned} \tag{18}$$

i.e., the differentiated index modeling value at day t is driven by the difference of the real temperature of the last day and the predicted temperature for the last, the difference of the temperature forecasts for the next six days, and the difference between the temperature forecast $\hat{T}_{t+7,n+1}(t)$ and the average of the temperature indices of the same day of the past n years. Thus, assuming the market follows an index modeling approach, we do not expect the differentiated market prices ∇M_t to become very high. We therefore consider all differentiated market prices where the daily change is above three times the standard deviation $\hat{\sigma}_{\nabla M} = 16.65$ as outliers (which occurs at 105 of 6291 data points) and exclude these from further analysis.[3]

Next, we calculate the cross-correlation between the differentiated market prices and the differentiated index modeling prices. Since daily temperatures exhibit a strong autocorrelation for the first few days (Jewson and Caballero, 2003a), we also expect the differentiated index modeling price time series (18) to exhibit autocorrelation for the first few days. Therefore, we need to pre-whiten the series $\nabla \bar{I}^{10}_t$ (Box et al., 1994). We do this by removing an autoregressive process that minimizes the corrected Akaike information criterion $\mathrm{AIC_C}$ (Hurvich and Tsai, 1989). However, for computational simplicity, we do not use the maximum likelihood estimator for the white noise variance of the AR-process, but the least squares residual variance as suggested in Tsay (1984). We find that for the series series $\nabla \bar{I}^{10}_t$ an AR(29)-process minimizes the $\mathrm{AIC_C}$. While we can calculate index modeling prices for each day, market prices are not available for certain days, such as weekends or days when there was no trade due to the illiquidity of the weather market. When dealing with incomplete time series, there are various methods to compute empirical cross-correlations (Little and Rubin, 1987). We use the complete cases approach, which considers for lag l only those $\nabla_{\mathrm{pw}} \bar{I}^{10}_{t+l}$, $t \in [\tau_1 + 1, \tau_2 - l]$, where ∇M_t, $t \in [\tau_1 + 1, \tau_2 - l]$, is defined.

[3] Note that the three sigma rule is valid for all unimodal distributions (cf. Vysochanskiĭ and Petunin, 1980).

The empirical cross-correlation is then computed as

$$\hat{\rho}(l) = \frac{\sum_t (\nabla M_t - \overline{\nabla M})(\nabla_{\text{pw}} \bar{I}^{10}_{t+l} - \overline{\nabla_{\text{pw}} \bar{I}^{10}})}{\hat{\sigma}_{\nabla M} \hat{\sigma}_{\nabla_{\text{pw}} \bar{I}^{10}}},$$

where $\overline{\nabla M}$ and $\overline{\nabla_{\text{pw}} \bar{I}^{10}}$ denote the mean of the differentiated market prices and the pre-whitend differentiated index modeling prices, respectively, and $\hat{\sigma}_{\nabla M}$, $\hat{\sigma}_{\nabla_{\text{pw}} \bar{I}^{10}}$ denote the corresponding empirical standard deviations (all depending on l due to the incomplete time series mentioned above).[4]

Table 3 lists the computed cross-correlations of the differentiated CME prices with pre-whitened differentiated index modeling prices using 10 years without detrending. Note that the series exhibit highly significant positive cross-correlations up to a lag of three to four days.

Table 3: Cross-correlation of differentiated CME prices with pre-whitened differentiated index modeling prices using 10 years without detrending. *, **, and *** represent significance on a 10%, 5%, and 1% level, respectively.

Lag	Correlation	Lag	Correlation	Lag	Correlation
0	0.2838***	10	0.0058	20	0.0088
1	0.1676***	11	−0.0120	21	−0.0281
2	0.0829***	12	−0.0366**	22	0.0179
3	0.0712***	13	0.0034	23	0.0699***
4	0.0271**	14	0.0123	24	0.0595***
5	0.0028	15	0.0132	25	0.0196
6	0.0290**	16	−0.0014	26	0.0053
7	−0.0005	17	0.0151	27	0.0292
8	0.0138	18	0.0212	28	0.0037
9	−0.0213	19	0.0026	29	−0.0177

In addition to the seven day temperature forecasts that we include into the index modeling prices, this suggests that the market participants have reliable forecasts available for the next 10–11 days. The correlations of lags 23 and 24 indicate that the prices are also influenced by weather forecasts for the next 30–31 days, which are exactly one month ahead forecasts.

McCollor and Stull (2009) study the quality of mid-term weather forecasts for western North American stations using various verification measures. As equation (18) indicates, the daily change in the price of weather derivatives is driven by the changes in the temperature forecasts for the next seven days plus the difference of the actual temperature on the particular day and the long-term average temperature, which is also referred as climatological temperature. Meteorologists call this difference *temperature anomaly* on the particular day. Using the continuous ranked

[4] Unfortunately, the volume data provided by CME does not include block trades. Due to the illiquidity of the weather market, we cannot guarantee that contracts were actually traded on days t with $\nabla M_t = 0$, i.e., days of which the settlement price provided by CME does not change. In order to ensure that only traded prices were considered, we exclude all market prices M_t where $\nabla M_t = 0$ from our analysis. This also ensures the (weakly) stationarity of the series ∇M_t.

probabilistic skill score (CRPSS), McCollor and Stull (2009) conclude that the forecast skill for temperature anomalies diminishes between one and twelve days in advance, while the forecast quality for 13 through 15 days is negligible. This ties in very well with the observations from above. McCollor and Stull (2009) also show that the CRPSS monotonously increases the nearer the forecast day comes, i.e., the weather forecast becomes sounder with each day. When analyzing the RMSE of weather forecasts, Szunyogh and Toth (2002) come to the same conclusion. This coincides with the fact that all cross-correlations for lags less than 4 are significantly positive in Table 3.

A similar observation can be found in Roll (1984). Analyzing the prices of orange juice futures, Roll found that the prices incorporate weather predictions exceeding the three day forecasts that were issued by the NWS at these times.

We have seen clear indication in this section that market prices of weather derivatives can be explained well using an index modeling approach with 10 years of data and no detrending and include weather forecasts up to 11 days ahead. Yet, the results of subsection 2.2 show that this pricing method is biased.

4. Trading strategies

In this section we develop trading strategies for weather futures relying on the fact that their pricing is biased and check whether the strategies would have paid off in the past.

Although according to equation (2) the futures values are set to be zero at the beginning of the contract, market participants have to deposit a certain margin of the notional value at the clearinghouse. The margin requirements for the different monthly weather futures are listed in Table 4.

For all trading strategies, we use the following general setting:

Firstly, we only use monthly contracts. Due to the additivity of HDD and CDD indices, the price of a seasonal contract must equal the sum of the price of the corresponding monthly contracts for the market to be arbitrage free.

Secondly, for each specific contract, we take the average of all CME prices between 10 and 20 days prior to the measurement period. We use the minimum of 10 days ahead of the measurement period to guarantee that the influence of weather forecasts remains low; we choose the maximum of 20 days because according to an index modeling approach the prices should remain constant ahead of the measurement period. Due to discounting effects there is no reason to buy the contracts too early. Because of the relatively high illiquidity of the weather market not all historical contracts were traded in the specified time frame. In our analysis, we henceforth exclude the contracts that were not traded in that period. This yields 283 different monthly contracts. From these contracts, we discard all contracts whose market price M differs from the index modeling price $\bar{I}^{10}$ by more than 20%. Greater differences are probably caused by single market imbalances, and it is highly improbable that a speculator could exploit those imbalances by buying considerable amounts of such contracts in practice. This is the case for 9 contracts,

Table 4: Margin requirements for monthly weather futures at CME as of July 2008. The initial margins are 135% of the maintenance margins.

	HDD		CDD	
City	**Initial**	**Maintenance**	**Initial**	**Maintenance**
Atlanta	8.10%	6.00%	7.56%	5.60%
Baltimore	6.21%	4.60%	9.45%	7.00%
Boston	4.19%	3.10%	6.48%	4.80%
Chicago	6.48%	4.80%	14.85%	11.00%
Cincinnati	7.02%	5.20%	12.15%	9.00%
Dallas	10.80%	8.00%	5.40%	4.00%
Des Moines	6.08%	4.50%	12.15%	9.00%
Detroit	7.02%	5.20%	12.15%	9.00%
Houston	11.75%	8.70%	5.40%	4.00%
Kansas	6.75%	5.00%	9.45%	7.00%
Las Vegas	7.29%	5.40%	5.40%	4.00%
Minneapolis	5.54%	4.10%	12.15%	9.00%
New York City	5.40%	4.00%	9.45%	7.00%
Philadelphia	6.21%	4.60%	9.45%	7.00%
Portland	6.35%	4.70%	22.95%	17.00%
Sacramento	6.48%	4.80%	12.15%	9.00%
Salt Lake City	10.80%	8.00%	14.85%	11.00%
Tuscon	9.59%	7.10%	5.40%	4.00%

leaving 274 different contracts for our trading analysis.

Thirdly, to each month we allocate a certain wealth to be invested. We choose the contracts in which we invest depending on the particular trading strategy and invest our monthly wealth equally in the selected contracts before the beginning of the measurement period. The futures contracts are then usually held until expiration. However, in order to limit possible losses, we square a contract once a loss exceeds 100%. Note that due to the high leverage in the futures contracts, it can still be possible to generate losses far exceeding 100% if the difference of the prices between two trading days is sufficiently high.

For each month, we calculate the return on the futures contracts as the average of the returns on margin of each single contract, which is

$$\text{ROM} = \frac{\text{final settlement price} - \text{initial market price}}{\text{initial margin} \times \text{initial market price}}. \tag{19}$$

Calculating the return with equation (19) simplifies the real world by neglecting the possibility of margin calls. However, this approximation is not critical as long as the margin account has more cash inflows than outflows.

In order to evaluate the different outcomes of the trading strategies, we compute the monthly Sharpe ratio (Sharpe, 1994), which each strategy performed in the past. We use yields on Treasury nominal securities as provided by Thomson Datastream. As a benchmark for a futures trading strategy, we compute the monthly Sharpe ratio of futures on the S&P 500 index using a fixed initial margin of 112.5 (the margin as of March 2009 at CME) and again neglecting possible margin calls. In the period from June 2002 through September 2006, the S&P 500 futures yields a monthly Sharpe ratio of 4.2%. Investing in monthly S&P 500 futures yields an average monthly return of 1.34% with a monthly standard deviation of 28.0%, and a total return over the period from June 2002 to September 2006 of 69.9%.

4.1. Buy/sell and hold all

Given that there is a general warming trend displayed in temperatures at all weather stations, the simplest possible strategy would be the following: A warming trend in temperatures leads to more CDD in the summer and less HDD in the winter. Because the market prices weather futures without taking this effect into account and the no detrending model describes market prices best, we expect CDD futures to be underpriced and HDD futures to be overpriced. Therefore, the trading strategy consists of buying all available CDD futures slightly before the beginning of the measurement period and selling all available HDD futures before the beginning of the measurement period.

Table 5 shows the monthly returns of this trading strategy. The strategy yields a mean monthly return of 54.1% with a standard deviation of 222.7%. Since in three of the 52 months there was no trade observed in the regarded period, we have a total return of 2,652.7%. When testing, whether this strategy yields an excess return to the S&P 500 index, we use the hypothesis $H_0 : \delta\mu \leq 0$ against $H_1 : \delta\mu > 0$. Due to the relatively small sample, we use the adjusted bootstrap percentile method from Efron (1987), and obtain a p-value of 1.7%, proving that the strategy generates significant excess returns. The strategy yields a monthly Sharpe ratio of 24.3%.

4.2. Buy/sell and hold expected wins

As we show in subsection 2.2, we can obtain an unbiased estimator for the temperature index by using index modeling I^{30} prices with 30 years of data and detrending. Usually for CDD indices, we expect the market price M, to be lower than the I^{30} price, because they follow to a large degree the index modeling $\bar{I}^{10}$ price with 10 years of data without detrending, as has been shown in section 3. Similarly, for HDD indices, we expect the market price M to be higher than the I^{30} price. However, there are cases in which the market price is above the I^{30} price for CDD indices and below the I^{30} price for HDD indices. These are cases in which we would expect to lose money when buying CDD futures and selling HDD futures. Therefore we refine strategy 4.1 by buying only those CDD futures with $M < I^{30}$ and selling only those HDD futures with $M > I^{30}$.

Table 6 shows the monthly returns of this trading strategy. The strategy yields a mean monthly return of 70.2% with a standard deviation of 250.7%. For this strategy, there were 8 months without suitable trading opportunities, which yields a total return of 3,089.3%. Testing again for

Table 5: Mean returns of a buy/sell and hold all strategy. The values in parentheses indicate the number of contracts traded for each month.

	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
2002						144.4% (3)	155.6% (5)	−40.0% (7)	90.9% (4)		−56.6% (9)	−59.7% (9)
2003	−40.8% (8)	−103.3% (8)	−121.9% (7)	56.9% (4)	−52.2% (3)	−84.1% (8)	−27.4% (5)	−13.9% (4)	−154.7% (4)	1000.1% (3)	−114.5% (4)	35.8% (8)
2004	−56.0% (5)	−142.3% (5)	348.7% (3)	−112.1% (2)	301.0% (5)	−114.3% (9)	−107.8% (1)	−84.5% (4)	180.6% (7)		122.8% (4)	−37.0% (4)
2005	−96.0% (5)	53.0% (6)	−97.7% (5)	102.7% (2)	−121.6% (2)	331.6% (7)	49.8% (5)	62.8% (1)	648.6% (2)		188.3% (3)	−111.7% (8)
2006	426.9% (13)	−75.3% (12)	19.6% (12)	439.0% (5)	−34.0% (6)	−25.6% (8)	100.5% (9)	43.8% (5)	−166.0% (6)			

Table 6: Mean returns of a buy/sell and hold expected win strategy. The values in parentheses indicate the number of contracts traded for each month.

	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
2002						−11.5% (1)	121.2% (3)	−10.0% (5)	96.8% (2)		−27.1% (6)	−101.0% (6)
2003	−40.8% (8)	−110.0% (7)	−121.9% (7)	74.4% (2)	−52.2% (3)	−128.8% (5)	−5.3% (4)	−85.0% (2)	−154.7% (4)	1000.1% (3)		144.4% (1)
2004	−56.0% (5)	−142.3% (5)	512.4% (1)		242.4% (3)	−162.9% (2)		−147.0% (3)	180.6% (7)		−100.2% (1)	−58.4% (2)
2005	−96.0% (5)	53.0% (6)	−68.2% (3)		−137.7% (1)	329.1% (6)	216.5% (1)	62.8% (1)	707.4% (1)		183.7% (2)	−111.7% (8)
2006	457.5% (12)	−57.2% (11)	75.0% (5)	559.0% (3)	−6.6% (2)	37.4% (3)	194.4% (2)		−166.8% (1)			

an excess return, we get a p-value of 1.3%, proving that the strategy generates significant excess returns. The strategy yields a monthly Sharpe ratio of 28.0%.

4.3. Buy/sell and hold significant wins

In strategy 4.2 we invest long in all CDD futures whose I^{30} price is above the market price and short in all HDD futures whose I^{30} price is below the market price. This strategy uses a point estimate for the I^{30} price. When using the notation of section 2, it is well known that the standardized error of the forecast is distributed as

$$\frac{y_0 + \varepsilon_{n+1} - \hat{y}_0}{\hat{\sigma}\sqrt{1 + \boldsymbol{x}_0'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_0}} \sim \mathrm{t}(n-2),$$

where $\hat{\sigma}^2 = \frac{1}{n-2}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})$ is the unbiased estimate of σ^2 and $\boldsymbol{Y}$, $\boldsymbol{X}$, $\boldsymbol{\beta}$, and $\boldsymbol{x}_0$ are specified as in equation (4). Defining

$$t_M = \frac{M - \hat{y}_0}{\hat{\sigma}\sqrt{1 + \boldsymbol{x}_0'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_0}}$$

to be the t-value of the standardized market price in the I^{30}-model, we can define $p_M = \mathrm{T}_{n-2}(t_M)$, where T_{n-2} denotes the cumulative distribution function of the t distribution with $n-2$ degrees of freedom. Note that the one-sided prediction interval with confidence level $1-\alpha$ for the future observation $y_0 + \varepsilon_{n+1}$ in the model (3) is given by

$$\mathbb{P}\left(\frac{y_0 + \varepsilon_{n+1} - \hat{y}_0}{\hat{\sigma}\sqrt{1 + \boldsymbol{x}_0'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_0}} \leq \mathrm{T}_{n-2}^{-1}(1-\alpha)\right) = 1 - \alpha.$$

Plugging in $1 - \alpha = p_M$ yields

$$p_M = \mathbb{P}\left(\frac{y_0 + \varepsilon_{n+1} - \hat{y}_0}{\hat{\sigma}\sqrt{1 + \boldsymbol{x}_0'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_0}} \leq t_M\right) = \mathbb{P}(y_0 + \varepsilon_{n+1} \leq M),$$

so p_M can be interpreted as the probability that the realized index value $y_0 + \varepsilon_{n+1}$ is below the market price M. The strategy now consists of choosing a critical probability p^* in a first step and then buying all CDD contracts with $p_M < p^*$ and selling all HDD contracts with $p_M > 1 - p^*$. Notice that choosing $p^* = 1$ corresponds to the buy/sell all strategy 4.1, whereas choosing $p^* = .5$ equals the buy/sell expected wins strategy 4.2.

The choice of p^* is crucial to the success of the trading strategy and yet it yields a classical trade-off between the expected return and the availability of the contracts. When making p^* too large, we buy contracts where our model expects only a small return (or even a negative return for contracts with $p^* > .5$). By making p^* too small the total return of the strategy may be limited due to the lack of matching contracts.

To find an adequate value for p^*, we start using the same virtual contracts as in subsection 2.2. For each virtual contract, we calculate the I^{30} price and the $\bar{I}^{10}$ price and use the latter as a proxy for the virtual market price. We then calculate the total returns of trading strategies with $p^* = 0, .01, .02, \ldots, 1$ for the period from 1983 through 1999. As Table 7 shows, this analysis yields a maximum total return from 1983 to 1999 of 6,095.6% which can be achieved by selecting $p^* = .44$.

Table 7: Total return from virtual temperature futures from 1983 to 1999 as described in subsection 2.2 for different values of p^* using strategy 4.3. The values in parentheses indicate the number of months in which trading has occurred. We only indicate the returns for $p^* \leq .5$, since contracts with higher p^*-values would indicate expected losses and thus should not be included in a successful trading strategy.

p^*	**Total return**	p^*	**Total return**
$\leq$ 0.27	0.00% (0)	0.39	1,429.36% (98)
0.28	225.93% (1)	0.40	747.93% (121)
0.29	272.42% (2)	0.41	3,038.23% (136)
0.30	−101.07% (4)	0.42	5,213.49% (149)
0.31	−277.90% (4)	0.43	4,153.99% (159)
0.32	339.30% (8)	0.44	6,095.65% (163)
0.33	981.07% (17)	0.45	5,906.11% (166)
0.34	375.63% (22)	0.46	5,978.98% (168)
0.35	626.65% (34)	0.47	5,173.30% (168)
0.36	1,739.05% (47)	0.48	4,416.66% (169)
0.37	2,574.15% (61)	0.49	4,408.88% (170)
0.38	35.35% (79)	0.50	5,162.10% (170)

Having calibrated $p^* = .44$ with these virtual contracts, Table 8 shows the monthly returns of the trading strategy with the CME prices. The strategy yields a mean monthly return of 102.7% with a standard deviation of 483.0%. Although the strategy is rather selective when choosing fitting contracts and did not trade in 17 of the 52 months, it still generates a total return of 3,593.1%. When testing again for a positive return, we obtain p-value of 1.6%, proving again that the strategy generates significant excess returns. The strategy yields a monthly Sharpe ratio of 21.3%.

4.4. Comparison

Table 9 summarizes the results of the different trading strategies. Compared with trading traditional stock index futures, weather derivatives can generate enormous returns for speculators. Although these returns are traded off by high volatility, weather derivatives' speculation can still be of interest from a reward-to-variability ratio perspective, which exceeds the stock index futures' by four to six times. When testing whether the Sharpe ratios of the trading strategies exceed the S&P 500 Sharpe ratio, we rely on the one-sided Ledoit and Wolf (2008) test, which is robust against the underlying distribution and rather strict compared to the traditional Jobson and

Table 8: Mean returns of a buy/sell and hold significant strategy with $p^* = 0.44$. The values in parentheses indicate the number of contracts traded for each month.

	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
2002						−11.5% (1)	121.2% (3)	−22.5% (1)	45.2% (1)		−12.0% (5)	−127.4% (1)
2003	−31.1% (7)	−113.6% (4)	−79.5% (4)	74.4% (2)	−186.8% (2)	−134.9% (4)	34.4% (3)	−106.1% (1)	−159.5% (2)	2674.9% (1)		
2004	−56.0% (5)	−136.8% (4)	512.4% (1)		418.2% (2)	−187.3% (1)		−91.0% (1)	126.7% (5)			
2005	−93.1% (4)	130.8% (1)			−137.7% (1)	374.6% (3)		62.8% (1)			183.7% (2)	−111.7% (8)
2006	457.5% (12)	−51.8% (10)	−11.4% (2)			112.7% (2)	125.0% (1)					

Table 9: Overview of the different trading strategies for the period of June 2002–September 2006. The values in parentheses indicate the p-values of the one-sided Ledoit and Wolf (2008) test whether the corresponding Sharpe ratio is less or equal to the S&P 500 Futures Sharpe ratio. ** indicates significant excess returns on a 5% level.

	S&P 500 Futures	**Strategy 4.1** $p^* = 1$	**Strategy 4.2** $p^* = .5$	**Strategy 4.3** $p^* = .44$
Mean Monthly Return	1.3%	54.1%**	70.2%**	102.7%**
Standard Deviation	28.0%	222.7%	288.1%	483.0%
Sharpe Ratio	4.2%	24.3% (10.3%)	28.0% (13.3%)	21.3% (14.3%)
Total Return	69.9%	2,652.7%	3,089.3%	3,593.1%

Korkie (1981) test. Unfortunately, the tests yield slight insignificances for the outperformance of the Sharpe ratios, which is most probably due to the relatively short time series in our analysis.

When using the Treynor (1965) ratio as a performance measure, it is clear from the previous results that an investor can increase a portfolio's performance by including weather derivatives. Although some weather properties like cloudiness might influence stock markets returns (see e.g. Hirshleifer and Shumway, 2003; Chang et al., 2008), we can generally assume a Beta of approximately zero for weather derivatives with the market portfolio.[5] Thus, adding weather derivatives according to strategy 4.1, 4.2, or 4.3 decreases a portfolio's Beta and probably increases its expected return at the same time, yielding an even more distinct outperformance of the S&P 500 futures strategy for the Treynor ratio upon being compared to the Sharpe ratio.

However, the opportunity to make money with such a strategy is limited because the number of traded contracts is generally rather low for a specific single contract within our investigation period. Moreover, our time series ends in 2006. Since then the market has become more mature and has increased in volume. Thus, the extent of mispricing identified may have changed.

None of our trading strategies invests money in the months in which there is no suitable weather contract available. For testing purposes, we alter the strategies by investing in treasury bonds in these months. This leads to reduced mean monthly returns, reduced standard deviations of the monthly returns and slightly lower Sharpe ratios. However, the main results from above remain valid.

5. Summary and conclusion

This paper is the first one of its kind to analyze the prices of CME temperature (HDD and CDD) futures contracts for major U.S. cities from a theoretical and an empirical point of view. The results of our study can be summarized as follows.

[5] Moreover, the causation of the weather whatsoever for the observed seasonal patterns in stock market returns has been challenged recently, see e.g. Jacobsen and Marquering (2008); Kamstra et al. (2009); Jacobsen and Marquering (2009).

Since it is the most likely approach to be observed in practice, we restrict ourselves to index modeling as a pricing mechanism. We assume a linear trend over time in the temperature indices under consideration as the correct model, including the possible case of a trend equal to zero. We then theoretically analyze two variants of index modeling, namely simple averaging over the last n years without detrending and an estimation using a linear detrending routine. For both variants the number of years n used is parametric and can be varied in principle. However, there are two limitations to increasing n arbitrarily, the first being the limited availability of data, and the second being the fact that by dating too far back the trend, if existent, may not be linear anymore. Generally, such a trend tends toward higher temperatures implying less HDD in winter months and more CDD in summer months. From our theoretical analysis we derive the result that detrending is better than no-detrending with respect to the MSE the more significant the slope of the trend is compared to the standard deviation and the higher the number of years n used. However, the latter is subject to the afore-mentioned restrictions on n. Since we have no a priori knowledge of the size of the trend, we carry out an empirical analysis using virtual weather futures, to determine whether the linear model is suitable for those weather stations whose real futures are traded on the CME. It appears that the model is highly suitable. Albeit the trend is, on average, not strong enough to make detrending a clearly superior model from an MSE perspective in general, detrending is clearly the preferred model with respect to the mean error since it leads to a bias of approximately zero whereas no-detrending exhibits a significant bias.

In this paper we also investigate which index model can explain real market prices. With an average error of only 1.8%, the parsimonious index modeling approach with no detrending processing 10 years of data explains the market behavior exceptionally well; this is a model that also performs very well in the theoretical analysis from an MSE perspective. The proper fit of the model to the market prices also gives an ex-post justification for the strict expectations hypothesis established at the beginning. Moreover, market participants appear to use meteorological temperature forecasts up to 10-11 days ahead. This has strong implications for further research when estimating the market price of risk in a daily simulation model. Using historical weather and market data while neglecting weather forecasts could result in severe misinterpretations.

Finally, we investigate whether one can devise a profitable trading strategy from the fact that the futures are priced with a bias. We develop three differently sophisticated trading strategies based on unbiased pricing using detrending and on investing every month in monthly futures. Each of these three strategies not only yields high overall returns; they also perform well on a risk-adjusted basis if compared with a stock-index futures trading strategy. Therefore, we conclude that the observed pricing according to index modeling without detrending and with 10 years of data leads to the possibility of generating abnormal returns if it is traded against.

References

Alaton, P., Djehiche, M., and Stillberger, D., 2002. On modelling and pricing weather derivatives. Applied Mathematical Finance 9, 1–20.

Benth, F. E. and Šaltytė-Benth, J., 2005. Stochastic modelling of temperature variations with a view towards weather derivatives. Applied Mathematical Finance 12, 53–85.

Benth, F. E. and Šaltytė-Benth, J., 2007. The volatility of temperature and pricing of weather derivatives. Quantitative Finance 7, 553–561.

Benth, F. E., Šaltytė-Benth, J., and Koekebakker, S., 2007. Putting a price on temperature. Scandinavian Journal of Statistics 34, 746–767.

Boissonnade, A. C., Heitkemper, L. J., and Whitehead, D., 2002. Weather data: Cleaning and enhancement. Climate Risk and the Weather Market, Risk Books 73–98.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C., 1994. Time Series Analysis. Forecasting and Control. Prentice Hall: Englewood Clifs.

Brody, D. C., Syroka, J., and Zervos, M., 2002. Dynamical pricing of weather derivatives. Quantitative Finance 2, 189–198.

Caballero, R., Jewson, S., and Brix, A., 2002. Long memory in surface air temperature: Detection, modeling, and application to weather derivative valuation. Climate Research 21, 127–140.

Campbell, S. D. and Diebold, F. X., 2005. Weather forecasting for weather derivatives. Journal of the American Statistical Association 100, 6–16.

Cao, M. and Wei, J., 2000. Pricing the weather. Risk 67–70.

Cao, M. and Wei, J., 2004. Weather derivatives valuation and market price of weather risk. Journal of Futures Markets 24, 1065–1089.

Chang, S.-C., Chen, S.-S., Chou, R. K., and Lin, Y.-H., 2008. Weather and intraday patterns in stock returns and trading activity. Journal of Banking & Finance 32, 1754–1766.

Cotton, W. R. and Pielke, R. A., 2007. Human Impacts on Weather and Climate. Cambridge University Press: Cambridge.

Davis, M., 2001. Pricing weather derivatives by marginal value. Quantitative Finance 1, 1–4.

Dischel, B., 1998. At last: A model for weather risk. Energy & Power Risk Management 11, 20–21.

Dornier, F. and Queruel, M., 2000. Caution to the wind. Energy & Power Risk Management 13, 30–32.

Dutton, J. A., 2002. Opportunities and priorities in a new era for weather and climate services. Bulletin of the American Meteorological Society 83, 1303–1311.

Efron, B., 1987. Better bootstrap confidence intervals. Journal of the American Statistical Association 82, 171–185.

Hamisultane, H., 2009. Utility-based pricing of weather derivatives. The European Journal of Finance , forthcoming.

Härdle, W. K. and López Cabrera, B., 2009. Implied market price of weather risk. SFB 649 Discussion Paper 2009, from: http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2009-001.pdf.

Harrod, M., Larsen, P. H., Lazo, J. K., and Waldman, D. M., 2007. Sensitivity of the U.S. economy to weather variability. NCAR Societal Impacts Program, Boulder, Colorado working paper .

Hirshleifer, D. A. and Shumway, T., 2003. Good day sunshine: Stock returns and the weather. The Journal of Finance 58, 1009–1032.

Hurvich, C. M. and Tsai, C.-L., 1989. Regression and time series model selection in small samples. Biometrika 76, 297–307.

Intergovernmental Panel on Climate Change, 2007. Fourth assessment report: Climate change 2007: The AR4 synthesis report. IPCC: Geneva, from: http://www.ipcc.ch/ipccreports/ar4-wg1.htm.

Jacobsen, B. and Marquering, W., 2008. Is it the weather? Journal of Banking & Finance 32, 526–540.

Jacobsen, B. and Marquering, W., 2009. Is it the weather? Response. Journal of Banking & Finance 33, 583–587.

Jewson, S. and Brix, A., 2005. Weather Derivative Valuation: The Meteorological, Statistical, Financial and Mathematical Foundations. Cambridge University Press: Cambridge.

Jewson, S. and Caballero, R., 2003a. Seasonality in the statistics of surface air temperature and the pricing of weather derivatives. Meteorological Applications 10, 367–376.

Jewson, S. and Caballero, R., 2003b. The use of weather forecasts in the pricing of weather derivatives. Meteorological Applications 10, 377–389.

Jewson, S. and Penzer, J., 2004. Closed form expressions for the uncertainty from linear detrending, and the pricing of weather derivatives. SSRN eLibrary .

Jewson, S., Ziehmann, C., and Brix, A., 2002. Use of meteorological forecasts in weather derivative pricing. Climate Risk and the Weather Market, Risk Books 169–183.

Jobson, J. D. and Korkie, B. M., 1981. Performance hypothesis testing with the Sharpe and Treynor measures. The Journal of Finance 36, 889–908.

Kamstra, M. J., Kramer, L. A., and Levi, M. D., 2009. Is it the weather? Comment. Journal of Banking & Finance 33, 578–582.

Ledoit, O. and Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio. Journal of Empirical Finance 15, 850–859.

Little, R. J. A. and Rubin, D. B., 1987. Statistical Analysis with Missing Data. Wiley: New York.

Lucas, R. E., 1978. Asset prices in an exchange economy. Econometrica 46, 1429–1445.

McCollor, D. and Stull, R., 2009. Evaluation of probabilistic medium-range temperature forecasts from the North American ensemble forecast system. Weather and Forecasting 24, 3–17.

Oetomo, T. and Stevenson, M., 2005. Hot or cold? A comparison of different approaches to the pricing of weather derivatives. Journal of Emerging Market Finance 4, 101–133.

Pirrong, C. and Jermakyan, M., 2008. The price of power: The valuation of power and weather derivatives. Journal of Banking & Finance 32, 2520–2529.

PricewaterhouseCoopers, 2006. Weather risk management association 2006 survey results. From: http://www.wrma.org/wrma/library/PwCResultsJune222006.ppt.

Roll, R., 1984. Orange juice and weather. The American Economic Review 74, 861–880.

Schiller, F., Seidler, G., and Wimmer, M., 2008. Temperature models for pricing weather derivatives. SSRN eLibrary .

Sharpe, W. F., 1994. The Sharpe ratio. The Journal of Portfolio Management 21, 49–58.

Szunyogh, I. and Toth, Z., 2002. The effect of increased horizontal resolution on the NCEP global ensemble mean forecasts. Monthly Weather Review 130, 1125–1143.

Treynor, J. L., 1965. How to rate management of investment funds. Harvard Business Review 43, 63–75.

Tsay, R. S., 1984. Order selection in nonstationary autoregressive models. The Annals of Statistics 12, 1425–1433.

Vysochanskiĭ, D. F. and Petunin, Y. I., 1980. Justification of the 3σ rule for unimodal distributions. Theory of Probability and Mathematical Statistics 21, 25–36.

Zapranis, A. D. and Alexandridis, A., 2008. Modelling the temperature time-dependent speed of mean reversion in the context of weather derivatives pricing. Applied Mathematical Finance 15, 335–386.