PRICING OF WEATHER DERIVATIVES

Shih-Ying Lee

A Thesis
In
The John Molson School of Business

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Science in Administration at
Concordia University
Montreal, Quebec, Canada

April 2006

Your file Votre référence
ISBN: 0-494-14370-3
Our file Notre référence
ISBN: 0-494-14370-3

ABSTRACT

Pricing of Weather Derivatives

Shih-Ying Lee

Values of weather derivatives depend on weather outcomes, such as temperature or precipitation. Academics have differed on how to value this type of financial instrument, since weather is not tradeable and the No-Arbitrage Pricing Theory cannot be applied. Cao and Wei (2004) propose a valuation model and I test its predicting accuracy by comparing simulated futures prices to market prices, for cumulative Heating/Cooling Degree Day futures for New York City for contracts offered by the Chicago Mercantile Exchange.

The simulation of weather futures prices requires assumptions of values for the risk aversion parameter. Following Cao and Wei, the values -2, -10 and -40 are used. The simulation requires values for the speed of mean reversion of aggregate dividends. Following Cao and Wei, the values of 0.8, 0.9 and 0.99 are used. Due to the lack of a sufficiently long time series data to determine the daily correlation between temperature and aggregate dividends, Cao and Wei assume that the aggregate dividends depend on either the contemporaneous temperature or the 30 lagged temperatures. There are consequently 18 simulation settings.

Results indicate that Cao and Wei's (2004) model is useful in predicting weather derivative prices, especially when the risk aversion parameter is -10. Forecast accuracy is very sensitive to the risk aversion parameter, followed by the number of temperature lags that aggregate dividends depend on. The speed of mean reversion of aggregate dividends is not found to be a crucial parameter.

ACKNOWLEDGEMENT

I would like to thank Dr. Latha Shanker for her supervision and expertise on derivatives throughout the research. Dr. Fassil Nebebe's knowledge on statistics and help on programming greatly aided me. Special thanks are dedicated to Dr. Ian Rakita for his valuable comments.

To my family I dedicate my thesis. Without their support, I would not be able to finish this work.

And last, I would like to express my gratitude to the John Molson School of Business for providing me with the research grant to help facilitate this study.

TABLE OF CONTENTS

LIST OF APPENDICES

EXECUTIVE SUMMARY

A major risk of an economy comes from weather. Utility companies face the risk of decreasing income when winters are warmer than expected. On a rainy day entertainment resorts receive fewer tourists, yet, umbrella producers have more sales. It is estimated that 14% to 20% of the US economy is susceptible to weather risks.[1] Investors now can hedge them after the introduction of weather derivatives in the 1990s.

Unanticipated weather outcomes in 2005 have increased the industrial needs for weather derivatives. Financial News of Yahoo!Finance reported on November 9, 2005 that the nominal value of weather contracts has doubled from $4.3 billion in 2004 to $8.4 billion in 2005[2]. Chicago Mercantile Exchange (CME) added three new U.S. cities, Baltimore, Detroit and Salt Lake City, to the 15 existing U.S. cities to which temperature derivatives are indexed. Products are also expanded to include frost day derivatives. With their increasing demand, both academics and finance practitioners need to understand more about weather derivatives.

Academics have not agreed on how to value this type of financial instrument. Given that the underlying index of weather derivatives is not tradeable, we can not price weather derivatives by applying the No-Arbitrage Pricing Theory. Cao and Wei

[1] Challis (1999) and http://www.cme.com/trading/prd/env/abtwthder2766.html
[2] I extract information from http://biz.yahoo.com/prnews/051109/phw011.html?.v=35

(2004) propose to price them by applying the Lucas (1978) asset-pricing model on the forecasted temperature indexes.

The purpose of this thesis is to verify the accuracy of Cao and Wei's (2004) pricing model. This thesis is also the first one dedicated to examine the forecast accuracy among weather derivatives pricing models. Given that cumulative Heating/Cooling Degree Day futures are most frequently traded, this research will: (1) address this type of derivative traded on the CME for New York City from June 18, 2002 to July 28, 2005, (2) forecast temperature outcomes as of the valuation date with temperatures observed during the 20-year period preceding the valuation date, (3) forecast the aggregate dividend rate as of the valuation date with aggregate dividend rates observed during the 20-year period preceding the valuation date, (4) forecast futures prices, and (5) verify the accuracy of the simulated prices with both parametric and non-parametric tests.

Results indicate that Cao and Wei's model is useful in simulating prices close to the market prices of temperature futures. Because the previous literature does not indicate a precise value for the public's risk aversion parameter, Cao and Wei assume it to take on the values -2, -10 or -40. Academics are unsure of the speed of mean reversion of the aggregate dividends, and Cao and Wei assume that they can take on the values of 0.8, 0.9 or 0.99. Due to lack of a sufficiently long time series data to determine the daily correlation between temperatures and aggregate dividends, CAO AND WEI assume that aggregate dividends depend on either the contemporaneous temperature or the 30 lagged temperatures. As a result, there are in total 18 simulation settings. Simulated prices are found to be closest to the market prices when the risk aversion parameter is -10. Forecast accuracy is found to be very sensitive to the risk aversion parameter, followed by the number of lagged temperatures that aggregate dividends depend on. The mean-reverting speed of the aggregate dividends is not a critical parameter.

1. INTRODUCTION

1.1 Brief History

Both individuals and businesses are susceptible to weather risks. Severe weather outcomes take lives and damage properties. For a long time the public could hedge weather risks by buying insurance contracts. Insurers pay policyholders values of the damaged properties or the amount of money needed to repair the damaged properties when they are destroyed by various weather events.

As individuals can hedge catastrophic risks with insurance contracts, insurance companies, which have acquired tremendous amounts of risk from their clients, can only share risk with re-insurers. Given that there are only a few re-insurers in the world and they have limited capital, they are unable to absorb all of the insurers' risks. Hurricane Katrina in 2005 proved this point. It incurred $34.4 billion of insured losses [3], as estimated in October 2005, and made 11 companies insolvent [4]. Catastrophes of this magnitude make re-insurers hesitate, in addition to being unable, to absorb risks. Insurers need to seek a secondary risk buffer.

In 1992 the Chicago Board of Trade (CBoT) launched the catastrophe insurance derivative. Its payoffs depended on the ratio of a quarter's insured

---

[3] http://katrinainformation.org/disaster2/facts/katrina_faq/

[4] http://www.informationweek.com/news/showArticle.jhtml?articleID=171203928

catastrophic loss over a quarter's property premium. Although they functioned as reinsurance contracts, state insurance departments did not recognize this. Instead, catastrophe derivatives were viewed as risky assets and insurers had to maintain additional capital to absorb risk inherent in them. Catastrophe derivatives had, therefore, low trading volume and the CBoT ceased offering them in 1999[5].

While insurance can help mitigate catastrophic risks, it is not useful to companies subject to non-catastrophic weather risks. Enron Corp. and Koch Energy, facing increased competition after the mid-1997 deregulation of the energy sector, entered a custom-made swap in which Enron (Koch) agreed to pay Koch (Enron) $10,000 per each degree the temperature fell (above) below normal in the winter 1997 – 1998[6]. The CME launched Heating/Cooling Degree Day derivatives to provide industries a means to hedge non-catastrophic weather risks in September 1999.

1.2 Cumulative Heating/Cooling Degree Day Futures

Among non-catastrophic weather derivatives, cumulative Heating/Cooling Degree Day futures are most frequently traded. They are thus the focus of this thesis. Payoffs to this type of weather derivative depend on the cumulative sum of Heating/Cooling Degree Day during the contract month. The CME defines

[5] *Catastrophe Exposures and Insurance Industry Catastrophe Management Practices*, prepared by American Academy of Actuaries, June 11, 2001
[6] http://www.bizjournals.com/houston/stories/1999/11/22/story7.html

Heating/Cooling Degree Day as the number of degrees Fahrenheit that a daily average temperature is below/above 65°F in a one-day period.

$$HDD = Max(65°F - DailyAvgTemp, 0)$$

$$CDD = Max(DailyAvgTemp - 65°F, 0)$$

A daily average temperature is defined as the arithmetic average of the daily maximum and the daily minimum temperatures. Temperature indexes which Heating/Cooling Degree Day futures are settled at are

Settled index of a Heating Degree Day futures contract

$$= \sum_{t=1}^{\#\cdot of\cdot days\cdot in\cdot a\cdot contract\cdot month} Max(65°F - DailyAvgTemp_{Contract\cdot Month,t}, 0)$$

Settled index of a Cooling Degree Day futures contract

$$= \sum_{t=1}^{\#\cdot of\cdot days\cdot in\cdot a\cdot contract\cdot month} Max(DailyAvgTemp_{Contract\cdot Month,t} - 65°F, 0)$$

Traders taking long positions in the cumulative Heating Degree Day futures of, say, January 2005 in New York City at a price of F_0 are paid

$$Payoff = US\$20 \times (\sum_{t=1}^{31} Max(65°F - DailyAvgTemp_{January\cdot t,2005}, 0) - F_0),$$

where F_0 is the cumulative Heating Degree Days in January 2005 expected on the day traders buy this futures contract. Similarly traders buying the cumulative Cooling Degree Day futures of, say, July 2005 in New York City at a price of F_0 are paid

$$Payoff = US\$20 \times (\sum_{t=1}^{31} Max(DailyAvgTemp_{July\cdot t,2005} - 65°F, 0) - F_0),$$

where F_0 is the cumulative Cooling Degree Days in July 2005 expected on the day traders buy this futures contract. Tick size for temperature futures is set to US$20.00. The CME specifies the locations where daily temperatures are measured for each future contract. Table A-1 shows the specified stations for 18 American cities.

[Please insert Table A-1 about here.]

The Heating (Cooling) Degree Day futures merit their names because consumers are observed to turn on the heat (the air conditioning) when temperatures drop below (increase above) 65°F. Contract months of cumulative Heating Degree Day Futures are January, February, March, April, October, November and December. On the other hand, contract months of cumulative Cooling Degree Day futures range from April to October. We call October and April shoulder months since temperatures in these two months fluctuate around 65°F. Both Heating and Cooling Degree Day derivatives are offered for these two contract months.

The CME specifies that the futures are to be settled in cash on the first exchange business day that is at least two calendar days after the end of the contract month. The minimum price change is set at one Degree Day Index point. Futures contracts are traded on the electronic platform Globex, whereas prices of options on futures are negotiated on the CME floor through the traditional outcry system.

1.3 Major Players

Weather derivatives offered by the CME are temperature indexed, and companies subject to temperature risks are thus targeted hedgers. They include energy-related businesses, agricultural firms, restaurants, companies involved in tourism and travel, and OTC weather derivative traders[7]. These entities enter temperature derivatives to hedge against risks that are either inherited in their business or acquired from doing business. In the next paragraph how hedgers hedge with Heating/Cooling Degree Day derivatives will be illustrated.

Suppose that a utility firm is concerned about a warm January 2006 in New York City and its revenue will drop $50,000 in a day for each degree Fahrenheit higher than $65^{\circ}F$. As Heating Degree Day decreases as temperatures increase, this utility company can lock in revenues by selling 2,500 HDD futures indexed to New York at, say, 1090[8]. If the cumulative sum of Heating Degree Days in January 2006 turns out to be 1065, this firm will lose $1.25 million dollars from its own business but receive $1.25 million dollars from its short derivative position.

Speculators also exist in this market. Consider that a meteorologist who wants to take advantage of her professional knowledge. Assume that she observes that frequent hurricanes in Mexican Bay often lead to cold winters in the New England

---

[7] An Introduction to THE CME Weather Products published by Chicago Mercantile Exchange

[8] 1090 is the average cumulative Heating Degree Day in January observed in the LaGuardia Airport during the period from 1986 to 2005.

region. She can buy a call on Seasonal Heating Degree Day futures indexed to New York in 2005 - 2006. (She chooses New York because it is closest to New England among the 18 cities to which temperature derivatives are indexed.) If the temperature in New York City from November 2005 to March 2006 is lower, she will profit by \$20 for each HDD lower. If it is warmer, she will lose \$20 for each HDD higher.

Arbitrageurs, the last category of derivative traders, in the weather derivative market are investment banks, hedge funds, and traditional weather derivative market makers. Exploiting price differences between contracts traded in the exchanges and the OTC market, several cases of arbitrage were believed to have taken place[9].

1.4 Market Structure

An investor who wants to trade temperature futures will start by calling her broker, and this person will enter the desired price and the desired amount into the electronic trading platform Globex. This platform connects to market makers all over the world, and trades can be executed at the best price. Nevertheless, temperature derivatives are very illiquid and traders pay high bid-asked spreads. Furthermore, the mark-to-market mechanism is a black-box. The CME marks-to-market with an econometric model with a pricing rule[10]. Market makers for temperature derivatives

---

9 http://www.environmental-finance.com/2000/featmar2.htm

10 http://www.astin2004.no/papers/Roustant.pdf

are mostly utility companies and insurers, and they are led in the CME by Wolverine Trading. Options on temperature futures, as discussed before, are traded with the traditional outcry system, and their traders are as well subject to high transaction costs.

Temperatures needed to calculate HDD/CDD indexes are collected by Earth Satellite Corporation (EarthSat). Specializing in remote sensing, it collects data with the so-called Automated Surface Observation System. Earth Satellite Corporation provides extremely accurate data and there have been few discrepancies between their temperatures and the official records posted by the National Climate Data Centre (NCDC) of the U.S. government[11].

1.5 Accounting, Regulation, and Taxes

Because weather derivatives derive their values from non-commodity weather variables, they do not fall under the jurisdiction of Commodity Exchange Act and are not regulated by the Commodity Futures Trading Commission (CFTC). In 2003 the CFTC obtained regulatory jurisdiction for some weather derivatives, and their traders, if they meet the hedging criteria, will be allowed by the Internal Revenue Service (IRS) to use hedge accounting.

---

[11] http://www.financewise.com/public/edit/energy/weather99/wthr99-exchangep.htm

While derivatives regulated by the CFTC are treated as derivatives, other weather derivatives are in the grey zone. State insurance commissions have argued that weather derivatives are a disguised form of insurance and they should have jurisdiction. Given that derivatives and insurance have very different accounting treatments and tax consequences, traders of non-CFTC regulated weather derivatives need to consult professionals on this point.

2. LITERATURE REVIEW

Weather derivatives are traded in an incomplete market and the No-Arbitrage pricing theory can not be applied to price them. Given that weather derivatives derive their values from the underlying weather indexes, pricing them involves two stages: forecasting the weather indexes and calculating derivative prices. These two stages require researchers to develop the temperature model and the pricing model, which will be reviewed in the next section.

2.1 Temperature Forecasting Models

The intuitive source of temperature forecasting models is meteorological works. Models producing temperature forecasts which we see everyday on TV, although acceptably accurate in three-dimensional space, are of a short-term nature. They can mostly produce five-day forecasts. Meteorologists believe it is a mission impossible to produce accurate long-term predictions.

Financial professionals and scholars, on the other hand, are determined to conquer this challenging task. Lacking geo-science background, they believe that weather comes in patterns and it can be predicted by inputting past observations into models. Two groups of models are proposed: autoregressive models (AR) and Brownian motion models (BM). AR models are suggested because they are applied to predict future values with a variable's past values, whereas BM models are suggested because of temperatures' randomness and their mean-reverting nature.

AR-type temperature models consist of four parts: the trend, the seasonality, the autoregressive lag part, and the cyclical error variance. The trend part is designed to capture the well-known observation that temperatures increase with time, in other words the Green House Effect. As a result, this part is a function of time. Platen and West (2004) and Campbell and Diebold (2005) start with a polynomial function and find that only the linear term is significant. Cao and Wei and Roustant et al (2003), on the other hand, aim to streamline the temperature forecast model and fit their data with a linear specification.

The second part of AR-type temperature models is the seasonality, and it is added to control the fact that temperatures are seasonal. They increase in spring, reach the peak in summer, decrease in fall, and plunge in winter. Because Fourier series provide smooth seasonal pattern and numerical stability in estimation, Campbell and Diebold, and Platen and West fit their data with them. Fourier series are the sum of series of sine and cosine functions which are designed to capture phenomena repeating every one, two, three …years. The adequate frequency of the Fourier series represents the longest period that the statistically significant phenomena repeat. If the adequate frequency is three, this means that phenomena repeating from every year to every three years are statistically significant. Platen and West find that their Australian temperatures have an adequate frequency of two, whereas Campbell and Diebold assume it to be three. Cao and Wei, in contrast, want to deal with a simpler model and control seasonality via daily average temperatures. Davis (2001) further smoothes daily average temperatures to remove the fluctuating pattern.

The next item addressed in the AR-type temperature models is the autoregressive lag part. It is used because warm/cold days are often observed to follow warm/cold days. Temperatures observed in Atlanta, Chicago, Dallas, New York and Philadelphia from 1979 to 1998 indicate to Cao and Wei that the optimal lag is three for all five cities. Roustant et al reach the same result with their Chicago and Paris temperatures. Campbell and Diebold, referencing Bloomfield's (1992) results, do not test the data but assume the lag to be 25. Neither does Davis attempt to find the optimal lag but directly assumes it to be one.

Because temperatures are more volatile in winter than in summer, Campbell and Diebold and Cao and Wei control for this fact by conditioning temperature errors. Given that the error variance can be decomposed into the seasonal contribution and the cyclical contribution, Campbell and Diebold use a Fourier series to capture the first part and a generalized autoregressive conditional heteroscedasticity (GARCH) model to capture the second one in their 2005 paper. Cao and Wei do this job with a straightforward trigonometric sine function, whereas Roustant et al construct a model using both sine and cosine functions.

Some researchers, on the other hand, propose BM models to model temperatures. They include Alaton et al, Benth and Saltyte-Benth, Brody et al, and Davis. Brownian motion is suggested to control temperatures' randomness and their mean-reverting nature. Interestingly, Brody et al (2002) simply propose a model and do not test its appropriateness with their temperature dataset. Davis proposes a BM-type temperature model but fits his data with an AR model. BM-type temperature models can be decomposed into four parts: the long-term average, the mean-reverting speed, the error variance, and the error process.

The first component of BM-type models is the long-term average, and its significance is that temperatures return to their long-term means. Alaton et al propose $A + B \times t + C \times \sin(\frac{2\pi}{365}t + \varphi)$ to capture it, where A is the minimum temperature in a year, B is the trend, t is the time, C is the difference between the minimum and the maximum temperatures in a year, and φ is to adjust the fact that the coldest day in a year is often not January 1. Alaton et al find a small trend in Sweden. Benth and Benth-Saltyte fit their Norwegian data with a cosine function and they also find an insignificant trend. Brody et al portray how close simulated temperatures are to the real ones when long-term averages follow $A + C \times \sin(\frac{2\pi}{365}t + \varphi)$. Davis does not discuss this term in detail.

The second element of the BM-type temperature models is the mean-reverting speed which measures how fast temperatures will return to their long-term averages. Alaton et al and Benth and Benth-Saltyte define it to be constant. Brody et al present the closeness of their simulated temperatures to the actual ones when the mean-reverting speed is 0.95. Davis suggests a Geometric Brownian Motion approach so this term is absent in his function.

The next is the error variance, the third component of BM-type temperature models. Because temperatures are stable in summer and volatile in winter, error variance is conditional on time. Alaton et al observe that Swedish temperatures have different variances in different months, and they do not vary much within each month. They consequently recommend a step function to model error variances. Benth and Benth-Saltyte do not like the fluctuating feature of the step function and come up with the smoothed daily temperature variances. According to their model, $smoothed \cdot temp \cdot \mathrm{var} \cdot on \cdot Jan \cdot 3 = \exp\{\frac{1}{3}[\log(temp \cdot \mathrm{var} \cdot on \cdot Jan \cdot 1) + \log(temp \cdot \mathrm{var} \cdot on \cdot Jan \cdot 2) + \log(temp \cdot \mathrm{var} \cdot on \cdot Jan \cdot 3)\}$. Brody et al, once again, present how perfectly matched the simulated temperatures are to the actual ones when temperature variances follow a sine-function. Davis does not discuss this term in detail.

The final part of BM-type temperature models is the error process. This term is crucial because temperatures are random. Davis and Alaton et al suggest that temperatures follow a Brownian process. Brody et al later claim that temperatures have a long-term dependency and should be modeled with fractional Brownian process. Benth and Benth-Saltyte verify with their Norwegian temperatures and find that errors are neither normal nor fractional Brownian. Instead the fat tail presented in the histogram leads them to model temperatures with the Levy process[12].

Campbell and Diebold compare the out-of-sample forecast accuracy of their model to two other models, using the criterion of u statistics. The u statistics is defined as $\sqrt{\frac{\sum(\hat{T}^{CD}_{t+h,t}-T_{t+h})^2}{\sum(\hat{T}^{M_i}_{t+h,t}-T_{t+h})^2}}$, where $\hat{T}^{CD}_{t+h,t}$ is the h-day ahead point forecast based on the information known on day t and the Campbell and Diebold's model, T_{t+h} is the actual temperature outcome, and $\hat{T}^{M_i}_{t+h,t}$ is the h-day ahead point forecast based on the information known on day t and the i[th] contrasting model. Temperatures in the first contrasting model are assumed to follow a normal distribution after removing the trend and the daily average temperatures, whereas temperatures in the second model are assumed to be their daily averages in the 20-year period preceding the computation date. The authors find that their complicated model does an excellent

[12] Errors are specified to follow a generalized hyperbolic Levy process because Norwegian temperatures possess semi-heavy tails and skewness. Furthermore, this class of Levy process is selected because its density function is explicitly known and has the normal distribution as a limiting case.

predicting job if the forecast horizon is within eight days. Beyond eight days, it is possible to obtain accurate predictions as the Campbell and Diebold's 2005 model using instead the Climatological model.

Because pricing temperature futures requires accurate forecasts, it is more important to obtain correct out-of-sample predictions than controlling for well-known temperature phenomena. In addition, since the algorithm used to calculate the parameters of Cao and Wei's model did not converge with the current dataset, it was decided that futures would be priced with the Climatological Model.

To sum up, financial researchers have proposed autoregressive (AR) models and Brownian motion (BM) to model ground temperatures. They add various parts into models to control for well-known weather phenomena. Researchers have proposed AR-type temperature models and have tried to fit their temperature datasets. This is not the case for researchers who have proposed BM-type models. Furthermore, Campbell and Diebold compare the forecast accuracy of three models. The Climatological model in particular does an excellent prediction job when the forecast horizon is beyond eight days. We will therefore apply this model to forecast temperatures required in pricing cHDD/cCDD futures contracts.

2.2 Valuation Methods

After coming up with temperature predictions, the second step in valuing weather derivatives is to calculate the temperature derivative prices. Documented in Roustant et al and Zeng, practitioners price them with the Actuarial Method. This is because one group of market makers, insurers, use this to value their traditional products. Prices of weather derivatives are the sum of the expected payoffs and a cushion to absorb unusual deviations. Given that each insurer has a different cost of capital and different risk considerations, they discount payoffs with different interest rates. In addition, they do not compute the cushion amounts in the same way. As a result, different practitioners will have different prices for the same weather derivatives.

Scholars, on the other hand, argue that the equilibrium prices of weather derivatives can be derived by discounting the expected payoffs. The question becomes: how much risk do weather derivatives have and what are their risk-adjusted interest rates?

Most traders enter the derivative market for one of two purposes: hedging or speculating. Hedgers, holding a long position in the spot market, want to lock in the price of their commodities in the future. Speculators wish to earn extra return by taking risks. If hedgers do not sell derivative contracts, they face the risk of selling their commodities in the spot market at a lower price. On the other hand, it will do no harm to speculators if they do not buy derivatives. They simply earn lower returns with their less-risky portfolios. The derivative market is thus a buyer-dominated market. Hedgers have to lower derivative prices so speculators can earn a decent profit. The amount that hedgers are willing to forgo depends on the protection derivatives can offer. The more protection they can offer to hedgers, the higher the risk premium is. Derivatives offer strong protection when their cash flows closely correlate with the underlying commodities. If selling derivatives provides hedgers no protection, hedgers will not want to lower derivative prices. Derivative sellers will want to keep the price high so buyers will not get profits, and buyers will want to buy them low so they can earn money. In other words, derivative prices will be equal to the expected spot price of the underlying assets if there is no risk premium.

Because weather derivative prices observed in the market are in equilibrium, one observes the market-wide average of the protection offered by weather derivatives to hedgers. In other words, this protection can be quantified by the correlation between weather derivative payoffs and the growth rate of the whole economy. Platen and West find that weather derivative payoffs are not correlated with the growth rate of the MSCI World, a proxy for the world economy, and therefore the appropriate discount rates for weather derivative payoffs are the risk-free interest rates.

Roustant et al reach the same conclusion after they assume that temperatures follow an autoregressive model and that interest rates are deterministic. Zeng discounts weather derivative payoffs with risk-free rates but does not provide his rationale for this decision.

Another group of researchers, however, holds a different position on this issue. These researchers include Brody et al, Alaton et al, Benth and Saltyte-Benth, Cao and Wei, and Davis. Authors of the first three papers just propose the existence of a weather risk premium but do not discuss them in detail. Conversely, Cao and Wei and Davis price weather derivatives by maximizing utilities. The risk premium is thus the difference between the price from the utility-maximizing model and the price derived by discounting with the risk-less interest rate.

Most researchers suggest pricing weather derivatives with numerical procedures. Brody et al are one exception to this practice. They develop a partial differential equation (PDE) for a call whose cash flows are a combination of several weather derivatives.

To sum up, weather derivatives can be evaluated with the Actuarial Method, the Discounted Payoffs Method, and the Utility Maximization Method. The Actuarial Method does not produce equilibrium prices, since each market maker uses it to calculate his/her lowest profitable price. Pricing with the Discounted Payoffs Method, on the other hand, requires the ex-ante knowledge of the weather risk premium. Given that the weather market is young and there is insufficient data, researchers can not reliably obtain a weather risk premium based on past values. Cao and Wei's (also Davis's) utility-maximizing pricing formula is the only approach that provides a clearer testable model. Therefore we choose to test the prediction accuracy of this model.

3. METHODOLOGY

In this section we will explain how Cao and Wei evaluate weather derivatives, followed by a discussion of the assumptions necessary to carry out the simulation. Because the current pricing practice is to gauge futures prices based on 20 years of historical data, both temperatures and aggregate dividend rates will be simulated with 20-year observations on these variables preceding the valuation dates.

3.1 The Basic Valuation Framework

Since the underlying index of weather derivatives is not tradeable, the No-Arbitrage Pricing Theory can not be used. Cao and Wei price weather derivatives by extending the pure-exchange economy of Lucas (1978). In this economy, uncertainties come from two sources: the temperature Y_t and the aggregate dividend rate of the economy δ_t. Given that equilibrium in this economy occurs when the total consumption is equal to the total output, a contingent claim with a payoff q_t at a future time T will have the equilibrium time-t price $X(t,T)$ given as:

(1) $$X(t,T) = \frac{E[U'(\delta_T,T)\times q_T]}{U'(\delta_t,t)}, \quad \forall t \in (0,T)$$

where $U'(\delta_T,T)$ is the first derivative of a representative investor's utility function with respect to the aggregate dividend rate. Following the standard in the utility literature, a typical agent is assumed to exhibit a constant relative risk aversion and his/her utility function follows:

$$(2) \qquad U(\delta_t, t) = e^{-\rho \times t} \frac{\delta_t^{\gamma+1}}{\gamma+1},$$

with the rate of time preference $\rho > 0$ and the risk parameter γ in the range of $(-\infty, 0]$. Weather derivative prices can be obtained using equation (1) after specifying the temperature process, the aggregate dividend rate process, and the agent's preference.

3.2 The Temperature Model

Recall that the Climatological Model is a straightforward model which produces good predictions beyond an eight-day horizon. Temperatures in this model are assumed to follow

$$(3) \qquad T_t = \beta_0 + \beta_1 \times t + \sum_{i=1}^{364} C_i \times D_{it} + d \times \xi_t, \quad 1 \le t \le 20 \times 365 = 7300 \text{ and } 1 \le i \le 364$$

$$d \times \xi_t \sim Normal(0, d^2)$$

where T_t is the daily average temperature on day t and D_{it} is a dummy variable which takes on a value of 1 if T_t is on the ith day of a year. The error term ξ_t is the standardized temperature innovation. We omit temperatures on February 29 to keep the forecasting process simple, and we will not simulate futures prices for any 29-day Februaries. Given that the estimation window is 20-years long, t ranges from 1 to

7300. Notice that there are 364 dummy variables, and temperatures for December 31 will have all dummy variables as zero.

3.3 The Aggregate Dividend Growth Rate Model

Cao and Wei employ Marsh and Merton's (1987) result that the aggregate dividend growth rates follow a mean-reverting process. Thus:

$$\text{(4)} \qquad \ln \delta_t = \alpha + \mu \times \ln \delta_{t-1} + \nu_t, \qquad \forall \mu \le 1$$

where $1-\mu$ measures the speed of mean reversion. The error term ν_t takes the following form:

$$\text{(5)} \qquad \nu_t = \sigma \times \varepsilon_t + \sigma \left[\left(\frac{\kappa}{\sqrt{1-\kappa^2}} \xi_t + \sum_{i=1}^{m} \eta_i \xi_{t-i} \right) \right], \qquad 0 \le m < +\infty$$

where ε_t is a i.i.d. standard normal variable representing non-temperature-related factors and $\xi_{t-i}, i = 1,2,\ldots,m$, are standardized temperature innovations defined in equation (3). As the whole economy can not vary without limit and it does not depend entirely on temperatures, $\sum_{i=1}^{m} \eta_i^2$ is finite. Moreover, Cao and Wei construct equation (5) such that the correlation between the contemporaneous temperature shock ξ_t and the aggregate dividend growth rate innovation ν_t is κ.

3.4 Valuing Cumulative Heating/Cooling Degree Day Futures

Denoting HDD(T_1, T_2) as the cumulative heating degree days between time T_1 and T_2, the time-t price of a cHDD futures contract with a tick size of $1, a strike temperature index of K, and an effective period from T_1 to T_2 will be:

(6) $$f_{HDD}(t,T_1,T_2,K) = E_t\left\{\frac{U'(\delta_{T_2},T_2)}{U'(\delta_t,t)}\times\left[HDD(T_1,T_2)-K\right]\right\}$$
$$= e^{-\rho(T_2-t)}E_t\left\{\frac{\delta_{T_2}^{\gamma}}{\delta_t^{\gamma}}\times\left[HDD(T_1,T_2)-K\right]\right\}$$

The futures price will be the value of K which makes $f_{HDD} = 0$. That is:

(7) $$F_{HDD}(t,T_1,T_2) = \frac{E_t[\delta_{T_2}^{\gamma}\times HDD(T_1,T_2)]}{E_t(\delta_{T_2}^{\gamma})}$$

Similar expressions can be derived for cumulative Cooling Degree Day futures.

3.5 The Simulation Design

To streamline the simulation process, Cao and Wei took advantage of results from previous work by other researchers and make assumptions on some parameters necessary in equations (2) through (7).

The first assumption they make is to set the rate of preference ρ to be 0.03. This is to reflect the long-term average of real risk-free interest rates.

The second assumption is to set the mean reversion parameter of the dividend process μ to be 0.8, 0.9 and 0.99. Shiller's (1983) empirical work shows it to be

0.807, while Marsh and Merton (1987) find it to be 0.945. When μ is 0.99, aggregate dividend rates follow a random walk.

As the volatility of the stock market index is around 20%, Cao and Wei assume that aggregate dividend growth rates should have similar variability. Thus the standard deviation of v_t is set to 20%. With equation (5) one can easily deduce σ.

The fourth assumption of Cao and Wei is that the risk aversion parameter γ is -2, -10 or -40. Most researchers find the public to have risk aversion between 0 and -2. However, Mehra and Prescott (1985) find that the high equity premium observed in the 20th century is only explainable when the public risk aversion is between -30 and -40. That is why Cao and Wei also postulate that γ should take on values -10 and -40.

Fifthly, Cao and Wei determine the average aggregate dividend growth rate α and the initial dividend rate ξ_t such that the long-term risk-free interest rate is maintained at 6%, i.e. $e^{-r(t,T)\times(T-t)} = E_t\left(\frac{U'(\delta_T)}{U'(\delta_t)}\right) = e^{-6\%\times(T-t)}$

As the prevalent pricing practice is to have a 20-year estimation window, we estimate parameters with all consumption and temperatures observed during the 20 years preceding the valuation dates. This implies that, for a cHDD December 2002 New York futures traded on June 18, 2002, we will perform an analysis on data observed covering the period from June 18, 1982 to June 17, 2002.

3.6 The Simulation Process

The simulation process can be decomposed into six steps. First of all, parameters in the temperature model (3), the temperature shocks ξ_t, and their standard deviation d are determined. Secondly, temperature shocks' coefficients in equation (5), i.e. κ and all η_i, are assessed. The third step is to determine the volatility parameter σ in equation (5) and to simulate the future consumption shock v_T. In the next step α in equation (4) is located and the future consumption rate δ_T is simulated. In the fifth step all information required in equation (7) is plugged in and a futures price is calculated. Finally, the above five steps are repeated 5000 times and the average of all simulated prices is calculated. This arithmetic average is the final simulated price for a contract as of the valuation date.

In the following paragraphs the first and the second steps in the simulation process will be clarified. The remaining steps can be easily deduced based on the previously-discussed assumptions.

The temperature model parameters are determined by applying the method of least squares. Errors are further standardized to derive ξ_t.

In the second step, we assess the correlation between temperatures and the whole economy is assessed. As one quantifies the growth rate of the whole economy with aggregate dividend growth rates, we can proxy them with GNP or aggregate consumption rates. Cao and Wei use the second choice and we follow their method. Given that aggregate consumption rates of the United States are not available as frequently as the ground temperatures, the most we can do is to assess their monthly correlation. We estimate μ with the least-squares method and obtain the innovations of aggregate consumption rates ν_t. Next we obtain monthly temperature shocks. This is done by inputting monthly temperature averages, i.e. the average of T_t's in a month, into equation (3). With a 20-year estimation window, one has $20 \times 12 = 240$ monthly averages. With both temperature shocks and consumption innovations, κ is therefore the correlation between them. Because Cao and Wei find that only the contemporaneous temperature shocks exert significant influences on consumption innovations, we assume that the dataset would herein exhibit the same characteristic. Our second reason to support this decision is that Cao and Wei's dataset includes a common test city, i.e. New York.

Since a month has mostly 31 days and it is impossible to determine when lagged temperatures lose their influence on consumptions with a monthly dataset, Cao and Wei assume the optimal lag in equation (5) to be either 0 or 30. When consumption is assumed to depend on 30 lagged temperatures, Cao and Wei postulate that lagged temperatures lose their influences by a constant q. Moreover, they design the last lagged temperature shock to have a constant correlation with consumption. Namely:

$$(8) \qquad \begin{aligned} &\eta_i = q^i \times \kappa, \qquad 0 < q < 1 \\ &|\eta_{30}| = 0.0001 \end{aligned}$$

A point to notice is that the correlation between temperatures and consumption is season-dependent. Consumers consume more while temperatures increase in summer or decrease in winter. Stated differently, the correlation is positive in CDD months and negative in HDD ones. The correlation magnitude is small when one does not recognize this fact. This correlation underestimation will lead to pricing temperature futures with smaller risk premia. Cao and Wei get around this issue by computing the contemporaneous correlation κ with only data from CDD months when the settlement date is within a CDD month. (Cao and Wei assume the two shoulder months April and October to be HDD months.) The lagged correlations will take on the reverse sign if lagged temperature shocks are from HDD months. For example, If we values a April cHDD futures settled on May 3 under the simulation setting of $m = 30$, κ, η_1 and η_2 will be positive and η_i, $3 \le i \le 30$ will be negative. A similar rationale is applied when the futures contract is settled in October.

After simulating temperatures and determining the correlation between temperatures and consumption, one is able to calculate all η_i, σ, α, ν_T and δ_T.

Plugging them into equation (7), one obtains the price for a cHDD futures contract as of one valuation date.

3.7 Examining Forecast Accuracy

After valuating futures prices with Cao and Wei model, one needs to determine their accuracy. Both parametric and non-parametric tests are conducted to assess the quality of these simulated prices.

Let S denote a simulated price and M a market price. Assume that there exists a linear relationship between them.

(9) $$M = \alpha_0 + \alpha_1 \times S + \varepsilon$$

If forecasts are close to market prices, then $\alpha_0 = 0$ and $\alpha_1 = 1$. These two hypotheses can be tested with two t-statistics, $t_{\alpha_0} = \frac{\hat{\alpha}_0}{\hat{s}_{\alpha_0}}$ and $t_{\alpha_1} = \frac{\hat{\alpha}_1 - 1}{\hat{s}_{\alpha_1}}$.

Insignificant t-statistics indicate that the simulated prices are statistically no different from the market prices, and Cao and Wei's model is useful. If we rewrite equation (9) as

$$M - S = \alpha_0 + (\alpha_1 - 1) \times S + \varepsilon .$$

We can do an F test to confirm that $\alpha_1 = 1$. This test is parametric because one has to assume that there exists a linear relationship between the two price sets and errors ε follow a normal distribution.

The second parametric test that one can do to test whether simulated prices are different from market prices is the z-test. Although the distribution which price differences follow is not known, one can assume with the large dataset that the price differences follow a Normal distribution. Because we assume that Cao and Wei model is correct, we will test the hypothesis that the population has a zero mean. If the mean of price differences is farther from zero, we reach the conclusion that Cao and Wei model does not simulate accurately. This test is parametric because one has to assume that price differences follow a normal distribution.

Since the price dataset is composed of prices of different temperature futures, they have different means and different variations. For example, January is often colder and has more volatile temperatures than February. Therefore, the cumulative temperature index in January should be larger and less stable. This implies that we can not assume a distribution that all prices follow. Non-parametric tests, requiring no ex-ante distribution assumption, are therefore better to test forecast accuracy of simulated prices. Three non-parametric tests are performed, including the Sign test, the Signed-Rank test, and the Wilcoxon test.

Under the Sign test, one tests differences between simulated prices and market prices. If these two price sets are close to each other, the number of negative differences will be similar to the number of positive ones. Moreover, the number of positive differences will follow a Normal distribution with mean equal to $\frac{275}{2}$ and variance equal to $\frac{275}{4}$, where 275 is the number of price differences. Insignificant z-statistics indicate that Cao and Wei's model produces accurate forecasts.

The Signed-Rank test is similar to the Sign test. After computing differences between simulated prices and actual ones, differences are sorted in ascending order and ranks are assigned. The smallest difference has rank 1 and the largest difference has rank 275. The sum of ranks of positive differences will follow a Normal distribution with mean as $\frac{275\times(275+1)}{4}$ and variance as $\frac{275\times(275+1)\times(2\times 275+1)}{24}$, provided that the simulated prices are close to the market counterparts. Like the Sign test, insignificant z-statistics imply accurate forecasts.

Under the Wilcoxon test, one mixes two price sets and sorts them in ascending order. Ranks are later assigned. If there is no discrepancy between these two sets, the sum of ranks of simulated prices should behave the same as the sum of ranks of actual counterparts. In fact, after subtracting $\frac{275\times(275+1)}{2}$ from these two sums, they follow a Normal distribution with mean as $\frac{275^2}{2}$ and variance as $\frac{275^2\times(2\times 275+1)}{12}$. Once again, statistically insignificant results for z-statistics imply accurate pricing.

4. DATA

4.1 Market Prices of the Monthly Cumulative Heating/Cooling Degree Day Futures for New York City Offered by the Chicago Mercantile Exchange

Daily closing prices of cHDD/cCDD futures indexed to New York City from the product inception date to July 28, 2005 are obtained from Bloomberg. After removing prices with zero trading volumes and the February 2004 cHDD contract, we have totally 275 prices ranging from June 18, 2002 to July 28, 2005. Futures contracts included in the dataset stretch from December 2002 cHDD futures to August 2005 cCDD futures. Table A-2 illustrates the exact price distribution among contracts.

[Please insert Table A-2 about here.]

4.2 Daily Temperatures of New York City

Given that my prices are for June 18, 2002 to July 28, 2005, we need daily temperatures observed from the period of June 18, 1982 to July 27, 2005 are necessary to fulfill the requirement of a 20-year estimation window. We purchased these temperature data from the National Climate Data Centre (NCDC) of the National Oceanic and Atmospheric Administration (NOAA) of the U.S. government.

4.3 Monthly Consumption Rates for the United States of America

As discussed in a previous section, Cao and Wei approximate the aggregate dividend growth rates with the aggregate consumption rates. Their practice is followed here and the monthly personal consumption rates are downloaded from the Federal Reserve Bank of St. Louis (http://www.stls.frb.org/fred/). In order to estimate parameters with a 20-year window, consumption rates for the period of July 1982 to July 2005 were extracted.

5. RESULTS

We explain in the sections of Simulation Design and Simulation Process that Cao and Wei make assumptions about parameters on which either academics have inconclusive opinions or the data indicates uncertain characters. These parameters include, firstly, the number of lagged temperatures that consumption depends on (m); secondly, the risk aversion parameter (γ); and finally, the aggregate dividends' mean-reverting speed (μ). Table A-3 summarizes these 18 simulation settings.

[Please insert Table A-3 about here.]

[Please insert Table A-4 about here.]

Table A-4 shows the basic statistics of percentage errors, which are the price differences divided by the corresponding actual prices and multiplied by 100%. If simulated prices coincide with the market ones, one expects the mean, the variance, the maximum, and the minimum of percentage errors to be zero. The numbers of positive or negative percentage errors are anticipated to be around half of the sample size, i.e. 137 or 138.

Inspecting the mean of percentage errors, we notice that the absolute value of the mean of percentage errors of the simulation setting with $(\gamma, m, \mu) = (-10, 0, 0.99)$ is the smallest, 0.54. This implies that simulated prices of this setting are only 0.54% different from the market counterparts. Note that the means of percentage errors do not vary much whether the parameter μ takes on values 0.8, 0.9 or 0.99. This insensitivity of forecast accuracy to the parameter μ will also be observed in other tests. When (γ, m) equal to $(-40,0)$ the means of percentage errors increase 58 times. When m is set to 30, this increment is more astonishing. Simulated prices are 118% more than the actual ones. This indicates that the weather futures market is not so risk averse as $\gamma = -40$ or the economy does not depend on as much temperature risk as $m = 30$. Note that when $\gamma = -2$ the means of percentage errors are of the similar range. This makes intuitive sense. Given that investors are assumed to be less risk conscious and the difference between $m = 0$ and $m = 30$ means little to them, they value temperature futures in the same fashion.

Examining the standard deviation of percentage errors, the simulation setting with $(\gamma, m, \mu) = (-2,0,0.99)$ produces the most accurate prices. The standard deviation of percentage errors is 20.23%. However, given that the variances of percentage errors for the simulation settings with $(\gamma, m, \mu)=(-2,0,0.8)$, $(\gamma, m, \mu)=(-2,0,0.9)$, $(\gamma, m, \mu)=(-10,0,0.8)$, $(\gamma, m, \mu)=(-10,0,0.9)$, and $(\gamma, m, \mu)=(-10,0,0.99)$ are only slightly larger, these simulation settings are also good, in the sense of small variance of percentage errors. This phenomenon may indicate that the variance of percentage errors is not sensitive when γ is greater than -10. When γ takes on the value -40, we see that the variance of percentage errors increases. The increment is especially pronounced when we set m to 30. We therefore have additional evidence to support the argument that $\gamma = -40$ is a poor simulation parameter. Please note that the variance of percentage errors does not vary much whether μ takes on the value of 0.8, 0.9 or 0.99. The speed of mean reversion of consumption rate is not crucial to the variance of percentage errors.

Observing that the maxima of percentage errors for the first 15 simulation settings are of similar magnitude, we trace that these maxima all correspond to the cCDD May 2005 futures traded on May 17, 2005. Examining the cumulative Cooling Degree Day (cCDD) index, the 20-year average of May cCDD is 65.6 and its standard deviation is 40.2. This implies that the cCDD index in May 2005 of 16.5 is 1.22 standard deviations below the mean. In other words, May 2005 was colder than anticipated. However, is the cold May 2005 the sole reason to cause the 300% overpricing? We suspect that temperatures in May are around 65°F and consumers enjoy this comfortable month such that they do not need to turn on the heat or the air-conditioning. The economy, therefore, may not depend on temperature risk. If this suspicion is correct, the magnitude of overpricing when $m = 0$ or $\gamma = -2$ should be smaller. Since this is not the case, the amount of temperature risk that the economy depends on is not the cause. We then inspect temperatures observed in May 2005 and find that cCDD index was zero from May 1, 2005 to May 14, 2005. Futures traders saw this pattern of the cCDD index and might have over-reacted. To conclude, the colder than expected May 2005 and traders' over-reaction might cause the 300% over-pricing. On the other hand, the maxima of percentage errors of simulation settings with $(\gamma, m) = (-40,30)$ correspond to the October 2003 cCDD contract. We analyze this contract as we do for the May 2005 cCDD futures. We come to the same conclusion that colder than anticipated October 2003 and traders' over-reaction caused the 600% over-pricing. The exact reasoning can not be deduced since percentage errors of this extreme magnitude are few.

The minima of percentage errors do not correspond to the same contract valued on the same day. Given that the under-pricing magnitude is not astonishing, these minima should be natural in the simulation process.

When $\gamma = -10$ we note that the distribution of percentage errors is most symmetrical. 47% of the percentage errors are positive, while 53% of them are negative. This implies that simulators face less under-(over-)pricing risk if they set γ to -10. Note that the distribution of percentage errors is not symmetrical around zero when $\gamma = -2$. Around 66% of percentage errors are negative and this implies that simulated prices are less than their market counterparts. On the other hand, 87% of the simulated prices are greater than the market prices when $\gamma = -40$. We are likely to over-price futures when we assume temperature futures traders to have a risk aversion parameter as -40. Inspecting numbers of positive and negative percentage errors, we observe again the forecasts' insensitivity to the speed of mean-reversion of consumption rate, m.

To sum up, price forecasts from the settings with $(\gamma, m) = (-10,0)$ are most accurate because these three settings have more statistics that met the expectations. We can forecast accurate prices with Cao and Wei's (2004) pricing model when we assume the futures traders have the risk aversion parameter as -10 and the economy depends on only the contemporaneous temperature.

5.1 Accuracy Checks with Parametric Tests

[Please insert Table A-5 about here.]

Table A-5 displays results of the parametric tests discussed previously. The second column shows the t-statistics used to test whether α_0 in equation (9) is zero. Listed in the third and the fourth columns are t-statistics and F-statistics that are used to examine the null that α_1 in equation (9) is one. Finally, the z-statistics are recorded in the last column of Table A-5.

Recall that statistically insignificant statistics signify that accurate futures prices can be simulated with the Cao and Wei's pricing model. Except for the simulation settings with $\gamma = -40$, t-statistics of all the other settings are insignificant at 95% confidence level. This evidence supports the argument that an intercept term is not

not required in Cao and Wei's model when the risk aversion parameter is moderate or low.

We use statistics listed in the third and fourth columns of Table A-5 to test whether α_1 equals to one in equation (9). Simulation settings with $(\gamma, m) = (-10, 0)$ produce better results. We can say, with 95% of confidence, that simulated prices are equal to the actual ones. Forecasted prices of the other simulation settings, however, have linear, but not diagonal, relationships with the market futures prices. The assumption of $\alpha_1 = 1$ are rejected at both 95% and 99% levels. Another point of interest is that there are no discrepancies between decisions made based on the t-statistics and ones made based on the F-statistics. This is a good sign.

Based on the z-statistics listed in the last column in Table A-5, we find that the simulation settings with $(\gamma, m) = (-10, 30)$ produce more accurate prices. Their z-statistics are around 2.55 and the null hypothesis can be accepted at the 99% confidence level. This is to say that price simulations are the same as the market prices. Although this result is consistent with our finding based on the t-statistics in the second column, it contradicts with the discovery based on the two statistics in the third and the fourth columns. Overall, one forecast futures prices precisely by setting γ to -10, and the settings with $(\gamma, m) = (-10, 0)$ and the settings with $(\gamma, m) = (-10, 30)$ are equally favoured.

Forecast accuracy is sensitive to the risk aversion parameter γ. As γ decreases from -2, -10, to -40, we change our decision whether to reject the null hypothesis. The dependence of forecast accuracy on m, the number of lagged temperatures that consumption depends on, seems to depend on γ. When $\gamma = -2$, statistics are not much different whether $m = 0$ or $m = 30$. As explained previously, this is intuitive since less-risk-averse investors will not value futures too differently whether the economy depends on no lagged temperatures or 30 lagged temperatures. They value temperature futures in the same fashion. However, when $\gamma = -10$ or $\gamma = -40$, investors are more sensitive to risks and differences between two groups of statistics show. The speed of mean reversion of consumption rate, μ, does not play a role in statistics listed in Table A-5.

5.2 Accuracy Checks with Non-Parametric Tests

[Please insert Table A-6 about here.]

Recall that three non-parametric tests are conducted to check the validity of simulated prices, including the Sign Test, the Signed-Rank Test, and the Wilcoxon Test. Results appear in Table A-6.

The Sign Test indicates that the simulation settings with $\gamma = -10$ produce more accurate price forecasts. The null hypothesis that price forecasts and market prices are the same is accepted with 95% of confidence. Given that we examine the number of number of positive price differences in the Sign Test, negative (positive) z-statistics indicate an under- (over-) pricing issue. We therefore observe again the under-pricing phenomenon when γ is set to -2 and the over-pricing effect when γ is -40.

Results of the Signed-Rank Test are consistent with the previous Sign Test. We forecast more accurately when $\gamma = -10$. Though, this time the null hypothesis is accepted at the 99% confidence level. We find that the simulated prices of the other settings differ from their market counterparts at both 95% and 99% of confidence levels. This signifies that the Signed-Rank test does not find forecasted prices of the other simulation settings accurate.

The non-parametric Wilcoxon test shows the most support for Cao and Wei model. As can be seen in Table A-6, the Wilcoxon test indicates that 12 settings produce accurate prices, unlike the previous tests that only 3 or 6 settings. In addition, they are all insignificant at the 95% confidence level.

With respect to the sensitivity of forecast accuracy to simulation parameters, we find that the risk aversion parameter plays the most important role. Our acceptance of the null hypothesis changes when γ takes on different values. We notice that, with aids of the three non-parametric tests, the sensitivity of forecast accuracy to m depends on γ. Different from the results of the parametric tests, m only impacts the z-statistics when investors exhibit extreme risk aversion (i.e. $\gamma = -40$). The mean-reverting speed of consumption rate seems to have a more critical role on the z-statistics of the Sign Test when the risk aversion parameter is -10. The z-statistic of the setting $(\gamma, m, \mu) = (-10,30,0.8)$ differs from the z-statistic of the simulation setting $(\gamma, m, \mu) = (-10,30,0.9)$ by 0.6. However, overall, the mean-reverting speed of consumption rate is not a decisive parameter.

To summarize, we find simulation settings with $\gamma = -10$ produce the most promising outcomes. One can say with at least 99% of confidence that price forecasts are the same as the market prices. However, one can not decide whether the simulation settings with $(\gamma, m) = (-10,0)$ or the settings with $(\gamma, m) = (-10,30)$ are better. They are equally favoured by the statistical tests. Forecast accuracy is most sensitive to the risk aversion parameter, since with different γ one either accepts or rejects the null hypothesis. The sensitivity of forecast accuracy to the number of lagged temperatures that consumption depends on seems to hinge on the risk aversion parameter. When $\gamma = -2$ statistics do not vary when one sets m = 0 or 30, but they change substantially when $\gamma = -40$. The mean-reverting speed of the consumption rate does not have a significant influence on forecast accuracy. Different μ values do not lead to opposite conclusions regarding the null hypothesis.

6. CONCLUSIONS AND FUTURE RESEARCH

A well-known fact is that weather has an impact upon the economy. Excessive rain causes mudslides, and, yet, volcanic eruptions bring nourishing ash to plants. With the 1999 inception of weather derivatives on the Chicago Mercantile Exchange, investors can now hedge weather risks. However, given that the underlying index is not a tradeable asset, one can not apply the No-Arbitrage Pricing Theory to value them. Cao and Wei propose a pricing framework and its validity is tested in this thesis.

Cao and Wei design their pricing simulation to be flexible enough to accommodate parameters that either academics do not have consensus on or the data fails to indicate a precise character. Prices are simulated under 18 different combinations of parameters. When the risk aversion parameter is -10, Cao and Wei's pricing framework provides promising results. Its prediction accuracy is verified with both parametric and non-parametric tests, and all tests indicate that simulated prices are the same as the market ones. Forecast accuracy is found to be very sensitive to the risk aversion parameter, followed by the number of lagged temperature shocks that consumption innovations depend on. Consumption's mean-reversion speed minimally impacts prediction precision.

After confirming that Cao and Wei's model is useful in pricing weather derivatives, future researchers can examine weather derivatives' risk premia. As discussed in the Literature Review section, many scholars agree that weather derivative traders demand risk premia but no one knows how they behave. It is also interesting to investigate whether the behaviour of the risk premia vary between cities.

Another appealing topic would be to investigate whether the parameters producing the most accurate predictions are conditioned on time. We hypothesize that the best risk aversion parameter will increase from -10 when the weather market matures and traders pay smaller liquidity costs. Moreover, researchers can expand to other cities to check whether the best set of parameters are the same in other cities.

Furthermore, we can compare the forecast accuracy of various proposed pricing models. Which model provides the most accurate forecasts? Are their accuracies drastically different? Is the most accurate model location-dependent, such that they provide poor forecasts for some cities?

Finally, we can investigate the impact of generalizing the simple economy used in Cao and Wei's pricing model on the weather derivatives pricing. As discussed before that the simple economy depends on only two variables. Will the price forecasts become more accurate if we design the economy more like what we have now?

7. REFERENCES

Alaton, P., Djehiche, B. and Stillberger, D., *On Modelling and Pricing Weather Derivatives*, Applied Mathematical Finance, Vol. 9, Iss. 1, 1 – 20, 2002

Benth, F. E. and Saltyte-Benth, J, *Stochastic Modelling of Temperature Variations with a View towards Weather Derivatives*, Statistical Research Report No. 1, ISBN 0806-3842, January 2004

Bloomfield, P., *Trends in Global Temperature*, Climate Change, Vol. 21, 1 - 16

Brody. D, Syroka, J. and Servos, M., *Dynamical Priding of Weather Derivatives*, Quantitative Finance, Vol. 2, 189 – 198, 2002

Campbell, S. D. and Diebold, F. X., *Weather Forecasting for Weather Derivatives*, Journal of the American Statistical Association, Vol. 100, Iss. 469, 6 – 16, 2005

Cao, M. and Wei, J., *Weather Derivatives Valuation and Market Price of Weather Risk*, the Journal of Futures Markets, Vol. 24, No. 11, 1065 – 1089, 2004

Cao, M. and Wei, J. , *Watching the Weather Report*, Canadian Investment Review, Summer 2004, 27 – 33

Davis, M, *Pricing Weather Derivatives by Marginal Value*, Quantitative Finance, Vol. 1, 305 - 308, 2001

Lucas, R. E., *Asset Prices in an Exchange Economy*, Econometrica, Vol. 46, No. 6, 1429 – 1445, 1978

Marsh, T. A. and Merton, R., *Dividend Behaviour for the Aggregate Stock Market*, Journal of Business, Vol. 60, 1 – 40, 1987

Mehra, R. and Prescott, E., *The Equity Premium, A Puzzle*, Journal of Monetary Economics, Vol. 15, 145 – 161, 1985

Müller, A. and Grandi, M., Swiss Re., *Weather derivatives for protection against weather risks*, The Geneva Papers on Risks and Insurance, April 2000

Platen, E. and West, J., *Fair Pricing of Weather Derivatives*, University of Technology, Working Paper, July 2, 2004

Roustant, O., Laurent, J-P, Bay, X. and Carraro, L., *A Bootstrap Approach to the Price Uncertainty of Weather Derivatives*, École des Mines de Saint-Etienne, Working Paper, December 2003

Shiller, R. J., *Do Stock Market Prices Move too much to be Justified by Subsequent Changes in Dividends? Reply*, American Economic Review, Vol. 73, 236 – 237, 1983

Zeng, L., *Pricing Weather Derivatives*, Journal of Risk Finance, Spring 2000, 72 – 78

8. APPENDICES

8.1 APPENDIX A

Table A-1 - Cities for which the CME Offers Temperature Futures and the Stations at which Temperatures are Measured

City	Station
Atlanta, Georgia	Hartsfield International Air Port
Baltimore, Maryland	Washington International Air Port
Boston, Maine	Logan International Air Port
Chicago, Illinois	O'Hare International Air Port
Cincinnati, Ohio	Northern Kentucky Air Port
Dallas, Texas	Fort Worth International Air Port
Des Moines, Iowa	Des Moines International Air Port
Detroit, Michigan	Detroit Metropolitan Air Port
Houston, Texas	Bush Intercontinental Air Port
Kansas City, Oklahoma	Kansas City International Air Port
Las Vegas, Arizona	McCarran International Air Port
Minneapolis, Minnesota	St. Paul International Air Port
New York, New York	LaGuardia Air Port
Philadelphia, Pennsylvania	Philadelphia International Air Port
Portland, Oregon	Portland International Air Port
Sacramento, California	Sacramento Executive Air Port
Salt Lake City, Utah	Salt Lake City International Air Port
Tucson, Arizona	Tucson International Air Port

Table A-2 - Price Distribution of the Market Price Dataset among Contracts

Contract (Code)	Number of Prices
December 2002 cHDD (NFZ2)	7
January 2003 cHDD (NFF3)	11
February 2003 cHDD (NFG3)	10
March 2003 cHDD (NFH3)	11
April 2003 cHDD (NFJ3)	12
October 2003 cCDD (NAV3)	1
October 2003 cHDD (NFV3)	5
November 2003 cHDD (NFX3)	10
December 2003 cHDD (NFZ3)	15
January 2004 cHDD (NFF4)	15
March 2004 cHDD (NFH4)	19
April 2004 cHDD (NFJ4)	4
July 2004 cCDD (NAN4)	13
August 2004 cCDD (NAQ4)	12
October 2004 cHDD (NFV4)	7
November 2004 cHDD (NFX4)	11
December 2004 cHDD (NFZ4)	13
January 2005 cHDD (NFF5)	14
February 2005 cHDD (NFG5)	25
March 2005 cHDD (NFH5)	27
April 2005 cHDD (NFJ5)	6
May 2005 cCDD (NAK5)	1
July 2005 cCDD (NAN5)	23
August 2005 cCDD (NAQ5)	3
Total	275

Table A-3 - Varying Parameters in the Simulation Settings

Varying Parameter	Symbol	Values
Number of Lagged Temperatures that consumptions depend on	m	0 or 30
Risk Aversion	γ	2-, -10, or -40
Mean-Reverting Speed of Consumption Rates	μ	0.8, 0.9, or 0.99

Table A-4 – Basic Statistics of Percentage Errors

Simulation Setting	Percentage Errors: Mean (%)	Standard Deviation (%)	Maximum (%)	Minimum (%)	Number of Positive Percentage Errors	Number of Negative Percentage Errors
$m = 0$						
$\gamma = -2$						
$\mu = 0.8$	-0.87	20.45	311.20	-16.44	94	181
$\mu = 0.9$	-0.88	20.44	310.75	-16.41	91	184
$\mu = 0.99$	-0.96	20.23	307.61	-16.41	91	184
$\gamma = -10$						
$\mu = 0.8$	0.56	20.53	312.61	-15.65	130	145
$\mu = 0.9$	0.55	20.38	309.86	-16.15	127	148
$\mu = 0.99$	0.54	20.37	309.63	-15.97	128	147
$\gamma = -40$						
$\mu = 0.8$	28.86	36.37	321.35	-11.88	220	55
$\mu = 0.9$	28.75	36.12	320.42	-11.98	218	57
$\mu = 0.99$	28.94	36.25	318.17	-12.09	219	56
$m = 30$						
$\gamma = -2$						
$\mu = 0.8$	-0.82	20.38	310.04	-16.10	95	180
$\mu = 0.9$	-0.84	20.26	307.86	-16.14	94	181
$\mu = 0.99$	-0.81	20.67	314.82	-16.15	94	181
$\gamma = -10$						
$\mu = 0.8$	2.78	20.68	309.96	-14.61	147	128
$\mu = 0.9$	2.78	20.64	308.27	-13.98	152	123
$\mu = 0.99$	2.79	20.68	309.90	-14.71	147	128
$\gamma = -40$						
$\mu = 0.8$	118.41	145.37	610.54	-8.91	260	15
$\mu = 0.9$	118.55	146.21	663.63	-8.24	261	14
$\mu = 0.99$	117.53	143.75	585.16	-9.55	261	14

* The percentage error is defined as the price difference divided by the corresponding market price and multiplied by 100%. If simulated prices are no different from market prices, the mean, the variance, the maximum and the minimum of percentage errors are expected to be zero. The number of positive percentage errors and the number of negative ones are anticipated to be around half of the data size, i.e. 137 or 138.

Table A-5 - Forecast Accuracy Check with the Parametric Tests

Simulation Setting	$H_0 : \alpha_0 = 0$	$H_0 : \alpha_1 = 1$		H_0: *Diff* = 0
	t-statistics	F Statistics	t-statistics	z-statistics
$m = 0$				
$(\gamma, \mu) = (-2, 0.8)$	-1.65	20.18**	4.47**	-7.30**
$(\gamma, \mu) = (-2, 0.9)$	-1.63	19.97**	4.47**	-7.31**
$(\gamma, \mu) = (-2, 0.99)$	-1.64	20.23**	4.50**	-7.34**
$(\gamma, \mu) = (-10, 0.8)$	-0.51	3.13	1.77	-3.33**
$(\gamma, \mu) = (-10, 0.9)$	-0.52	3.19	1.78	-3.34**
$(\gamma, \mu) = (-10, 0.99)$	-0.45	2.93	1.71	-3.34**
$(\gamma, \mu) = (-40, 0.8)$	13.61**	758.44**	-27.54**	12.66**
$(\gamma, \mu) = (-40, 0.9)$	13.44**	741.50**	-27.23**	12.73**
$(\gamma, \mu) = (-40, 0.8)$	13.49**	750.39**	-27.39**	12.75**
$m = 30$				
$(\gamma, \mu) = (-2, 0.8)$	-1.62	19.44**	4.41**	-7.17**
$(\gamma, \mu) = (-2, 0.9)$	-1.65	19.84**	4.45**	-7.20**
$(\gamma, \mu) = (-2, 0.99)$	-1.60	19.17**	4.38**	-7.15**
$(\gamma, \mu) = (-10, 0.8)$	1.61	7.16**	-2.68**	2.54*
$(\gamma, \mu) = (-10, 0.9)$	1.63	7.38**	-2.72**	2.57*
$(\gamma, \mu) = (-10, 0.99)$	1.61	7.16**	-2.68**	2.53*
$(\gamma, \mu) = (-40, 0.8)$	26.88**	9906.47**	-99.53**	12.10**
$(\gamma, \mu) = (-40, 0.9)$	26.96**	10003.9**	-100.02**	12.07**
$(\gamma, \mu) = (-40, 0.8)$	26.98**	10004.5**	-100.02**	12.08**

1. Denote S and M as the simulated price and the market price, respectively. Assume that there exists a linear relationship between S and M such that $M = \alpha_0 + \alpha_1 \cdot S + \varepsilon$. If S and M are the same, one expects that $\alpha_0 = 0$ and $\alpha_1 = 1$. The first hypothesis is tested with the t-statistics listed in the second column, while the second hypothesis is tested with the t-statistics in the third column and the F-statistics in the fourth column. Insignificant statistics indicate that one forecast precise futures prices with Cao and Wei's model.
2. If S and M are the same and the central limit theorem applies, (S-M) follows a normal distribution with mean as zero. z-statistics is used to test whether the sample does have a zero mean. Once again, insignificant-statistics imply accuracy of Cao and Wei's pricing model.
3. * and ** indicate that statistics are 5% and 1% significant, respectively.

Table A-6 - Forecast Accuracy Check with the Non-Parametric Tests

Simulation Setting	The Sign Test[1]	The Sign-Rank Test[2]	The Wilcoxon Test[3]
$m = 0$			
$(\gamma, \mu) = (-2, 0.8)$	-5.25**	-6.45**	1.43
$(\gamma, \mu) = (-2, 0.9)$	-5.61**	-6.47**	1.44
$(\gamma, \mu) = (-2, 0.99)$	-5.61**	-6.50**	1.44
$(\gamma, \mu) = (-10, 0.8)$	-0.90	-2.39*	0.47
$(\gamma, \mu) = (-10, 0.9)$	-1.27	-2.40*	0.47
$(\gamma, \mu) = (-10, 0.99)$	-1.15	-2.41*	0.49
$(\gamma, \mu) = (-40, 0.8)$	9.95**	11.69**	-5.23**
$(\gamma, \mu) = (-40, 0.9)$	9.71**	11.66**	-5.25**
$(\gamma, \mu) = (-40, 0.8)$	9.83**	11.68**	-5.32**
$m = 30$			
$(\gamma, \mu) = (-2, 0.8)$	-5.13**	-6.32**	1.42
$(\gamma, \mu) = (-2, 0.9)$	-5.25**	-6.36**	1.43
$(\gamma, \mu) = (-2, 0.99)$	-5.25**	-6.31**	1.41
$(\gamma, \mu) = (-10, 0.8)$	1.15	2.13*	-0.30
$(\gamma, \mu) = (-10, 0.9)$	1.75	2.21*	-0.29
$(\gamma, \mu) = (-10, 0.99)$	1.15	2.09*	-0.28
$(\gamma, \mu) = (-40, 0.8)$	14.77**	14.06**	-9.27**
$(\gamma, \mu) = (-40, 0.9)$	14.89**	14.09**	-9.35**
$(\gamma, \mu) = (-40, 0.8)$	14.89**	14.07**	-9.26**

1. To test whether the simulated prices are the same to the market counterparts, one tests their differences under the Sign Test. If these two price sets are the same, the number of positive differences will behave the same as the number of negative differences. In fact, these two numbers will follow a normal distribution with mean as 275/2 and variance as 275/4. 275 is the sample size of the price dataset. The standardized z-statistics of the Sign Test are shown in the second column. Insignificant statistics imply that the Cao and Wei's pricing model produces correct price forecasts.

2. The Signed-Rank Test is similar to the Sign Test. Price differences are sorted in ascending order and assigned ranks. If simulated prices are no different from the market counterparts, the sum of the ranks of positive price differences will follow a normal distribution with mean as 275•(275+1)/4 and variance as 275•(275+1) •(2•275+1)/24. The z-statistics of the Signed-Rank Test are shown in the third column. Again, insignificant statistics indicate the validity of Cao and Wei's model.

3. Under the Wilcoxon Test simulated prices are mixed with the market ones. They are further sorted in ascending order and assigned ranks. 275•(275+1)/2 is subtracted from the sum of the ranks of simulated prices. If simulated prices behave in the same fashion as the market counterparts, this variable will follow a normal distribution with mean as $275^2/2$ and variance as 275^2 •(2•275+1)/12. The z-statistics of the Wilcoxon Test are shown in the fourth column and insignificant statistics serve as supporting evidences for Cao and Wei's pricing model.

4. * and ** indicate that statistics are 5% and 1% significant, respectively.

8.2 APPENDIX B

Appendix B-1: SAS Code for Correlation Calculation

```
******CODE OF MACRO QKSigma******************************************

  1.   k = Contemporaneous Correlation b/w Standardized Tmp Residuals and
Standardized Consumption Residuals
  2.   q = Geometric Decay Factor
  3.   sigma = Variance of Non-temperature Related Part of Consumption
Residuals;

%Macro QKSigma(ValIdx,StartIdx,LastIdx,Setle_M,Setle_D,SetleIdx);

* Extract consumption rates of the right period;
Data ConsReg;
  Set OrigCons;
  If tIndex lt (&ValIdx-7300) or tIndex ge &ValIdx then Delete;
  lnConsL = lag1(lnCons);
  If lnConsL='.' then delete;
  Keep Year Month Day lnConsL lnCons tIndex;
Run;

* Compute the shocks in equation (4);
Proc Reg Data=ConsReg noprint;
   model lnCons = lnConsL;
   output out=ConsReg student=vt;
Run;

* Extract temperatures of the right period;
Data MthlyTmpReg;
 Set Mthly_NY;
 If tIndex lt (&ValIdx-7300) or tIndex ge &ValIdx then Delete;
Run;

* Compute the shocks in equation (3);
Proc Reg data=MthlyTmpReg noprint;
   model MthTmp = tIndex D1-D11;
   output out=MthlyTmpReg student=xi_t;
Run;
```

```
* Keep shocks from HDD(CDD) months if contracts are settled in a HDD(CDD)
month;
Data Correlation;
   Merge MthlyTmpReg ConsReg;
   By tIndex;
   Drop D1-D11;
   If vt='.' or xi_t='.' then delete;
   If &Setle_M le 4 or &Setle_M ge 10. then Do;
       If Month ge 5 and Month le 9 then delete;
   End;
   Else Do;
       If Month le 4 or Month ge 10 then delete;
   End;
Run;

* Compute k, q and sigma;
Data Correlation;
  Set Correlation End=Last;
  sum_vt + vt;
  sum_xit + xi_t;
  If Last then call symput('nobs',_n_);
Run;
%put ***nobs = &nobs ***;

Data Correlation;
  Set Correlation End=Last;
  avg_vt = sum_vt / &nobs;
  avg_xit = sum_xit / &nobs;
  If Last then call symput('avg_vt',avg_vt);
  If Last then call symput('avg_xit',avg_xit);
Run;
%put ***avg_vt = &avg_vt ***;
%put ***avg_xit = &avg_xit ***;

Data Correlation;
  Set Correlation End=Last;
  var_vt + (vt - &avg_vt)**2.;
  var_xit + (xi_t - &avg_xit)**2.;
```

```
 Numerator + (vt-&avg_vt)*(xi_t-&avg_xit);
 k = Numerator / ((var_vt*var_xit)**.5);
 q = (0.0001/abs(k))**(1./30.);                    *equation (8);
 sigma = (0.04 / ( 1+ k**2./(1-k**2.) +
q**2.*k**2.*(1-q**60.)/(1-q**2.) ))**.5;
 If Last then call symput('k',k);
 If Last then call symput('q',q);
 If Last then call symput('sigma',sigma);
Run;
%Put ****k = &k***;
%Put ****q = &q***;
%Put ****sigma = &sigma***;

*Determination of coefficients of 31 temperature residuals;
Data Coefficient;
   Set New_York;
   If tIndex lt (&setleIdx-30.) or tIndex gt &SetleIdx then Delete;
   Do i=1 to 30;
      If &SetleIdx - tIndex =i then Coefficient = &q**i*&k;
   End;
   If tIndex = &SetleIdx then Coefficient = &k * ((1-&k**2.)**-.5);
   If &Setle_M=5 or &Setle_M=10 then Do;
      If tIndex le &SetleIdx-&Setle_D then Coefficient = -Coefficient;
   End;
   Keep tIndex Coefficient;
 Run;

*Determination of 31 standardized temperature residuals;
%If &ValIdx gt %Eval(&SetleIdx-30) %then
%Do;
Proc Reg data=ParaEsti noprint;
 model tAvg = tIndex D1-D364;
 output out=TmpRsdlM30 student=student;
Run;

Data TmpRsdlM30;
   Set TmpRsdlM30;
   Keep tIndex student;
```

```
   If tIndex lt (&SetleIdx-30.) then Delete;
Run;

Data TmpRsdlM30_1;
  Set New_York;
  If tIndex lt &ValIdx or tIndex gt &SetleIdx then delete;
  Array xi{&NumSimul} X1-X&NumSimul;
  Do i=1 to &NumSimul;
    xi{i} = Rand('Normal');
  End;
  Keep tIndex X1-X&NumSimul;
Run;

*Combine coefficients with lagged temperature shocks;
Data TmpRsdlM30;
   Merge TmpRsdlM30 TmpRsdlM30_1 Coefficient;
   By tIndex;
   Array xi{&NumSimul} X1-X&NumSimul;
   Array v{&NumSimul} v1-v&NumSimul;
   Array Pre_v{&NumSimul} Pre_v1-Pre_v&NumSimul;
   Array TRF_v{&NumSimul} TRF_v1-TRF_v&NumSimul;

   Do i=1 to &NumSimul;
      If tIndex lt &ValIdx then Pre_v{i} + Coefficient * student;
       Else Pre_v{i} + Coefficient * xi{i};
      TRF_v{i} = &sigma*Rand('Normal');
       v{i} = TRF_v{i} + &sigma*Pre_v{i};
   End;
   If tIndex lt &SetleIdx then Delete; *Need the last sum;
   Drop tIndex X1-X&NumSimul Pre_v1-Pre_v&NumSimul Coefficient i;
Run;

%End;
%Else %Do;

Data TmpRsdlM30;
   Set New_York;
   If tIndex lt (&SetleIdx-30.) or tIndex gt &SetleIdx then Delete;
```

```
  Array xi{&NumSimul} X1-X&NumSimul;
  Do i=1 to &NumSimul;
    xi{i} = Rand('Normal');
  End;
  Keep tIndex X1-X&NumSimul;
Run;

Data TmpRsdlM30;
   Merge TmpRsdlM30 Coefficient;
   By tIndex;
   Array xi{&NumSimul} X1-X&NumSimul;
   Array v{&NumSimul} v1-v&NumSimul;
   Array Pre_v{&NumSimul} Pre_v1-Pre_v&NumSimul;
   Array TRF_v{&NumSimul} TRF_v1-TRF_v&NumSimul;

   Do i=1 to &NumSimul;
     TRF_v{i} = &sigma*Rand('Normal');
      Pre_v{i} + Coefficient * xi{i};
      v{i} = TRF_v{i} + &sigma*Pre_v{i};
   End;
   If tIndex lt &SetleIdx then Delete; *Need the last sum;
   Drop tIndex X1-X&NumSimul Pre_v1-Pre_v&NumSimul Coefficient i;
Run;
%End;

Proc Datasets Library=Work;
  Delete ConsReg MthlyTmpReg Correlation TmpRsdlM30_1 Coefficient;
Run;

%Mend QKSigma;
```

Appendix B-2: SAS Code to Simulate a cHDD Futures Price

```
*********MACRO STATEMENT PRICEHDD**********;

%Macro
PriceHDD(Val_M,Val_D,ValIdx,StartIdx,LastIdx,Setle_M,Setle_D,SetleIdx
,n);

*Isolate temperatures within the 20-year period preceding the valuation
date and temperatures recorded within the contract month;
   Data ParaEsti Price&n;
      Set New_York;
      If &ValIdx < &StartIdx then
         Do;
           If tIndex < (&ValIdx-7300) then Delete;
           Else If tIndex < &ValIdx then Output ParaEsti;
             Else If tIndex < &StartIdx then delete;
           Else if tIndex <= &LastIdx then Output Price&n;
           Else Delete;
          End;
      Else  If &ValIdx = &StartIdx then
         Do;
           If tIndex < (&ValIdx-7300) then Delete;
           Else If tIndex < &ValIdx then Output ParaEsti;
           Else if tIndex <= &LastIdx then Output Price&n;
           Else Delete;
          End;
      Else If &ValIdx < &LastIdx then
         Do;
           If tIndex < (&ValIdx-7300) then Delete;
           Else If tIndex < &StartIdx then Output ParaEsti;
           Else if tIndex < &ValIdx then output ParaEsti Price&n;
             Else if tIndex <= &LastIdx then output Price&n;
             Else delete;
          End;
       Else If &ValIdx = &LastIdx then
         Do;
           If tIndex < (&ValIdx-7300) then Delete;
```

```
        Else If tIndex < &StartIdx then Output ParaEsti;
        Else if tIndex < &ValIdx then output ParaEsti Price&n;
          Else if tIndex = &ValIdx then output Price&n;
          Else delete;
      End;
  Else
    Do;
        If tIndex <= (&LastIdx-7300) then Delete;
        Else If tIndex < &StartIdx then output ParaEsti;
        Else If tIndex <= &LastIdx then output ParaEsti Price&n;
          Else If tIndex < &ValIdx then output ParaEsti;
          Else delete;
      End;
Run;

*PARAMETER ESTIMATION FOR TEMPERATURE MODEL (3);
 Proc Model noprint;
   Parms b0 b1 c1-c364;
   Array Coeffict{364} C1-C364;
   Array DummyVar{364} D1-D364;

   tAvg = b0 + b1*tIndex + C1*D1 + C2*D2 + C3*D3 + C4*D4 + C5*D5 + C6*D6
+ C7*D7 + C8*D8 + C9*D9 + C10*D10 + C11*D11 + C12*D12 + C13*D13 + C14*D14
+ C15*D15 + C16*D16 + C17*D17 + C18*D18 + C19*D19 + C20*D20 + C21*D21 +
C22*D22 + C23*D23 + C24*D24 + C25*D25 + C26*D26 + C27*D27 + C28*D28 + C29*D29
+ C30*D30 + C31*D31 + C32*D32 + C33*D33 + C34*D34 + C35*D35 + C36*D36 +
C37*D37 + C38*D38 + C39*D39 + C40*D40 + C41*D41 + C42*D42 + C43*D43 + C44*D44
+ C45*D45 + C46*D46 + C47*D47 + C48*D48 + C49*D49 + C50*D50 + C51*D51 +
C52*D52 + C53*D53 + C54*D54 + C55*D55 + C56*D56 + C57*D57 + C58*D58 + C59*D59
+ C60*D60 + C61*D61 + C62*D62 + C63*D63 + C64*D64 + C65*D65 + C66*D66 +
C67*D67 + C68*D68 + C69*D69 + C70*D70 + C71*D71 + C72*D72 + C73*D73 + C74*D74
+ C75*D75 + C76*D76 + C77*D77 + C78*D78 + C79*D79 + C80*D80 + C81*D81 +
C82*D82 + C83*D83 + C84*D84 + C85*D85 + C86*D86 + C87*D87 + C88*D88 + C89*D89
+ C90*D90 + C91*D91 + C92*D92 + C93*D93 + C94*D94 + C95*D95 + C96*D96 +
C97*D97 + C98*D98 + C99*D99 + C100*D100 + C101*D101 + C102*D102 + C103*D103
+ C104*D104 + C105*D105 + C106*D106 + C107*D107 + C108*D108 + C109*D109
+ C110*D110 + C111*D111 + C112*D112 + C113*D113 + C114*D114 + C115*D115
+ C116*D116 + C117*D117 + C118*D118 + C119*D119 + C120*D120 + C121*D121
```

+ C122*D122 + C123*D123 + C124*D124 + C125*D125 + C126*D126 + C127*D127
+ C128*D128 + C129*D129 + C130*D130 + C131*D131 + C132*D132 + C133*D133
+ C134*D134 + C135*D135 + C136*D136 + C137*D137 + C138*D138 + C139*D139
+ C140*D140 + C141*D141 + C142*D142 + C143*D143 + C144*D144 + C145*D145
+ C146*D146 + C147*D147 + C148*D148 + C149*D149 + C150*D150 + C151*D151
+ C152*D152 + C153*D153 + C154*D154 + C155*D155 + C156*D156 + C157*D157
+ C158*D158 + C159*D159 + C160*D160 + C161*D161 + C162*D162 + C163*D163
+ C164*D164 + C165*D165 + C166*D166 + C167*D167 + C168*D168 + C169*D169
+ C170*D170 + C171*D171 + C172*D172 + C173*D173 + C174*D174 + C175*D175
+ C176*D176 + C177*D177 + C178*D178 + C179*D179 + C180*D180 + C181*D181
+ C182*D182 + C183*D183 + C184*D184 + C185*D185 + C186*D186 + C187*D187
+ C188*D188 + C189*D189 + C190*D190 + C191*D191 + C192*D192 + C193*D193
+ C194*D194 + C195*D195 + C196*D196 + C197*D197 + C198*D198 + C199*D199
+ C200*D200 + C201*D201 + C202*D202 + C203*D203 + C204*D204 + C205*D205
+ C206*D206 + C207*D207 + C208*D208 + C209*D209 + C210*D210 + C211*D211
+ C212*D212 + C213*D213 + C214*D214 + C215*D215 + C216*D216 + C217*D217
+ C218*D218 + C219*D219 + C220*D220 + C221*D221 + C222*D222 + C223*D223
+ C224*D224 + C225*D225 + C226*D226 + C227*D227 + C228*D228 + C229*D229
+ C230*D230 + C231*D231 + C232*D232 + C233*D233 + C234*D234 + C235*D235
+ C236*D236 + C237*D237 + C238*D238 + C239*D239 + C240*D240 + C241*D241
+ C242*D242 + C243*D243 + C244*D244 + C245*D245 + C246*D246 + C247*D247
+ C248*D248 + C249*D249 + C250*D250 + C251*D251 + C252*D252 + C253*D253
+ C254*D254 + C255*D255 + C256*D256 + C257*D257 + C258*D258 + C259*D259
+ C260*D260 + C261*D261 + C262*D262 + C263*D263 + C264*D264 + C265*D265
+ C266*D266 + C267*D267 + C268*D268 + C269*D269 + C270*D270 + C271*D271
+ C272*D272 + C273*D273 + C274*D274 + C275*D275 + C276*D276 + C277*D277
+ C278*D278 + C279*D279 + C280*D280 + C281*D281 + C282*D282 + C283*D283
+ C284*D284 + C285*D285 + C286*D286 + C287*D287 + C288*D288 + C289*D289
+ C290*D290 + C291*D291 + C292*D292 + C293*D293 + C294*D294 + C295*D295
+ C296*D296 + C297*D297 + C298*D298 + C299*D299 + C300*D300 + C301*D301
+ C302*D302 + C303*D303 + C304*D304 + C305*D305 + C306*D306 + C307*D307
+ C308*D308 + C309*D309 + C310*D310 + C311*D311 + C312*D312 + C313*D313
+ C314*D314 + C315*D315 + C316*D316 + C317*D317 + C318*D318 + C319*D319
+ C320*D320 + C321*D321 + C322*D322 + C323*D323 + C324*D324 + C325*D325
+ C326*D326 + C327*D327 + C328*D328 + C329*D329 + C330*D330 + C331*D331
+ C332*D332 + C333*D333 + C334*D334 + C335*D335 + C336*D336 + C337*D337
+ C338*D338 + C339*D339 + C340*D340 + C341*D341 + C342*D342 + C343*D343
+ C344*D344 + C345*D345 + C346*D346 + C347*D347 + C348*D348 + C349*D349

```
+ C350*D350 + C351*D351 + C352*D352 + C353*D353 + C354*D354 + C355*D355
+ C356*D356 + C357*D357 + C358*D358 + C359*D359 + C360*D360 + C361*D361
+ C362*D362 + C363*D363 + C364*D364;

      Fit tAvg / data=ParaEsti outest=Tmp_Parm out=Resid outall;
   Run;

*Compute d and standardized temperature residuals in equation (3);
   Data Resid;
     Set Resid;
     If mod(_n_,3) = 0 then Sum_Rsdl_Sq + tAvg**2.;
    *each record has 3 data lines: _type_=ACTUAL, PREDICTED, RESIDUAL;
     sigma_t = (Sum_Rsdl_Sq/7300)**.5;
     If _n_ < 21900. then delete;
     Do tIndex=&StartIdx to &LastIdx;
        output Resid;
     End;
     Keep tIndex sigma_t;
   Run;

   Data Tmp_Parm;
     Set Tmp_Parm;
     if _n_ gt 1 then delete;
     Do tIndex=&StartIdx to &LastIdx;
        output;
     End;
   Run;

*Simulate temperature index;
    Data Price&n;
     Merge Tmp_Parm Price&n Resid;
     By tIndex;

     Array Coeffict{364} C1-C364;
     Array DummyVar{364} D1-D364;
     Array Sim_Tmp{&NumSimul} ST1-ST&NumSimul;
      Array Sim_DD{&NumSimul} SH1-SH&NumSimul;
      Array cumTmpID{&NumSimul} CTI1-CTI&NumSimul;
```

```
Act_Tmp = tAvg;
 Act_DD = max(0., 65.-Act_Tmp);                * HDD contract;
 Sum_ActDD + Act_DD;

Do i=1 to &NumSimul;
  Do j=1 to 364;
    SumDummy + Coeffict{j}*DummyVar{j};
  End;
  Sim_Tmp{i} = b0+b1*tIndex+SumDummy+sigma_t*Rand('Normal');
  SumDummy=0.;

  Sim_DD{i} = max(0., 65.-Sim_Tmp{i});        * HDD contract;
   If tIndex lt &ValIdx then cumTmpID{i} + Act_DD;
  Else cumTmpID{i} + Sim_DD{i};
 End;

 Format CTI1-CTI&NumSimul Sum_ActDD 10.5;
If _n_ lt (&LastIdx-&StartIdx+1.) then delete;*need the final sum;
 Keep CTI1-CTI&NumSimul Sum_ActDD;
Run;

*Isolation of previous consumption rate;
 Data Consmptn;
   Set OrigCons;

   If &Val_D eq 1 then Do;
     If &Val_M eq 1 then Do;
       If Month eq 12 and tIndex ge (&ValIdx-365) and tIndex lt &ValIdx
then output;
        Else delete;
      End;
      Else If &Val_M gt 1 then Do;
        If Month eq (&Val_M-1) and tIndex ge (&ValIdx-365) and tindex
lt &ValIdx then output;
        Else delete;
      End;
  End;
```

```
      Else If &Val_D gt 1 then Do;
        If Month eq &Val_M and tIndex ge (&ValIdx-365) and tindex lt
&ValIdx then output;
         Else delete;
      End;

     Drop Cons;
   Run;

*Call macro to compute k, q, sigma;
   %QKSigma(&ValIdx,&StartIdx,&LastIdx,&Setle_M,&Setle_D,&SetleIdx)

*Simulate the consumption rate following the equation (6);
   Data Consmptn;
     Merge Consmptn TmpRsdlM30;

    Array v{&NumSimul} v1-v&NumSimul;
    Array TRF_v{&NumSimul} TRF_v1-TRF_v&NumSimul;
     Array ln_Cons{&NumSimul} lc1 - lc&NumSimul;
     Array Cons{&NumSimul} c1 - c&NumSimul;
     Array ln_Cons_TRF{&NumSimul} lcTRF1 - lcTRF&NumSimul;
     Array Cons_TRF{&NumSimul} cTRF1 - cTRF&NumSimul;

    a = -&Rho*(&SetleIdx-&ValIdx)/(&Gamma*365.) - (&Mu-1.)*lnCons;

    Do i=1 to &NumSimul;
      ln_Cons{i} = a + &Mu*lnCons + v{i};
      Cons{i} = exp(ln_Cons{i});
      ln_Cons_TRF{i} = a + &Mu*lnCons + TRF_v{i};
      Cons_TRF{i} = exp(ln_Cons_TRF{i});
    End;
     Keep c1 - c&NumSimul cTRF1 - cTRF&NumSimul;
   Run;

*Simulate a price following the equation (7);
   Data Price&n;
     Merge Price&n Consmptn;
```

```
      Array cumTmpID{&NumSimul} CTI1 - CTI&NumSimul;
      Array Cons{&NumSimul} c1 - c&NumSimul;
      Array Cons_TRF{&NumSimul} cTRF1 - cTRF&NumSimul;
      Array PremPrice{&NumSimul} Price1 - Price&NumSimul;
      Array PremPrice_ActDD{&NumSimul} P_ActDD1 - P_ActDD&NumSimul;

      Do i=1 to &NumSimul;
        PremPrice{i} = ((Cons{i}**&Gamma)*cumTmpID{i})   /
((Cons_TRF{i}**&Gamma)*1.);
          SumPrice + PremPrice{i};

        PremPrice_ActDD{i} = ((Cons{i}**&Gamma)*Sum_ActDD)   /
((Cons_TRF{i}**&Gamma)*1.);
          sumPrice_ActDD + PremPrice_ActDD{i};
      End;

      Price_SimDD = SumPrice/&NumSimul;     *Price w Simulated Temps;
      Price_ActDD = SumPrice_ActDD/&NumSimul; *Price w Actual Temps;

      Format Price_SimDD Price_ActDD 11.5;
      OBS = &n;                          *To merge w Thesis.Set_i;
      Keep Price_SimDD Price_ActDD OBS;
    Run;

   Proc Datasets library=Work;
     Delete ParaEsti Tmp_Parm Consmptn Resid TmpRsdlM30;
   Run;
%Mend PriceHDD;
```

Appendix B-3: SAS Code to Invoke the Pricing Macro

```
*SETTINGS OF GLOBAL VARIABLES;
  %let NumSimul = 5000;          * # of Simulation per price;
  %let Rho = 0.03;                    * Rate of Preference;
  %Global k sigma PrSimDD PriceActDD SumActDD;
  %let avg_vt=;
  %let avg_xit=;
  %let nobs;
  %let avg_vt=;

*MACRO STATEMENT CONTRACT;
%Macro Contract(Gamma, Mu);

 Data _null_;
   Set MktPrice End=Last;
   If Last then call symput('n',_n_);
 Run;

*Assign macro parameters into the Macro PriceHDD;
 %Do i=1 %To &n;
   Data _null_;          *to solve the timing issue of call execute;
     Set MktPrice;
      If _n_=&i then Do;
        If Indicator='F' then call Execute('%PriceHDD(' || Val_ID || ','
|| Start_ID || ',' || Last_ID || ',' || Setle_ID || ')');
        Else call Execute('%PriceCDD(' || Val_ID || ',' || Start_ID ||
',' || Last_ID || ',' || Setle_ID || ')');
      End;
    Run;

*Output the price;
   Data Thesis.Set_9_MktPrice;
     Set MktPrice;
     If _n_=&i then Do;
       Price_SimDD = &PrSimDD;
        Price_ActDD = &PriceActDD;
```

```
      Sum_ActDD = &SumActDD;
      k = &k;
      sigma = &sigma;
    End;
   Run;
 %End;

*Compute the percentage errors;
Data Thesis.Set_9_MktPrice;
   Set Thesis.Set_9_MktPrice;
    PriceDiff = Price_SimDD - ActPrice;
   DiffPcnt = PriceDiff / ActPrice;
   Format DiffPcnt percent7.2;
    Format PriceDiff 10.5;
 Run;

 Proc Print data=Thesis.Set_9_MktPrice;
 Run;
%Mend Contract;
```