No Chills or Burns from Temperature Surprises: An Empirical Analysis of the Weather Derivatives Market

Ludwig Chincarini*

This version: May 19, 2009.

Abstract

This article examines the efficiency of the weather futures market traded on the CME in both HDD and CDD futures contracts in 18 cities across the United States. Efficiency is examined in three ways. Firstly, by comparing the market's implied forecasts for the weather against other forecasts. Secondly, by looking at whether market's overreact or underreact to temperature *surprises*. Thirdly, by looking at weather derivative patterns across cities. We find that generally the market seems very efficient despite its lack of liquidity. We also find significant risk premia that do not seem to reflect the standard compensation for speculators.

JEL Classification: G0, G14
Key Words: Weather derivatives, weather forecasting, market efficiency, HDD, CDD

*I would like to thank John Wick and Yuxiang Wu for research assistance. I thank Daehwan Kim for valuable suggestions, Jan Carpenter, Kathryn Gilbert, Paul Knight, Christopher Hyde for details of the official HDD calculation, Heidi Stonehill, Ken Heideman, Jamie Meier, Scott Stephens, for aid in obtaining weather forecasting data, Bloomberg, the FEI Institute for access to their data resources. I also thank an anonymous referee, Amir Barnea, Pierangelo DePace, Eric Hughson, John Jurewitz, Slavi Slavov, Robert F. Whitelaw, and Fan Yu for suggestions. We also thank the National Climatic Data Center for historical weather data. Contact: Ludwig Chincarini, CFA, Ph.D., is an assistant professor at the Department of Economics, Pomona College, 425 N. College Avenue #211, Claremont, CA 91711. Email: `chincarinil@hotmail.com`. Phone: 909-621-8881. Fax: 909-621-8576.

I Introduction

The weather futures market is a relatively new market. Weather is different than many other types of financial instruments in that it has no underlying measurable value. Weather futures contracts are used by producers to hedge their business risks due to factors that may affect the demand for their products. For example, a natural gas producer or oil producer's quantity demanded will be effected by the weather in a particular city. A mild winter will cause the demand for natural gas and oil to decline. On the other hand, a farmer might use weather derivatives to hedge the future yield of a crop. If the weather is poor, which might depend on a variety of factors, including rainfall, temperature, frost, etc., then using weather derivatives might aid in hedging the quantity supplied or available for the farmer. On the other side of this transaction would be the speculator. Presumably, the speculator would receive some compensation for his activities of providing liquidity to producers to hedge business risk. In this paper, we examine this as well as study the efficiency of the weather futures market.

It has been shown that even in very established and liquid markets, like the US stock market and other developed markets around the world, there seems to be inefficiencies due to investor overreaction (DeBondt and Thaler (1985, 1987), Zarowin (1989), Jegadeesh and Titman (1995), Loughran and Ritter (1996), Chiao and Cheng (2005), and Zhu (2007)). Given the inefficiency in established markets due perhaps to investor biases or other reasons, we might expect to see such inefficiencies to a larger extent in the weathers' market, which is not only less liquid, but also does not have any fundamental pricing model at its heart. There is a short history of weather derivatives and even a shorter history of literature in the area. Much of the previous literature is concerned with building models to price weather risk. Weather risk is slightly more complicated to price, since there is no underlying instrument to hedge with, like there are with other futures contracts and options. Because risk-neutral pricing models a la Black-Scholes will not work in these markets, various articles have been written on pricing weather derivatives (Garman et al. (2000), Brody et al. (2002), Cao and Wei (2004), Richards et al. (2004), Geman (2005), Taylor and Buizza (2006)). A good summary of weather derivative valuation can be found in Jewson et al. (2007). Another strand of research in this area has focused on the design of optimal weather contracts for producers of commodities. These papers discuss, design, and examine the effectiveness of different types of weather derivatives (Brockett et al. (2005)and Leggio (2007)). Other research has focused on the usefulness of weather forecasts and other types of models to forecast the weather (Zeng (2000), Campbell and Diebold (2005), Benth and Benth (2006), and Brix et al. (2002)). There are some papers associated with using weather derivatives for hedging, such as Rolfo (1980). For a collection of chapters on many aspects about the weather markets, see Dischel (2002).

To our knowledge, none of the prior literature has focused itself on the testing of the efficiency of the weather futures market. In this paper we examine the efficiency of the weather futures market by examining the accuracy of national weather forecasts, the accuracy of simple models of the weather and the accuracy of the actual market prices of weather futures. We also borrow techniques from the overreaction literature to test whether there seems to be overreaction in the weather futures market.

The paper is organized as follows: section II discusses the details about the weather futures market; section III discusses the data that we use in our research; section IV discusses the methodology which we use to test for the efficiency of the weather futures market; section V discusses the results from our empirical investigation of the weather futures market; and section VI concludes the paper.

II The Weather Derivatives Market

Weather derivatives were introduced on the Chicago Mercantile Exchange (CME) in 1999. The growth in weather derivatives usage has been large. By September 2005, the notional value of weather contracts stood at $22B with over 630,000 contracts being traded on the CME. CME weather products are temperature-based indices of futures and options. These derivatives exist for 18 U.S., nine European, and two Asia-Pacific cities.[1]

The derivatives trade based on a measured value of the temperature in each city. The daily indices, upon which the futures and options are based upon, are the calculated as HDD and CDD. The HDD stands for heating degree days and the CDD stands for cooling degree days. The HDD and CDD are computed each *calendar day* for each city upon which contracts trade. The HDD is computed as $\max[65 - x, 0]$ where 65 represents a fixed reference temperature of 65 degrees Fahrenheit and x represents the daily average temperature in that city defined as the arithmetic average of the daily maximum and minimum temperatures.[2] The CDD is computed as $\max[x - 65, 0]$, where variables are defined above.[3]

Many contracts trade based upon the accumulation of HDD or CDD. Thus, the monthly index value of HDD and CDD are important for determination of a contracts final payout. The monthly HDD and CDD index are simply the sum of the values of the daily HDD and CDD values for that particular month. Thus, if there were 5 values for HDD for the month of January with 25, 15, 20,

[1] In particular, there are contracts for Atlanta, Baltimore, Chicago, Cincinnati, Dallas, Des Moines, Detroit, Houston, Kansas City, Las Vegas, Minneapolis-St. Paul, New York, Philadelphia, Portland, Sacramento, Salt Lake City, Tucson for the United States. In Europe, there are products for Amsterdam, Netherlands, Barcelona, Berlin, Essen, London, Madrid, Paris, Rome, and Stockholm. For the Asian region, there are products for Tokyo and Osaka.

[2] The maximum and minimum values of temperature in that city are computed from midnight to midnight each day. In the United States, they are determined by the Earth Satellite Corporation which uses temperatures obtained from the National Climate Data Center (NCDC). For European and Asian cities, Earth Satellite Corporation also provides settlement values.

[3] The reason that HDD is zero for values above 65 and CDD is equal to zero for values below 65, is that heating degree days refers to days in which one would need to use a heater, while CDD refers to days one would need to use an air conditioner. Of course, the rationale for the creation of these contracts is immaterial to their application.

10, and 15, then the monthly HDD value would be 85. To get the value of the specific contract, one must multiply the HDD or CDD by the contract size. Table 1 contains the major features of the weather contracts in the United States. A simple example using one of the specific city contracts might makes things clearer. On February 28, 2005, the monthly HDD contract for Atlanta closed at 305. This indicated that the market's fair value for the sum of HDD daily values in Atlanta was 305. The weather for the HDD contract is shown in Table 2 along with the daily closing prices of the March 2006 futures contract. For this month, the market underestimated the actually final value of weather which turned out to be 349. Of course, the final settlement of the contract was equal to 349 on the first business day of the following month. Someone who purchased the HDD contract on February 28, 2008 would have paid $6,100 for one contract and had they held until expiration would made $880 ($6980-$6100).

[INSERT TABLE 1 ABOUT HERE]

[INSERT TABLE 2 ABOUT HERE]

In addition to the basic HDD and CDD futures contract for individual months and the option contracts based upon these months (both European and American), there are also HDD and CDD Seasonal strips. These are futures contracts which trade on multiple months rather than just one specific month. For HDD Seasonal strips, there are contracts with a minimum of two consecutive months and a maximum of seven consecutive months. For CDD Seasonal strips, there are a minimum of two consecutive months and a maximum of six consecutive months. There are options on all of the seasonal strips which trade European style only. For example, an H2VJ6 is an HDD Seasonal strip that trades based upon the values of HDD in Chicago (2) for the months from October (V) to the April (J) in the year 2006 (6). There are numerous strips with varying time spans available for users to trade weather in a series of months versus one individual month.

The HDD seasonal strips have start months from October - March, while CDD seasonal strips have start months from April - September. Our paper will focus only on the HDD and CDD monthly futures contracts.

The main types of users in the weather market include energy traders, hedge funds, institutional banks and re insurers. Many energy companies, such as heating oil distributors use weather derivatives to help them smooth their profits from fluctuations in demand in heating needs during winter. There is a very high correlation between HDDs and gallons of heating oil consumption in winter months. Thus, weather derivatives, including futures, options, and custom-made varieties help these companies hedge the quantity demanded of their product, in addition to the hedging the price of their product which can be done with other instruments.

III The Data

In order to test the efficiency of the weather futures market we collected the following data. For weather futures data, we collected the daily closing prices of each weather contract in each of the U.S. cities in which weather trades since the opening of exchange-traded weather contracts in 1999 from Bloomberg. We also collected daily volume and open interest information from Bloomberg. For each city, we collected both the HDD and CDD monthly futures contracts. Our futures data spans the period January 1, 1999 to July 8, 2008. Table 3 summarizes the futures contracts for all of the cities. For each of these contracts, the first contract available for trading is listed. For example, for Atlanta HDD, the first available contract for trading was the October 1999 HDD contract based on weather for October, 1999. The next column indicates the average daily trading volume for each contract. This average is computed across all trading days for which there is a volume measure. For example, for the Atlanta HDD contract, the average daily trading volume was 52.45. Thus, when there is trading, about 52 contracts trade per day on average. The last column represents the number of days of which there is volume information for any of the contracts for that particular city. For example, for Atlanta HDD, there were 934 trading days with volume for the entire sample period.

[INSERT TABLE 3 ABOUT HERE]

Our historical actual weather data was obtained from the archives of the National Climatic Data Center, which is the government agency which stores all official records of temperatures and HDD and CDD values on which the futures contracts are based. For all of our 18 major cities, we have daily maximum and minimum temperatures for each 24-hour period going back as far as data was collected for that particular city. For some cities, we have daily data since 1880. For the purposes of our work, we only use data from January 1, 1990 onwards. The summary statistics for the weather in each of these cities is contained in the supplemental appendix.

In addition to historical weather data, we obtained historical weather forecast data as produced by National Weather Service Model Output Statistics (MOS) Global Forecast System (GFS) guidance model. This model produces forecasts of the daily maximum and minimum temperature for weather stations around the United States for up to 7 days forward. The archived forecasts begin in September 2005. Thus, we have daily 7-day forecasts for all our cities since September, 2005. One important difference between the forecast data and the temperature data should be pointed out. While the daily max and min temperatures are for a 24-hour period from midnight to midnight at each local station, the GFS forecasts are different. The extended GFS forecasts are produced at 00:00 Greenwich Mean Time (GMT) and at 12:00 GMT. The forecasts for the maximum temperature is for period from 7 A.M. to 7 P.M. local time of that particular weather station. The forecasts for minimum temperature are from 7 P.M. to 8 A.M. local time for that particular station. Thus, the historical data and forecast data do not coincide, but overlap.

Figure 1 may help illustrate the issue. This particular timeline is constructed for the a city in the same time zone as Atlanta. All of the other cities should be adjusted according to their time zone, but roughly the same concepts apply. At 00:00 GMT, the NWS releases their MOS forecasts for all cities. In the case of Atlanta, the minimum forecast for the *next day*, day t, would be for the measurement period from 7:00 PM Atlanta time on day $t-1$ to 8 A.M. Atlanta time on day t. Thus technically the forecast for the minimum temperature occurs over a period covering both day $t-1$ and day t for Atlanta. The maximum temperature forecast released at 00:00 is for the period illustrated in the figure as *MOS Forecast Period for Max on Day* t, which would be from 7 A.M. to 7 P.M. Atlanta local time on day t. The actuall HDD and CDD measurement for the derivative contracts are based upon the 24-hour period covering only day t for Atlanta (see the area in the figure labeled *Measurement of Daily Max. and Min. for HDD and CDD on Day* t.

[INSERT FIGURE 1 ABOUT HERE]

While it is true, that most of the time, the minimum temperature will occur from midnight to 7 AM and that the maximum temperature will occur from 7 A.M. to 7 P.M. and the two measures will for all practical purposes correspond, this not need be the case. For instance, when there are extreme weather conditions, the minimum temperature may occur later in the day, and forecasters might wish to diverge from the GFS model. We explain more about this in other parts of the paper. A summary of the weather forecast data is presented in the Supplemental Appendix.

Finally, we remove any closing prices without volume information for that particular day. We do this since we worry about the integrity of the price information when there are days with no volume and hence no trades took place at those prices.[4] Also, in order to remove potentially bad

[4]Although some may argue that these are legitimate quotes without any volume, we felt it was worse to leave them in then to remove them.

data, we drop all closing prices that are either two-times the historical HDD or CDD for that particular month in that particular city or are 1/2 of that particular month for that particular city. We did use alternative cutoffs and it did not affect the nature of the results.

IV The Efficiency of the Weather Futures Markets

In this section, we discuss what the theoretical price of weather futures might be and we describe the three ways that we use to measure the efficiency of the weather derivatives market.

A Futures Premia and Pricing

The typical pricing of a futures contract relies on the futures-spot price relationship based upon a cost-of-carry model. Thus, in many futures markets,

$$F_t = S_t e^{r(T-t)} \tag{1}$$

where F_t is the futures price at date t, S_t is the spot price at date t, r is the interest rate, and $T-t$ is the time until expiration. Unfortunately, this relationship is not entirely useful for the weather market, because there is no underlying spot weather for a given futures contract.[5]

Thus, rather than describe the forward premia theoretically, we calculate it empirically by computing the normal backwardation or contango of the market.[6] Thus, we could look at the forward premium as:

[5]In fact, Campbell and Diebold write that "...standard approaches to arbitrage-free pricing are irrelevant in weather derivative contexts, and so the only way to price options reliably is again by modeling and forecasting the underlying weather variable."

[6]Recently, more complicated models for weather derivative pricing have been proposed, but are not entirely satisfactory for our purposes either. See Hardle and Cabrera (2009).

$$FP_{im,t} = E_t[F_{imt} - S_{im,t+1}] \tag{2}$$

where F_{imt} is the futures price of the HDD (or CDD) contract in city i on day t for contract of month m and $S_{im,t+1}$ is the realized HDD (or CDD) in city i for the given particular month, m. According to Keynes, speculators should be given a premium for taking on the role of liquidity provider. If we think of a typical scenario, where energy distributors are attempting to hedge energy demand in a particular city, then this entity will typically wish to be short the contracts. That is, if winter temperatures tend to be higher than normal, there will be less demand for energy and the distributor would wish to compensate for this by being short HDD contracts. The same will be true for summer months. Thus, the speculator will be net long the weather contracts and hence we would expect according to a liquidity argument that $F_{it} < E(S_{i,t+1})$ for any city i, and we might also expect that $F_{it} - E(S_{i,t+1})$ will be larger for cities with more volatile temperatures in any given month.

Table 4 shows the average returns to holding both HDD and CDD futures contracts from day of purchase until month-end for all of the US cities. The table shows that while forward premia are insignificant for most cities, they are significant for Baltimore, Chicago, Dallas, Minnesota, and New York for HDD and significant for Des Moines, Minnesota, and New York for CDD. In these specific cases for HDD the premia are all positive, contrary to what we postulated in the earlier section. The return to buying these HDD futures is negative and isn't accounted for merely from interest rates. The only exception is New York, where CDD contracts provide a negative forward premia.

[INSERT TABLE 4 ABOUT HERE]

Table 5 shows the realized forward premia by day of month. From this table, we still observe that most premia are positive but insignificant with the exception of day 6, which might simply be an aberation. Overall, the results indicate that there do not seem to be significant premia to speculators in weather derivatives, which is puzzling.

[INSERT TABLE 5 ABOUT HERE]

B The Accuracy of Market Predictions of Weather

In this section, we are concerned with how precise the markets forecast weather. Our approach is to construct several models of weather forecasts, static as well as dynamic, and to compare them to the weather forecasts implicit in the market price of the weather derivatives.

Let's define a few variables. Let $T_{i,y,m,d}$ be the daily average temperature of city i on year y, month m, and day d. As noted earlier, $T_{i,y,m,d}$ is a simple average of the daily maximum temperature $H_{i,y,m,d}$ and the daily minimum temperature $L_{i,y,m,d}$, i.e.

$$T_{i,y,m,d} \equiv \frac{H_{i,y,m,d} + L_{i,y,m,d}}{2} \tag{3}$$

We define the daily HDD as

$$HDD_{i,y,m,d} \equiv \max(65 - T_{i,y,m,d}, 0) \tag{4}$$

The monthly HDD is the sum of daily HDDs, i.e.:

$$HDD_{i,y,m} \equiv \sum_{d=1}^{D(y,m)} HDD_{i,y,m,d} \tag{5}$$

where $D(y,m)$ indicates the number of days in the month. We define corresponding quantities for

CDD in a similar way.

The quantity of our primary interest is the monthly HDD $HDD_{i,y,m}$ or the monthly CDD $CDD_{i,y,m}$. We will work with six models that forecast this quantity.

B.1 Static Models

The first static model, which we call Static Historical Model, is a simple model that uses an historical average of monthly HDD or CDD as its forecast. We construct the historical average of monthly HDD by taking a simple average of the HDD in that particular month going back historically as far back as 30-years, but not including the current month.[7] For example, for March 2005, the value of the historical average of monthly HDD would be the average of all the previous March HDD including March 2004 HDD, but excluding March 2005 HDD. Denoting the historical average monthly HDD by $\overline{HDD}^{HIST}_{i,y,m,}$,

$$\overline{HDD}^{HIST}_{i,y,m} \equiv \frac{1}{y-y_0}\sum_{y'=y_0}^{y-1} HDD_{i,y',m} \tag{6}$$

where y_0 is the first year in our data set.

The second static model, which we call Static MOS Model, combines the historical average of monthly HDD or CDD with MOS weather forecasts. It is only constructed once on the day before the month begins. We describe the model for HDD here. The model for CDD is similar.

On each day, the NWS MOS produces seven-day forecasts for the max and min temperatures in each city. Let us denote these forecasts as $H^{MOS}_{i,y,m,d,1}, \cdots, H^{MOS}_{i,y,m,d,7}$ and $L^{MOS}_{i,y,m,d,1}, \cdots, L^{MOS}_{i,y,m,d,7}$, where the last subscript indicates the number of days ahead for which the forecast is made. The average of the max forecast and the min forecast is our forecasts for the average daily temperature,

[7]Practitioners seem to use a 30-year window or a 10-year window to forecast weather temperatures (see Dischel p. 268). We begin with a 40-year window and increase the window as the years progress from 2000 to 2008.

which we denote as $T^{MOS}_{i,y,m,d,1}, \cdots, T^{MOS}_{i,y,m,d,7}$. We construct the MOS forecasted daily HDD as[8]

$$\overline{HDD}^{MOS}_{i,y,m,d,j} \equiv \max(65 - T^{MOS}_{i,y,m,d,j}, 0), \quad j = 1, \cdots, 7. \tag{7}$$

Note that we are using the adjustment factor for the 7th day of the month ($d = 7$), which is what is needed given the MOS forecasts.

At the beginning of the month, we have the MOS forecast daily HDD for the first seven days of the month. We convert these figures into monthly HDD using the HDD adjustment factor.

$$\overline{HDD}^{MOS}_{i,y,m} \equiv \sum_{j=1}^{7} \overline{HDD}^{MOS}_{i,y,m-1,D(y,m-1),j} + HADJ_{i,y,m,7} \tag{8}$$

where

$$HADJ_{i,y,m,d} \equiv \frac{1}{y - y_0} \sum_{y'=y_0}^{y-1} \left[\sum_{d'=1}^{D(y,m)} HDD_{i,y,m,d'} - \sum_{d'=1}^{d} HDD_{i,y,m,d'} \right] \tag{9}$$

This forecast for the entire month uses the first 7 days forecast and then uses the historical average of temperature for that city and that month for the rest of the month as the monthly forecast for HDD.[9]

The third static model, which we call Static Market Model, uses the monthly HDD implied by the closing futures price. We use the closing prices of the contract on the day before the month begins. That is, for city i, year y, and month m, we use the closing prices of the futures on the last day of month $m - 1$, which we denote as $F_{i,y,m-1,D(y,m-1)}$.

One might argue that these alternative models are too easy to beat. For the historical model,

---

[8]The corresponding value for CDD is $\overline{CDD}^{MOS}_{i,y,m,d,j} \equiv \max(T^{MOS}_{i,y,m,d,j} - 65, 0), \quad j = 1, \cdots, 7.$

[9]Although not reported here, another adjustment factor was used that converted the first 7-days of the forecast into a monthly forecast by multiplying that the ratio of the entire month's HDD to the first seven days by the 7-days forecast. In all cases, the MOS forecasts were inferior to the market forecasts and in some cases inferior to the historical forecasts.

one might argue that it doesn't incorporate the future. For the NWS MOS forecasts, one might argue that it only contains 7-days of forecast data and the rest of the month's prediction is based upon the historical averages. While these are legitimate concerns with the benchmarks, a few comments are in order. First, this is a first attempt to examine the efficiency of weather market forecasts and to compare against some simple alternatives. Even if it is true that they are simplistic, we will learn something depending on if and by how much the market forecasts improve upon them. Second, there are no other publicly available government numerical forecasts for temperatures in these cities other than the government's GFS operational forecasts, which go out 15 days. Thus, data limitations limit the ability to produce a more accurate model forecast.[10] Third, studies using slightly longer forecasts beyond 8 days do not perform particularly well anyway, which might not help improve these models vis-a-vis the market's forecast.[11]

B.2 Dynamic Models

The first dynamic model, which we call Dynamic Historical Model is the actual HDD through day d of the month plus the historical average HDD for the rest of the month. Thus,

$$\overline{HDD}^{HIST}_{i,y,m,d} \equiv \left(\sum_{d'=1}^{d} HDD'_{i,y,m,d} \right) + HADJ_{i,y,m,d} \tag{10}$$

The second dynamic model, which we call the Dynamic MOS Model computes the 7-day HDD from the 7-day temperature forecasts and combines this with the historical value of HDD for the rest of the month. This dynamic forecast is given by:

[10] Recently, the European Center for Medium-Range Weather Forecasts (http://www.ecmwf.int/) has created an ensemble series of 32-day forecasts for many cities around the world, however they only have 8-months of historical data at the time of the writing of this paper. In the near future, it certainly would be of interest to expand on this current research with this longer-dated forecast horizon.

[11] For example, Campbell and Diebold (2005) note that time series models are not as good as NWP forecasts produced by EarthSat up to a horizon of 8 days, but after that all models performed equally well.

$$\overline{HDD}^{MOS}_{i,y,m,d} \equiv \sum_{d'=1}^{d} HDD_{i,y,m,d} + \sum_{j=1}^{k} \overline{HDD}^{MOS}_{i,y,m,d,j} + HADJ_{i,y,m,d} \cdot \Phi(d) \tag{11}$$

where where $\Phi(d)$ is an indicator variable that takes on a value of 1 if $D - d > 7$ and is equal to 0 otherwise, $k = 7$ if $D(y,m) - d >= 7$, and $k = D(y,m) - d$ if $D(y,m) - d < 7$. This captures the idea that as we approach the end of the month, we cannot use all of the 7-day forecasts for the forecast of the remainder of the month when there are less than 7-days left until the end of the month.

The third dynamic model, which we call Dynamic Market Model, uses the last day's closing futures price as the forecast for the rest of the month, i.e. $F_{i,y,m,d}$.

C Market Surprise and Overreaction

One method to test whether markets over-react or under-react to weather developments is to create an index for the surprise of temperature on a particular day. The surprise measure is defined as $\hat{S}_{i,y,m,d} = \frac{T_{i,y,m,d} - E(T_{i,y,m,d})}{SD(T_{i,y,m,d})}$, where $T_{i,y,m,d} - E(T_{i,y,m,d})$ is the actual temperature on a given day minus the expected value of the daily average temperature, and $SD(T_{i,y,m,d})$ is the standard deviation of temperatures on that particular day in that particular city historically.[12] For $E(T_{i,y,m,d})$, we use three different measures described below.

The first measure of $E(T_{i,y,m,d})$ is to use the historical average of that temperature in the past on that given day. That is, $\frac{1}{y-y_0} \sum_{y'=y_0}^{y-1} T_{i,y',m,d}$.

The second measure is to use the latest MOS forecast for that particular day. That is, $T^{MOS}_{i,y,m,d-1,1}$.

[12] Dividing by the standard deviation of historical temperatures on that day is one way to normalize surprises across cities and a similar measure is used in other finance literature like analyst surprise forecasts.

The third is to use the actual weather future's closing prices. The future prices are based upon monthly HDD, and we will infer the daily average temperature from the monthly HDD implied by the future price. On any given day of the month, the HDD contract closing value $F_{i,y,m,d}$ is related to the daily HDDs in the following way:

$$F_{i,y,m,d} = \sum_{d'=1}^{d} HDD_{i,y,m,d'} + \sum_{d'=d+1}^{D(y,m)} \overline{HDD}_{i,y,m,d'}^{MKT} \tag{12}$$

where $\overline{HDD}_{i,y,m,d'}^{MKT}$ is the daily HDD forecast of the market, which is not directly observable. In order to uncover the expected value of the market for HDD on day $d+1$, we construct another adjustment factor. This adjustment factor is given by:

$$HADJ2_{i,y,m,d} \equiv \frac{1}{y-y_0} \sum_{y'=y_0}^{y-1} \frac{\frac{1}{D(y,m)} \sum_{d'=1}^{D(y,m)} HDD_{i,y,m,d'}}{\frac{1}{d} \sum_{d'=1}^{d} HDD_{i,y,m,d'}} \tag{13}$$

This is just a calcuation of the ratio of the average HDD over a month in a given city divided by the average HDD to day d in that month. For our recovery of the market's expectation of HDD on day $d+1$, we assume that the adjustment factor is relevant so that

$$\frac{\frac{1}{D(y,m)}(\sum_{d'=1}^{d} HDD_{i,y,m,d'} + \sum_{d'=d+1}^{D(y,m)} \overline{HDD}_{i,y,m,d'}^{MKT})}{\frac{1}{d+1}(\sum_{d'=1}^{d} HDD_{i,y,m,d'} + \overline{HDD}_{i,y,m,d+1}^{MKT})} = HADJ2_{i,y,m,d+1} \tag{14}$$

which implies

$$\begin{aligned}\overline{HDD}^{MKT}_{i,y,m,d+1} =& \frac{1}{HADJ2_{i,y,m,d+1}} \frac{d+1}{D(y,m)} \\ & \left(\sum_{d'=1}^{d} HDD_{i,y,m,d'} + \sum_{d'=d+1}^{D(y,m)} \overline{HDD}^{MKT}_{i,y,m,d'} \right) \\ & - \sum_{d'=1}^{d} HDD_{i,y,m,d'} \\ =& \frac{1}{HADJ2_{i,y,m,d+1}} \frac{d+1}{D(y,m)} F_{i,y,m,d} - \sum_{d'=1}^{d} HDD_{i,y,m,d'}.\end{aligned} \tag{15}$$

Finally, we convert the daily HDD forecast into the daily average temperature forecast:

$$\bar{T}^{MKT}_{i,y,m,d} = 65 - \overline{HDD}^{MKT}_{i,y,m,d} \tag{16}$$

Note that $\overline{HDD}^{MKT}_{i,y,m,d}$ could be negative, which is not allowed in the original definition of the daily HDD.[13]

Once these series are created, we collect the value of all of these for all days and then sort them by decile. We then compute the returns of buying the contract at the close of business of the *next day* and holding until month's end.[14] We can then draw histograms of the decile results and test for significant difference in mean returns. If on days of high positive surprise (i.e. the temperature is higher than people expected) we see lower returns, then there was over-reaction because market traders reacted to the surprise by raising their estimate of the future month's HDD by too much and vice versa. We would expect the opposite for CDD contracts, that is on days of high positive surprise, overreaction would be given by subsequent higher returns.

Since the HDD contracts trade based upon the entire month of average daily weather temper-

[13] A similar calculation is constructed for CDD contracts.

[14] We also compute the results for purchase on the same day, which are contained in the Supplemental Appendix.

atures, there is also a dynamic component to the under- or over-reaction of market participants. Thus, overreaction or underreaction may be different conditional on past weather surprises. There may be *market learning*.[15]

In order to determine whether the differences in returns between the highest and lowest quintile are significant, we consider two statistical measures; a two-sample t-test and a Mann-Whitney U test for the difference in means between deciles 1 and 5.[16]

D The Inter-Market Behavior of Weather and Futures Markets

Our final approach to investigate the efficiency of the weather markets is to look at the cross-correlation of daily weather changes across cities and compare that to the changes in the daily average temperature implied by the market to get an idea if there seems to be some kind of cross-city inefficiencies.

As a heuristic test of the inter-market efficiency, we implement the following trading strategy. The strategy compares the historical differences in HDD among cities and compares them to the

[15] Although not done in this paper, one idea for testing market learning might be XXX.

[16] The first measure is a two-sample t-test for the differences in means between two samples. The test statistic is:

$$U = \frac{Z_1}{[Z_2/(m+n-2)]^{1/2}} \tag{17}$$

where $Z_1 = \frac{\bar{R}_1 - \bar{R}_5}{\left(\frac{1}{m}+\frac{1}{n}\right)^{1/2}\sigma}$, $Z_2 = \frac{S^2_{R_1}+S^2_{R_5}}{\sigma^2}$, $\bar{R}_1$ represents the mean return from quintile 1, $\bar{R}_5$ represents the mean return from quintile 5, σ is the population standard deviation, m is the number of observations from quintile 1 and n is the number of observations from quintile 5, $S^2_{R_1} = \sum_{i=1}^{m}\left(R_{1i} - \bar{R}_1\right)^2$ and $S^2_{R_5} = \sum_{i=1}^{n}\left(R_{5i} - \bar{R}_5\right)^2$. Given certain assumptions, U will be distributed as a t-distribution with $m+n-2$ degrees of freedom (i.e. $U \sim t_{m+n-2}$).

Since many of our deciles have a very small amount of observations, we also computed the Mann-Whitney U test to test for differences in the means with a few number of data points (Mann and Whitney (1947)). The procedure for the test is as follows:

1. Rank all observations in quintile 1 and quintile 5 from smallest to largest.
2. Sum the ranks of quintile 1 and call this SR_1.
3. Compute the Mann-Whitney test statistic as: $U = mn + \frac{m(m+1)}{2} - SR_1$.
4. Large values of this statistic suggest that the samples are drawn from different populations where quintile 1 has smaller means.
5. Compute z-statistics for difference in means for larger samples. For small samples, use a statistical table of the Mann-Whitney test to determine whether decile 1's location parameter is great than that of quintile 5.

current differences in HDD implied by market prices. Then we choose 3 pairs of cities whose current differences exceed the historical differences most and create zero-investment portfolios out of these pairs. Similarly, we choose 3 pairs of cities whose current differences fall below the historical differences most and create zero-investment portfolios out of these pairs. If there is no inter-market inefficiencies, these strategies should not produce any abnormal returns. Of course, not finding any abnormal returns is not enough to prove that these market are efficient, but is consistent with market efficiency.

We provide the details of the strategy. First, for each month in the data set, we calculate the historical HDD difference matrix. The historical HDD difference matrix is a $N-1$ by $N-1$ upper diagonal matrix, where N is the number cities:

$$\Delta^{Hist}_{y,m} = \begin{pmatrix} \overline{HDD}^{HIST}_{1,y,m} - \overline{HDD}^{HIST}_{2,y,m} & \overline{HDD}^{HIST}_{1,y,m} - \overline{HDD}^{HIST}_{3,y,m} & \cdots \\ & \overline{HDD}^{HIST}_{2,y,m} - \overline{HDD}^{HIST}_{3,y,m} & \cdots \\ & & \ddots \end{pmatrix} \tag{18}$$

Next, for each day in the data set, we calculate the market HDD difference matrix, which is the diffrences in the monthly HDD implicit in the market prices:

$$\Delta^{MKT}_{y,m,d} = \begin{pmatrix} F_{1,y,m,d} - F_{2,y,m,d} & F_{1,y,m,d} - F_{3,y,m,d} & \cdots \\ & F_{2,y,m} - F_{3,y,m} & \cdots \\ & & \ddots \end{pmatrix} \tag{19}$$

For each day, we take the differences in these two matrices $\Delta^{MKT}_{y,m,d} - \Delta^{Hist}_{y,m}$ and find the three cells with the largest absolute values. Let us denotes these cells as $(i_1, j_1), (i_2, j_2), (i_3, j_3)$. These are the pairs of cities for which we suspect "mis-pricing." To be more specific, we suspect the prices for j_1, j_2, j_3 are too low and the prices for i_1, i_2, i_3 are too high. Thus, we buy j_1, j_2, j_3 and sell

i_1, i_2, i_3. Each of long and short positions are equally weighted. This creates a long-short portfolio to exploit a potential "mispricing" opportunity. Then we calculate the return of these portfolios until the end of the month. Significant positive returns indicate cross-market inefficiencies.

V Empirical Results

A The Accuracy of Market Predictions of Weather

A.1 Static Models

Given the simple models for forecasting the weather, we first started with three models to forecast the monthly HDD and CDD. The period of study was from September 2005 to June 2008. Although data exists for the derivative contracts for many cities going back to 1999 and for historical weather going back even further, we begin our data sample for when the first NWS MOS forecasts are available so as to keep everything symmetric. One drawback to this symmetry is that we lose many observations, nevertheless it's the only fair comparison of the models.[17]

Table 6 contains the root-mean-squared error (RMSE) and mean-absolute error (MAE) for each of the three models for both HDD contracts and CDD contracts. It also contains a column entitled MAE Dif which computes the percentage difference in MAE between the historical and MOS forecast model versus the market model. For HDD and CDD contracts, the market prices of the weather at the beginning of the month are a better predictor of the eventual month's HDD (as well as weather) then either a simple historical forecast or an interpolated NWS MOS forecast. In fact, for HDD contracts the increase in accuracy performance is 9% and 16% with respect to the MOS forecasts and historical forecasts. This is quite an interesting result, especially in a market where issues of inside information and other frictions are minimal.

[17] In the supplemental appendix, we look at the efficiency over longer horizons.

[INSERT TABLE 6 ABOUT HERE]

Table 7 shows the performance of these models by city and by contract type. When we look at the data this way, we find that generally the MOS forecasts and the market forecasts do better than historical forecasts. In some cases, the accuracy is substantially better, like in Dallas, Texas where the market improves over the MOS forecasts by 118%. However, for HDD in 3 of 18 cities, the market's price is less informative about future weather than the government forecasts. For CDD, this is true in 4 of 18 cases.

Table 8 presents the performance results aggregated by month of the year. A similar pattern emerges, the market does substantially better than the MOS and historical forecasts. In some cases, the historical forecasts are better than MOS but in most cases, it's the reverse. However, in all but three cases (February-HDD, August and September-CDD), the market's forecasts are much more reliable.

[INSERT TABLE 7 ABOUT HERE]

[INSERT TABLE 8 ABOUT HERE]

A.2 Dynamic Models

The dynamic models are slightly more complex, since we have a prediction for every day of the month for the month's final settlement value of the contract. Thus, when aggregating the results across days and across cities, we might expect earlier days in the month to have larger forecast errors than later days in the month. Overall, these effects should average out.[18] Table 9 contains the results for the dynamic forecasts aggregated. The results are similar to the static case. That

[18] Another way to control for this would be to average the results by first dividing the MAE or RMSE by number of days left until the end of the month. This consequently makes all of the MAE and RMSE much smaller relative to the static results.

is, the market still seems to be the best predictor of eventual temperature and HDD for a month when information is updated daily. Tables 9 and 10 show more results of the dynamic case with different methods of data aggregation. In the dynamic case, the general results are the same. The MOS forecasts do better than the historical and the markets do better than both.[19]

[**INSERT TABLE 9 ABOUT HERE**]

[**INSERT TABLE 10 ABOUT HERE**]

Overall the results of the static and dynamic models are quite supportive of a relatively efficient market place in weather derivatives. One could criticize the above analysis for a variety of reasons mentioned in the previous section, but given our data limitations we find it at least consistent with an efficient market in weather derivatives.

B Overreaction

Before turning to the results, we restate what we should expect if there is overreaction or under-reaction in the weather derivatives market. For HDD contracts, we would expect that a story consistent with overreaction would be higher average returns as the quintiles increase. The logic is that, when there is a lower-that-expected temperature, the market overreacts by raising the HDD too much, thus a strategy of buying HDD and holding until expiration would lead to lower returns. The opposite result would occur for days with high positive temperature surprises. For CDD contracts, we would expect the opposite, since CDD pays when temperatures are high. Thus, on days with negative temperature surprises, the overreaction would coincide with CDD contracts trading lower than necessary and consequently making returns to buying CDD contracts higher

[19]The acute reader might notice that there are average errors by the market on the final day of trading. In fact, sometimes, the closing prices on the final day of trading are not equal to the settlement value of the contract. In speaking with the CME, they say that they quotes are legitimate, even though they make little sense. I have left them in for completeness, but they are puzzling.

than normal. Thus, we would expect that returns should decrease for CDD contracts as we move from lower to higher quintiles if there is overreaction in the markets.[20]

Figures 2 - 3 show the aggregated returns across days and cities for surprise measure 3.[21] In some sense, we believe this is the most reliable measure of surprise, since it's the temperature deviation from the market's implied expected temperature based on the previous day's market prices. For both HDD and CDD there is not strong evidence consistent with an overreaction hypothesis, except maybe for the CDD contracts.

[**INSERT FIGURES 2 AND 3 HERE**]

In order to examine the issue of overreaction without some of the potential biases of aggregation, we show the results of overreaction for each measure by *day of month*.[22] Thus, the overreaction returns are computed by looking at all quintile 1 surprises for each city in isolation on day 1, day 2, and so on and computing returns to the end of the month.

Tables 11 shows the averaged returns from each decile for surprise measure 3 using data from 2005-2008 and returns as measured from the next business day's closing prices.[23] For none of the surprise measures does there seem to be any consistent pattern emerging in terms of overreaction or underreaction to weather surprises.

[**INSERT TABLE 11 ABOUT HERE**]

Table 12 contains the statistical tests for differences in means of quintile 1 and quintile 5 returns. Almost all of the t-statistics and Mann-Whitney tests fail to reject the hypothesis that the returns

[20]The results would be the reverse for an underreaction story.

[21]The results for the other surprise measures are qualitatively the same and contained in the Supplemental Appendix.

[22]First, aggregate measures combine the overreaction returns for different days of the month which might blur the results. However, Quintile 1 could have a mixture of data points from Day 1 of the month or any other day. Thus, when we average across the quintile, we are mixing effects. That is, presumably a surprise on day 1 of the month might have a much larger impact than if the surprise occurs on day 30th of the month. Dividing by days left in the month might help to make the results less distorted.

[23]The same tables for surprise measures 1 and 2 are contained in the Supplemental Appendix.

on any given surprise day for quintile 1 and quintile 5 are the same. For surprise measure 3, the highest level of significance is on day 4 with a t-statistic of 3.51 and a z-statistic of 2.20.

[**INSERT TABLE 12 ABOUT HERE**]

All of the overreaction tests were also done when the investors were allowed to purchase contracts on the em next day of the surprise rather than the same business day (see the Tables in the Supplemental Appendix) using data from 2005-2008.[24] In addition to this, the overreaction tests were done with a longer sample period from 1999-2008 using t for computing the returns (see the Tables in the Supplemental Appendix). In all cases, the qualitative results are the same. There not appear to be overreaction or underreaction in these markets, except for occassional sporadic cases.

Table 13 we presents the overreaction results by city. While some cities seem to exhibit patterns of overreaction and others underreaction, there is no consistent pattern across cities.

[**INSERT TABLE 13 ABOUT HERE**]

C Inter-Market Behavior

The study of intermarket potential inefficiencies was examined for the base period 2005-2008 and for the whole period from 1999-2008. The results for the long, short, and long-short portfolios are contained in Table 14.

[**INSERT TABLE 14 ABOUT HERE**]

Unfortunately, there are very few observations for both periods. The return are in the direction of inefficiency for the whole sample period but the t-statistics for difference in returns is insignificant.

[24]We originally computed the returns using the next trading day so as to avoid trading on information that might have been not been known, however in our study it is very likely that by the close of trading the maximum and minimum temperatures of the day are already well known and that trading on that information by day's end is ok.

Thus, for this particular test for cross-city weather market inefficiency, again the weather markets look quite efficient.

VI Conclusion

The weather derivatives market is a relatively new market. It has traded on the CME since 1999. Recently, there has been quite a lot of skepticism in the efficiency of markets. In fact, in the equity markets, many anomalies have been documented that bring into question the efficiency of markets. The weather derivatives market stands apart from many markets in that the symmetry of information between agents is very high. That is, there is no possibility for inside information, since the weather is truly exogenous to our system. The weather derivatives market is also a market where contracts only live for a relatively short-period of time, unlike equity markets. In this paper, the efficiency of the weather derivatives market was examined in a variety of ways. Overall, despite its lack of sufficient depth, one fails to reject the hypothesis that this market trades very efficiently. We find this when comparing the prediction implied in weather futures prices versus historical models of the temperature and government model forecasts of the weather.

One criticism of this comparison of models is that the alternatives are too easy to defeat. For the MOS forecasts, the criticism is that the forecasts are only for seven days and historical measures are combined with them. This could indicate that the market is using a longer-term forecast from a private agency and hence is more accurate because of this. In order to address this criticism, we mentioned studies showing that weather forecasting is very poor at horizons longer than a week and secondly, we examine the efficiency of this market from another perspective, by studying the over-or-under reaction of prices to surprise weather information. From this perspective, weather futures prices seem to be consistent with efficiency in that there is no consistent overreaction or underreaction to temperature surprises. We also find this when examining weather derivative prices across cities to determine whether there might be some inter-market inefficiencies.

The weather market's efficiency might be due to many factors, including the low volatility of weather surprises, or the symmetry of information (i.e. lack of inside information), or the short-term nature of this market. One can imagine that a market which lacks the potential for informed traders will be more efficient, since there might not be a guessing game by non-informed traders on movements in prices. Given the perhaps puzzling values for the forward premia in this market, further research on the forward premia might be interesting. Also, further research into the market's efficiency and the role of information would be interesting. Also, it might be interesting to study whether other derivative contracts can span the set of weather derivative contracts or they truly are an invaluable hedging instrument for companies wanting to hedge energy demand.

VII Appendix

A Tables

Table 1: CME Weather Contract Specifications

Trading Hours	*Futures products* trade electronically only on CME Globex from Monday through Friday from 3:45 PM to 3:15PM (central time) of the following day, and on Sundays from 5:30 PM to 3:15 PM. On the last trading day they trade until 9 AM. *Option products* trade only Monday through Friday from 8:15 AM to 3:15 PM on the CME trading floor.
Contract Size	$20 times the monthly index. The monthly index is provided by the Earth Satellite Corporation.
Minimum Tick Fluctuation	One degree day index point.
Settlement	Cash settled. All contracts that remain open at the termination of trading of a particular contract shall be settled using the respective CME Degree Days Index for that city and that contract season, using the methodology in effect on that date, on the first Exchange business day that is at least two calendar days after the derivatives contract month.
Maximum Order Size	10,000 contracts net long or short in all contract months combined.
Trading Venue	Only options can be traded via open outcry; the futures products are traded exclusively on the CME Globex electronic trading platform.

Note: This table was copied from the CME website. The contract size changed from $100 per contract to $20 per contract on March 8 and April 12, 2004.

Table 2: Example of Weather and Weather HDD Futures for Atlanta in March, 2006

Date	Tmax	Tmin	Tavg	HDD	Cumulative HDD	HDD March Futures Closing Prices
3/1/2006	71	50	60.5	4.5	4.5	290
3/2/2006	70	55	62.5	2.5	7	275
3/3/2006	56	41	48.5	16.5	23.5	235
3/4/2006	61	32	46.5	18.5	42	NA
3/5/2006	66	38	52	13	55	NA
3/6/2006	69	48	58.5	6.5	61.5	210
3/7/2006	60	38	49	16	77.5	245
3/8/2006	66	41	53.5	11.5	89	240
3/9/2006	69	44	56.5	8.5	97.5	257
3/10/2006	76	53	64.5	0.5	98	250
3/11/2006	78	61	69.5	0	98	NA
3/12/2006	80	61	70.5	0	98	NA
3/13/2006	75	61	68	0	98	300
3/14/2006	69	44	56.5	8.5	106.5	325
3/15/2006	61	37	49	16	122.5	320
3/16/2006	65	41	53	12	134.5	320
3/17/2006	69	45	57	8	142.5	315
3/18/2006	59	39	49	16	158.5	NA
3/19/2006	53	44	48.5	16.5	175	NA
3/20/2006	55	39	47	18	193	335
3/21/2006	68	40	54	11	204	345
3/22/2006	54	35	44.5	20.5	224.5	342
3/23/2006	52	38	45	20	244.5	343
3/24/2006	51	38	44.5	20.5	265	337
3/25/2006	51	33	42	23	288	NA
3/26/2006	54	30	42	23	311	NA
3/27/2006	61	34	47.5	17.5	328.5	350
3/28/2006	56	48	52	13	341.5	350
3/29/2006	73	48	60.5	4.5	346	350
3/30/2006	74	50	62	3	349	351
3/31/2006	77	55	66	0	349	350
4/3/2006	NA	NA	NA	NA	NA	349

Note: The official weather data is for Atlanta Hartsfield International Airport 13874 as computed by the National Climatic Data Center.

Table 3: Summary Statistics of Weather Futures Contracts

City and Contract Type	First Contract Traded	Average Daily Volume	Number of Trading Days w/ Volume
Atlanta HDD	10:1999	52.45	934
Atlanta CDD	7:1999	42.43	659
Baltimore HDD	11:2005	24.84	73
Baltimore CDD	6:2006	18.27	29
Boston HDD	3:2003	470.51	499
Boston CDD	6:2003	49.95	145
Chicago HDD	10:1999	49.06	586
Chicago CDD	6:2002	60.69	335
Cincinnati HDD	10:1999	34.74	427
Cincinnati CDD	5:2002	39.70	307
Dallas HDD	10:2002	25.78	283
Dallas CDD	6:2000	39.76	290
Des Moines HDD	3:1999	78.48	866
Des Moines CDD	7:2000	31.90	237
Detroit HDD	4:2008	0.17	2
Detroit CDD	12:2012	0.00	0
Houston HDD	3:2003	68.34	180
Houston CDD	9:2003	400.91	1427
Kansas City HDD	10:2003	62.71	258
Kansas City CDD	10:2003	40.61	169
Las Vegas HDD	11:2002	15.94	144
Las Vegas CDD	6:2000	31.95	173
Minneapolis HDD	10:2003	44.91	307
Minneapolis CDD	10:2003	44.45	414
New York HDD	11:1999	66.82	660
New York CDD	5:2002	73.06	494
Philadelphia HDD	1:2002	26.05	739
Philadelphia CDD	4:1999	49.09	702
Portland HDD	1:2002	17.66	110
Portland CDD	4:1999	1370.59	3113
Sacramento HDD	1:2003	16.54	659
Sacramento CDD	6:2003	45.01	248
Salt Lake City HDD	10:2006	9.42	6
Salt Lake City CDD	7:2006	6.09	9
Tucson HDD	11:2002	15.22	114
Tucson CDD	5:2000	26.79	148

Note: First Contract Traded is the month and year in which the first contract for a particular city and type of contract traded on the CME. The average daily volume is computed as the average daily volume traded for that particular contract and city conditional on volume existing for that contract. The number of trading days with volume is the number of observations for a particular contract month and city that there were contracts traded. The data for the entire sample period (1999-2008) was used for these summary statistics.

Table 4: Realized Percentage Forward Premia in the Weather Derivative Futures

	HDD						CDD						
City	Mean	t-stat	S.D.	Max	Min	nobs	Mean	t-stat	S.D.	Max	Min	nobs	σ
at	-0.71	-0.81	14.40	71.74	-66.19	275.00	-0.62	-0.45	18.19	45.71	-76.67	172.00	6.77
ba	2.49	2.81	6.26	15.00	-10.92	50.00	8.91	1.85	16.66	41.60	-9.67	12.00	7.63
bo	0.21	0.34	7.50	25.36	-21.64	149.00	0.12	0.05	22.26	76.67	-153.85	93.00	7.57
ch	0.89	2.17	7.89	21.31	-43.20	371.00	-2.10	-0.88	35.12	91.67	-180.00	217.00	8.81
ck	0.53	0.87	10.64	25.28	-49.02	304.00	1.51	0.87	23.48	65.88	-87.86	184.00	8.53
da	4.60	3.86	16.86	62.94	-55.95	200.00	0.61	0.70	11.55	35.35	-61.87	171.00	7.30
dm	0.90	1.31	10.07	23.90	-45.66	217.00	3.01	2.15	17.20	51.89	-63.75	151.00	9.24
de	10.75	1.04	14.68	21.13	0.38	2.00	.	.	.	.	.	0.00	8.19
ho	0.27	0.12	19.26	67.50	-45.36	73.00	-0.71	-0.89	6.24	8.71	-20.90	60.00	6.41
kc	1.01	1.22	11.16	33.25	-35.38	181.00	2.05	1.11	18.34	61.43	-53.81	99.00	9.01
lv	0.50	0.25	17.95	66.51	-39.32	84.00	-0.21	-0.31	6.91	34.00	-28.44	100.00	6.18
mn	1.46	3.13	6.83	25.11	-24.12	214.00	6.75	3.27	21.67	78.00	-68.75	110.00	9.40
ny	2.93	6.02	9.81	48.65	-27.95	406.00	-8.87	-5.42	25.69	85.96	-159.00	246.00	7.34
ph	-1.81	-2.01	10.73	34.05	-67.46	142.00	-1.27	-0.65	21.34	86.67	-85.83	119.00	7.39
po	-0.96	-0.91	9.01	20.44	-32.35	73.00	0.94	0.21	33.84	55.66	-127.27	57.00	5.16
sa	2.14	1.32	9.59	38.18	-16.17	35.00	-4.10	-1.29	33.75	47.12	-101.20	113.00	5.20
sl	-6.12	-1.07	12.79	6.02	-25.34	5.00	5.33	.	.	5.33	5.33	1.00	7.72
tu	5.65	1.87	24.35	85.00	-23.58	65.00	1.63	1.87	8.08	21.36	-39.23	86.00	5.84

Note: The realized forward premia are computed as: $\frac{F_t - S_{t+1}}{F_t}$ and are expressed in percentage terms. The t-stats are for the mean return being different than 0. All term premia are from contract purchase until end of the month. The premia are computed over the full span of data from 1999 - 2008. The return from buying the contract is the negative of the forward premium percentage in the table.

Table 5: Realized Percentage Forward Premia in the Weather Derivative Futures by Day of Month

Day	Mean	Median	SD	t-stat	Max	Min	nobs	Mean	Median	SD	t-stat	Max	Min	nobs
1.00	4.93	3.80	16.08	0.64	25.58	-14.92	92.00	17.24	10.93	39.78	0.47	80.94	-25.31	96.00
2.00	2.50	1.80	14.31	0.58	20.87	-14.00	111.00	22.32	17.72	42.66	0.88	76.52	-21.04	95.00
3.00	-4.24	-3.84	15.25	-0.71	13.31	-27.04	132.00	3.12	0.27	32.96	0.36	36.15	-28.00	58.00
4.00	-1.22	-0.94	12.26	-0.26	15.52	-19.20	122.00	7.58	3.00	35.89	1.14	47.86	-27.05	60.00
5.00	3.18	2.65	13.45	0.34	22.48	-15.24	139.00	14.48	10.33	35.03	1.23	60.04	-20.88	85.00
6.00	4.91	6.20	13.47	4.11	21.06	-16.27	126.00	3.64	-1.80	30.96	0.49	53.00	-27.51	93.00
7.00	-0.79	-1.07	10.12	-0.60	13.83	-14.52	140.00	18.83	15.40	34.78	0.56	74.36	-17.96	116.00
8.00	-0.09	-1.91	10.19	-0.02	16.69	-10.30	113.00	10.32	7.83	41.38	0.49	63.49	-47.57	103.00
9.00	1.80	1.49	13.00	0.71	17.97	-13.18	119.00	21.94	16.29	38.26	0.79	68.10	-20.76	102.00
10.00	1.27	1.16	11.36	-0.75	16.06	-13.27	124.00	18.22	16.58	35.94	0.27	60.17	-20.28	84.00
11.00	1.91	2.16	10.80	0.57	16.87	-12.34	112.00	8.11	2.95	28.43	0.31	47.69	-26.21	71.00
12.00	1.13	1.17	12.08	0.10	17.03	-14.75	119.00	14.36	6.85	35.07	0.41	63.89	-23.99	92.00
13.00	1.52	0.65	11.52	0.47	15.46	-10.67	118.00	11.85	3.03	33.80	0.60	69.39	-22.14	98.00
14.00	0.13	-0.95	10.81	0.12	13.63	-14.19	130.00	10.41	2.19	30.08	0.77	61.66	-19.59	85.00
15.00	0.28	-1.82	8.99	0.79	14.52	-8.45	117.00	1.69	1.91	18.64	0.32	27.82	-27.19	95.00
16.00	0.36	0.36	7.00	0.39	9.42	-7.36	108.00	8.19	5.50	25.30	0.74	48.07	-23.21	96.00
17.00	0.29	-0.14	11.07	0.56	14.31	-12.47	110.00	-2.01	-0.70	24.16	-0.09	32.39	-33.60	87.00
18.00	2.93	1.31	11.84	0.88	21.90	-8.17	112.00	1.69	2.42	23.72	0.36	28.85	-28.32	73.00
19.00	0.54	1.48	5.70	0.41	5.53	-6.36	82.00	7.54	2.53	21.65	1.02	44.15	-10.48	78.00
20.00	-0.90	-0.84	5.90	-0.30	4.84	-8.61	90.00	10.60	2.14	24.88	1.26	49.78	-7.13	80.00
21.00	1.81	2.15	4.82	-0.16	6.05	-3.32	81.00	5.36	1.86	17.90	-0.42	32.47	-5.88	64.00
22.00	4.16	1.83	9.60	0.93	17.10	-4.50	105.00	11.50	2.07	30.87	0.71	57.04	-8.12	72.00
23.00	5.84	4.59	8.92	0.61	14.53	-0.33	71.00	8.02	3.72	27.67	0.30	39.68	-13.61	55.00
24.00	2.04	1.72	3.84	1.48	6.11	-1.29	55.00	7.03	3.62	18.83	0.08	29.82	-4.81	58.00
25.00	1.31	0.80	3.70	0.03	4.79	-1.20	46.00	15.36	11.68	23.70	0.28	39.28	2.30	58.00
26.00	-0.02	-0.06	3.05	0.13	2.42	-2.31	49.00	9.21	6.28	23.87	0.60	28.33	-2.56	38.00
27.00	0.53	0.63	2.18	0.47	1.85	-0.96	37.00	22.53	14.77	32.61	.	50.98	5.99	38.00
28.00	-1.10	-1.11	2.91	-0.40	0.82	-2.98	34.00	3.21	3.24	12.41	0.91	8.59	-2.07	29.00
29.00	2.15	1.73	5.61	0.76	6.06	-0.74	31.00	13.81	11.29	26.59	0.83	23.07	7.01	23.00
30.00	0.90	0.89	2.36	-0.65	1.70	0.12	21.00	0.70	0.70	0.96	0.65	0.93	0.48	12.00
31.00	.	.	.	.	.	.	0.00	.	.	.	.	.	.	0.00

Note: The realized forward premia are computed as: $\frac{F_t - S_{t+1}}{F_t}$ and are expressed in percentage terms. The premia are computed for each day of each month for each city and then averaged across cities. All variables are averaged across cities, except for nobs which represent the total number of observations across all cities. The premia are computed over the full span of data from 1999 - 2008

Table 6: Static Models' Forecast Performance

		Historical		MOS Forecasts		Market	
Contract Type	Statistics	RMSE	MAE	RMSE	MAE	RMSE	MAE
HDD	Mean	96.71	79.18	92.80	73.09	77.97	60.66
	Max	165.05	128.81	162.64	119.49	141.10	113.18
	Min	29.19	18.23	24.48	15.62	16.27	15.33
	nobs	141.00	141.00	141.00	141.00	141.00	141.00
CDD	Mean	64.19	56.76	49.39	40.90	39.98	34.30
	Max	116.84	105.81	95.51	80.07	91.13	66.20
	Min	16.38	16.38	12.24	12.24	19.50	19.25
	nobs	75.00	75.00	75.00	75.00	75.00	75.00

Note: Since the NWS MOS forecasts, the beginning date for these statistics is set to September 01, 2005 and the ending date is June 30, 2008. This gives symmetry between measures. The results for the market statistics are done conditioning on volume. That is, the market prices are only used even if there existed trading volume on that particular day. RMSE is for root mean squared error and MAE is for mean absolute error. The results are computed only over observations which exist for all three forecast methodologies, which is limited by the market statistics availability. The sample period is constrained to days in which data exist for all three measures.

Table 7: Static Models' Forecast Performance By City

	Historical			MOS Forecasts			Market		
City	RMSE	MAE	MAE Dif	RMSE	MAE	MAE Dif	RMSE	MAE	nobs
				HDD					
at	107.83	75.70	26.59	100.82	69.41	16.07	88.09	59.80	10.00
ba	105.29	85.08	33.39	92.34	75.81	18.85	81.73	63.79	7.00
bo	118.32	87.33	8.94	114.93	80.81	0.80	108.69	80.17	9.00
ch	165.05	128.81	28.91	158.56	119.49	19.57	135.37	99.93	14.00
ck	149.32	120.84	35.17	141.15	108.07	20.89	123.92	89.40	10.00
da	87.17	75.12	148.05	70.36	65.59	116.58	32.61	30.29	7.00
dm	144.92	117.96	9.39	162.64	116.82	8.34	138.42	107.83	12.00
de	.	.	.	.	.	.	.	.	0.00
ho	60.43	59.08	285.32	42.92	38.31	149.82	16.27	15.33	3.00
kc	130.81	107.94	-4.63	151.74	105.50	-6.78	141.10	113.18	11.00
lv	79.63	67.44	23.86	77.71	68.85	26.47	62.74	54.44	9.00
mn	127.40	106.21	56.95	136.15	114.37	69.01	97.24	67.67	9.00
ny	137.50	110.76	52.49	135.42	105.87	45.76	99.81	72.63	15.00
ph	130.45	109.81	53.28	103.13	86.12	20.21	92.53	71.64	7.00
po	37.40	32.08	-39.56	47.77	46.31	-12.75	59.96	53.08	6.00
sa	29.19	18.23	-34.51	24.48	15.62	-43.87	37.48	27.83	6.00
sl	49.56	49.56	37.67	48.46	48.46	34.61	36.00	36.00	1.00
tu	80.59	73.27	49.83	61.74	50.15	2.55	51.52	48.90	5.00
				CDD					
at	111.38	84.56	27.73	95.51	70.53	6.55	91.13	66.20	5.00
ba	26.39	26.39	-9.02	12.24	12.24	-57.79	29.00	29.00	1.00
bo	16.38	16.38	-16.01	27.79	27.79	42.50	19.50	19.50	1.00
ch	56.89	50.28	120.99	45.95	38.52	69.30	26.18	22.75	6.00
ck	116.84	94.22	51.23	91.65	70.32	12.87	81.90	62.30	5.00
da	88.33	75.51	48.24	74.76	64.43	26.49	58.00	50.94	8.00
dm	56.99	45.92	130.76	28.82	22.65	13.83	24.83	19.90	5.00
de	.	.	.	.	.	.	.	.	0.00
ho	63.29	54.65	183.89	50.24	38.54	100.19	23.86	19.25	4.00
kc	67.15	58.24	45.00	47.72	46.94	16.87	40.63	40.17	3.00
lv	110.85	105.81	239.51	49.64	45.06	44.58	34.26	31.17	3.00
mn	62.08	55.36	125.95	55.37	37.50	53.07	30.84	24.50	5.00
ny	63.50	59.42	73.66	52.43	39.80	16.31	36.60	34.21	7.00
ph	53.36	51.29	107.65	40.08	28.95	17.22	25.20	24.70	5.00
po	22.63	22.02	-11.92	29.94	23.66	-5.35	26.19	25.00	3.00
sa	58.53	46.72	-27.40	55.77	46.80	-27.28	72.91	64.36	7.00
sl	104.05	104.05	181.22	80.07	80.07	116.41	37.00	37.00	1.00
tu	76.88	70.78	52.22	50.99	42.46	-8.69	61.67	46.50	6.00

Note: Since the NWS MOS forecasts, the beginning date for these statistics is set to September 01, 2005 and the ending date is June 30, 2008. This gives symmetry between measures. The results for the market statistics are done without conditioning on volume. That is, the market prices are used even if there might have not been trading on that particular day. Later we investigate these results conditioned on volume. The sample period is constrained to days in which data exist for all three measures.

Table 8: Static Models' Forecast Performance By Month

	Historical			MOS Forecasts			Market		
Month	RMSE	MAE	MAE Dif	RMSE	MAE	MAE Dif	RMSE	MAE	nobs
				HDD					
1.00	151.72	147.47	15.50	133.65	128.06	0.30	137.31	127.68	11.00
2.00	110.95	99.60	3.71	75.29	67.38	-29.84	104.75	96.04	28.00
3.00	81.87	74.51	42.50	89.91	75.40	44.20	62.68	52.29	34.00
4.00	100.98	100.98	158.08	79.84	79.84	104.06	39.13	39.13	4.00
5.00	.	.	.	.	.	.	.	.	0.00
6.00	.	.	.	.	.	.	.	.	0.00
7.00	.	.	.	.	.	.	.	.	0.00
8.00	.	.	.	.	.	.	.	.	0.00
9.00	.	.	.	.	.	.	.	.	0.00
10.00	109.12	109.12	19.26	108.81	108.81	18.92	91.50	91.50	4.00
11.00	67.55	62.97	51.44	64.15	56.56	36.02	44.27	41.58	27.00
12.00	104.32	91.34	16.00	124.19	101.43	28.81	95.88	78.74	33.00
				CDD					
1.00	.	.	.	.	.	.	.	.	0.00
2.00	.	.	.	.	.	.	.	.	0.00
3.00	.	.	.	.	.	.	.	.	0.00
4.00	113.52	113.52	489.71	88.50	88.50	359.76	19.25	19.25	2.00
5.00	67.62	67.62	65.34	76.12	76.12	86.10	40.90	40.90	5.00
6.00	54.26	52.35	81.39	40.93	39.31	36.20	30.52	28.86	21.00
7.00	62.48	59.19	68.88	46.86	45.45	29.68	35.49	35.05	13.00
8.00	83.14	82.62	38.87	54.57	54.25	-8.82	59.68	59.50	15.00
9.00	57.31	55.35	50.68	38.79	36.01	-1.97	39.35	36.73	19.00
10.00	.	.	.	.	.	.	.	.	0.00
11.00	.	.	.	.	.	.	.	.	0.00
12.00	.	.	.	.	.	.	.	.	0.00

Note: Since the NWS MOS forecasts, the beginning date for these statistics is set to September 01, 2005 and the ending date is June 30, 2008. This gives symmetry between measures. The results for the market statistics are done without conditioning on volume. That is, the market prices are used even if there might have not been trading on that particular day. The sample period is constrained to days in which data exist for all three measures.

Table 9: Dynamic Models' Forecast Performance

		Historical		MOS Forecasts		Market	
Contract Type	Statistics	RMSE	MAE	RMSE	MAE	RMSE	MAE
HDD	Mean	88.32	65.29	72.65	51.28	57.19	42.90
	Max	149.12	100.05	121.93	80.83	91.52	72.00
	Min	42.67	29.56	39.31	25.59	28.46	21.69
	nobs	1830.00	1830.00	1818.00	1818.00	1830.00	1830.00
CDD	Mean	40.24	31.18	33.45	26.40	30.45	23.68
	Max	76.85	61.04	54.88	44.49	50.12	37.65
	Min	6.88	6.88	16.54	16.16	8.00	8.00
	nobs	938.00	938.00	934.00	934.00	938.00	938.00

Note: Since the NWS MOS forecasts, the beginning date for these statistics is set to September 01, 2005 and the ending date is June 30, 2008. The sample period is constrained to days in which data exist for all three measures. The results for the market statistics are done conditioning on volume. That is, the market prices are only used even if there existed trading volume on that particular day. RMSE is for root mean squared error and MAE is for mean absolute error. The results are computed only over observations which exist for all three forecast methodologies, which is limited by the market statistics availability. The dynamic forecasts are computed for every day of the month and compared against final realized future values for that particular month.

Table 10: Dynamic Models Forecast Performance By Day of Month

	Historical			MOS Forecasts			Market		
Day	RMSE	MAE	MAE Dif	RMSE	MAE	MAE Dif	RMSE	MAE	nobs
HDD									
1.00	88.87	74.11	-0.48	93.65	75.59	1.51	88.72	74.47	56.00
2.00	96.77	84.56	15.02	86.33	72.28	-1.69	85.92	73.52	57.00
3.00	106.52	84.22	27.67	84.18	64.77	-1.82	77.83	65.97	78.00
4.00	128.00	96.18	35.83	109.76	79.10	11.70	86.31	70.81	68.00
5.00	123.71	98.98	53.76	110.82	88.00	36.70	77.83	64.38	87.00
6.00	127.34	102.02	56.21	110.89	91.19	39.61	76.11	65.31	86.00
7.00	91.42	75.36	47.51	76.22	60.91	19.21	60.99	51.09	84.00
8.00	93.08	73.16	57.66	76.69	57.59	24.11	55.76	46.40	78.00
9.00	99.69	77.05	66.76	76.56	58.54	26.70	57.12	46.20	82.00
10.00	87.33	68.01	74.43	63.87	50.45	29.38	50.33	38.99	70.00
11.00	101.72	81.96	110.59	75.08	60.88	56.42	48.44	38.92	65.00
12.00	103.82	83.80	75.00	77.88	64.27	34.23	54.98	47.88	73.00
13.00	93.74	72.79	61.58	70.36	59.13	31.25	52.04	45.05	81.00
14.00	63.49	51.03	35.90	50.04	41.98	11.81	44.60	37.55	95.00
15.00	66.10	51.07	47.14	58.35	46.57	34.19	40.17	34.71	85.00
16.00	55.24	43.51	56.17	52.00	39.78	42.81	33.83	27.86	77.00
17.00	50.28	41.11	45.45	40.18	33.59	18.83	32.41	28.27	75.00
18.00	62.77	51.13	91.08	41.01	33.13	23.81	30.69	26.76	68.00
19.00	81.83	70.17	133.40	55.09	45.50	51.36	34.16	30.06	57.00
20.00	82.37	72.94	190.45	54.82	47.99	91.10	28.93	25.11	62.00
21.00	60.20	48.46	106.47	36.26	29.68	26.45	26.67	23.47	51.00
22.00	57.69	46.91	135.10	35.43	29.64	48.54	22.98	19.95	72.00
23.00	49.95	40.00	99.11	28.45	24.54	22.18	22.79	20.09	47.00
24.00	37.37	31.67	104.81	11.37	9.76	-36.92	17.00	15.47	31.00
25.00	18.89	16.54	42.18	10.00	8.47	-27.18	12.76	11.63	30.00
26.00	45.15	38.65	385.67	10.13	9.42	18.43	8.87	7.96	30.00
27.00	51.58	46.32	419.62	11.43	10.02	12.44	10.08	8.91	24.00
28.00	21.82	21.16	260.85	6.86	6.04	2.99	6.13	5.86	18.00
29.00	24.52	22.37	306.00	6.94	6.14	11.47	6.08	5.51	25.00
30.00	8.81	7.85	177.18	2.39	2.08	-26.47	2.86	2.83	11.00
31.00	0.00	0.00	-100.00	0.00	0.00	-100.00	2.33	2.30	7.00
CDD									
1.00	93.77	85.98	104.73	56.17	48.94	16.53	48.51	42.00	43.00
2.00	52.52	48.44	31.69	41.87	37.70	2.49	40.58	36.78	28.00
3.00	55.83	50.80	35.73	40.76	38.18	2.02	39.32	37.43	30.00
4.00	47.98	43.66	1.14	41.78	37.01	-14.27	45.83	43.17	27.00
5.00	40.77	37.03	29.20	32.93	29.88	4.26	32.63	28.66	41.00
6.00	45.42	40.77	14.33	41.64	37.08	3.98	38.76	35.66	44.00
7.00	52.07	45.40	17.76	43.15	38.19	-0.94	43.68	38.55	44.00
8.00	48.93	43.43	15.35	39.23	35.55	-5.57	40.71	37.65	40.00
9.00	42.43	38.07	-2.84	40.36	36.54	-6.74	42.75	39.18	31.00
10.00	48.93	42.60	5.04	42.66	38.03	-6.21	44.60	40.55	40.00
11.00	37.79	34.89	14.10	34.33	30.94	1.20	34.05	30.57	29.00
12.00	43.05	39.90	43.75	36.30	33.46	20.58	31.03	27.75	38.00
13.00	34.79	30.77	23.37	33.06	30.32	21.55	28.13	24.95	37.00
14.00	38.41	34.31	53.74	26.19	23.46	5.12	25.85	22.32	36.00
15.00	36.96	32.09	16.84	34.05	30.36	10.56	30.51	27.46	41.00
16.00	36.18	31.21	8.38	37.68	33.68	16.95	31.15	28.80	40.00
17.00	40.18	34.34	25.41	35.74	31.96	16.70	30.85	27.38	44.00
18.00	32.30	26.90	9.47	31.69	27.11	10.32	28.72	24.57	36.00
19.00	29.30	28.05	35.14	32.46	30.80	48.39	22.04	20.76	27.00
20.00	29.68	26.46	72.45	22.51	20.04	30.60	17.67	15.35	33.00
21.00	22.27	20.24	29.28	21.97	19.78	26.39	17.05	15.65	33.00
22.00	28.37	25.69	123.67	25.99	24.33	111.84	13.47	11.49	30.00
23.00	30.85	29.25	135.14	26.99	26.17	110.41	12.93	12.44	23.00
24.00	27.45	25.70	102.44	14.92	13.84	9.05	13.35	12.69	25.00
25.00	22.76	20.70	100.65	11.03	9.92	-3.79	12.13	10.31	31.00
26.00	18.31	17.76	167.03	7.83	7.62	14.54	6.88	6.65	18.00
27.00	13.21	12.37	96.39	8.11	7.61	20.84	6.81	6.30	13.00
28.00	15.16	14.36	254.94	6.01	5.70	41.01	4.21	4.05	19.00
29.00	12.94	12.64	519.11	2.54	2.50	22.45	2.04	2.04	7.00
30.00	11.95	11.95	233.56	3.17	3.17	-11.63	3.58	3.58	6.00
31.00	0.00	0.00	-100.00	0.00	0.00	-100.00	2.01	1.67	4.00

Note: Since the NWS MOS forecasts, the beginning date for these statistics is set to September 01, 2005 and the ending date is June 30, 2008. The sample period is constrained to days in which data exist for all three measures. The results for the market statistics are done conditioning on volume. That is, the market prices are only used even if there existed trading volume on that particular day. RMSE is for root mean squared error and MAE is for mean absolute error. The results are computed only over observations which exist for all three forecast methodologies, which is limited by the market statistics availability. The dynamic forecasts are computed for every day of the month and compared against final realized future values for that particular month. Thus, day 1 represents the forecasts of the dynamic model after one day of the month for the entire month's realized HDD or CDD. Day 2 represents the forecasts of the dynamic model after two days of realized values for the month and so on and so forth. The performance for each city and each contract is aggregated and shown for the day of month in which it occurs.

Table 11: Overreaction Returns in HDD and CDD Contracts by Day of Month using Surprise Measure 3

	HDD					CDD				
	Quintile (Lowest to Highest)					Quintile (Lowest to Highest)				
Day	1	2	3	4	5	1	2	3	4	5
1	-2.71	2.05	-1.56	.	.	-16.92	.	37.80	.	.
2	-0.15	11.68	4.65	8.66	-4.78	17.05	.	-11.66	1.24	-4.45
3	6.87	6.62	8.41	31.70	4.33	5.16	43.49	-7.48	0.45	-3.67
4	10.79	7.54	-3.86	-6.25	-10.46	-7.47	112.12	4.66	-9.22	-8.25
5	-4.92	-3.35	-10.68	-5.44	-3.57	-6.47	.	6.41	22.39	.
6	-3.99	-9.25	-2.75	-3.29	6.24	-14.82	0.37	-14.47	-11.11	18.84
7	-3.91	-10.66	-2.70	4.84	4.66	4.88	-5.06	-20.99	10.91	27.46
8	-11.95	-1.52	1.86	4.12	4.00	0.57	7.09	2.00	6.19	.
9	0.32	-8.77	-2.87	0.67	-3.91	53.61	0.99	-21.23	8.66	-7.16
10	4.46	-7.75	1.20	-10.61	-3.26	5.11	19.39	-9.92	1.21	7.57
11	-3.97	5.48	-4.97	-0.72	-2.17	49.11	-2.35	-3.89	-14.45	8.89
12	-1.63	3.62	-7.17	-14.70	-2.87	27.09	12.35	-1.80	6.69	-19.10
13	-2.17	-11.68	2.56	-7.77	-3.53	-7.33	0.67	7.76	2.60	-2.13
14	-0.59	-5.42	-10.41	2.86	-0.39	-11.05	-4.20	3.10	3.96	.
15	-1.20	-1.37	0.52	-5.14	-2.35	-6.25	-12.83	33.44	-5.99	-2.31
16	0.69	-0.64	-6.65	-3.90	1.26	31.08	12.22	-4.40	-6.56	14.73
17	-2.26	-9.66	-4.24	0.32	-1.26	12.82	3.05	-6.71	16.91	13.72
18	0.52	-9.90	-2.69	-3.44	-1.17	25.55	-0.47	-4.45	9.82	5.91
19	0.47	-1.64	3.48	-2.36	-1.02	8.98	-2.47	-3.57	-3.62	.
20	-2.20	-2.45	0.50	-3.05	-5.23	1.32	-8.34	-16.95	-2.61	.
21	-1.22	-4.02	-3.35	0.68	4.28	2.27	-6.81	1.30	4.77	.
22	-5.56	4.59	-2.99	-2.01	-1.22	-2.41	-4.08	-2.66	-5.25	-9.04
23	-9.44	-28.36	1.72	-3.75	-10.25	13.99	7.69	-2.83	0.89	.
24	8.97	-0.05	-4.42	-1.56	-6.45	2.72	14.84	3.38	18.84	-3.67
25	-0.64	-7.87	-0.96	-1.29	-2.88	0.61	5.65	-2.72	-4.99	.
26	-0.54	.	-0.39	-4.62	1.63	1.27	9.57	-3.76	.	.
27	-1.87	0.00	0.37	0.21	.	0.79	0.90	-1.20	.	.
28	-0.92	-0.84	-0.16	.	-0.58	0.85	.	0.41	.	.
29	-0.12	0.06	.	0.70	.	-0.86	.	0.46	-0.30	.
30	-0.71	.	-1.34	-14.29	.	.	.	.	.	.
31	.	.	.	.	.	.	.	.	.	.

Note: Overreaction is measured over each quintile. Thus, for any given day for any given contract and any given city, the surprises for that particular day are ordered from lowest to highest. Thus, a low surprise measure means that the temperature that day was much lower than the expected value. The values in the table are the returns from purchasing a futures contract at the close of that day and holding until expiration at the end of the month. These returns are computed for all surprise measures and then the results are aggregated and averaged by the quintile. Signs of overreaction would be increasing returns from lowest to highest quintile for HDD contracts and decreasing returns from lowest to highest quintile for CDD contracts.

Table 12: Overreaction Test Statistics for Difference in Means of Quintile 5 and Quintile 1 in HDD and CDD Contracts by Day of Month using Surprise Measure 3

Day	HDD								CDD							
	Sample Size		t-statistics				Mann-Whitney Tests		Sample Size		t-statistics				Mann-Whitney Tests	
	N_1	N_2	t-stat	p_u	p_l	p	U	z-statistics	N_1	N_2	t-stat	p_u	p_l	p	U	z-statistics
1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2	9.00	2.00	0.50	0.68	0.32	0.63	8.00	0.24	4.00	2.00	1.37	0.88	0.12	0.24	0.00	1.85
3	9.00	3.00	0.18	0.57	0.43	0.86	13.00	0.09	8.00	1.00	.	.	.	.	0.00	1.55
4	10.00	3.00	3.51	1.00	0.00	0.00	2.00	2.20	3.00	1.00	.	.	.	.	2.00	-0.45
5	11.00	4.00	-0.36	0.36	0.64	0.72	24.00	-0.26	.	.	.	.	.	.	.	.
6	19.00	1.00	.	.	.	.	15.00	-0.95	10.00	1.00	.	.	.	.	10.00	-1.58
7	14.00	6.00	-1.10	0.14	0.86	0.28	63.00	-1.73	11.00	1.00	.	.	.	.	10.00	-1.30
8	11.00	7.00	-3.35	0.00	1.00	0.00	71.00	-2.94	.	.	.	.	.	.	.	.
9	7.00	6.00	1.39	0.90	0.10	0.19	12.00	1.29	4.00	2.00	1.14	0.84	0.16	0.32	2.00	0.93
10	2.00	10.00	1.76	0.95	0.05	0.11	2.00	1.72	6.00	2.00	-0.13	0.45	0.55	0.90	7.00	-0.33
11	8.00	6.00	-0.50	0.31	0.69	0.63	26.00	-0.26	5.00	2.00	0.74	0.75	0.25	0.49	4.00	0.39
12	9.00	2.00	-0.02	0.49	0.51	0.99	10.00	-0.24	5.00	1.00	.	.	.	.	0.00	1.46
13	13.00	3.00	0.24	0.59	0.41	0.82	17.00	0.34	6.00	2.00	-0.44	0.34	0.66	0.68	7.00	-0.33
14	17.00	7.00	0.03	0.51	0.49	0.97	51.00	0.54	.	.	.	.	.	.	.	.
15	10.00	9.00	0.55	0.70	0.30	0.59	40.00	0.41	6.00	1.00	.	.	.	.	5.00	-1.00
16	11.00	9.00	-0.03	0.49	0.51	0.97	49.00	0.04	7.00	1.00	.	.	.	.	2.00	0.65
17	4.00	9.00	-0.32	0.38	0.62	0.76	22.00	-0.62	8.00	1.00	.	.	.	.	5.00	-0.39
18	10.00	1.00	.	.	.	.	5.00	0.00	7.00	2.00	0.50	0.68	0.32	0.63	8.00	-0.29
19	9.00	6.00	0.34	0.63	0.37	0.74	27.00	0.00	.	.	.	.	.	.	.	.
20	13.00	2.00	0.62	0.73	0.27	0.55	10.00	0.51	.	.	.	.	.	.	.	.
21	7.00	4.00	-1.92	0.04	0.96	0.09	22.00	-1.51	.	.	.	.	.	.	.	.
22	9.00	2.00	-2.06	0.03	0.97	0.07	15.00	-1.41	5.00	1.00	.	.	.	.	1.00	0.88
23	9.00	6.00	-0.02	0.49	0.51	0.98	15.00	1.41	.	.	.	.	.	.	.	.
24	1.00	2.00	.	.	.	.	0.00	1.22	6.00	1.00	.	.	.	.	0.00	1.50
25	6.00	1.00	.	.	.	.	1.00	1.00	.	.	.	.	.	.	.	.
26	2.00	1.00	.	.	.	.	2.00	-1.22	.	.	.	.	.	.	.	.
27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
28	4.00	1.00	.	.	.	.	1.00	0.71	.	.	.	.	.	.	.	.
29	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Note: Overreaction is measure over each quintile. Thus, for any given day for any given contract and any given city, the surprises for that particular day are ordered from lowest to highest. Thus, a low surprise measure means that the temperature that day was much lower than the expected value. The values in the table are the standard deviation of the returns in the previous table.

Table 13: Overreaction in HDD and CDD Contracts by City

	HDD					CDD				
	Quintile (Lowest to Highest)					Quintile (Lowest to Highest)				
Day	1	2	3	4	5	1	2	3	4	5
Atlanta	4.14	-4.21	4.81	0.04	3.19	0.58	0.75	0.31	0.03	0.33
Baltimore	-0.90	0.36	-2.54	1.81	-7.04	11.81	43.94	7.95	-0.31	1.07
Boston	2.98	0.01	-2.33	-3.38	-2.32	0.99	.	.	.	.
Chicago	-3.62	-1.55	3.96	3.24	-4.00	7.36	2.94	9.03	9.97	-19.55
Cincinnati	-2.78	-4.14	4.89	-2.15	-1.29	37.26	16.46	-7.18	-7.34	-12.15
Dallas	-12.45	-3.53	-13.38	-6.26	-0.27	13.90	5.50	-1.38	9.87	15.04
Des Moines	-9.47	2.84	4.73	-2.87	-2.17	-4.07	-2.00	27.78	-2.12	0.86
Detroit	.	.	.	.	.	0.21	-4.26	-10.52	-2.82	-0.00
Houston	-6.80	5.95	2.00	-28.65	-4.60	.	.	.	.	.
Kansas City	-10.22	-0.19	0.15	0.82	-2.02	-2.13	0.75	1.76	2.54	3.26
Las Vegas	-0.95	2.27	12.12	-21.26	-4.79	16.10	-24.58	-13.69	-6.79	-8.72
Minnesota	-1.91	3.88	2.06	-4.44	-4.62	4.99	5.00	0.71	-6.24	1.62
New York	-2.76	-4.57	-6.13	-5.93	-4.56	1.40	-5.41	-14.62	-7.94	-8.65
Philadelphia	6.64	2.35	8.70	0.13	-1.54	13.01	7.57	11.04	6.72	4.75
Portland	-10.06	0.47	.	5.52	-1.52	79.84	11.74	2.07	-11.26	.
Sacramento	-3.24	-1.47	-1.49	.	-3.25	65.63	-15.40	-15.91	.	-10.10
Salt Lake City	.	.	.	-6.02	.	-15.35	-15.72	11.24	4.17	14.69
Tuscon	-2.85	9.59	12.02	-29.81	-41.60	.	.	.	.	.
Aggregate	-3.39	0.50	1.97	-6.20	-5.15	14.79	1.47	0.17	-1.22	-1.77

Note: Overreaction is measured over each quintile. Thus, for any given day for any given contract and any given city, the surprises for that particular day are ordered from lowest to highest. Thus, a low surprise measure means that the temperature that day was much lower than the expected value. The values in the table are the returns from purchasing a futures contract at the close of that day and holding until expiration at the end of the month. These returns are computed for all surprise measures and then the results are aggregated and averaged by the quintile. Signs of overreaction would be increasing returns from lowest to highest quintile for HDD contracts and decreasing returns from lowest to highest quintile for CDD contracts. These results are aggregated by city and individual returns are divided by time until month's end in order to make surprises from different days of the month more comparable.

Table 14: Tests for Intermarket Efficiency

Sample	r_L	r_S	$r_L - r_S$	nobs	t-stat
2005-2008	0.40	-2.32	2.71	10.00	0.56
1999-2008	-4.05	-6.26	2.21	18.00	0.47

Note: r_L is the return for the long portfolio, r_S is the return for the short portfolio, $r_L - r_S$ is the return for the long minus short portfolio, and t-stats represents the t-statistic for the hypothesis that the difference in returns equals 0.

VIII Figures

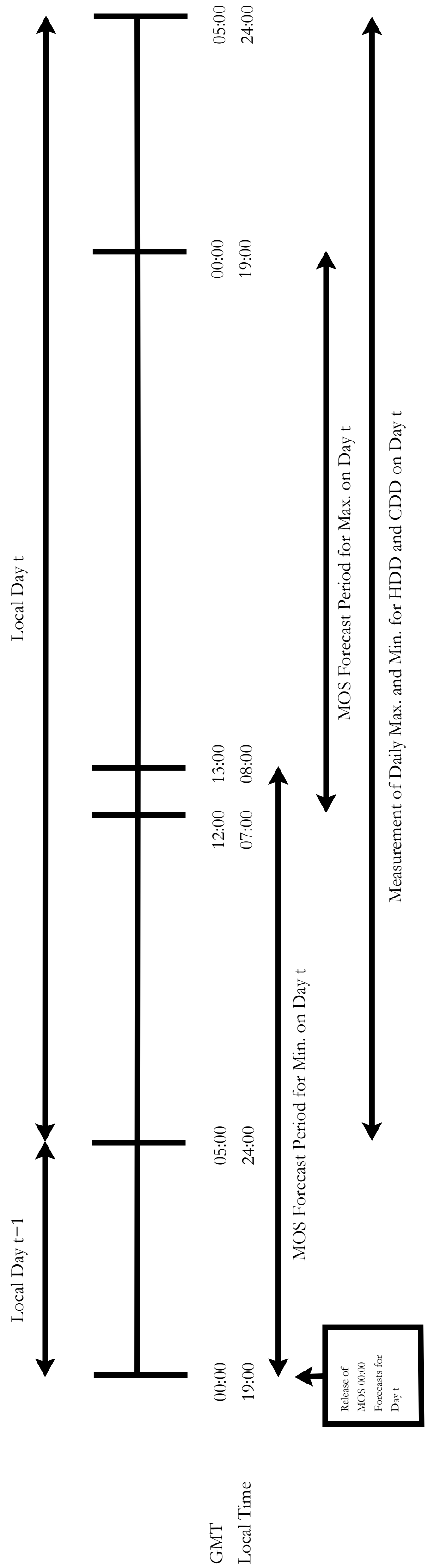

Figure 1: Example of Relationship Between MOS Forecasts at GMT Time and the Local Time for the City of Atlanta

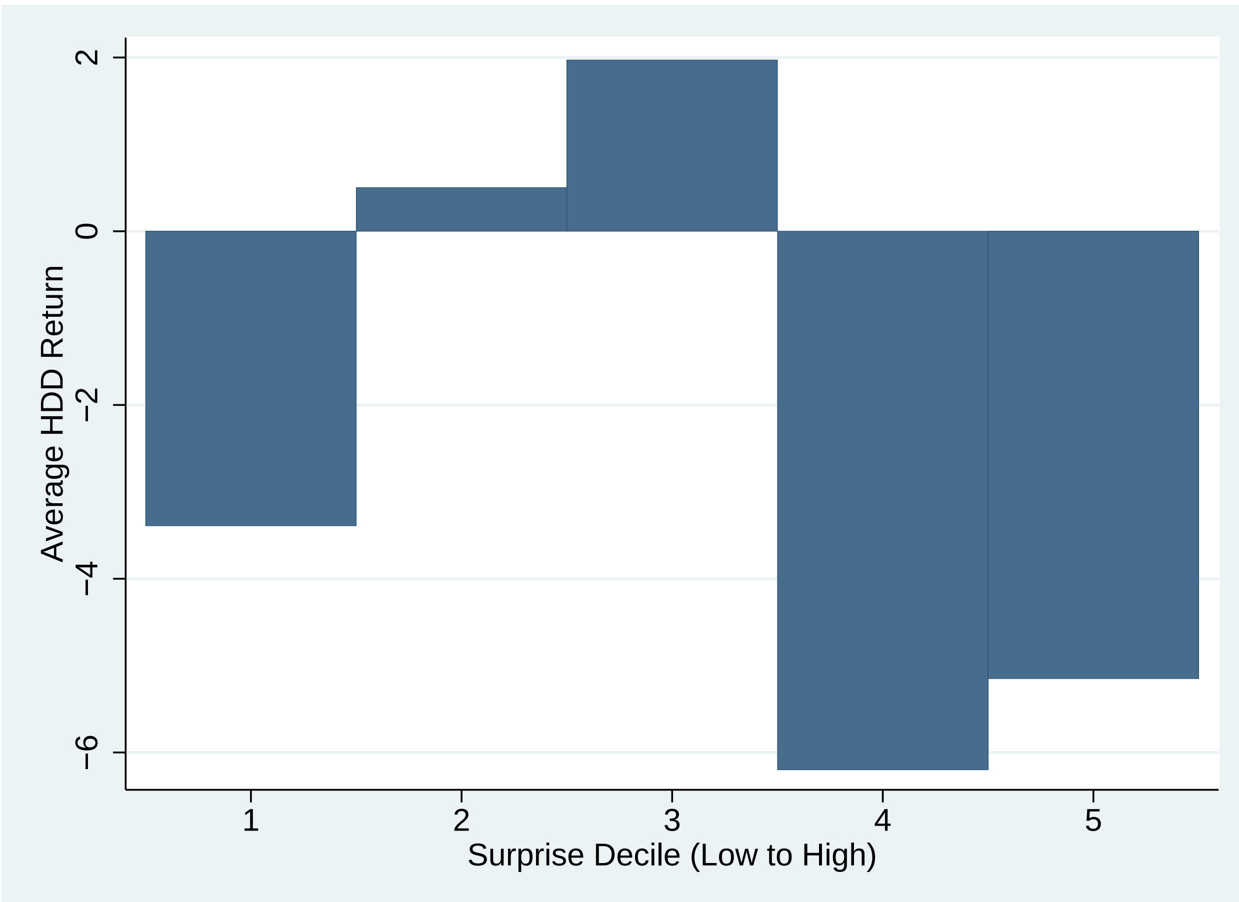

Figure 2: Overreaction to Weather Surprises by Quintile for HDD and Surprise Measure 3

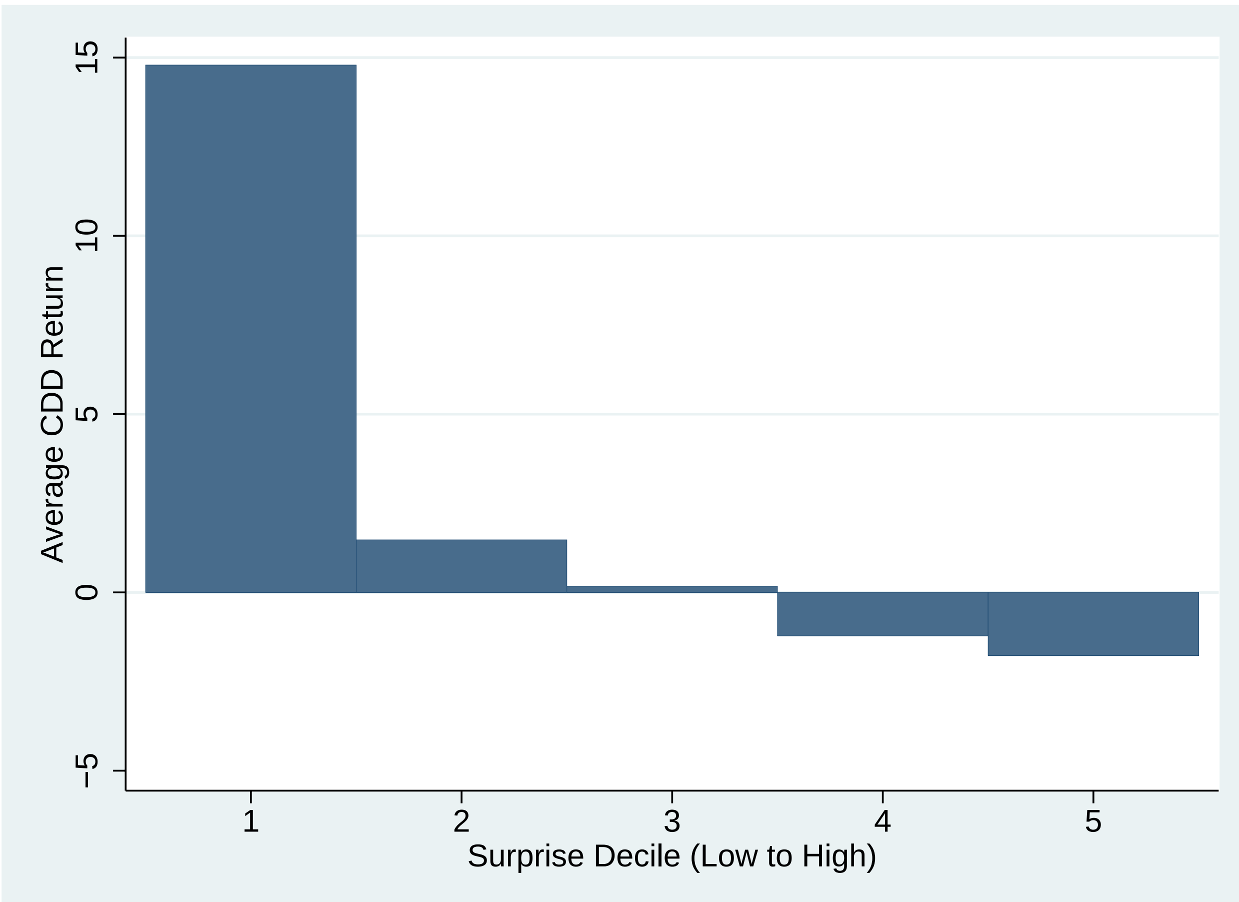

Figure 3: Overreaction to Weather Surprises by Quintile for CDD and Surprise Measure 3

References

[1] Benth, Fred Espen and Jurate Saltyte Benth. "The Volatility of Temperature and Pricing of Weather Deriviatives". *Quantitative Finance*, 7:553–561, October 2007.

[2] Boudoukh, Jason, Richardson, Matthew, Shen YuQing, and Robert F. Whitelaw. "Do Asset Prices Reflect Fundamentals? Freshly Squeezed Evidence from the OJ Market". *Journal of Financial Economics*, 83:397–412, 2007.

[3] Brockett, Patrick L., Wang, Mulong, and Chuanhou Yang. "Weather Derivatives and Weather Risk Management". *Risk Management and Insurance Review*, 8:127–140, Spring 2005.

[4] Brody, Dorje C, Syroka, Joanna, and Mihail Zervos. "Dynamical Pricing of Weather Derivatives". *Quantitative Finance*, 2:189–198, May 2002.

[5] Campbell, Sean D. and Francis X. Diebold. "Weather Forecasting for Weather Derivatives". *Journal of the American Statistical Association*, 100:6–16, March 2005.

[6] Cao, Melanie and Jason Wei. "Weather Derivatives Valuation and Market Price of Weather Risk". *Journal of Futures Markets*, 24:1089–1065, November 2004.

[7] Chiao, Chaoshin, Chengand, David C. and Hung Welfeng. "Overreaction after Controlling for Size and Book-to-Market Efects and its Mimicking Portfolio in Japan". *Review of Quantitative Finance and Accounting*, 51:65–91, December 2005.

[8] DeBondt, Werner F.M. and Richard Thaler. "Does the Stock Market Overreact?". *The Journal of Finance*, 40:793–805, July 1985.

[9] DeBondt, Werner F.M. and Richard Thaler. "Further Evidence on Investor Overreaction and Stock Market Seasonality". *The Journal of Finance*, 42:557–581, July 1987.

[10] Dischel, Robert S. (Editor). *Climate Risk and the Weather Market.* Risk Waters Group, 2002.

[11] Garman, Mark, Blanco, Carlos and Rovert Erikson. "Seeking a Standard Pricing Model". *Environmental Science*, March 2000.

[12] Garman, Mark, Blanco, Carlos and Rovert Erikson. "Chapter 8. Weather Derivative Modelling and Valuation: A Statistical Perspective". *Climate Risk and the Weather Market: Financial Risk Management with Weather Hedges*, 2002.

[13] Jegadeesh, Narasimhan and Sheridan Titman. "Overreaction, Delayed Reaction, and Contrarian Profits". *The Review of Financial Studies*, 8:973–993, Winter 1995.

[14] Jewson, Stephen, Brix, Anders, and Christine Ziehmann. *Weather Derivative Valuation.* Cambridge University Press, 2007.

[15] Longstaff, Francis A. and Ashley W. Wang. "Electricity Forward Prices: A High-Frequency Empirical Analysis". *Journal of Finance*, 59:1877–1900, August 2004.

[16] Loughran, Tim and Jay R. Ritter. "Long-Term Market Overreaction: The Effect of Low-Priced Stocks". *Journal of Finance*, 51:1959–1970, December 1996.

[17] Mann, H.B. and D.R. Whitney. "On a Test of Whether One or Two Random Variables is Stochastically Larger Than the Other". *Annals of Mathematical Statistics*, 18:50–60, 1947.

[18] Richards, Timothy J., Manfredo, Mark R., and Dwight R. Sanders. "Pricing Weather Derivatives". *American Journal of Agricultural Economics*, 86:1005–1017, November 2004.

[19] Rolfo, Jacques. "Optimal Hedging Under Price and Quantity Uncertainty: The Case of a Cocoa Producer". *Journal of Political Economy*, 88:100–116, Feb 1980.

[20] Roll, Richard. "Orange Juice and Weather". *The American Economic Review*, 74:861–880, December 1984.

[21] Taylor, James W. and Roberto Buizza. "A Comparison of Temperature Density Forecasts from GARCH and Atmospheric Models". *Journal of Forecasting*, 23:3337–355, 2004.

[22] Vedenov, Dmitry V. and Barry J. Barnett. "Efficiency of Weather Derivatives as Primary Crop Insurance Instruments". *Agricultural and Resource Economics*, 29:387–403, December 2004.

[23] Zarowin, Paul. "Long-Term Market Overreaction: The Effect of Low-Priced Stocks". *Journal of Financial and Quantitative Analysis*, 25:113–125, March 1990.

[24] Zeng, Lixin. "Pricing Weather Derivatives". *Journal of Portfolio Management*, pages 72–78, Spring 2000.

[25] Zhu, Jianzhou. "Do Investors Over- or Under-React? Evidence from Hong Kong Stock Market". *Journal of American Academy of Business*, 11:58–63, September 2007.