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

LUDWIG CHINCARINI*

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. First, by comparing the market's implied forecasts for the weather against other forecasts. Second, by looking at whether market's overreact or under-react to temperature surprises. Third, 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 risk premia that seem to vary across cities and over time. © 2010 Wiley Periodicals, Inc. Jrl Fut Mark 31:1–33, 2011

I thank John Wick and Yuxiang Wu for research assistance. I thank Daehwan Kim for valuable suggestions; Jan Carpenter, Kathryn Gilbert, Paul Knight, and Christopher Hyde for details of the official HDD calculation; Paul Peterson, Heidi Stonehill, Ken Heideman, Jamie Meier, and Scott Stephens for aid in obtaining weather forecasting data; Bloomberg and the FEI Institute for access to weather futures data and the National Climatic Data Center for historical weather data. In addition, I thank Neer Asherie, Amir Barnea, Pierangelo DePace, Eric Hughson, John Jurewitz, Slavi Slavov, Robert F. Whitelaw, Fan Yu, two anonymous referees, and the editor, Robert Webb, for helpful comments.

*Correspondence author, Department of Economics, Pomona College, 425 N. College Avenue #208, Claremont, CA 91711. Tel: 909-621-8881, Fax: 909-621-8576, e-mail: chincarinil@hotmail.com

Received August 2009; Accepted February 2010

■ *Ludwig Chincarini is an Assistant Professor at the Department of Economics, Pomona College, Claremont, California.*

The Journal of Futures Markets, Vol. 31, No. 1, 1–33 (2011)

Published online April 1, 2010 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/fut.20463

INTRODUCTION

The weather futures market is a relatively new market. Weather futures contracts are used by producers to hedge their business risks due to factors that may affect the demand for their products. The term given to this quantity risk is *volumetric risk*. For example, a natural gas producer or oil producer's quantity demanded will be affected 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. The main types of users of weather derivatives include energy traders, hedge funds, institutional banks, and re-insurers. In this study, we examine 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 seem to be inefficiencies due to investor overreaction (Chiao, Cheng, & Hung, 2005; DeBondt & Thaler, 1985, 1987; Jegadeesh & Titman, 1995; Loughran & Ritter, 1996; Zarowin, 1989; 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 weather market, which is not only less liquid but also does not have any fundamental pricing model at its heart.

Much of the previous literature on weather derivatives 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. For example, with a typical commodity futures contract, such as oil, there is an underlying storable good. This enables the formulation of a pricing relationship between the spot price for the commodity and the futures price of the commodity that is based upon the no-arbitrage pricing approach. Weather is not a storable and tradable good, making the standard futures pricing framework unusable. Because of this aspect of weather derivatives, various articles have attempted to build weather derivative pricing models (Brody, Joanna, & Zervos, 2002; Cao & Wei, 2004; Garman, Carlos, & Erikson, 2000; Richards, Manfredo, & Sanders, 2004; Taylor & Buizza, 2004). A good summary of weather derivative valuation can be found in Jewson, Anders, and Ziehmann (2007). Another strand of research in this area has focused on the design of optimal weather contracts for producers of commodities. These studies discuss, design, and examine the effectiveness of different types of weather derivatives (Brockett, Mulong, & Yang, 2005; Leggio, 2007). Other research has focused on the usefulness of weather forecasts and other types of models to forecast the weather (Benth & Benth, 2007; Brix, Stephen, & Ziehmann, 2002; Campbell & Diebold, 2005; Zeng, 2000). There are some studies that discuss weather derivatives in the context of 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 study 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 is overreaction or under-reaction in the weather futures market.

The study is organized as follows: the next section discusses the details about the weather futures market; the third section discusses the data that we use in our research; the fourth section discusses the methodology which we use to test for the efficiency of the weather futures market; the fifth section discusses the results from our empirical investigation of the weather futures market; and the last section concludes.

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 index futures and options that are geared to seasonal and monthly weather in 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 HDD and CDD. The HDD stands for heating degree days and 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

[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.

TABLE I
CME Weather Contract Specifications

Trading hours	*Futures products* trade electronically only on CME Globex from Monday through Friday from 3:45 P.M. to 3:15 P.M. (central time) of the following day, and on Sundays from 5:30 P.M. to 3:15 P.M. On the last trading day they trade until 9 A.M. *Option products* trade only from Monday through Friday from 8:15 A.M. to 3:15 P.M. 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 Settlement	One degree day index point Cash settled. All contracts that remain open at the termination 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.

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 the 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 five values for HDD for the month of January with 25, 15, 20, 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 I contains the major features of the weather contracts in the United States. A simple example using one of the specific city contracts might make things clearer. On February 28, 2006, the March monthly HDD contract for Atlanta closed at 305. This indicated that the market's fair value for the sum of March HDD daily values in Atlanta was 305. The weather for the HDD contract is shown in Table II along with the daily closing prices of the March 2006 futures contract. For this month, the market underestimated

[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 0 for values below 65 is that heating degree days (HDD) refers to days in which one would need to use a heater, while cooling degree days (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.

TABLE II
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.

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, 2006 would have paid $6,100 for one contract and had they held the contract until expiration would have made $880 ($6,980–$6,100).

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 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 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 to March, while CDD seasonal strips have start months from April to September. This study will focus only on the HDD and CDD monthly futures contracts.

THE DATA

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 span the period September 22, 1999 to June 30, 2008.[4] Table III summarizes the futures contracts for all 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 32.85. Thus, when there is trading, about 33 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 409 trading days with volume for the entire sample period.

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 the 18 major cities, we have daily maximum and minimum temperatures for each 24-hr period going back as far as data

[4]Weather futures were introduced on the CME in 1999 for the following cities: Atlanta, Chicago, Cincinnati, Dallas, Denver, Las Vegas, New York, Philadelphia, Portland, and Tucson. The Denver contract was eventually replaced with Des Moines. Contracts for other cities were introduced in subsequent years. Despite being launched in September 1999, some of these contracts did not actually start trading until much later.

TABLE III
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	32.85	409
Atlanta CDD	5:2002	39.14	341
Baltimore HDD	11:2005	24.84	73
Baltimore CDD	6:2006	18.27	29
Boston HDD	3:2003	36.71	224
Boston CDD	5:2004	49.56	134
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	39.53	288
Des Moines HDD	10:2002	30.11	320
Des Moines CDD	7:2000	31.90	237
Detroit HDD	4:2008	0.17	2
Detroit CDD	N.A.	N.A.	0
Houston HDD	10:2003	30.21	92
Houston CDD	9:2003	30.91	98
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	42.77	187
New York HDD	11:1999	66.82	660
New York CDD	5:2002	73.06	494
Philadelphia HDD	1:2002	23.23	251
Philadelphia CDD	5:2002	46.13	242
Portland HDD	1:2002	17.66	110
Portland CDD	6:2002	19.48	97
Sacramento HDD	10:2003	14.22	59
Sacramento CDD	10:2003	38.18	149
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 price is available 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) were used for these summary statistics.

were collected for that particular city. For the purposes of our work, we use only data from January 1, 1960 onwards.[5] The summary statistics for the weather in each of these cities is available from the author on request.

[5]The historical weather data was cutoff in 1960 since most practitioners use either a 10-year or 30-year window of historical weather data to forecast the weather. Since our weather futures data begin in 1999, this gives us slightly more data than the typical period.

In addition to historical weather data, we obtained historical weather forecast data as produced by the 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.[6] A summary of the weather forecast data is available from the authors on request.

Finally, any closing prices without volume information for that particular day are removed from the data set. 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.[7] Also, in order to remove potentially bad 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. This led to a total of 13 dropped observations for the analysis period of this study from 2005 to 2008. We did use alternative cut-offs and it did not affect the nature of the results.

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 used to measure the efficiency of the weather derivatives market.

Futures Premia and the Market Price of Risk

The typical pricing of a futures contract relies on the futures-spot price relationship based on 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 useful for the weather market because there is no underlying spot weather for a given futures contract.[8] Although some models for weather derivatives

[6]There is a slight difference between the period forecasted by the GFS and the actual measurement of daily maximum and minimum temperatures, which is described in Chincarini (2009).

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

[8]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."

have been proposed (Campbell & Diebold, 2005; Hardle & Cabrera, 2009), they are not arbitrage-free models or models based on sound economic theory. The models involve fitting the weather data to some complicated functional forms. Since these models are not based on strong theoretical foundations, rather than describe the forward premia theoretically, we follow the approach of Longstaff and Wang (2004) and calculate the forward premia empirically by computing the normal backwardation or contango of the market. Thus, we measure the forward premium in percentage terms as:

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

where F_{imt} is the futures price of the HDD (or CDD) contract in city i on day t for the 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. A similar type of logic would apply to electricity or energy suppliers in the summer with the goal of hedging volumetric risk.[9] 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.[10]

Although we have speculated on what values the premia might have, it is actually important to understand them. To test market efficiency, it would help to know the market price of risk. For instance, suppose that HDD trade at a discount of 5% to their expected settlement prices because that is the expected return required to compensate speculators for the risk of holding these contracts. If we ignored this when using the price of the HDD to determine how well the market predicted the actual settlement of HDD, we might incorrectly conclude that the market on average was incorrect by 5% although the market was perfectly predicting the weather, but was extracting a 5% premium. Given our lack of reliable models, one might choose to use a sub-sample of the data to

[9]This might not be the only direction this would go. It really depends on the weight of various hedgers in this market. For example, if the primary hedgers are electricity suppliers that are concerned with extremely hot summer temperatures, which might drive up electricity demand to the point where suppliers were forced to use very high cost methods (e.g. purchasing electricity off the national grid at current spot market prices), then these hedgers might take long positions in CDD contracts or in call option contracts.

[10]Of course, if local utility companies were the primary hedgers, then this result would be exactly the opposite.

extract the market price of risk in the weather markets. However, this might be problematic for several reasons. First, the premia might be different for different cities. Second, the market price of risk might be time-varying or even have changed since the introduction of weather futures' contracts due perhaps to the changing composition of the different players in this market. Third, estimating a premium will require dividing an already small market in two other sub-sample periods, which will leave us with fewer data points to estimate it. Fourth, any empirical measure of the risk premia might contain both market risk and liquidity risk components.

To get an empirical idea of the market risk premia, we chose to look at the five most traded cities from 1999 to 2004. These cities are Atlanta, Chicago, Cincinnati, Dallas, and New York. Table IV shows the estimated premia for each individual city and for the top five most liquid cities from 1999 to 2004 and from 2005 to 2008.[11]

The market risk premia are all normalized to be comparable as a monthly return. The premium for HDD is 0.47% and insignificant for the most liquid cities from 1999 to 2004. The CDD premium is 5.76% and significantly different than zero. This premium is positive, unlike what our discussion above implied. Also, it is mainly being driven by Atlanta, Chicago, and Cincinnati. Dallas has an insignificant premium and New York is actually significantly negative. A quick look at the other cities not included in the top five shows that some have positive premia for HDD, whereas others have negative premia. This is also true for the CDD contracts. In the latter period (2005–2008), the HDD premia is positive and significant and the CDD premia are negative and significant for the top five cities. Thus, the CDD average premia for the top five cities actually changes sign from the 1999–2004 period to the 2005–2008 period. Many of the other cities do not have risk premia significantly different from zero. And in the cases of significance, sometimes the signs are different. The bottom of Table IV computes the risk premia for the five most liquid cities for each year. Once again, this premia changes from year-to-year and it is hard to find a consistent value. This inconsistency in the realized forward premia might be due to a different mix of hedgers across markets and across time. For example, New York City might be a market dominated by local utility hedgers, whereas other cities like Minneapolis are not.

Although not shown here, we investigated the forward premia by day of month. Overall, the results indicated that there did not seem to be significant

[11]The reader will notice that the number of observations drops significantly as compared to Table III. This is not due to the elimination of outliers. This is because in Table III every day that a contract has volume is counted, including when a contract for a particular month trades in prior months. In the empirical analysis of this study, we only consider the returns of contracts during the actual contract month. Thus, many observations are lost due to this, but not in a way to impart any particular bias to the analysis.

TABLE IV

Realized Percentage Forward Premia in the Weather Derivative Futures

	HDD						CDD						
City	*Mean*	*t-stat*	*SD*	*Max*	*Min*	*nobs*	*Mean*	*t-stat*	*SD*	*Max*	*Min*	*nobs*	σ
Period: 1999–2004													
at	−0.96	−0.39	26.25	96.69	−71.43	113.00	9.37	1.75	42.41	214.29	−121.09	63.00	6.77
ba	–	–	–	–	–	0.00	–	–	–	–	–	0.00	7.63
bo	1.16	0.49	14.87	26.49	−25.56	40.00	17.29	2.38	30.01	114.95	−25.12	17.00	7.57
ch	2.40	2.04	12.04	56.53	−32.01	105.00	12.74	2.07	39.97	135.96	−100.45	42.00	8.81
ck	3.85	1.84	19.11	55.63	−54.27	83.00	24.43	5.42	34.31	107.66	−97.27	58.00	8.53
da	1.51	0.32	38.60	81.19	−155.00	67.00	3.97	1.29	23.80	45.66	−106.55	60.00	7.30
dm	−2.23	−0.92	18.19	23.08	−58.97	56.00	6.77	1.24	36.33	91.28	−68.30	44.00	9.24
de	–	–	–	–	–	0.00	–	–	–	–	–	0.00	8.19
ho	−3.75	−0.19	48.36	72.16	−60.00	6.00	−0.32	−0.04	23.01	39.80	−39.35	8.00	6.41
kc	7.79	2.03	12.72	24.70	−19.48	11.00	16.99	1.27	46.39	82.65	−40.48	12.00	9.01
lv	0.16	0.04	19.36	68.73	−20.71	23.00	0.23	0.07	17.35	72.86	−30.40	31.00	6.18
mn	1.36	0.52	14.74	18.01	−58.65	32.00	23.51	1.47	62.14	124.00	−89.67	15.00	9.40
ny	−0.30	−0.21	13.76	30.04	−28.47	98.00	−9.64	−2.15	35.53	106.41	−167.57	63.00	7.34
ph	−5.99	−2.77	15.73	22.19	−74.68	53.00	2.00	0.54	27.24	110.71	−74.06	54.00	7.39
po	−0.24	−0.03	32.15	33.83	−125.37	21.00	−11.16	−0.91	53.21	128.40	−142.50	19.00	5.16
sa	17.65	1.18	25.92	47.35	−0.42	3.00	−9.78	−0.41	47.99	38.75	−67.17	4.00	5.20
sl	–	–	–	–	–	0.00	–	–	–	–	–	0.00	7.72
tu	−0.48	−0.09	27.04	57.08	−32.89	27.00	−1.28	−0.39	19.88	33.99	−90.53	36.00	5.84
Top 5 Cities	0.47	0.34	22.11	96.69	−155.00	266.00	5.76	2.11	36.79	214.29	−167.57	182.00	

(*Continued*)

TABLE IV *(Continued)*

	HDD						CDD						
City	*Mean*	*t-stat*	*SD*	*Max*	*Min*	*nobs*	*Mean*	*t-stat*	*SD*	*Max*	*Min*	*nobs*	σ
Period: 2005–2008													
at	−2.33	−1.27	23.28	75.00	−93.78	162.00	−7.77	−2.84	28.54	59.92	−125.09	109.00	6.77
ba	3.46	2.42	10.08	30.00	−14.24	50.00	10.22	1.38	25.70	56.72	−21.40	12.00	7.63
bo	2.11	1.41	15.69	86.51	−25.30	109.00	−6.02	−1.38	37.97	81.95	−207.36	76.00	7.57
ch	1.93	2.25	13.98	43.06	−47.83	266.00	−8.34	−1.84	59.92	88.26	−262.31	175.00	8.81
ck	0.23	0.20	17.04	37.49	−80.57	221.00	−8.80	−2.77	35.67	100.68	−112.52	126.00	8.53
da	14.67	3.05	55.47	442.86	−61.95	133.00	0.52	0.31	17.45	71.43	−39.14	111.00	7.30
dm	3.83	3.06	15.87	57.54	−38.60	161.00	3.61	1.75	21.33	60.00	−50.38	107.00	9.24
de	11.95	1.12	15.12	22.64	1.25	2.00	–	–	–	–	–	0.00	8.19
ho	12.36	1.43	64.74	420.71	−80.05	56.00	−1.32	−0.84	11.25	22.82	−28.79	52.00	6.41
kc	2.23	1.53	19.01	86.11	−64.52	170.00	0.56	0.20	26.18	50.81	−111.21	87.00	9.01
lv	3.81	0.66	44.79	193.75	−81.25	61.00	1.31	1.12	9.76	29.37	−16.74	69.00	6.18
mn	3.70	3.60	13.86	90.82	−33.22	182.00	10.65	4.65	22.30	89.56	−42.83	95.00	9.40
ny	7.60	8.27	16.12	87.44	−30.95	308.00	−16.89	−4.90	46.67	134.78	−219.73	183.00	7.34
ph	0.09	0.06	14.68	67.33	−31.69	89.00	−6.70	−1.53	35.34	95.95	−110.87	65.00	7.39
po	−1.73	−0.93	13.38	42.94	−26.21	52.00	3.37	0.23	91.16	77.65	−493.18	38.00	5.16
sa	5.16	1.55	18.87	65.92	−25.00	32.00	5.42	1.64	32.93	69.41	−139.29	100.00	5.20
sl	−5.25	−0.73	16.01	11.64	−27.09	5.00	6.15	–	–	6.15	6.15	1.00	7.72
tu	20.70	1.58	80.95	322.92	−86.80	38.00	6.67	4.40	10.72	27.32	−15.14	50.00	5.84
Top 5 Cities	4.51	4.54	19.94	220.91	−46.18	403.00	−8.10	−3.36	40.68	134.78	−192.47	284.00	–
Top5:2000	−23.25	−2.70	27.24	21.53	−66.61	10.00	–	–	–	–	–	0.00	–
Top5:2001	–	–	–	–	–	0.00	–	–	–	–	–	0.00	–
Top5:2002	−3.96	−0.78	33.26	47.17	−155.00	43.00	−3.08	−0.46	37.92	46.62	−167.57	32.00	–
Top5:2003	0.62	0.30	20.16	64.27	−71.43	94.00	15.30	3.01	43.09	214.29	−121.09	72.00	–
Top5:2004	2.87	1.78	16.24	96.69	−32.01	102.00	0.58	0.19	27.33	57.15	−106.55	78.00	–
Top5:2005	0.35	0.26	14.07	58.62	−44.19	113.00	−0.52	−0.17	29.34	71.43	−83.85	93.00	–
Top5:2006	5.95	4.24	15.06	33.37	−46.18	115.00	−8.61	−1.58	52.03	134.78	−192.47	91.00	–
Top5:2007	10.65	3.69	29.87	220.91	−17.72	107.00	−16.33	−4.29	37.11	49.86	−170.28	95.00	–
Top5:2008	−0.65	−0.46	11.75	41.77	−26.75	68.00	16.43	2.07	17.74	31.07	−14.15	5.00	–

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 normalized to a monthly premium equivalent. The premia are computed over the two sub-samples of data from 1999 to 2004 and 2005 to 2008. The return from buying the contract is the negative of the forward premium percentage in the table. σ represents the average of the standard deviation of daily temperature across months in each city. The top five cities are Atlanta, Chicago, Cincinnati, Dallas, and New York.

premia to speculators in weather derivatives for particular days of the month, which might be due to the varying concentrations of the types of hedgers in different cities over time. Because of the inconsistency of a reliable market price of risk in these markets from our sub-sample analysis, we chose to use a zero risk premium in the remainder of the study.

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 us mathematically define a few variables discussed earlier. 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.

Static models

The first static model, which we call Static Historical Model, is a simple model that uses a 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 40 years, but not including the current month.[12] 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 $H^{MOS}_{i,y,m,d,1}, \ldots, H^{MOS}_{i,y,m,d,7}$ and $L^{MOS}_{i,y,m,d,1}, \ldots, L^{MOS}_{i,y,m,d,7}$, 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 forecast for the average daily temperature, which we denote as $T^{MOS}_{i,y,m,d,1}, \ldots, T^{MOS}_{i,y,m,d,7}$. We construct the MOS forecasted daily HDD as[13]

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

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

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

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. This forecast for the entire month uses the first 7 days forecast and then uses the historical average

[12]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. In other work by the author, certain cities contained some trend in average temperatures over time. Future research might consider using a more elaborate historical model, which estimates a linear trend model and then makes forecasts based on such a model. Jewson et al. (2007) discuss some of the problems with estimating simple trend models.

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

of temperature for that city and that month for the rest of the month as the monthly forecast for HDD.[14]

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, one might argue that it does not 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 on 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. These are not useful for our study because they do not give point estimates for the weather. Thus, data limitations limit the ability to produce a more accurate model forecast.[15] 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.[16]

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,

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

[14]Although not reported here, another adjustment factor was used that converted the first 7-days of the forecast into a monthly forecast by multiplying 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.

[15]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 had 8-months of historical data at the time of writing of this study. In the near future, it certainly would be of interest to expand on this current research with this longer-dated forecast horizon.

[16]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.

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:

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

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 $D(y, m) - d \geq 7$, 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 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. $F_{i,y,m,d}$.

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 $\hat{S}_{i,y,m,d} = (T_{i,y,m,d} - E(T_{i,y,m,d}))/SD(T_{i,y,m,d})$, 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.[17] 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, $(1/(y - y_0)) \sum_{y'=y_0}^{y-1} T_{i,y',m,d}$.

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

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:

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

[17]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.

where $\overline{HDD}^{MKT}_{i,y,m,d'}$ 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 calculation 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 the recovery of the market's expectation of HDD on day $d + 1$, we assume that the adjustment factor is relevant so that

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

which implies

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

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

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

Note that $\overline{HDD}^{MKT}_{i,y,m,d}$ could be negative, which is not allowed in the original definition of the daily HDD.[18]

Once these series are created, we collect the value of all these for all days and then sort them by quintile. We then compute the returns of buying the contract at the close of business of the *next day* and holding until month's end.[19] We can then draw histograms of the quintile results and test for significant

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

[19]We also computed the results for purchase on the same day, which are available from the author on request.

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

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 quintiles 1 and 5. 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

$$Z_1 = \frac{\overline{R}_1 - \overline{R}_5}{(\frac{1}{m} + \frac{1}{n})^{1/2}\sigma},\ Z_2 = \frac{S_{R_1}^2 + S_{R_5}^2}{\sigma^2}$$

where $\overline{R}_1$ represents the mean return from quintile 1, $\overline{R}_5$ 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, $S_{R_1}^2 = \sum_{i=1}^{m} (R_{1i} - \overline{R}_1)^2$ and $S_{R_5}^2 = \sum_{i=1}^{n} (R_{5i} - \overline{R}_5)^2$. 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 quintiles 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 & Whitney, 1947). The procedure for the test is as follows:

Rank all observations in quintile 1 and quintile 5 from smallest to largest.

Sum the ranks of quintile 1 and call this SR_1.

Compute the Mann–Whitney test statistic as:

$$U = mn + \frac{m(m + 1)}{2} - SR_1.$$

*We do not directly test for market learning in this paper, but this might be an idea for future research.

Large values of this statistic suggest that the samples are drawn from different populations where quintile 1 has smaller means.

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 quintile 1's location parameter is greater than that of quintile 5.

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 whether 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 current differences in HDD implied by market prices. Then, we choose three pairs of cities whose current differences exceed the historical differences most and create zero-investment portfolios out of these pairs. Similarly, we choose three 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 markets are efficient, but it 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 an $N - 1$ by $N - 1$ upper diagonal matrix, where N is the number of cities:

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

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

$$\Delta_{y,m,d}^{MKT} = \begin{pmatrix} F_{1,y,m,d} - F_{2,y,m,d} & F_{1,y,m,d} - F_{3,y,m,d} & \cdots \\ & F_{2,y,m} - F_{3,y,m} & \cdots \\ & & \ddots \end{pmatrix} \quad (19)$$

For each day, we take the differences in these two matrices $\Delta_{y,m,d}^{MKT} - \Delta_{y,m}^{Hist}$ and find the three cells with the largest absolute values. Let us denote 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 "mispricing." 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 the 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.

EMPIRICAL RESULTS

The Accuracy of Market Predictions of Weather

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 is the only fair comparison of the models.[20] The empirical analysis is done for all cities since it makes a study on this new market more complete, although some cities have very few observations.

Table V 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) than 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 non-existent.

Table VI shows the performance of these models by city and 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

[20]Although not reported here, the efficiency of this market was examined over longer horizons and the qualitative nature of the results did not change.

TABLE V
Static Models' Forecast Performance

Contract Type	Statistics	Historical RMSE	Historical MAE	MOS Forecasts RMSE	MOS Forecasts MAE	Market RMSE	Market 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 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 that exist for all three forecast methodologies, which are limited by the market statistics availability. The sample period is constrained to days in which data exist for all three measures.

accuracy is substantially better, like in Dallas, Texas where the market improves over the MOS forecasts by 117%. 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 VII 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 is the reverse. However, in all but three cases (February-HDD, August, and September-CDD), the market's forecasts are much more reliable.

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.[21] Table VIII contains the results for the dynamic forecasts aggregated. The results are similar to the static case. That is, the market still seems to be the best predictor of eventual

[21]Another way to control 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 the MAE and RMSE much smaller relative to the static results.

TABLE VI
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 conditioning on volume. That is, the market prices are only used if there existed trading volume on that particular day. The sample period is constrained to days in which data exist for all three measures.

TABLE VII
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 conditioning on volume. That is, the market prices are only used if there existed trading volume on that particular day. The sample period is constrained to days in which data exist for all three measures.

temperature and HDD for a month when information is updated daily. Tables VIII and IX 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.[22]

Overall, the results of the static and dynamic models are quite supportive of a relatively efficient market place in weather derivatives.

[22]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 said that the quotes are legitimate. Although they make little sense, I have left them in for completeness.

TABLE VIII
Dynamic Models' Forecast Performance

		Historical		*MOS Forecasts*		*Market*	
Contract Type	*Statistics*	*RMSE*	*MAE*	*RMSE*	*MAE*	*RMSE*	*MAE*
HDD	Mean	88.31	65.18	72.46	51.02	57.19	42.90
	Max	149.11	99.87	121.52	80.28	91.52	72.00
	Min	42.67	29.56	39.31	25.59	28.46	21.69
	Nobs	1,830.00	1,830.00	1,830.00	1,830.00	1,830.00	1,830.00
CDD	Mean	40.23	31.13	33.39	26.30	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	938.00	938.00	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 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 that exist for all three forecast methodologies, which are 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.

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-than-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 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.[23]

Figures 1 and 2 show the aggregated returns across days and cities for surprise measure 3.[24] In some sense, we believe this is the most reliable measure of surprise, since it is the temperature deviation from the market's implied expected

[23]The results would be the reverse for an under-reaction story.
[24]The results for the other surprise measures are qualitatively the same and available from the authors on request.

TABLE IX
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

(Continued)

TABLE IX *(Continued)*

	Historical			*MOS Forecasts*			*Market*		
Day	*RMSE*	*MAE*	*MAE Dif*	*RMSE*	*MAE*	*MAE Dif*	*RMSE*	*MAE*	*nobs*
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 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 that exist for all three forecast methodologies, which are 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.

temperature based on the previous day's market prices. For both HDD and CDD, there is no strong evidence consistent with an overreaction hypothesis.

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*.[25] 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 X shows the averaged returns from each quintile for surprise measure 3 using data from 2005 to 2008 and returns as measured from the same business day's closing prices.[26] For none of the surprise measures does there

[25]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 30 of the month. Dividing by days left in the month might help to make the results less distorted.
[26]The same tables for surprise measures 1 and 2 are available from the author upon request.

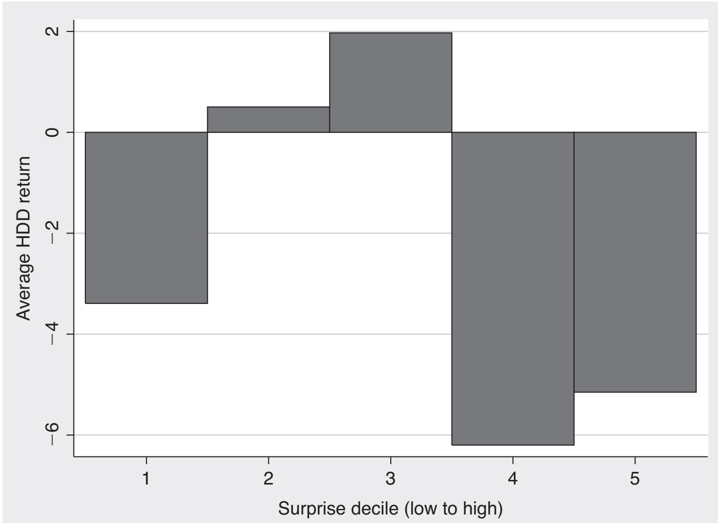

FIGURE 1
Overreaction to weather surprises by quintile for HDD and surprise measure 3.

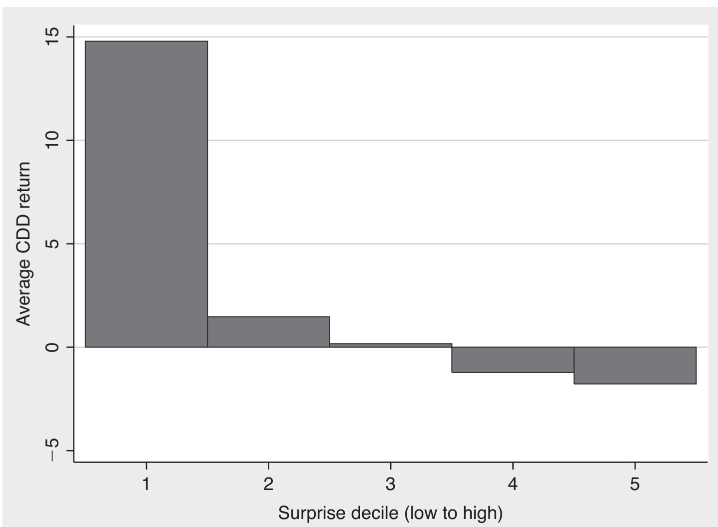

FIGURE 2
Overreaction to Weather Surprises by quintile for CDD and surprise measure 3.

seem to be any consistent pattern emerging in terms of overreaction or underreaction to weather surprises.

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

TABLE X
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.

All the overreaction tests were also done when the investors were allowed to purchase contracts on the *next day* of the surprise rather than the same business day using data from 2005 to 2008.[27] In addition to this, the overreaction tests

[27]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 realistic.

TABLE XI

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

	HDD								CDD							
	Sample Size			t-Statistics			Mann-Whitney Tests		Sample Size			t-Statistics			Mann-Whitney Tests	
Day	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 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.

TABLE XII
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	0.22	0.30	0.57	0.56	0.89
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 on 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. Individual returns are divided by time until month's end in order to make surprises from different days of the month more comparable.

were done with a longer sample period from 1999 to 2008 for computing the returns. In all cases, the qualitative results are the same. There does not appear to be overreaction or under-reaction in these markets, except for occasional sporadic cases.

In Table XII we present the overreaction results by city. While some cities seem to exhibit patterns of overreaction and others under-reaction, there is no consistent pattern across cities.

Inter-Market Behavior

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

TABLE XIII
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.

Unfortunately, there are very few observations for both periods. The return is in the direction of inefficiency for the whole sample period, but the *t*-statistics for difference in returns is insignificant. Thus, for this particular test for cross-city weather market inefficiency, again the weather markets look quite efficient.

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 about 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 study, the efficiency of the weather derivatives market was examined in a variety of ways. Overall, despite its lack of sufficient depth as measured by volume and open interest, 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 second, we examined 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 under-reaction 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 might be more efficient, since there might not be a guessing game by non-informed traders on movements in prices, which may lead to overreaction in prices from uninformed price movements. Given the varying forward premia for different cities and across time in this market, it might be interesting to further study the composition of natural hedgers in different cities and across time. Also, further research into the interaction between information and efficiency might be investigated. In addition, it might prove useful to study whether other derivative contracts can span the set of weather derivative contracts or whether they are truly an invaluable hedging instrument for companies wanting to hedge volumetric risk.

BIBLIOGRAPHY

Benth, F. E., & Benth, J. S. (2007). The volatility of temperature and pricing of weather derivatives. Quantitative Finance, 7, 553–561.

Brix, A., Stephen, J., & Ziehmann, C. (2002). Chapter 8. Weather derivative modelling and valuation: A statistical perspective. Climate risk and the weather market: Financial risk management with weather hedges. London: Risk Books.

Brockett, P. L., Mulong, W., & Yang, C. (2005). Weather derivatives and weather risk management. Risk Management and Insurance Review, 8, 127–140.

Brody, D. C., Joanna, S., & Zervos, M. (2002). Dynamical pricing of weather derivatives. Quantitative Finance, 2, 189–198.

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

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

Chiao, C., Cheng, D. C., Hung, W. (2005). Overreaction after controlling for size and book-to-market effects and its mimicking portfolio in Japan. Review of Quantitative Finance and Accounting, 51, 65–91.

Chincarini, L. B. (2009). No chills or burns from temperature surprises: An empirical analysis of the weather derivatives market. Available at SSRN: http://ssrn.com/abstract=1340761.

DeBondt, W. F. M., & Thaler, R. (1985). Does the stock market overreact? The Journal of Finance, 40, 793–805.

DeBondt, W. F. M., & Thaler, R. (1987). Further evidence on investor overreaction and stock market seasonality. The Journal of Finance, 42, 557–581.

Dischel, R. S. (Ed.) (2002). Climate risk and the weather market. London: Risk Books.

Garman, M., Carlos, B., & Erikson, R. (2000). Seeking a standard pricing model. Environmental Finance.

Hardle, W. K., & Lopez Cabrera, B. (2009). Implied market price of weather risk. SFB 649 Discussion Paper 2009, from: http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2009-001.pdf.

Jegadeesh, N., & Titman, S. (1995). Overreaction, delayed reaction, and contrarian profits. The Review of Financial Studies, 8, 973–993.

Jewson, S., Aders, B., & Ziehmann, C. (2007). Weather derivative valuation. Cambridge: Cambridge University Press.

Leggio, K. B. (2007). Using weather derivatives to hedge precipitation exposure. Management Finance, 33, 246–252.

Longstaff, F. A., & Wang, A. W. (2004). Electricity forward prices: A high-frequency empirical analysis. Journal of Finance, 59, 1877–1900.

Loughran, T., & Ritter, J. R. (1996). Long-term market overreaction: The effect of low-priced stocks. Journal of Finance, 51, 1959–1970.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one or two random variables is stochastically larger than the other. Annals of Mathematical Statistics, 18, 50–60.

Richards, T. J., Manfredo, M. R., & Sanders, D. R. (2004). Pricing weather derivatives. American Journal of Agricultural Economics, 86, 1005–1017.

Rolfo, J. (1980). Optimal hedging under price and quantity uncertainty: The case of a cocoa producer. Journal of Political Economy, 88, 100–116.

Taylor, J. W., & Buizza, R. (2004). A comparison of temperature density forecasts from GARCH and atmospheric models. Journal of Forecasting, 23, 3337–3355.

Zarowin, P. (1990). Long-term market overreaction: The effect of low-priced stocks. Journal of Financial and Quantitative Analysis, 25, 113–125.

Zeng, L. (2000). Pricing weather derivatives. Journal of Risk Finance, 1, 72–78.

Zhu, J. (2007). Do investors over- or under-react? Evidence from Hong Kong Stock Market. Journal of American Academy of Business, 11, 58–63.