# Weather-Induced Mood, Institutional Investors, and Stock Returns

**William N. Goetzmann**
Yale University

**Dasol Kim**
Case Western Reserve University

**Alok Kumar**
University of Miami

**Qin Wang**
University of Michigan at Dearborn

This study shows that weather-based indicators of mood impact perceptions of mispricing and trading decisions of institutional investors. Using survey and disaggregated trade data, we show that relatively cloudier days increase perceived overpricing in individual stocks and the Dow Jones Industrial Index and increase selling propensities of institutions. We introduce stock-level measures of investor mood; investor optimism positively impacts stock returns among stocks with higher arbitrage costs, and stocks experiencing similar investor mood exhibit return comovement. These findings complement existing studies on how weather impacts stock index returns and identify another channel through which it can manifest. (*JEL* D84, G11, G12, G14, G23)

A number of recent studies show that weather patterns in major financial centers influence stock index returns and suggest that investor mood influences asset prices. For example, using data from international stock exchanges, Hirshleifer and Shumway (2003) show that stock market returns are higher on days when the weather is sunny, which is presumably when market participants are in a good mood.[1] The findings from this finance literature are consistent with evidence in the psychology literature, in that individuals misattribute mood

We thank two anonymous referees, Jawad Addoum, David Hirshleifer (the editor), George Korniotis, and Lisa Kramer for helpful discussions and valuable comments. We thank seminar participants at the University of Miami, Case Western Reserve University, University of Arkansas, and Western Finance Association. We are responsible for all remaining errors and omissions. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Dasol Kim, Department of Banking and Finance, Weatherhead School of Management, Case Western Reserve University, Cleveland, OH 44106-7235, USA. E-mail: dasol.kim@case.edu.

[1] Other studies that examine the impact of cloud cover on stock market index returns include Saunders (1993) and Goetzmann and Zhu (2005). See Section 1 for additional details.


doi:10.1093/rfs/hhu063 Advance Access publication September 23, 2014

induced by weather as information when making assessments about objects that should be otherwise unrelated (e.g., sunny days with positive assessments about the market).[2]

In this paper, we focus on the economic mechanism through which weather-induced mood affects asset prices. In particular, we investigate how mood affects the trading behavior of a group of investors that play a key role in the price formation process: institutional investors. Despite commonly held assumptions about investor sophistication, there is evidence that institutional investors and other sophisticated market participants are also susceptible to cognitive biases.[3] Motivated by these findings, we conjecture that weather patterns would affect the mood of institutional investors, in turn impacting their trading decisions and stock prices congruently. We test this conjecture using financial data from multiple sources that are matched to ZIP code-level weather station data.

Similar to previous studies, we use deseasonalized cloud cover as our investor mood proxy. In addition, we use survey data from the Yale International Center for Finance that provide institutional investor perceptions on stock market investment. Using the survey data, we are able to construct a variety of measures that capture *perceived* investor mispricing. We also employ disaggregated data that contain the daily trades of institutional investors. Finally, using weather and institutional holdings data, we construct a novel, stock-level measure of investor mood to assess its impact on individual stock prices. Our ability to observe the locations of investors allows us to precisely measure the weather conditions that they experience at each point in time.

Our findings can be summarized as follows. First, using the survey data, we show that deseasonalized cloud cover increases the likelihood of perceived overpricing in both individual stocks and the Dow Jones Industrial Average (DJIA) among institutional investors. Second, using the trade data, we find evidence consistent with the survey results, showing that institutional investors with lower exposure to deseasonalized cloud cover exhibit a greater propensity to buy. The impact of weather on beliefs and trading behavior is significant both statistically and economically. In particular, a one-standard-deviation increase in deseasonalized cloud cover increases the likelihood of perceived overpricing by approximately 3%, which represents more than 13% of the sample mean. Further, differences in trade imbalances between investors in the top and bottom

---

2 These studies show that individuals misattribute their affective states for information, resulting in mood-congruent judgments. Supportive evidence using cloud cover as a mood priming device has been found in a variety of contexts, including tipping (Cunningham 1979), life satisfaction (Schwarz and Clore 1983), and responsiveness to persuasion (Clore, Schwarz, and Conway 1994).

3 For example, Coval and Shumway (2005) document trading patterns that are consistent with loss aversion among traders in the Chicago Board of Trade. Frazzini (2006) finds evidence of the disposition effect among U.S. mutual fund managers, whereas Barber et al. (2007), using international data, show consistent evidence in a broader set of investors. More closely related to our study, Goetzmann and Zhu (2005) present evidence suggesting that New York City weather patterns affect New York Stock Exchange market makers, while showing little effect on retail investors.

10th percentile in deseasonalized cloud cover represent more than 9% of its total sample variation.

Third, we assess how our stock-level proxy for investor mood relates to individual stock returns. Because U.S. stock ownership is generally dispersed across investors in different locations, tests using cloud cover in a single location to proxy for mood may be biased due to poor identification. Our proxy for investor mood uses institutional investor holdings data to incorporate information about weather conditions faced by different investors in the same stock. We find evidence that the stock-level mood proxy has a strong effect on the daily returns of stocks associated with higher arbitrage costs. Finally, we document return comovement attributable to the investor mood proxy and find that these patterns are generally short lived.

These empirical findings contribute to the growing theoretical literature that identifies the channels through which mood affects economic and financial market outcomes.[4] Studies that focus on the weather effect are motivated by findings in the psychology literature that shows that affective states can generate mood-congruent biases in risk-based decisions.[5] Accordingly, investor behavior may be subconsciously influenced by factors affecting mood unrelated to market information and may in turn impact financial market outcomes (Mehra and Sah 2002; Bodoh-Creed 2013; Kamstra et al. 2014).[6] The salience of cloud cover as a mood-priming device allows us to both verify the impact of mood on beliefs and trading behavior of institutional investors and trace price effects directly attributable to institutional investors using our stock-level mood proxies.

Recent experimental studies also examine how mood-inducing cues related to weather affect risk preferences and financial decisions of individuals and document evidence consistent with mood congruency in assessments. For example, Bassi, Colacito, and Fulghieri (2013) provide experimental evidence suggesting that cloud cover has a strong impact on measures of risk tolerance. Similarly, Kramer and Weber (2012) use experimental evidence to establish a link between seasonal affective disorder (SAD) and risk attitudes. Whereas these studies use subjects from the general population, professional investors

[4] Edmans, Garcia, and Norli (2007) and Agarwal, Duchin, and Sosyura (2012) examine sporting and singing competition outcomes. Dougal et al. (2012) and Garcia (2013) examine the impact of sentiment reflected in financial news articles on stock market activity.

[5] Schwartz (1990), Clore and Parrott (1991), Wilson and Schooler (1991), and Clore, Schwarz, and Conway (1994) discuss the general role of mood and emotion in decision making. Loewenstein (2000) and Loewenstein et al. (2001) discuss the role of emotion in judgment.

[6] Mehra and Sah (2002) demonstrate that even small perturbations in preference parameters related to mood can have a measurable impact on stock returns and volatility. Bodoh-Creed (2013) shows that informed market participants can be more susceptible to mood-related biases. Motivated by the empirical regularities linking seasonal forms of depression to equity and Treasury returns, Kamstra et al. (2014) construct a model that incorporates seasonal dependence in both risk aversion and the intertemporal elasticity of substitution, showing that it can match features of the data.

may not necessarily be subject to the same dynamics.[7] We offer direct examination of how deseasonalized cloud cover impacts institutional investor perceptions using survey data and provide evidence that investor optimism, as measured by lower levels of deseasonalized cloud cover, decreases the likelihood of perceived investor overpricing.

Other previous studies examine the effect of weather on investor trading behavior. Loughran and Schultz (2004) and Goetzmann and Zhu (2005) use disaggregated data to identify the impact of weather conditions across different geographies. Loughran and Schultz (2004) show cloud cover around the firm's location has little impact on stock trading volume attributable to mood, except in cases of extreme weather conditions. Goetzmann and Zhu (2005) match retail investor trade data to cloud cover in five major U.S. cities from 1991 to 1996. They find the impact of cloud cover on investor propensities to buy and sell is statistically insignificant. However, they find a strong relation between bid-ask spreads of NYSE stocks and NYSE index returns with New York City cloud cover. They interpret these findings as evidence of a weather effect on market makers.[8]

To our knowledge, our study is the first to directly test for the weather effect in institutional investor trading behavior. We provide evidence that investor optimism increases buy-sell trade imbalances. Further, these results complement our findings on investor perceptions of mispricing.

We provide evidence that investor optimism increases daily returns of individual stocks associated with higher arbitrage costs. Binding short-sale constraints may allow for mispricing to persist (Nagel 2005), as these stocks are more costly to arbitrage. We acknowledge that our estimates may be understated due to measurement error in the investor mood proxy. Specifically, we cannot observe the daily holdings of institutional investors, as we rely on the most recent quarterly snapshots for construction. As a result, the mood proxy may not incorporate all investors trading in a stock between the reporting dates. However, we also find evidence of return comovement attributable to investor mood and find that these effects do not persist beyond one month.

Taken as a whole, our results provide strong evidence on how mood influences how institutional investors form their beliefs, affecting their trading decisions. Together with evidence from Goetzmann and Zhu (2005), our findings suggest that the weather effect is detectable among sophisticated market participants. More importantly, we demonstrate how investor mood can influence various aspects of the price formation process.

---

[7] For example, see Camerer et al. (1997) for violations of the law of supply on wage data for New York taxi drivers.

[8] Linnainmaa and Rosu (2009) show similar evidence using weather conditions around the location of the Helsinki Stock Exchange. In contrast with Goetzmann and Zhu (2005), their evidence cannot be explained by the role of market makers, suggesting that their results are driven by investors.

## 1. Theoretical Motivation

A number of studies document a robust relationship between stock index returns and local weather patterns, referred to here as the “weather effect.” Saunders (1993) shows that sky cloud cover over New York City has a strong, negative association with New York Stock Exchange (NYSE) index returns. Because his findings appear to relate to the local weather conditions around the stock exchange, the results are plausibly attributable to factors related to investor mood rather than to direct effects on firm fundamentals. Hirshleifer and Shumway (2003) provide additional evidence using an international sample of twenty-six stock exchanges. They find a similar relation between cloud cover in the locations of the stock exchanges and the stock market index returns, indicating that the relation between weather and stock market returns is pervasive and not a result of unintended data mining.

Despite strong evidence of the weather effect on stock index returns, establishing plausibility in mood-based explanations relies in part on distinguishing which group of investors drives the weather effect. Given the pervasiveness of these effects, the influence of mood should be detectable at the investor level. In particular, by employing localized proxies for mood of investors located in different regions, mood-based explanations of the weather effect imply that we should observe similar patterns between mood and investor trading behavior.

The use of sunshine, or sky cloud cover, as a proxy for mood is appealing, as it is difficult to predict accurately and its influence on mood has been well documented in the social psychology literature. A broad assertion in that literature is that individuals can misattribute affective states for information, which, in turn, affects judgment (Cunningham 1979; Schwarz and Clore 1983). Several studies use sunshine as a priming device to induce positive affective states and provide evidence that it corresponds with mood-congruent choices in cases unrelated to clinical depression.[9] The evidence has been documented in a number of settings, including tipping (Cunningham 1979; Rind 1996), life satisfaction (Schwarz and Clore 1983), and responsiveness to persuasion (Clore, Schwarz, and Conway 1994).

Prior research also supports the idea that mood can have an intertemporal effect on decision making. That is, judgment may not only be influenced by immediate affective states but also by a composite of affective states

[9] The effect of sunshine as an antidepressant has been well documented, though its effect on clinically depressed individuals is thought to influence individual decision making in different ways (Wyer, Clore, and Isbell 1999). Light therapy, which simulates sunshine, has been shown to have antidepressant benefits in individuals suffering from seasonal (Rosenthal et al. 1984) and nonseasonal (Kripke 1998) forms of clinical depression. However, these effects are distinct from mood attribution. In their seminal study, Schwarz and Clore (1983) find that individuals generally respond more positively when asked about life satisfaction on sunnier days. Their key finding is that the effect weakens when attention is drawn to the weather or their current mood before the start of the interview. They find similar evidence in a different experiment within the same study that uses a nonweather priming device. On the other hand, clinical depression is arguably a condition that is generally recognized by the patient.

experienced over time. Affective states can influence how individuals encode information into memory; when information is subsequently recalled, evaluation may be biased congruently according to the composite of prior moods (Fishbein 1963).[10]

Additionally, the psychology literature provides evidence that emotion can strongly influence responses when an individual is asked to extrapolate future events (Johnson and Tversky 1983), as well as when asked to make assessments on abstract objects in the presence of limited information (Clore, Schwarz, and Conway 1994).[11] Professional investors are likely to use information collected over time to make financial decisions about a particular investment. Therefore, the composite of affective states over time should be more informative about how investor mood affects belief formation and trading behavior than the investor's affective state at a single point in time.

Turning to the clinical literature, studies that examine the impact of simulated sunlight on seasonal and nonseasonal depression suggest that repeated treatments over time, ranging from one to five weeks, are required to produce antidepressant benefits.[12] In their seminal study on seasonal depression, Rosenthal et al. (1984) show that light treatment simulating sunshine during clear sky days has stronger antidepressant benefits than does treatment simulating sunshine in fully overcast days. When subjects were withdrawn from light therapy, they showed that the subject conditions generally reversed after one week. Similar evidence is in clinical studies on nonseasonal depression (Goel et al. 2005).

The influence of cloud cover on mood also may be related to seasonal variation in the number of daylight hours through its regulating role on sunlight exposure, which is a well-documented factor responsible for symptoms associated with SAD, as well as milder forms of seasonal depression (e.g., "winter blues"). Kamstra, Kramer, and Levi (2003) examine the influence of seasonal forms of depression on stock market activity. They find strong evidence that the number of daylight hours has a positive impact on stock index returns using a broad set of stock exchanges in different locations spanning the Southern and Northern hemispheres. Their results also support the role of mood on stock market activity, though such effects are more likely to be related to disruptions

[10] Fishbein (1963) argues that the composite effect of impressions formed on various components of the object over time influences judgments. Adaval (2001) provides supportive, experimental evidence, demonstrating that affective reactions to distinguishable aspects of a product together influence purchasing decisions, even when those reactions may be unrelated to the object of assessment. Additionally, Kida, Smith, and Maletta (1998) provide similar experimental evidence using a subject pool of experienced managers. The subjects are given information at an earlier period and are asked to describe that information at a later point in time. They show that subjects are better able to recall affective reactions at the initial exposure than the actual information presented.

[11] Johnson and Tversky (1983) show that individuals attribute greater likelihood of particular events congruently to mood, even when the event is unrelated to the cause of their mood. Clore, Schwarz, and Conway (1994) show that the influence of affective states on judgment is pronounced in assessments of objects that are abstract or when faced with limited information.

[12] For examples, see Goel et al. (2005) for nonseasonal depression and Eastman et al. (1998) for seasonal depression.

in biorhythm due to changes in seasons.[13] Nevertheless, because light therapy has been shown to be effective in treating both seasonal and nonseasonal forms of depression, the influence of cloud cover on mood may not be completely a seasonal phenomenon.

Our empirical design is motivated by these previous experimental and clinical studies. Specifically, we use cloud cover to prime investor mood and assess its impact on investor perceptions and trading behavior. Given that investors plausibly rely on complex sets of information collected over time for their decisions, we consider the intertemporal effects of cloud cover over several weeks, rather than on a single day. Additionally, to distinguish factors related to seasonal variation in cloud cover, we use deseasonalized cloud cover as the main priming instrument.

## 2. Data and Summary Statistics

### 2.1 Main data sources

Our empirical analysis relies on data from several sources. First, institutional investor survey data are collected from the Investor Behavior Project at Yale University. Since 1989, questionnaire survey data have been collected on the perceptions of investors in the United States about stock market investment.[14] Approximately 100 professional investors are surveyed per month from January 2005 to February 2007.[15] The response-level data include the date when the respondent completes the survey as well as his or her location. To ensure data consistency, we only use responses with values to the survey questions used in the analysis. Altogether, there are 1,543 responses in the sample before they are matched to the weather data described below.

Second, the weather data are collected from the Integrated Surface Database (ISD) and are publicly available from the National Oceanic and Atmospheric Administration Web site (www.ncdc.noaa.gov). The ISD database contains hourly weather observations from over 20,000 weather stations worldwide, of which 11,000 are currently active. For each weather station, we are able to observe location and weather conditions, including sky cloud cover. We collect hourly value of the weather variables from all U.S. weather stations from January 1998 to December 2010.

Third, the institutional daily trading data are provided by ANcerno Ltd. (formerly the Abel Noser Corporation). The sample period is from January

---

13 A similar mechanism can be found related to daylight savings due to sleep deprivation. Kamstra, Kramer, and Levi (2000) document anomalous stock index returns patterns around daylight saving weekends in the United States, the United Kingdom, and Germany.

14 A detailed overview of the survey is described by Shiller, Kon-Ya and Tsutsui (1996) and Shiller (2000). Additional information can be found at the International Center for Finance web site (icf.som.yale.edu) under "Stock Market Confidence Indices."

15 Institutional investor respondents are randomly sampled from a directory of institutional investors found in the investment managers section of the "Money Market Directory of Pension Funds and Their Investment Managers." The survey is also conducted in the Japanese and Chinese markets.

1999 to December 2010. ANcerno is a widely recognized consulting firm that monitors equity trading costs of institutional investors, such as CalPERS, Putman Investments, and Lazard Asset Management.[16] The client manager code along with the institutional client code allows for identification of the investor. Additional fields used in the analysis include the stock historical CUSIP number, trade date, trade direction, quantity of shares traded, and trade execution price.

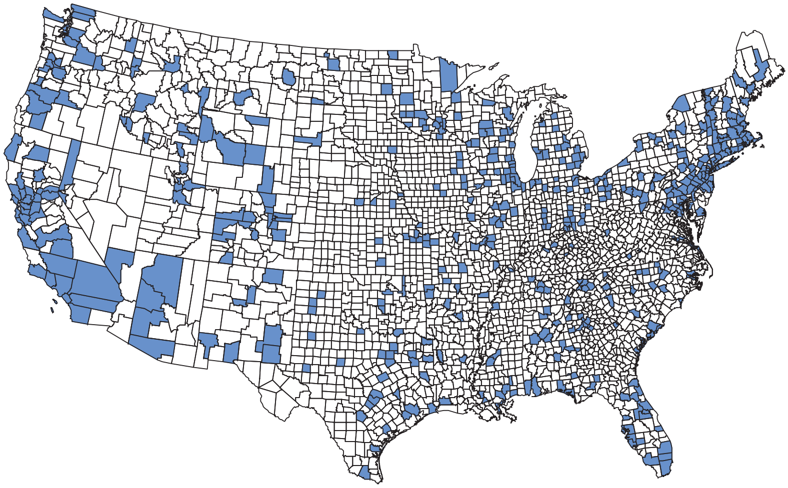
There are two unique features of the ANcerno data that are central to the trade-related tests. First, the ANcerno dataset allows us to observe the true direction (i.e., buy/sell) for all executed trades, eliminating the need to rely on the Lee and Ready (1991) algorithm. Second, the ANcerno dataset provides the identities of the institutions, allowing us to hand-collect the ZIP codes of their locations. Our analysis uses only institutions located in United States, which represents approximately 80% of all trades in the database.

Fourth, we obtain the 13(f) institutional holdings data from Thompson Reuters for the 1999:Q1 to 2010:Q4 sample period. The data provides quarterly snapshots of investor portfolio positions. The ANcerno database provides trading information for only a subset of the investors in the 13(f) data. We are able to hand-collect the ZIP codes of the institution's headquarters using the *Nelson's Directory of Investment Managers*.

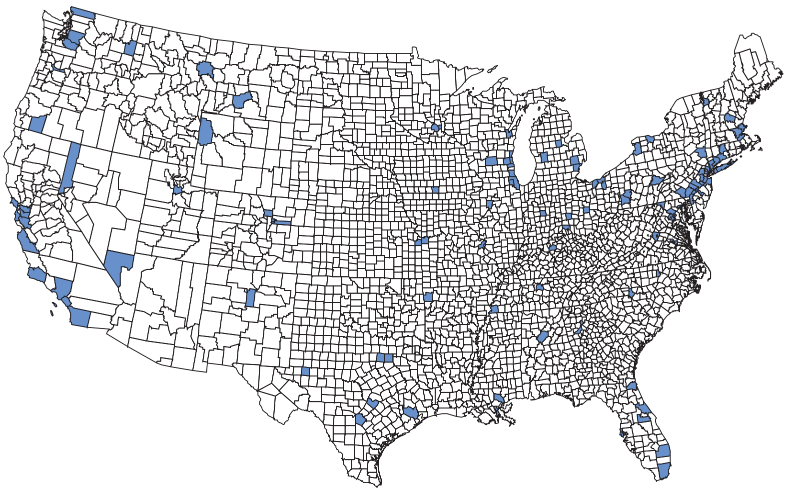
Other data sources used in the analysis are as follows. Stock characteristics are collected from the Center for Research in Security Prices (CRSP). Only common stocks (share code of 10 or 11) from January 1999 to January 2010 are included in the analysis. The county-level estimates of median household income and population are obtained from the Bureau of Economic Analysis (BEA).

Figure 1 presents the geographical distribution of institutional investors represented in the survey (panel A), trade (panel B), and holdings (panel C) datasets at the county level. Shaded counties correspond with regions in which at least one investor is represented. All three panels convey considerable geographical heterogeneity in each of the datasets. Panel A shows that the survey respondents are geographically well represented and are more heavily represented in regions with greater population. Panel B shows a similar pattern within the trade dataset, though this geographical distribution is relatively sparse relative to the survey data. This pattern is not surprising, given the limited number of investors in the trade database. Panel C shows that the holdings data have greater heterogeneity than do the trade data, which is again expected because investors in the trade data represent only a subset of investors in the holding data.

16 See Puckett and Yan (2011) for a detailed description of the ANcerno data.

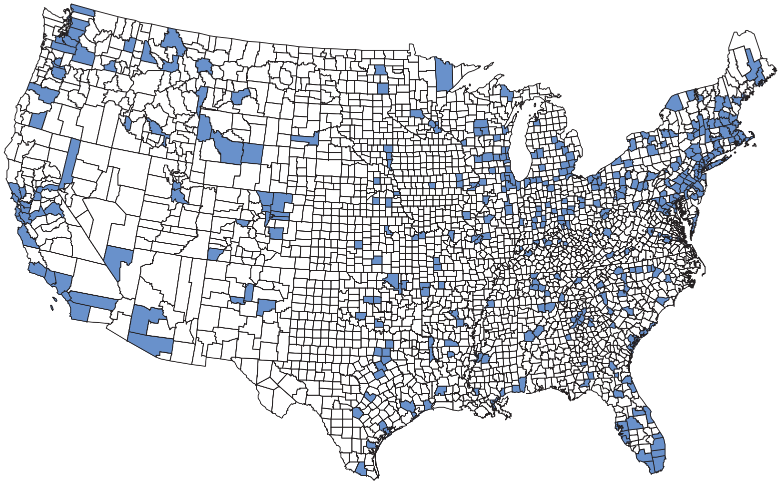


Panel A: Survey data



Panel B: Trade data

**Figure 1**
**Geographical dispersion of the survey and trade datasets**
This figure presents the geographical distribution of the survey in the United States (panel A), institutional investor trade (panel B), and institutional investor holding (panel C) data on the county level. Black areas denote the existence of at least one investor in the corresponding dataset at the county level.



Panel C: Holdings data

**Figure 1**
**Continued**

## 2.2 Variable construction

**2.2.1 Deseasonalized cloud cover measures.** The primary variable of interest from the weather data is the hourly sky cloud cover, which takes on integer values from zero (sky clear) to eight (full cloud cover). Similar data are used by Hirshleifer and Shumway (2003) and Goetzmann and Zhu (2005). The survey and trade data are merged with the weather data using criteria based on geographical distance. Specifically, using the location coordinates of each station and investor, we calculate their distances based on the haversine distance formula.[17] The average daily sky cloud cover is calculated using hourly values from 6 a.m. to 12 p.m., for each date and weather station, the time period during which investors are most likely to observe outdoor weather conditions. For each ZIP code, the average daily sky cloud cover is calculated using values from all weather stations within a 50-kilometer radius of the investor's ZIP code centroid.[18]

Because mood may have an intertemporal effect on individuals, we compute rolling averages of the cloud measure in each ZIP code based on information from $x$ days before to the response or trade date. Because the average amount of

---

[17] The haversine formula calculates the distance between locations *1* and *2* as $d_{1,2}=2\times R\times$ arcsin(min(1, $\sqrt{A}$, where $R$ is the earth's radius (approximately 6,371 kilometers), $A=\sin^2(\Delta lat/2)+\cos(lat_1)+\cos(lat_2)\times\sin^2(\Delta lon/2)$. In this expression, $\Delta lat = lat_2 - lat_1$ and $\Delta lon - lon_2 - lon_1$, where $lat_1$ and ($lat_2$ and $lon_2$) is the latitutde and longitude of location *1* (*2*), respectively.

[18] We also examine alternative distance thresholds. The results remain similar or strengthen in terms of statistical significance when using a 30-kilometer threshold, though slightly weaken when increasing the threshold to 100 kilometers.

sunlight is a decreasing, convex function in sky cloud cover, the sky cloud cover measure used in the analysis (*SKC*) is defined as the natural logarithm of one plus the rolling average of the ZIP code-level sky cloud cover.[19] Additionally, we require that each institutional investor has at least one matched weather station included in the analysis. We note that measurement error in observed cloud cover is likely to bias our tests from finding any effect.

To adjust for seasonality in the *SKC* measure, we consider two approaches. First, we calculate seasonal cloud cover as the average daily cloud cover for the same month over the entire sample period. *Seasonal SKC* is the natural logarithm of one plus the seasonal value and can be used to control for seasonality when assessing the impact of *SKC* in the tests. Second, deseasonalized *SKC*, or *DSKC*, is defined as the difference between *SKC* and *Seasonal SKC*.[20] Including *DSKC* measure in place of *SKC* and *Seasonal SKC* in an *OLS* regression model is equivalent to constraining the model coefficient on *Seasonal SKC* to be equal to the negative of the *SKC* coefficient. Because this representation may be too strong, we estimate the model using both approaches.

**2.2.2 Investor perceived mispricing measures.** We begin the empirical analysis by focusing on responses from the survey related to perceived mispricing. To directly assess whether investors believe that stocks are overpriced (or underpriced) relative to fundamentals, we examine the association between the weather measures and responses to the question, "Stock prices in the United States, when compared with measures of true fundamental value or sensible investment value, are: (a) Too low, (b) too high, (c) just right, and (d) do not know." An indicator variable, *TooHigh* is constructed, which takes the value of one to a response of "too high" and zero otherwise. Similarly, *TooLow* takes the value of one to a response of "too low" and zero otherwise.

Next, we construct a continuous measure of perceived mispricing using reported values of what the respondent believes to be the intrinsic level of the DJIA.[21] The responses provide estimates of the investors' perceptions on the intrinsic value of the DJIA and can be compared with the actual DJIA level around the survey date. A measure can be constructed representing the percentage mispricing, or *%DJIAMisPrc*, defined as the natural logarithm of the ratio between the survey response and the average DJIA level over the

[19] The logarithmic transformation is motivated by the non-linearity in the relationship between illuminance and sky cloud cover. However, similar results are obtained without the transformation.

[20] The procedure to seasonally adjust the cloud cover measure is similar to that of Goetzmann and Zhu (2005) and Hirshleifer and Shumway (2003). We find similar results when using the average cloud cover in the same month in the previous year for the seasonal adjustment.

[21] Specifically, the question asks, "What do you think would be a sensible level for the Dow Jones Industrial Average based on your assessment of U.S. corporate strength (fundamental)?"

past seven days. Lower values of *%DJIAMisPrc* are associated with relatively greater overpricing.[22]

In the sample of investors examined, we conjecture that pessimism should be positively related to perceived overpricing, as most of the survey respondents are more likely to be limited in their ability to short stocks. However, there are several key limitations affecting inferences with the *TooHigh* and *TooLow* variables. Tests on those measures must assume that the respondents apply similar criteria to assess mispricing. However, this assumption is likely to be too strong. To address this issue, we construct similar binary variables from the *%DJIAMisPrc* measure. Specifically, we code *DJIATooHigh* as a value of one if *%DJIAMisPrc* is below the bottom sample quartile. Similarly, *DJIATooLow* is coded with a value one if *%DJIAMisPrc* is in the top sample quartile.

The *DJIATooHigh* and *DJIATooLow* measures are based on a stock index rather than on individual stocks, making it more likely that the respondents are basing their responses on their perceptions of market trends rather than in the context of their own portfolio holdings. We assess these five measures of perceived mispricing to validate whether cloud cover is positively associated with bad mood states.

**2.2.3 Investor trade measures.** We construct two measures based on the institutional trade data. For both measures, we aggregate trades across all investors for a particular stock within the same ZIP code. First, we construct a measure of investor-level buy-sell imbalance, or *Investor BSI Ratio*, that aggregates all buy and sell trades across stocks and is defined as the difference between the daily dollar buy and sell volume across all investors in the dataset in a particular ZIP code, scaled by the total dollar volume in the same date and ZIP code. When there is no trade on a particular date, *Investor BSI Ratio* takes a value of zero. Second, we construct an analogous measure of daily buy-sell imbalance by investor and stock, or *StockInvestor BSI Ratio*. The measure is defined as the daily, buy minus sell dollar volume scaled by the total dollar volume for the same date, stock, and ZIP code. Because we cannot observe daily positions, we only use dates where investors trade in the stock.

We aggregate trades at the ZIP code level given that the trade and weather data are matched using investor ZIP code. Idiosyncratic trading behavior related to other factors may serve to increase noise in the BSI measures unrelated to geographical factors. To alleviate these issues, the analysis restricts the sample to only ZIP codes with at least three investors at each point in time and with at least half of the sample period available for the trade data. In addition, we remove extreme observations associated with ZIP codes in the top and bottom 1% of the total dollar buy-sell volume over the entire sample.

[22] The *%DJIAMisPrc variable* is expected to be negatively related to *TooHigh*, and indeed a paired $t$-test on the difference between the group means of *%DJIAMisPrc* across values of *TooHigh* is statistically significant ($t$-value = 28.81).

**2.2.4 Stock-level mood proxies.** We construct a stock-level measure of investor mood, or *StockDSKC*, measured approximately as the average *DSKC* across locations of institutional investors based on their portfolio holdings. Institutional investors in stock $i$ are identified using the most recent 13(f) holdings data. The weather data are linked to the ZIP code of the institution's location. *StockDSKC* is calculated as the logarithm of one plus the difference between the average *SKC* and *Seasonal SKC* of all institutional investors in stock $i$ on date $t$. *StockDSKC* is updated daily, using the most recent holdings data.[23]

**2.2.5 Seasonal affective disorder proxy.** In addition to the seasonally adjusted cloud cover measure, we also condition the tests on seasonal factors that may also affect mood. Our proxy for seasonal affective disorder, or *SAD*, corresponds with the number of nighttime hours, which can be approximated using a mathematical representation that takes into account the time of year and the locational coordinate. *SAD* equals zero during spring and summer months and the number of nighttime hours minus twelve during the other months. We use the coordinate centroids of each investor's ZIP code location to calculate the *SAD* measure, using the procedure described by Kamstra, Kramer, and Levi (2003).[24] We note that the measure contains significantly less variation in our sample than that of Kamstra, Kramer, and Levi (2003), given that our sample focuses on the contiguous United States. On the other hand, they use an international data sample. Additionally, the length of the sample period in our data sources is also significantly smaller.

## 2.3 Summary statistics

Table 1 provides the summary statistics for the weather, survey, and trade variables used in the analysis. Panel A describes the ZIP code level deseasonalized sky cloud cover measures across different estimation windows. The estimation window of $x$ days calculates average *DSKC* using data from days $t-x$ to $t$. Because there is little theory to help guide the selection of the estimation window used in the analysis, the comparison helps in pinning down an appropriate length that yields well-behaved estimates.

As the estimation window increases, the sample average decreases in absolute magnitude. The sample standard deviation also decreases monotonically, following a convex pattern in the length of the estimation window. In particular, the sample standard deviation of the measure that uses a one-day window, which is most similar to the weather measure used by Hirshleifer and Shumway (2003), is 0.695. The sample standard deviation decreases considerably up to

[23] Using weighted averages based on investor position size does not qualitatively alter the results.

[24] Kamstra, Kramer, and Levi (2003) also include a fall dummy variable in their tests to distinguish the impact of SAD during the fall and winter. Because our tests are focused on isolating nonseasonal effects, we instead include time fixed effects in some of the specifications.

**Table 1**
**Summary statistics on survey, trade, and holdings data**

| Variable | Mean | SD | 25th percentile | Median | 75th percentile |
|---|---|---|---|---|---|
| Panel A: DSKC by estimation window (survey data) | | | | | |
| 1 day | −0.204 | 0.695 | −0.755 | −0.086 | 0.376 |
| 3 days | −0.085 | 0.444 | −0.362 | −0.032 | 0.234 |
| 1 week | −0.067 | 0.346 | −0.267 | −0.044 | 0.173 |
| 2 weeks | −0.049 | 0.263 | −0.190 | −0.041 | 0.124 |
| 4 weeks | −0.045 | 0.223 | −0.183 | −0.039 | 0.097 |
| Panel B: Survey data sample | | | | | |
| TooHigh | 0.195 | 0.397 | | | |
| TooLow | 0.148 | 0.355 | | | |
| DJIATooHigh | 0.252 | 0.435 | | | |
| DJIATooLow | 0.253 | 0.435 | | | |
| %DJIAMisPrc | −0.004 | 0.121 | −0.031 | 0.014 | 0.060 |
| Daily DJIA Volatility (pts) | 0.620 | 0.153 | 0.486 | 0.610 | 0.698 |
| ln(Portfolio Size) | 14.541 | 5.136 | 12.899 | 14.732 | 17.910 |
| Earnings Growth | 0.071 | 0.087 | 0.040 | 0.060 | 0.080 |
| ln(Population) | 8.534 | 1.298 | 7.801 | 8.615 | 9.321 |
| ln(Income) | 6.022 | 0.284 | 5.823 | 5.984 | 6.178 |
| SAD | 0.812 | 0.997 | 0.000 | 0.158 | 1.544 |
| Panel C: Trade data sample | | | | | |
| DSKC | −0.029 | 0.229 | −0.161 | −0.010 | 0.126 |
| Investor BSI ratio | 0.027 | 0.502 | −0.270 | 0.000 | 0.334 |
| Stock-Investor BSI ratio | 0.000 | 0.591 | −0.253 | 0.000 | 0.282 |
| Panel D: Holdings data sample | | | | | |
| StockDSKC | −0.014 | 0.125 | −0.095 | −0.012 | 0.070 |
| %IO | 0.459 | 0.239 | 0.267 | 0.483 | 0.649 |
| PosRet | 0.473 | 0.499 | | | |

This table reports summary statistics for the average sky-cloud coverage variables that are matched to the survey dataset (panel A), other characteristics linked to or from the survey dataset (panel B), variables from and linked to the ZIP code-level trade dataset (panel C), and variables from and linked to the ZIP code-level holding and stock returns dataset (panel D). Please refer to Section 3 for descriptions of each dataset. SKC is the natural logarithm of one plus the daily average sky-cloud coverage over 1 day to 4 weeks before the response or trade date. DSKC is the difference between *SKC* and the seasonal value for*SKC*, calculated as the average cloud cover for the same month over the entire sample. *Portfolio size* is the sum of the capital across asset classes reported by the investor respondent. Population is the county-level population for the response year. *Income* is the county-level median income for the response year. *Earnings Growth* is the percentage long-term growth rate estimate of the respondent. Adjusted nighttime hours per day (*SAD*) is calculated using the procedure outlined by Kamstra, Kramer, and Levi (2003), based upon the investors coordinate locations. *%DJIA MisPrc* is the natural logarithm of the ratio of the respondent estimated, intrinsic level of the DJIA to the average, actual DJIA level over the previous week of the response date. *TooHigh*(*TooLow*) takes the value one if the respondent answers that actual prices in the stock market are too high (too low) relative to fundamentals. *DJIA TooHigh* (*DJIA TooLow*) takes the value one if the *%DJIA MisPrc* for a respondent is above the 75th percentile (below the 25th percentile) of the overall sample. *Investor BSI* is the daily total, net buy minus sell dollar volume, scaled by the total dollar trading volume within the same ZIP code. *StockInvestor BSI* is the daily total, buy minus sell dollar volume for the same ZIP code and stock, scaled by the total dollar trading volume within the same stock. *StockDSKC* is the difference between the average SKC of institutional investors holding the same stock, less its seasonal value. %IO is the proportion of shares in the stock held by institutional investors. *PosRet*is an indicator variable that takes the value one if the stock experiences positive returns for that date and zero otherwise.

the two-week window (0.263) and is relatively stable when using a four-week window (0.223).

Panel B describes the primary variables used in the survey-based tests. The binary response variables are displayed in the first four rows. In our sample, slightly more investors indicate that stock prices are generally too high

(19.5%) than too low (14.8%) relative to fundamentals. By construction, the *DJIATooHigh* and *DJIATooLow* variables have a sample mean of approximately 25%. The volatility of daily DJIA returns over the past 30 days is represented in points.

Panel C provides descriptions on the trade sample. The *DSKC* measure is constructed over the trade sample and has a lower sample mean and standard deviation than those calculated over the survey sample. The *Investor* and *StockInvestor BSI* ratios are slightly positive in the sample period.

Finally, panel D provides statistical descriptions on the holdings sample. The changes in stock-level *DSKC* measures have a lower sample mean than either the survey or trade data samples. The sample standard deviations slightly half of the size less than that of the *DSKC* in the trade sample, which in part is due to averaging *DSKC* in stocks with a larger number of investors.

## 3. Weather-Induced Mood and Perceived Mispricing

The social psychology literature provides evidence that individuals provide mood-congruent assessments on objects unrelated to the cause of their affective states (Cunningham 1979; Schwarz and Clore 1983). In particular, Schwarz and Clore (1983) find that sunny days are associated with relatively more optimistic assessments. Recent experimental evidence also shows that weather conditions can influence risk attitudes through its impact on mood (Bassi, Colacito, and Fulghieri 2013).

In this section, we analyze the survey data where institutional investors are asked a series of questions related to their opinions on stock market investment. Our tests focus on perceived investor mispricing. Investors may view stocks as mispriced for a variety of reasons. While risk preferences across professional investors may be relatively homogenous, variation in weather patterns may generate mood-induced biases in investor beliefs about underlying fundamentals. In turn, its impact on beliefs may also affect their trading decisions, which we examine on trade data in subsequent tests. Because the survey does not draw attention to weather conditions during the interview, we view the analysis as a field experiment where deseasonalized weather conditions are randomly assigned to subjects.

### 3.1 Univariate estimates

Table 2 presents the Pearson correlation coefficients amongst the variables used in the regression models.[25] Panel A presents the results for the dependent variables and the conditioning variables. Panel B presents the results for the dependent variables and deseasonalized *SKC* measured over different estimation windows, defined similarly to panel A of Table 1. The correlation coefficients are accompanied with $p-$values, and are displayed in parentheses.

[25] The results are similar using nonparametric tests, such as rank-order correlations.

**Table 2**
**Pearson correlation matrix of survey and weather variables**

Panel A: Survey responses with conditioning variables

| | TooHigh | TooLow | %DJIA MisPrc | DJIA TooHigh | DJIA TooLow |
|---|---|---|---|---|---|
| TooLow | −20.55% | | | | |
| | (0.000) | | | | |
| %DJIAMisPrc | −58.55% | 27.21% | | | |
| | (0.000) | (0.000) | | | |
| DJIATooHigh | 67.80% | −19.93% | −72.42% | | |
| | (0.000) | (0.000) | (0.000) | | |
| DJIATooLow | −25.22% | 41.96% | 53.71% | −33.84% | |
| | (0.000) | (0.000) | (0.000) | (0.000) | |
| DJIA volatility | −0.42% | 4.25% | 6.65% | −0.65% | 11.00% |
| | (0.872) | (0.099) | (0.010) | (0.801) | (0.000) |
| Portfolio size | 2.23% | 0.65% | 3.06% | −0.28% | 2.37% |
| | (0.388) | (0.802) | (0.235) | (0.912) | (0.359) |
| Growth rate | −9.43% | 3.55% | 3.54% | −3.96% | 4.63% |
| | (0.000) | (0.169) | (0.170) | (0.124) | (0.073) |
| Population | 3.62% | 2.01% | −1.87% | 2.85% | −2.30% |
| | (0.160) | (0.437) | (0.470) | (0.269) | (0.372) |
| Median income | 1.20% | −2.01% | −2.03% | 0.43% | −4.80% |
| | (0.643) | (0.435) | (0.432) | (0.866) | (0.063) |
| Change in population | −0.41% | 0.82% | 0.38% | 0.03% | 1.65% |
| | (0.874) | (0.752) | (0.883) | (0.990) | (0.522) |
| Change in median income | −0.95% | 0.65% | −1.41% | −3.14% | −4.52% |
| | (0.712) | (0.801) | (0.585) | (0.223) | (0.079) |
| SAD | −0.43% | −3.88% | −2.20% | −1.32% | −3.50% |
| | (0.869) | (0.132) | (0.394) | (0.609) | (0.175) |

Panel B: Survey responses with deseasonalized SKC (DSKC)

| | DSKC estimation window | | | | |
|---|---|---|---|---|---|
| | 1 day | 3 days | 1 week | 2 weeks | 1 month |
| TooHigh | 2.25% | 2.60% | 1.70% | 4.74% | 4.95% |
| | (0.382) | (0.314) | (0.510) | (0.066) | (0.055) |
| TooLow | 0.73% | 2.06% | 2.59% | 2.04% | 2.00% |
| | (0.778) | (0.424) | (0.316) | (0.430) | (0.438) |
| %DJIA MisPrc | −1.88% | −2.81% | −2.98% | −5.93% | −5.22% |
| | (0.466) | (0.276) | (0.248) | (0.022) | (0.043) |
| DJIA TooHigh | 2.35% | 3.14% | 4.28% | 7.46% | 6.83% |
| | (0.362) | (0.223) | (0.097) | (0.004) | (0.008) |
| DJIA TooLow | 1.59% | −2.23% | −3.16% | −3.91% | −3.56% |
| | (0.537) | (0.387) | (0.221) | (0.130) | (0.168) |

The table presents the Pearson correlation coefficients between the mispricing measures and the conditioning variables in the regression models of Table 3 (panel A) and between the mispricing measures and the deseasonalized cloud cover (*DSKC*) measures across the estimation windows (panel B). The mispricing measures are *TooHigh*, *TooLow*, *%DJIA MisPrc*, *DJIA TooHigh*, and *DJIA TooLow*. The estimation windows for the *DSKC* measures are 1 day, 3 days, 1 week, 2 weeks, and 1 month. $p-$values on the correlation coefficients are reported in parentheses.

In panel A of Table 2, the correlation coefficients amongst the survey variables related to mispricing suggest internal consistency in the responses. The correlation between *TooHigh* and *DJIATooHigh* is 67.8% ($p$-value <0.001), and both are strongly, negatively associated with the *TooLow*, *DJIATooLow*, and *%DJIAMisPric* measures, respectively.

We next examine how other variables correlate with the mispricing measures. The volatility of DJIA returns has a positive and statistically significant correlation with measures related to perceived underpricing, and the correlation coefficient is higher with the *DJIATooLow* measure. The long-term corporate earnings growth rate estimates are signed consistently with received wisdom, though only the *DJIATooLow* correlation is statistically significant. Portfolio size does not have a statistically significant correlation coefficient with any of the mispricing measures. Neither levels nor changes in the county-level characteristics appear to be consistently correlated with the mispricing measures.

In panel B, the correlation coefficients between the mispricing measures and the *DSKC* are not statistically significant when using estimation windows of less than one week. When expanding the estimation windows to two weeks and beyond, the correlation coefficients become statistically significant for *TooHigh*, *%DJIAMisPrc* and *DJIATooHigh*. Both the *TooHigh* and *DJIATooHigh* coefficients are positive, while the *%DJIAMisPrc* coefficient is negative, as expected. Additionally, the *DJIATooHigh* coefficient is considerably larger than that of *TooHigh*. Finally, the correlation coefficients that use *TooLow* and *DJIATooLow* remain statistically insignificant, however, and only *DJIATooLow* has the predicted sign. Again, we cannot observe the criteria investors use to assess what is “too” high or low in the *TooHigh* and *TooLow* measures, respectively. In contrast, the *DJIATooLow* measure uses consistent criteria that we define, and so may explain differences in the signs between the two coefficients.

These results raise some interesting questions. First, the relation between the coefficients and the length of the *DSKC* estimation window may be due to measurement error. The respondent may take more than one day to fill out the survey, so that the relevant estimation window for the *DSKC* measure may be larger than just the day the survey was completed. Using this interpretation, the results would suggest that the respondent may have taken more than two weeks to ponder the survey questions, which does not appear to be plausible. On the other hand, investors may rely on information that has already been collected over time to determine their responses. The *DSKC* measure using longer estimation windows may better correspond with the composite affective reaction for the periods over which the relevant information is collected. Additionally, the results are also consistent with findings from clinical studies on light therapy and nonseasonal depression, showing that antidepressant benefits of simulated sunlight take hold only after repeated exposure over long treatment periods.

Second, while the correlation coefficients are large and statistically significant for the overpricing measures, the coefficients are statistically insignificant on the underpricing measures. An alternative interpretation of the findings on the *TooHigh* measure is that deseasonalized sunshine is positively related to responses of “too low” *or* “just right”. Bad moods may induce individuals to examine information with greater scrutiny (Schwartz 1990; Petty, Gleicher, and Baker 1991), so that good moods may not necessarily result in bullish responses over non-critical ones. The lack of statistically significant results associated with the *TooLow* measure is consistent with this interpretation. A similar pattern holds for the *DJIATooHigh* and *DJIATooLow* results.

These findings suggest that longer estimation windows used to construct the cloud cover measures may be important to capture mood effect on the mispricing measures. For the remainder of the analysis, we include only *DSKC* measured over a two-week estimation window given our findings from Table 2, and are consistent with the clinical studies discussed in Section 1. We next determine the robustness of the univariate results by conditioning on other factors that may also impact the survey responses.

### 3.2 Perceived mispricing regression specification

The OLS regression models of perceived mispricing are specified as follows:

$$\mathrm{Y}_{\mathrm{i,t}} = \mathrm{b}_0 + \mathrm{b}_1^{*}\mathrm{DSKC}_{\mathrm{i,t}} + \mathbf{b} * \mathbf{X}_{\mathbf{i,t}} + \varepsilon_{\mathrm{i,t}}$$

The dependent variables in the regression models are *TooHigh*, *TooLow, %DJIAMisPrc, DJIATooHigh, and DJIATooLow*. $DSKC_{\mathrm{i,t}}$ is constructed as a 14-day rolling average using information up to date $t$ of zip code-level cloud cover. $\mathbf{X}_{\mathrm{i,t}}$ represents a vector of other explanatory variables for respondent $i$ at date $t$.

We include the following conditioning variables in the models. First, proxies for local economic conditions are included, such as county-level population, changes in population, county-level median income, and changes in median income. This choice is motivated by the evidence in Coval and Moskowitz (1999) document local bias in investment choices of institutional investors. Economic conditions may affect the future prospects of local investment opportunities, which in turn may influence investor opinions on stock market investment.

Second, professional investors with greater assets under management may have systematically different opinions about equity market investment, and so the natural logarithm of one plus the size of the respondent's investment portfolio is included in the models. Third, large, recent fluctuations in stock prices may also influence responses, as periods of greater equity market volatility may negatively bias the responses. Accordingly, the natural logarithm of the sample volatility of the DJIA index returns over the past 30 days is included as a conditioning variable. Fourth, the respondent's estimate of the long-term corporate earnings growth is also included, as it may systematically influence the mispricing measures. Finally, we also include the *SAD* measure used in Kamstra, Kramer, and Levi (2003) to further distinguish from seasonal effects on mood.

For the binary dependent variables, the test coefficients may be influenced by the functional form of the estimator. To assess its effect, we also estimate the models using a probit estimator, which takes the form:

$$P(Y_{i,t}|Z_{i,t}) = \Phi(Z_{i,t})$$

Here, *Y* represents the binary variables associated with perceived mispricing, and $Z_{i,t} = c_0 + c_1 * DSKC_{i,t} + \mathbf{c} * \mathbf{X_{i,t}}$. As with the OLS specifications, the residuals across time and region are unlikely to be uncorrelated, biasing downward the standard error estimates. To address these issues, only two-way clustered standard errors on the zip code- and date-levels are reported.[26]

### 3.3 Perceived mispricing regression estimates

Table 3 presents the estimates. Panel A presents the OLS model estimates from all the mispricing models. Panel B presents the probit model estimates using only the binary dependent variables. Only marginal effects are reported in panel B to facilitate comparison with the OLS estimates.

Panel A shows that the OLS coefficients on *DSKC* are consistent with the univariate results, even after inclusion of the conditioning variables. The *DSKC* coefficient in model 1 is positive and statistically significant (estimate = 0.073, $t$-value = 1.87). In other words, higher deseasonalized cloud cover increases the likelihood that investors believe that stock prices are "too high" relative to fundamentals. When *DSKC* is replaced by *SKC* and *Seasonal SKC* in model 2, similar results obtain: the *SKC* coefficient is 0.087 ($t$-value = 2.25), while the *Seasonal SKC* coefficient is statistically insignificant (estimate = −0.039, $t$-value = −0.73).

In contrast, the *DSKC* coefficient in the *TooLow* model is not statistically significant (estimate = 0.029, $t$-value = 0.89). The *DSKC* coefficient in the *%DJIAMisPrc* model is negative (estimate = −0.025, $t$-value = −2.14), and shows that deseasonalized cloud cover decreases the continuous measure of perceived underpricing. However, most of the explanatory power of *DSKC* appears to be driven by responses associated with overpricing. The *DSKC* coefficient in the *DJIATooHigh* model is 0.120 ($t$-value = 3.04), while that of the *DJIATooLow* model is −0.045 ($t$-value = −0.84). The coefficients on the conditioning variables are statistically insignificant for the most part, and their signs are consistent with the univariate results from panel A of Table 2.

We next repeat the analysis using probit estimators. The results are reported in panel B, and are quite similar. The marginal effect of *DSKC* in the *TooHigh* model is slightly larger (estimate = 0.079, $z$-value = 1.94), and the result is similar in the *DJIATooHigh* model (estimate = 0.125, $z$-value = 3.07). To assess the economic magnitudes, a one-standard-deviation increase in *DSKC* translates to an increase of 0.021 and 0.033 in *TooHigh* and *DJIATooHigh* and represents 10.6% and 13.1% of the sample means, respectively.

[26] The standard error adjustments are performed using Jonah Gelbach and Douglas Miller's STATA code (cgmreg.ado), which is based on the procedures described by Cameron, Gelbach, and Miller (2011). Our estimates are identical for the two-way clustering cases in which we use Mitchell Petersen's STATA code (cluster2.ado). This procedure is described by Petersen (2009).

**Table 3**
**OLS regression models on survey responses**

Panel A: OLS regressions

| | Dependent variable | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | TooHigh | | TooLow | | %DJIA MisPrc | | DJIA TooHigh | | DJIA TooLow | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| DSKC | 0.073* | | 0.029 | | −0.025** | | 0.120*** | | −0.045 | |
| | (0.039) | | (0.033) | | (0.012) | | (0.040) | | (0.053) | |
| SKC | | 0.087** | | 0.035 | | −0.031** | | 0.139*** | | −0.046 |
| | | (0.039) | | (0.037) | | (0.012) | | (0.040) | | (0.055) |
| Seasonal SKC | | −0.039 | | −0.014 | | 0.008 | | −0.075 | | 0.043 |
| | | (0.053) | | (0.039) | | (0.013) | | (0.054) | | (0.062) |
| DJIA Vol | 0.002 | −0.003 | 0.055 | 0.053 | 0.031 | 0.034* | −0.011 | −0.018 | 0.204*** | 0.205*** |
| | (0.056) | (0.055) | (0.052) | (0.052) | (0.020) | (0.020) | (0.065) | (0.065) | (0.074) | (0.075) |
| Portfolio size | 0.001 | 0.001 | 0.000 | 0.000 | 0.001 | 0.001 | −0.001 | −0.001 | 0.002 | 0.002 |
| | (0.002) | (0.003) | (0.002) | (0.002) | (0.001) | (0.001) | (0.002) | (0.002) | (0.002) | (0.002) |
| Growth rate | −5.894*** | −5.908*** | 1.889 | 1.883 | 0.696 | 0.703 | −2.731 | −2.749 | 3.123* | 3.124* |
| | (1.441) | (1.453) | (1.412) | (1.415) | (0.807) | (0.811) | (1.837) | (1.854) | (1.875) | (1.876) |
| Population | 0.011 | 0.010 | 0.008 | 0.008 | −0.001 | −0.001 | 0.010 | 0.009 | −0.002 | −0.001 |
| | (0.009) | (0.010) | (0.008) | (0.008) | (0.003) | (0.003) | (0.009) | (0.009) | (0.010) | (0.009) |
| Income | −0.004 | −0.014 | −0.050 | −0.054 | −0.005 | 0.000 | 0.010 | −0.002 | −0.049 | −0.048 |
| | (0.047) | (0.048) | (0.038) | (0.039) | (0.015) | (0.015) | (0.053) | (0.053) | (0.042) | (0.044) |
| Population change | −0.121 | 0.001 | 0.222 | 0.274 | −0.007 | −0.065 | −0.279 | −0.117 | 0.012 | 0.006 |
| | (0.393) | (0.385) | (0.336) | (0.356) | (0.123) | (0.132) | (0.441) | (0.434) | (0.395) | (0.414) |
| Income change | −0.270 | −0.174 | 0.334 | 0.375 | −0.020 | −0.065 | −0.736 | −0.610 | −0.377 | −0.382 |
| | (0.514) | (0.520) | (0.403) | (0.413) | (0.161) | (0.170) | (0.568) | (0.569) | (0.496) | (0.521) |
| SAD | 0.003 | −0.001 | −0.006 | −0.008 | −0.000 | 0.002 | −0.003 | −0.009 | 0.007 | 0.007 |
| | (0.012) | (0.013) | (0.011) | (0.012) | (0.004) | (0.004) | (0.014) | (0.016) | (0.020) | (0.021) |
| N | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 |
| Adjusted $R^2$ | 1.374% | 1.463% | 0.629% | 0.649% | 1.074% | 1.290% | 1.020% | 1.149% | 1.917% | 1.917% |

(*continued*)

**Table 3**
**Continued**

Panel B: Probit regressions

| | Dependent variable | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | TooHigh | | TooLow | | TooHigh DJIA | | TooLow DJIA | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| DSKC | 0.079* | | 0.031 | | 0.125*** | | −0.042 | |
| | (0.041) | | (0.034) | | (0.041) | | (0.052) | |
| SKC | | 0.094** | | 0.037 | | 0.146*** | | −0.042 |
| | | (0.041) | | (0.039) | | (0.042) | | (0.054) |
| Seasonal SKC | | −0.046 | | −0.017 | | −0.078 | | 0.043 |
| | | (0.054) | | (0.039) | | (0.055) | | (0.061) |
| DJIA vol | 0.003 | −0.001 | 0.054 | 0.052 | −0.013 | −0.020 | 0.203*** | 0.202*** |
| | (0.056) | (0.055) | (0.050) | (0.050) | (0.066) | (0.065) | (0.073) | (0.073) |
| Portfolio size | 0.001 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.002 | 0.002 |
| | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Growth rate | −10.622** | −10.518** | 1.790 | 1.781 | −2.903 | −2.877 | 2.868* | 2.867* |
| | (5.150) | (5.142) | (1.146) | (1.149) | (2.108) | (2.110) | (1.613) | (1.614) |
| Population | 0.011 | 0.010 | 0.008 | 0.007 | 0.011 | 0.009 | −0.001 | −0.001 |
| | (0.009) | (0.010) | (0.008) | (0.008) | (0.009) | (0.009) | (0.009) | (0.009) |
| Income | −0.009 | −0.017 | −0.050 | −0.054 | 0.010 | 0.000 | −0.047 | −0.047 |
| | (0.048) | (0.049) | (0.039) | (0.040) | (0.053) | (0.053) | (0.044) | (0.045) |
| Population change | −0.071 | 0.069 | 0.219 | 0.279 | −0.199 | 0.021 | 0.174 | 0.179 |
| | (0.426) | (0.428) | (0.361) | (0.394) | (0.540) | (0.591) | (0.619) | (0.642) |
| Income change | −0.207 | −0.124 | 0.329 | 0.371 | −0.757 | −0.649 | −0.469 | −0.467 |
| | (0.527) | (0.536) | (0.411) | (0.421) | (0.580) | (0.590) | (0.530) | (0.549) |
| SAD | 0.004 | −0.000 | −0.007 | −0.008 | −0.003 | −0.009 | 0.006 | 0.006 |
| | (0.012) | (0.013) | (0.012) | (0.012) | (0.015) | (0.016) | (0.020) | (0.021) |
| N | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 |
| Pseudo $R^2$ | 1.767% | 1.860% | 0.643% | 0.654% | 0.873% | 0.979% | 1.636% | 1.638% |

The table reports the coefficients using OLS regression model (panel A)

$$Y_{i,t}=b_0+b_1 DSKC_{i,t}+b_2 DJIAVol_t+b_3 PortfolioSize_{i.t}+b_4 GrowthRate_{i,t}+b_5 Population_{i,t}+b_6 PopulationChange_{i,t}+b_7 Income_{i,t}+b_8 IncomeChange_{i,t}+b_9 SAD_{i,t}+\varepsilon_{i,t},$$

and using probit regression model (panel B)

$$\mathrm{P}(\mathrm{Y}_{i,t}|\mathrm{Z}_{i,t})=\Phi(\mathrm{Z}_{i,t}),$$

where $\mathrm{Z}_{i,t}=b_0+b_1 DSKC_{i,t}+b_2 DJIAVol_t+b_3 PortfolioSize_{i.t}+b_4 GrowthRate_{i,t}+b_5 Population_{i,t}+b_6 PopulationChange_{i,t}+b_7 Income_{i,t}+b_8 IncomeChange_{i,t}+b_9 SAD_{i,t}.$

Odd number models replace DSKC with SKC and Seasonal SKC. Panel B reports marginal effects from the probit models. The dependent variables (Y) in the models are reported in the top of each column and include a binary variable associated with whether individual stocks are perceived to be overpriced (*TooHigh*), a binary variable associated with whether individual stocks are perceived to be underpriced (*TooLow*), the percentage difference between the perceived DJIA level relative to a 7-day trailing average of the actual DJIA level (*%DJIA MisPrc*), a binary variable associated with whether *%DJIA MisPrc* is in the bottom sample quartile (*DJIA TooHigh*), and a binary variable associated with whether *%DJIA MisPrc* is in the top sample quartile (*DJIA TooLow*). Only binary dependent variable models are estimated in panel B. Robust standard errors clustered on ZIP code and date levels are reported in parentheses below. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

### 3.4 Perceived mispricing regression estimates: Robustness checks

One concern with the estimates in Table 3 is that they are susceptible to omitted variable biases that may be driven by seasonal or interregional factors. While the *DSKC* measure is constructed to purge seasonal variation, the seasonal adjustments may be insufficient so that time effects may overstate the effect of *DSKC*. Interregional factors invariant of time may also overstate the effect of *DSKC*, so that the results may be driven by a limited number of regions.

To assess the impact of these potential biases, we re-estimate the models from Table 3 after including state (panel A) or year-quarter (panel B) fixed effects, which should decrease the statistical power of the tests. Because of sample size limitations, using regional and time variables with greater granularity is infeasible. Finally, because of the inclusion of these fixed effects, we only estimate the models using OLS estimators, given the well-known, incidental parameters issues associated with fixed effect probit estimators.

In panel A, we show that the results from the state fixed effects models are similar to those in Table 3. All models include the conditioning variables from the models in Table 4, though the coefficients are not reported to conserve space. In the *TooHigh* model, the *DSKC* coefficient remains positive but becomes statistically insignificant at the 10% level (estimate = 0.067, $t$-value = 1.64). However, the *SKC* coefficient remains positive and statistically significant (estimate = 0.105; $t$-value = 2.54). The *DSKC* coefficients remain statistically significant in the *%DJIAMisPrc* (estimate = $-0.021$; $t$-value = $-1.85$) and *DJIATooHigh* (estimate = 0.112; $t$-value = 2.61) models. The *DSKC* coefficients are statistically insignificant for the *TooLow* and *DJIATooLow* models. The results for the *SKC* coefficients are similar. Altogether, the estimates are slightly smaller in absolute magnitude than those in Table 3, though most remain statistically significant.

Similar results obtain for the year-quarter fixed effects models in panel B. In the *TooHigh* model, the *DSKC* coefficient is positive and statistically significant (estimate = 0.066; $t$-value = 1.69), as is the *SKC* coefficient (estimate = 0.068; $t$-value = 1.77). The *DSKC* coefficient in the *%DJIAMisPrc* model remains negative but statistically insignificant at the 10% level (value = $-0.017$; $t$-value = $-1.54$), though the *SKC* coefficient is statistically significant (estimate = $-0.020$; $t$-value = $-1.73$). The *DSKC* coefficient in the *DJIATooHigh* model is positive and statistically significant (estimate = 0.093; $t$-value = 2.65). As with the state fixed effects models, the *DSKC* coefficients are statistically insignificant for the *TooLow* and *DJIATooLow* models.

Finally, we consider interaction terms in the perceived mispricing models related to *DSKC* and *SAD*. Because cloud cover is also negatively related to sunlight exposure, which, in the clinical literature, has been documented to be associated with seasonal affective disorder, it is plausible that cloud cover may amplify the effects of daylight on individual mood. We find that *SAD* coefficients across all the models we estimate are statistically insignificant at

**Table 4**
**Fixed effects regression estimates using the survey data**

| | Dependent variable | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | TooHigh | | TooLow | | %DJIA MisPrc | | DJIA TooHigh | | DJIA TooLow | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Panel A: State fixed effects | | | | | | | | | | |
| DSKC | 0.067 | | 0.017 | | −0.021* | | 0.112*** | | −0.026 | |
| | (0.041) | | (0.035) | | (0.011) | | (0.043) | | (0.049) | |
| SKC | | 0.105** | | 0.012 | | −0.033*** | | 0.148*** | | −0.033 |
| | | (0.042) | | (0.041) | | (0.013) | | (0.046) | | (0.053) |
| Seasonal SKC | | 0.010 | | −0.028 | | −0.002 | | −0.040 | | 0.012 |
| | | (0.055) | | (0.043) | | (0.012) | | (0.057) | | (0.060) |
| Conditioning variables | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| N | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 |
| Adjusted $R^2$ | 1.035% | 1.360% | −0.527% | −0.588% | 0.034% | 0.362% | −0.362% | −0.142% | 1.803% | 1.747% |
| Panel B: Year-quarter fixed effects | | | | | | | | | | |
| DSKC | 0.066* | | 0.038 | | −0.017 | | 0.093*** | | 0.007 | |
| | (0.039) | | (0.032) | | (0.011) | | (0.035) | | (0.049) | |
| SKC | | 0.068* | | 0.054 | | −0.020* | | 0.098*** | | 0.010 |
| | | (0.039) | | (0.036) | | (0.011) | | (0.035) | | (0.051) |
| Seasonal SKC | | −0.060 | | 0.001 | | 0.009 | | −0.082 | | 0.001 |
| | | (0.053) | | (0.037) | | (0.013) | | (0.050) | | (0.058) |
| Conditioning variables | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| N | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 | 1,505 |
| Adjusted $R^2$ | 1.822% | 1.758% | 0.496% | 0.559% | 2.298% | 2.277% | 2.604% | 2.546% | 4.880% | 4.819% |

The table reports the coefficients of the regression model:

$$Y_{i,t} = b_0 + b_1 DSKC_{i,t} + b_2 DJIAVol_t + b_3 PortfolioSize_{i,t} + b_4 GrowthRate_{i,t} + b_5 Population_{i,t} + b_6 PopulationChange_{i,t} + b_7 Income_{i,t} + b_8 IncomeChange_{i,t} + b_9 SAD_{i,t} + \varsigma_i + \tau_t + \varepsilon_{i,t}.$$

The dependent variables, $Y$, include *TooHigh*, *TooLow*, *%DJIA MisPrc*, *DJIA TooHigh*, and *DJIA TooLow*. Models 6 through 10 of each panel replace *DSKC* with *SKC* and *Seasonal SKC*. The regression models in panel A include state fixed effects, and in Panel B, they include year-quarter fixed effects. The estimates on the conditioning variables are not shown to conserve space. Robust standard errors clustered on the ZIP code and date levels are reported in parentheses below. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

the 10% level, and we posit that the lack of variation in nighttime hours in our data sample may be the explanation.

To evaluate these views, we include *DSKC* and *SAD* interaction terms in the OLS models from Table 3, and the results are displayed in Table A.1 in the Online Appendix. The *DSKC* coefficient remains statistically significant for the *TooHigh*, *%DJIA MisPrc*, and *DJIA TooHigh* models. In these models, the interaction term coefficients are of consistent sign as our conjecture, though only the one in the *TooHigh* model is statistically significant at the 10% level.[27] We conclude, as before, that the tests may not have sufficient statistical power given the lack of variation in the number of nighttime hours in our sample and leave further tests to future research.

## 4. Weather-Induced Mood and Institutional Investor Trading

The results from the previous section show that *DSKC* has a strong impact on perceived mispricing among institutional investors and imply that *DSKC* may also impact trading behavior. We evaluate this channel by constructing tests using daily institutional investor trades from the ANcerno database. Positive (negative) trade imbalances relate to institutional investors that are net buyers (sellers) for an observation date. The daily trade data are aggregated to the ZIP code level and are used to construct the *Investor* and *StockInvestor BSI* measures. *DSKC* is constructed identically to the procedure described in the survey-based tests.

### 4.1 Mood and trading behavior: Univariate results

We begin by using paired $t$-tests to assess whether *Investor BSI* is systematically higher on sunny relative to cloudy regions across time. We define sunny (cloudy) regions based on whether *DSKC* for a particular investor is in the top (bottom) $x$th percentile of the sample for each date. The paired $t$-tests assess the statistical significance of the difference in group means of *Investor BSI* across sunny and cloudy regions conditional upon $x$. We estimate cases in which $x$ takes on values of 50, 33, 25, and 10, respectively. For example, when $x = 10$, the paired $t$-test assesses the difference in group means of *Investor BSI* across investors in the top and bottom 10th percentile. In this manner, we are able to verify whether the difference in the group means increases using relatively more extreme *DSKC* values.

Table 5 displays the results. The thresholds used to define the investor groups are displayed in the first column. The average *Investor BSI* for the cloudy and sunny groups is displayed in Columns 2 and 3, respectively. The difference in the group means is displayed in the fourth column. The differences are accompanied by their standard errors and are displayed in parentheses.

[27] The results are similar with the inclusion of a dummy variable corresponding with fall months.

**Table 5**
**Univariate analysis of high-low *DSKC* on investor BSI**

| | High DSKC | Low DSKC | High DSKC - Low DSKC | |
|---|---|---|---|---|
| 50th percentile split | 0.021 | 0.034 | −0.013*** | (0.003) |
| 33rd percentile split | 0.018 | 0.037 | −0.018*** | (0.004) |
| 25th percentile split | 0.016 | 0.036 | −0.020*** | (0.005) |
| 10th percentile split | 0.013 | 0.059 | −0.046*** | (0.008) |

Group sample averages of the *Investor BSI* are reported for high and low *DSKC* investors. *Investor BSI* is defined as the total dollar buy less sell volumes, scaled by the total trading volume for ZIP-code $i$ and date $t$. Investors are ranked based on *DSKC* for each date. High (low) DSKC investor classification is based on whether the *DSKC* ranking is in the top (bottom) 50th, 33rd, 25th, or 10th percentile of that date. Standard errors on the difference in the group means are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

When defining the groups by the top (bottom) 50th percentile on *DSKC* for each sample date, the difference in *Investor BSI* is negative and statistically significant (estimate = −0.013; $t$-value = −3.71), as expected. Tightening the threshold serves to increase the group differences and appears to be primarily driven by increasing values of average *Investor BSI* in the sunny groups. For cloudy (sunny) groups based on the top (bottom) 10th percentile on *DSKC*, the difference in the group means is −0.046 ($t$-value = −5.48), which is approximately 9.2% of the sample standard deviation for *Investor BSI*. Altogether, these univariate tests show that *DSKC* has a negative impact on investor propensities to buy, as predicted. Additionally, the effect becomes pronounced when tightening the thresholds for the *DSKC* groupings.

## 4.2 Multivariate BSI regression specification

Next, we assess the robustness of the univariate results by using panel regression models to condition variation related to cross-sectional and time-series determinants of *Investor BSI*. The model estimated over the full sample is specified as follows:

$$\text{Investor BSI}_{z,t} = d_0 + d_1 * \text{DSKC}_{z,[t-14,t]} + \mathbf{d} * \mathbf{W}_{\mathbf{z},\mathbf{t}} + \varepsilon_{z,t}.$$

For ZIP code $z$ and date $t$, $\mathbf{W}_{\mathbf{z},\mathbf{t}}$represents a vector of variables that allows us to isolate the source of the explanatory power of *DSKC* in the models. The times-series variables include dummy variables for whether the sample date is in January or on a Monday, as done by Goetzmann and Zhu (2005), and the *SAD* measure, which is constructed using the ZIP code of the trading institution's location. As with the survey-based tests, we also control for county-level economic variables based on the institution's location. Finally, we also exploit within-ZIP code and within-date variation in the *DSKC* variable through inclusion of fixed effects as robustness checks after presenting the results on the baseline models. Table 6 displays the estimates from these models.

In the results reported in Table 7, the dataset is disaggregated further into a three-level panel, where the data is represented on the ZIP code ($z$), date ($t$), and stock ($j$) level. The dependent variable of those models is *StockInvestor BSI*. Additionally, unlike the *Investor BSI* measure, because the dataset does

**Table 6**
**Regression analysis on daily ZIP code-level investor BSI**

| | Dependent variable: Investor BSI ratio | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| DSKC | −0.024** | | −0.021** | | −0.031** | |
| | (0.011) | | (0.010) | | (0.014) | |
| SKC | | −0.023** | | −0.020** | | −0.031** |
| | | (0.010) | | (0.010) | | (0.014) |
| Seasonal SKC | | −0.029 | | −0.059** | | 0.017 |
| | | (0.057) | | (0.026) | | (0.070) |
| Monday | −0.014** | −0.014** | −0.014** | −0.014** | | |
| | (0.006) | (0.006) | (0.006) | (0.006) | | |
| January | 0.008 | 0.009 | 0.007 | 0.009 | | |
| | (0.012) | (0.012) | (0.012) | (0.012) | | |
| SAD | −0.003 | −0.003 | −0.002 | −0.001 | −0.091** | −0.085* |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.042) | (0.051) |
| Population | 0.001 | 0.001 | 0.508 | 0.507 | −0.004 | −0.003 |
| | (0.017) | (0.017) | (0.400) | (0.399) | (0.016) | (0.017) |
| Income | −0.005 | −0.003 | −0.039 | −0.039 | −0.005 | −0.004 |
| | (0.023) | (0.023) | (0.052) | (0.052) | (0.023) | (0.023) |
| Population change | −0.764 | −0.971 | 0.370 | 0.377 | −1.995* | −2.045* |
| | (0.962) | (0.915) | (0.824) | (0.823) | (1.137) | (1.073) |
| Income change | −0.009 | −0.016 | 0.060 | 0.059 | −0.174 | −0.174 |
| | (0.089) | (0.087) | (0.093) | (0.093) | (0.170) | (0.169) |
| N | 88,246 | 88,246 | 88,246 | 88,246 | 88,246 | 88,246 |
| Adjusted $R^2$ | 0.034% | 0.050% | 1.523% | 1.546% | 1.703% | 1.703% |
| ZIP code FEs | | | yes | yes | | |
| Date FEs | | | | | yes | yes |

The table reports panel regression estimates using a daily ZIP code-level institutional buy-sell imbalance ratio (*Investor BSI*) as the dependent variable:

$$InvestorBSI_{z,t} = d_0 + d_1 DSKC_{z,t} + d_2 Monday_t + d_3 January_t + d_4 Population_{z,t} + d_5 PopulationChange_{z,t}$$
$$+ d_6 Income_{z,t} + d_7 IncomeChange_{z,t} + d_8 SAD_{z,t} + \xi_z + \eta_t + \varepsilon_{z,t}.$$

Investor trades are aggregated to the date-ZIP code level. The explanatory variables of interest are the natural logarithm of one plus the average SKC over the past 14 days (*SKC*), the natural logarithm of one plus the average SKC over the same month over the entire sample period (*Seasonal SKC*), and the deseasonalized SKC defined as the difference between *SKC* and *Seasonal SKC* (*DSKC*). Though not displayed, some of the models include additional explanatory variables, as indicated. *Monday* takes the value of one if the observation date is a Monday and zero otherwise. *January* takes the value of one if the observation month is January and zero otherwise. County economic condition variables are as follows. *Population* is the logarithm of the county-level population. *PopulationChange* is the change over the previous year. *Income* is the logarithm of the county-level median household income. *IncomeChange* is the change in *Income* over the previous year. Adjusted nighttime hours per day (*SAD*) is calculated using the construction outlined by Kamstra, Kramer, and Levi (2003), based upon the investors coordinate locations. Models 3 and 4 include ZIP code fixed effects ($\xi$), and models 5 and 6 include date fixed effects ($\eta$). Coefficients on the fixed effects are not reported for viewing ease. Robust standard errors clustered on date and ZIP code levels are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

not reveal the investor's daily positions, the tests are conditional on whether the investor trades (e.g., either buys or sells) in the stock for a particular date.

The advantage of the three-level panel is that we can now account for stock-level factors. Specifically, we include in the regression models individual stock characteristics related to size, which is measured as the natural logarithm of the stock's market capitalization, and liquidity, which is measured as the inverse of the stock's share price. We also examine another specification that includes fixed effects for each stock-date pair, which is expected to significantly

**Table 7**
**Regression analysis on daily ZIP code-stock-level investor BSI**

| | Dependent variable: StockInvestor BSI | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| DSKC | −0.008* | | −0.007** | | −0.010** | | −0.007*** | |
| | (0.005) | | (0.003) | | (0.004) | | (0.003) | |
| SKC | | −0.008* | | −0.007** | | −0.009** | | −0.006** |
| | | (0.005) | | (0.003) | | (0.004) | | (0.003) |
| Seasonal SKC | | 0.014 | | 0.013 | | 0.056** | | 0.035*** |
| | | (0.028) | | (0.013) | | (0.028) | | (0.013) |
| Conditioning variables | | | | | YES | YES | YES | YES |
| N | 13,324,507 | 13,324,507 | 13,324,507 | 13,324,507 | 13,324,491 | 13,324,491 | 13,324,491 | 13,324,491 |
| Adjusted $R^2$ | 0.001% | 0.001% | 52.530% | 52.530% | 0.283% | 0.289% | 52.535% | 52.536% |
| Date and stock FEs | | | yes | yes | | | yes | yes |

This table reports OLS regression model estimates using the daily stock/investor-level BSI ratio (*StockInvestor BSI*) as the dependent variable:

$$StockInvestor\,BSI_{z,j,t} = f_0 + f_1 DSKC_{z,t} + f_2 Monday_t + f_3 January_t + f_4 Population_{z,t} + f_5 PopulationChange_{z,t} + f_6 Income_{z,t} + f_7 IncomeChange_{z,t}$$

$$+ f_8 SAD_{z,t} + f_9 MktCap_{j,t} + f_{10} StockLiq_{j,t} + \varphi_{j,t} + \varepsilon_{z,j,t}.$$

Investor trades are aggregated to the stock-date-ZIP code level. Though not displayed, some of the models include additional explanatory variables, as indicated. Models (1) through (4) include fixed effects based on date groups. Models (5) through (6) include fixed effects based on date-stock grouping pairs ($\phi$). Robust standard errors clustered on the stock, date, and ZIP code levels are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

decrease the statistical power of the tests. The fixed effects models limit the variation in *DSKC* to within-stock-date groupings, purging the estimators of any unobservable factors associated with a particular stock at each point in time.

The fixed effects model described above is specified as follows:

$$\text{StockInvestor BSI}_{z,j,t} = f_0 + f_1 * \text{DSKC}_{z,t} + \mathbf{f_2 * CountyCharacteristcs_{z,t}}$$
$$+ f_3 * \text{SAD}_{z,t} + \sum_j \sum_t f_{j,t} * D(\text{Stock} = j \bigcap \text{Date} = t) + \varepsilon_{z,j,t}.$$

The fixed effects model does not include a number of variables that are subsumed by the fixed effects terms, such as the stock characteristics and the time-series variables. Because the residuals in the three-level panel model are unlikely to be independent across the panel dimensions, the reported standard errors in the tables are adjusted for three-way clustering at the ZIP code, date, and stock levels.

### 4.3 Investor trading model estimates

The analysis proceeds by estimating the *Investor BSI* models. As shown in Table 6, models 1 and 2 represent the estimates without the fixed effects. Models 3 and 4 include ZIP code fixed effects. Models 5 and 6 include date fixed effects. The odd-numbered models provide estimates on *DSKC*, whereas the even-numbered models show estimates on *SKC* controlling for *Seasonal SKC*.

In model 1, the *DSKC* coefficient is negative and statistically significant (estimate = −0.024; $t$-value = −2.25), and the *SKC* coefficient in model 2 is also negative and statistically significant (estimate = −0.023; $t$-value = −2.25). Among the conditioning variables, only the *Monday* dummy variable is negative and statistically significant. When including the ZIP code fixed effects in models 3 and 4, the results are similar. The *DSKC* coefficient remains stable (estimate = −0.021; $t$-value = −2.03) relative to the estimates in model 1. Finally, similar results obtain when we include date fixed effects, though the *DSKC* coefficient is slightly higher (estimate = −0.031; $t$-value = −2.28). The *Monday* and *January* dummy variables are dropped from the model, as they are subsumed by the date fixed effects.

Because the *DSKC* measure is constructed in an overlapping manner given the panel data structure, one possible concern is that the standard error estimates are inconsistent, due to autocorrelated residual terms. We re-estimate the standard errors for the model coefficients from Table 6 using the procedure described by Driscoll and Kraay (1998), which accounts for serial and cross-sectional dependency in the regression residuals. The results are reported in Appendix Table A.2 and are similar to those in Table 6.

Table 7 presents the estimates for the *StockInvestor BSI* models. Models 3, 4, 7, and 8 include the stock-date fixed effects, whereas models 5 through 8 also include the other conditioning variables. The *DSKC* coefficient estimate in model 1 is negative and statistically significant (value = -0.008; $t$-value = −1.79). When we include fixed effects in model 3, the adjusted $R^2$ increases dramatically, from 0.001% to 52.530%. However, the *DSKC* coefficient remains stable and statistically significant (estimate = $-0.007$; $t$-value = $-2.34$), suggesting that the explanatory power of *DSKC* on *StockInvestor BSI* is not affected by stock-date factors. When including the control variables in model 7, the *DSKC* coefficient estimate remains similar (estimate = $-0.007$; $t$-value = $-2.60$). Finally, the results are similar when we replace *DSKC* with *SKC* and *Seasonal SKC*. Overall, our trading regression results are remarkably robust even when accounting for stock related factors.

In summary, consistent with our predictions, we find that mood as measured by cloud cover affects institutional trading behavior. Institutions are less likely to engage in buying activity during relatively cloudier periods. *DSKC* negatively impacts *Investor BSI* across a wide variety of specifications. Additionally, these relationships are robust to alternative tests that take into account stock-level factors.

## 5. Investor Mood, Stock Prices, and Return Comovement

Our results from the tests in the previous sections demonstrate that weather-based proxies of mood impact institutional investor beliefs and trading behavior. A related question is whether mood also impacts stock prices through its effect on investors, which is the focus of this section. Our key finding is that stock-level mood proxies that incorporate deseasonalized cloud cover across institutional investors, or *StockDSKC*, negatively impact returns of stocks with higher arbitrage costs. We also provide evidence of return comovement attributable to the mood proxies.

### 5.1 Investor mood and stock prices: Regression specification

Because the trading database only represents a fraction of the entire universe of institutional investors, our tests in this section utilize 13(f) holdings data to construct stock-level measures of investor mood. We restrict the sample to stocks represented in the 13(f) holdings data that can be matched to the weather data. *StockDSKC* is expected to have a negative relationship with daily stock returns.

Hirshleifer and Shumway (2003) posit that weather fluctuations are more likely to impact the sign of stock returns rather than the magnitude and so we employ similar tests. In particular, we estimate an OLS regression model in which the dependent variable of interest is a binary variable, or *PosRet*, which takes value one for the positive stock return days, and zero otherwise.

$$\mathrm{PosRet}_{i,t} = g_0 + g_1 * \mathrm{StockDSKC}_{i,t} + \mathbf{g} * \mathbf{X_{i,t}} + \varepsilon_{i,t}$$

Here, $\mathbf{X_{i,t}}$ represents a set of explanatory variables that includes the natural logarithm of the market capitalization from the end of the previous quarter, the proportion of shares held by institutional investors in the stock, the inverse of the share price from the end of the previous quarter, a Monday dummy variable, a January dummy variable, and a composite *SAD* measure that is constructed similarly to *StockDSKC*. The coefficient on the *StockDSKC* variable is expected to be negative, or $g_1 < 0$.

Several econometric issues are expected to bias our tests away from our predictions. First, the explanatory power of *StockDSKC* is expected to vary according to costs related to arbitrage. Mispricing arising from trades motivated by investor mood is expected to be quickly corrected by other investors in the absence of short-sale constraints. However, when these constraints are binding, the effects of mispricing are expected to persist (Nagel 2005). We measure arbitrage costs, or *ArbCosts*, associated with binding short-sale constraints as the inverse of one plus the proportion of shares held by institutional investors in a particular stock, so that higher *ArbCosts* corresponds with stocks that are more likely to have binding short-sale constraints. We expect the weather effect to be pronounced in stocks with higher *ArbCosts* and estimate effects across subsamples based on *ArbCost* rankings accordingly.

Second, the frequency of the holdings data does not allow us to observe investor positions within each quarter. In particular, we cannot identify investors that may actively trade in a stock but hold no position at the SEC filing dates. Measurement errors related to the observation frequency are unlikely to be systematic, which could lead to attenuation bias on the coefficients in our tests. Additionally, investors with short positions are not accounted for in *StockDSKC*, so that pessimistic investors may not be represented in the measure.

### 5.2 Investor mood and stock prices: Stock return regression estimates

Table 8 displays the OLS results from the *PosRet* models. Panel A reports the estimates when we include only the weather variables in the models, whereas panel B reports the estimates when all the conditioning variables are included. The coefficient estimates for the conditioning variables are not reported to conserve space.

The first two models report the pooled model estimates. The remaining eight columns report the model estimates on sample splits based on the *ArbCosts* rankings. Models 3 and 4 are estimated for the sample in which *ArbCosts* is the lowest 50th sample percentile. Models 5 and 6 are estimated on the sample in which *ArbCosts* is between the 50th and 80th sample percentiles. Models 7 and 8 are estimated on the sample in which *ArbCosts* is between the 80th and 95th sample percentiles. Models 9 and 10 are estimated on the sample in which *ArbCosts* is between the 95th and 100th sample percentiles.

The first row displays the estimates on the *StockDSKC* variable. In panel A, the *StockDSKC* coefficient in the pooled sample is statistically insignificant (estimate = $-0.003$; $t$-value = $-0.14$) with only the weather variables. However, the results from the regression models using subsamples based on *ArbCosts* rankings suggest that the weather effect is concentrated among stocks with higher arbitrage costs. For stocks with *ArbCosts* in the 95th to 100th sample percentile range, *StockDSKC* is negative and statistically significant (estimate

**Table 8**
**Stock-level SKC and daily stock returns**

| | All | | [0%,50%] | | [50%,80%] | | [80%,95%] | | [95%,100%] | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Panel A: Weather variables only | | | | | | | | | | |
| StockDSKC | −0.003 | | 0.025 | | 0.002 | | −0.030** | | −0.029*** | |
| | (0.024) | | (0.034) | | (0.023) | | (0.015) | | (0.009) | |
| StockSKC | | −0.002 | | 0.026 | | 0.004 | | −0.029* | | −0.028*** |
| | | (0.024) | | (0.034) | | (0.023) | | (0.015) | | (0.009) |
| Seasonal StockSKC | | 0.100** | | 0.057 | | 0.098** | | 0.080*** | | −0.028 |
| | | (0.048) | | (0.073) | | (0.048) | | (0.030) | | (0.017) |
| N | 8,590,220 | 8,590,220 | 4,295,101 | 4,295,101 | 2,577,086 | 2,577,086 | 1,288,516 | 1,288,516 | 429,517 | 429,517 |
| Adjusted $R^2$ | 0.000% | 0.017% | 0.003% | 0.013% | 0.000% | 0.016% | 0.007% | 0.011% | 0.009% | 0.024% |
| Panel B: Weather variables with conditioning variables | | | | | | | | | | |
| StockDSKC | 0.004 | | 0.024 | | 0.004 | | −0.026* | | −0.022** | |
| | (0.024) | | (0.034) | | (0.023) | | (0.015) | | (0.009) | |
| StockSKC | | 0.004 | | 0.026 | | 0.004 | | −0.026* | | −0.020** |
| | | (0.024) | | (0.034) | | (0.023) | | (0.015) | | (0.009) |
| Seasonal StockSKC | | 0.053 | | 0.074 | | 0.073 | | 0.010 | | −0.061*** |
| | | (0.053) | | (0.086) | | (0.054) | | (0.032) | | (0.018) |
| Conditioning variables | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| N | 8,590,220 | 8,590,220 | 4,295,101 | 4,295,101 | 2,577,086 | 2,577,086 | 1,288,516 | 1,288,516 | 429,517 | 429,517 |
| Adjusted $R^2$ | 0.235% | 0.239% | 0.098% | 0.108% | 0.160% | 0.167% | 0.166% | 0.166% | 0.205% | 0.234% |

The dependent variable in the OLS model below is an indicator variable (*PosRet*) taking the value of one for positive return days and 0 otherwise.

$$PosRet_{i,t} = g_0 + g_1 StockDSKC_{i,t} + g_2 Monday_t + g_3 January_t + g_4 StockSAD_{i,t} + g_5 MktCap_{i,t} + g_6 StockLiq_{i,t} + \varepsilon_{i,t}$$

The top matter of the table presents the subsample definition based on the *ArbCosts* rankings. The first two models present the pooled estimates. The remaining columns present the subsample estimates based on *ArbCosts* rankings for the entire 1999–2010 sample period: [0%,50%] in models 3 and 4; (50%,80%] in models 5 and 6; (80%,95%] in models 7 and 8; and (95%,100%] in models 9 and 10. We define *ArbCosts* as the inverse of the proportion of shares held by institutional investors in the stock. Panel A presents estimates from the models that only include the weather variables, whereas panel B presents estimates from the models that also include the conditioning variables. The estimates on the conditioning variables are not reported to conserve space. Deseasonalized stock-level SKC (*StockDSKC*) is defined as the difference in *StockSKC* and *Seasonal StockSKC*. Refer to the text for further details on construction. Robust standard errors clustered on the stock and date levels are displayed in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

= −0.029; $t$-value = −3.03). The coefficients before that point are relatively higher but mostly statistically insignificant. Using *StockSKC* and *Seasonal StockSKC* instead yields similar results. When we include the conditioning variables, panel B shows that the *StockDSKC* coefficients attenuate somewhat but remain statistically significant.

The *StockDSKC* coefficient is statistically significant in stocks within the top sample quintile based on *ArbCosts*. The average proportion of shares held by institutional investors within those stocks ranges from 2% to 20%. Although we cannot construct mood proxies for stocks not represented in the 13(f) holdings data, these stocks are expected to be characterized with relatively higher arbitrage costs. Accordingly, it may be reasonable to expect a detectable mood effect in stocks with no institutional ownership.

### 5.3 Investor mood and stock prices: Robustness checks

We consider several alternative specifications to assess the robustness of the OLS results. First, using probit instead of OLS regression models, we estimate similar tests in Table 9, which is formatted similarly to Table 8. As before, only marginal effects are reported for comparability with the OLS estimates. The marginal effects are virtually identical to the OLS estimates, though the standard error estimates are slightly larger.

Second, we re-estimate the standard errors for the OLS estimates to correct for serial dependency in the residual terms using the procedure described by Driscoll and Kraay (1998). Table A.3 from the Online Appendix reports the results and shows that the Driscoll-Kraay standard errors are generally smaller than those obtained using two-level clustering.

Third, we examine whether our results are sensitive to the measure of arbitrage costs used. In untabulated estimates, we obtain similar results when we redefine *ArbCosts* using other proxies, such as market capitalization or idiosyncratic volatility, though the results are slightly weaker.

Finally, we consider whether our results can be explained in part by the effect of mood fluctuations of market makers in NYSE stocks. Goetzmann and Zhu (2005) show that New York City cloud cover negatively impacts daily NYSE index returns and most of explanatory power is concentrated on days that experience the largest change in spreads in S&P 100 stocks. In untabulated results, we find similar results when we exclude NYSE stocks. We also explore alternative specifications that include New York City cloud cover as an additional conditioning variable and find that it has little impact on our findings.

### 5.4 Investor mood and return comovement

The results from the previous section confirm that the investor mood proxy has a negative impact on daily stock returns. Using *StockDSKC* as our proxy for investor mood, we next examine whether investor mood can induce return

**Table 9**
**Probit model estimates**

| | All | | [0%,50%] | | (50%,80%] | | (80%,95%] | | (95%,100%] | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Panel A: Weather variables only | | | | | | | | | | |
| StockDSKC | −0.003 | | 0.025 | | 0.002 | | −0.030* | | −0.029** | |
| | (0.024) | | (0.034) | | (0.023) | | (0.016) | | (0.011) | |
| StockSKC | | −0.002 | | 0.026 | | 0.004 | | −0.029* | | −0.028** |
| | | (0.024) | | (0.034) | | (0.023) | | (0.016) | | (0.011) |
| Seasonal StockSKC | | 0.100** | | 0.057 | | 0.098** | | 0.080** | | −0.028 |
| | | (0.048) | | (0.074) | | (0.048) | | (0.031) | | (0.021) |
| N | 8,590,220 | 8,590,220 | 4,295,101 | 4,295,101 | 2,577,086 | 2,577,086 | 1,288,516 | 1,288,516 | 429,517 | 429,517 |
| Pseudo $R^2$ | 0.000% | 0.012% | 0.002% | 0.010% | 0.000% | 0.011% | 0.005% | 0.008% | 0.007% | 0.018% |
| Panel B: Weather variables with conditioning variables | | | | | | | | | | |
| StockDSKC | 0.004 | | 0.024 | | 0.004 | | −0.026 | | −0.022** | |
| | (0.024) | | (0.034) | | (0.023) | | (0.016) | | (0.011) | |
| StockSKC | | 0.004 | | 0.026 | | 0.004 | | −0.026 | | −0.020* |
| | | (0.024) | | (0.034) | | (0.023) | | (0.016) | | (0.011) |
| Seasonal StockSKC | | 0.053 | | 0.074 | | 0.073 | | 0.010 | | −0.061*** |
| | | (0.053) | | (0.086) | | (0.054) | | (0.033) | | (0.022) |
| Conditioning variables | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| N | 8,590,220 | 8,590,220 | 4,295,101 | 4,295,101 | 2,577,086 | 2,577,086 | 1,288,516 | 1,288,516 | 429,517 | 429,517 |
| Pseudo $R^2$ | 0.235% | 0.239% | 0.098% | 0.108% | 0.160% | 0.167% | 0.166% | 0.166% | 0.205% | 0.234% |

The dependent variable in the probit model below is an indicator variable (*PosRet*) taking the value of one for positive return days and 0 otherwise.

$$P(PosRet_{i,t}|\mathrm{Z}_{\mathrm{i,t}})=\Phi(\mathrm{Z}_{\mathrm{i,t}}),$$

$$\text{where } \mathrm{Z}_{\mathrm{i,t}}=h_0+h_1StockDSKC_{i,t}+h_2Monday_t+h_3January_t+h_4StockSAD_{i,t}+h_5MktCap_{i,t}+h_6StockLiq_{i,t}.$$

Only marginal effects are reported. The top matter of the table presents the subsample definition based upon the *ArbCosts* rankings. The first two models present the pooled estimates. The remaining columns present the subsample estimates based on *ArbCosts* rankings for the entire 1999–2010 sample period: [0%,50%] in models 3 and 4; (50%,80%] in models 5 and 6; (80%,95%] in models 7 and 8; and (95%,100%] in models 9 and 10. We define *ArbCosts* as the inverse of the proportion of shares held by institutional investors in the stock. Panel A presents estimates from the models that only include the weather variables, whereas panel B presents estimates from the models that also include the conditioning variables. The estimates on the conditioning variables are not reported to conserve space. Deseasonalized stock-level SKC (*StockDSKC*) is defined as the difference in *StockSKC* and *Seasonal StockSKC*. Refer to the text for further details on construction. Robust standard errors clustered on the stock and date levels are displayed in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

comovement. We also examine the duration of these effects across different time horizons.

We adapt the return comovement tests from Green and Hwang (2009) and Kumar, Page, and Spalt (2013) for the analysis. Because we are interested in relatively shorter investment horizons than what is considered in those studies, we use daily stock returns data for estimation. First, we construct two portfolios based on daily *StockDSKC* rankings, using only stocks either in the lowest or in the highest 30th sample percentile for each date. The resulting two portfolios are defined as low *StockDSKC* (Optimistic) and high *StockDSKC* (Pessimistic). For each portfolio $j$, $Mood_j$ is the portfolio return in excess of the risk-free rate and is value weighted and rebalanced daily.[28]

Second, the coefficient on each *Mood* portfolio is estimated for each stock using rolling, time-series regressions that include the *Mood* portfolios, the Fama and French (1993) factors (MKTRF, HML, and SMB), and the Carhart (1997) momentum factor (UMD). If stock $i$ is included in either of the *Mood* portfolios at any point in time, the *Mood* portfolio is reconstructed to exclude stock $i$ to avoid any mechanical relation. The model takes the following form

$$\mathrm{R_{i,t}} - \mathrm{R_{f,t}} = \beta_0 + \sum_{\substack{j \in \{Pessimistic, \\ Optimistic\}}} \beta_{\mathrm{j}} \mathrm{Mood_{j,t}} + \beta_{\mathrm{MKTRF}} \mathrm{MKTRF_t} + \beta_{\mathrm{HML}} \mathrm{HML_t} + \beta_{\mathrm{SMB}} \mathrm{SMB_t} + \beta_{\mathrm{UMD}} \mathrm{UMD_t} + \varepsilon_{\mathrm{i,t}}$$

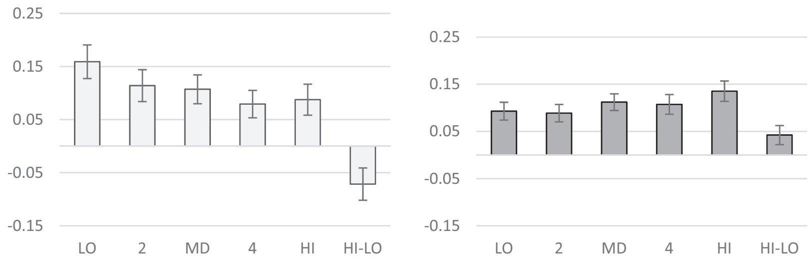
The comovement measures are the coefficients from the *Mood* portfolios. We use forward-looking 30- and 90-day estimation windows for the rolling regression models, each resulting in a panel of parameter estimates based on overlapping data. For inclusion in the sample, we require that each estimate is calculated using a sufficient number of returns observations.[29]

Finally, we compare the comovement estimates across stocks based on the *StockDSKC* rankings. Specifically, we calculate the value-weighted average of each portfolio loading for quintile groups based on *StockDSKC* rankings for each date. We predict that $\beta_{\mathrm{Pessimistic}}$ will be higher for stocks with higher *StockDSKC* values. Likewise, we predict that $\beta_{\mathrm{Optimistic}}$ will be lower for stocks with higher *StockDSKC* values. We calculate these estimates for both the 30-day and 90-day estimation windows. Because of the overlapping estimation windows used in the regression procedure, we adjust the standard errors using the procedure described by Newey and West (1994) to correct for heteroscedasticity and serial correlation.[30]

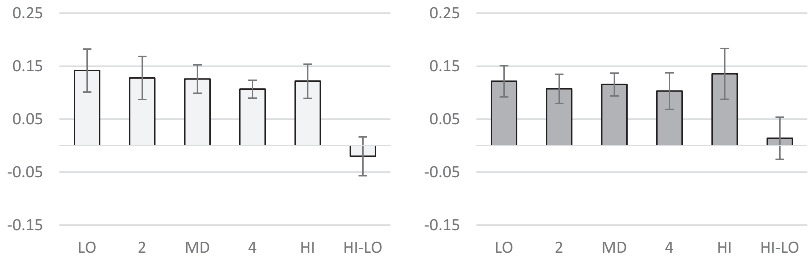
[28] We find similar results for equal weighted portfolios.

[29] Specifically, we require that the estimate use at least 10 and 30 days of nonmissing returns data for the 30- and 90-day estimation procedures, respectively.

[30] Lag selection for each specification is based upon the procedure of Newey and West (1994). We also consider alternative kernels to compute bandwidths and use the ones that provide the most conservative standard error estimates.




Panel A: 30-day estimation window




Panel B: 90-day estimation window

**Figure 2**
**Comovement estimates on *StockDSKC* portfolios**
This figure displays the comovement estimates on the factor portfolios associated with *Optimistic* (light bars, left) and *Pessimistic* (dark bars, right) *Mood* portfolios. For each bar graph, the sample average over the daily, value-weighted comovement estimates for each *StockDSKC* quintile group is reported, from the lowest (LO) to the highest (HI) daily *StockDSKC* rankings. If a stock also belongs to either the *Optimistic* or the *Pessimistic Mood* portfolio, the stock is removed from the factor portfolio before calculating the comovement estimate. The differences between the high and low quintile group estimates (HI-LO) are reported to the right. Panel A displays the graphs for the 30-day estimation window. Panel B displays the graphs for the 90-day estimation window. Ninety percent confidence intervals are overlaid on each series based on the Newey-West standard error estimates.

The results are graphically represented in Figure 2. The comovement estimates using a 30-day estimation window are represented in panel A, whereas those using the 90-day estimation window are reported in panel B.

Using the 30-day estimation window, we find that $\beta_{\text{Optimistic}}$ decreases in the *StockDSKC* rankings, whereas $\beta_{\text{Pessimistic}}$ increases in the *StockDSKC* rankings, as shown in panel A. To formally test our predictions, we calculate the difference in $\beta_{\text{Pessimistic}}$ and $\beta_{\text{Optimistic}}$ across the highest and lowest *StockDSKC* quintile groups. The difference is negative and statistically significant for $\beta_{\text{Optimistic}}$ (estimate = -0.071; $t$-value = -3.86) and is positive and statistically significant for $\beta_{\text{Pessimistic}}$ (estimate = 0.042; $t$-value = 2.08). The differences are slightly larger for the $\beta_{\text{Optimistic}}$ estimates. For the results from the 90-day estimation window reported in panel B, $\beta_{\text{Optimistic}}$ retains a similar pattern as before. However, the difference is no longer statistically significant at the 10% level (estimate = -0.020; $t$-value = -0.91). Likewise, for the $\beta_{\text{Pessimistic}}$ estimates, the differences remain positive but are also statistically insignificant (value = 0.014, $t$-value = 0.58).

In summary, the results provide evidence of return comovement attributable to the investor mood proxy. They also suggest that the weather effect is not long lived, given that weather patterns fluctuate more over longer time horizons. Consequently, the mood effect related to weather is unlikely to persist for long periods of time.

## 6. Conclusion

Several studies demonstrate that weather-based mood proxies explain variation in trading volume and prices of broad-based stock indexes. They use cloud cover in major stock market locations as a proxy for investor mood. In contrast to these earlier studies, we use disaggregated data on the locations of institutional investors to examine how weather-based measures of investor mood affect perceptions and trading behavior, as well as individual stock returns. To our knowledge, this study is the first to examine the direct impact of weather on the trading behavior of institutional investors.[31]

We find that the weather-based mood measures affect institutional investors' perceptions about market mispricing, thereby supporting mood-based explanations in the existing literature. Specifically, using survey data, we show that investor optimism, associated with lower values of deseasonalized cloud cover, is negatively associated with perceived overpricing in both individual stocks and the DJIA. Next, we use disaggregated trade data to provide additional evidence that investor optimism increases their propensities to buy. Finally, we also find that our stock-level investor mood measure explains a significant amount of variation in daily returns of stocks associated with higher arbitrage costs. In addition, we document return comovement attributable to investor mood. Collectively, these findings complement existing studies that document the effect of weather-induced mood on stock market returns.

In future work, it would be interesting to examine whether other groups of economic agents are also influenced by mood. For example, the decisions of firm managers that have greater autonomy in decision making, such as those of small- and medium-sized enterprises, may be more susceptible to mood-related biases, which, in turn, may significantly impact hiring and investment decisions. The forecasts of sell-side equity analysts also may be influenced by weather, which could affect how market participants incorporate new information into prices. We leave these interesting questions for future research.

---

[31] One notable exception is Loughran and Schultz (2004), who indirectly examine the effect of localized trading by firm headquarter locations.

## References

Adaval, R. 2001. Sometimes it just feels right: The differential weighting of affect-consistent and affect-inconsistent product information. *Journal of Consumer Research* 28:1–17.

Agarwal, S., R. Duchin, and D. Sosyura. 2012. *In the mood for a loan: The causal effect of sentiment on credit origination.* Working Paper, University of Michigan.

Barber, B., Y. Lee, J. Liu, and T. Odean. 2007. Is the aggregate investor reluctant to realize losses? *European Financial Management* 13:423–47.

Bassi, A., R. Colacito, and P. Fulghieri. 2013. ’O sole mio: An experimental analysis of weather and risk attitudes in financial decisions. *Review of Financial Studies* 26:1824–52.

Bodoh-Creed, A. 2013. Mood, memory, and the evaluation of asset prices. Working Paper, University of California at Berkeley.

Camerer, C., L. Babcock, G. Loewenstein, and R. Thaler. 1997. Labor supply of New York City cabdrivers: One day at a time. *Quarterly Journal of Economics* 112:407–41.

Cameron, A., J. Gelbach, and D. Miller. 2011. Robust inference with multiway clustering. *Journal of Business & Economic Statistics* 29:238–49.

Carhart, M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52:57–82.

Clore, G., and W. Parrott. 1991. In *Emotion and social judgments*. Ed. J. P. Forgas. Oxford: Pergamon Press.

Clore, G., N. Schwarz, and M. Conway. 1994. In *Handbook of social cognition*, vol. 1, pp. 107–23. Eds. R. S. Wyer and T. K. Srull. New York: Psychology Press.

Coval, J., and T. Moskowitz. 1999. Home bias at home: Local equity preference in domestic portfolios. *Journal of Finance* 54:2045–73.

Coval, J., and T. Shumway. 2005. Do behavioral biases affect prices? *Journal of Finance* 60:1–34.

Cunningham, M. 1979. Weather, mood, and helping behavior: Quasi-experiments with the sunshine Samaritan. *Journal of Personality and Social Psychology* 37:1947–56.

Dougal, C., J. Engelberg, D. García, and C. Parsons. 2012. Journalists and the stock market. *Review of Financial Studies* 25:639–79.

Driscoll, J. C., and A. C. Kraay. 1998. Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics* 80:549–60.

Eastman, C. I., M. A. Young, L. F. Fogg, L. Liu, and P. M. Meaden. 1998. Bright light treatment of winter depression: a placebo-controlled trial. *Archives of General Psychiatry* 55:883–9.

Edmans, A., D. Garcia, and Ø. Norli. 2007. Sports sentiment and stock returns. *Journal of Finance* 62:1967–98.

Fama, E., and K. French. 1993. Common risk factors in returns on stocks and bonds. *Journal of Financial Economics* 33:3–56.

Fishbein, M. 1963. An investigation of the relationships between beliefs about an object and the attitude toward that object. *Human Relations* 16:233–40.

Frazzini, A. 2006. The disposition effect and under-reaction to news. *Journal of Finance* 61:2017–46.

Garcia, D. 2013. Sentiment during recessions. *Journal of Finance* 68:1267–300.

Goel, N., M. Terman, J. Su Terman, M. M. Macchi, and J. W. Stewart. 2005. Controlled trial of bright light and negative air ions for chronic depression. *Psychological Medicine* 35:945–56.

Goetzmann, W., and N. Zhu. 2005. Rain or shine: Where is the weather effect? *European Financial Management* 11:559–78.

Green, C., and B. Hwang. 2009. Price-based return comovement. *Journal of Financial Economics* 93:37–50.

Hirshleifer, D., and T. Shumway. 2003. Good day sunshine: Stock returns and the weather. *Journal of Finance* 58:1009–32.

Johnson, E., and A. Tversky. 1983. Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology* 45:20–31.

Kamstra, M., L. Kramer, and M. Levi. 2000. Losing sleep at the market: The daylight saving anomaly. *American Economic Review* 90:1005–11.

———. 2003. Winter blues: A SAD stock market cycle. *American Economic Review* 93:324–43.

Kamstra, M., L. Kramer, M. Levi, and T. Wang. 2014. Seasonally varying preferences: Theoretical foundations for an empirical regularity. *Review of Asset Pricing Studies* 4:39–77.

Kida, T., J. Smith, and M. Maletta. 1998. The effects of encoded memory traces for numerical data on accounting decision making. *Accounting*, *Organizations and Society* 23:451–66.

Kramer, L., and J. Weber. 2012. This is your portfolio on winter seasonal affective disorder and risk aversion in financial decision making. *Social Psychological and Personality Science* 3:193–9.

Kripke, D. F. 1998. Light treatment for nonseasonal depression: speed, efficacy, and combined treatment. *Journal of Affective Disorders* 49:109–17.

Kumar, A., J. Page, and O. Spalt. 2013. Investor sentiment and return comovements: Evidence from stock splits and headquarters changes. *Review of Finance* 17:921–53.

Lee, C., and M. Ready. 1991. Inferring trade direction from intraday data. *Journal of Finance* 46:733–46.

Linnainmaa, J., and I. Rosu. 2009. Weather and time series determinants of liquidity in a limit order market. Working Paper, University of Chicago.

Lo, K. and S. Wu. 2011. The impact of seasonal affective disorder on financial analysts and equity market returns. Working Paper, University of British Columbia.

Loewenstein, G. 2000. Emotions in economic theory and economic behavior. *American Economic Review* 90:426–32.

Loewenstein, G., E. Weber, C. Hsee, and N. Welch. 2001. Risk as feelings. *Psychological Bulletin* 127:267–86.

Loughran, T., and P. Schultz. 2004. Weather, stock returns, and the impact of localized trading behavior. *Journal of Financial and Quantitative Analysis* 39:343–64.

Mehra, R. and R. Sah. 2002. Mood fluctuations, projection bias, and volatility of equity prices. *Journal of Economic Dynamics and Control* 26:869–87.

Nagel, S. 2005. Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics* 78:277–309.

Newey, W., and K. West. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61:631–54.

Petersen, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22:435–80.

Petty, R., F. Gleicher, and S. Baker. 1991. In *Emotion and social judgements*, pp. 181–200. Ed. J. P. Forgas. New York: Pergamon Press.

Puckett, A., and X. Yan. 2011. The interim trading skills of institutional investors. *Journal of Finance* 66:601–33.

Rind, B. 1996. Effect of beliefs about weather conditions on tipping. *Journal of Applied Social Psychology* 26:137–47.

Rosenthal, N. E., D. A. Sack, J. C. Gillin, A. J. Lewy, F. K. Goodwin, Y. Davenport, P. S. Mueller, D. A. Newsome, and T. A. Wehr. 1984. Seasonal affective disorder: A description of the syndrome and preliminary findings with light therapy. *Archives of General Psychiatry* 41:72–80.

Saunders, E. 1993. Stock prices and wall street weather. *American Economic Review* 83:1337–45.

Schwartz, G. 1990. In *Repression and dissociation*: *Implications for personality theory*, *psychopathology*, *and health*, pp. 405–34. Ed. J. Singer. Chicago: University of Chicago Press.

Schwarz, N., and G. Clore. 1983. Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology* 45:513–23.

Shiller, R. J., F. Kon-Ya, and Y. Tsutsui. 1996. Why did the Nikkei crash? Expanding the scope of expectations data collection. *Review of Economics and Statistics* 78:156–64.

Shiller, R. J. 2000. Measuring bubble expectations and investor confidence. *Journal of Psychology and Financial Markets* 1:49–60.

Wilson, T., and J. Schooler. 1991. Thinking too much: Introspection can reduce the quality of preferences and decision. *Journal of Personality and Social Psychology* 60:181–92.

Wyer, R. S. Jr., G. L. Clore, and L. M. Isbell. 1999. Affect and information processing. *Advances in Experimental Social Psychology* 31:1–77.