

Liquidity and Market Returns: A Natural Experiment in Pre-1903 Manhattan Summers*

Thummim Cho
Harvard University

December 2015

Abstract

What causes the daily reversals in a stock market index? Using exogenous temperature variation in Manhattan in the summers of 1889-1902, I provide evidence that non-informational trades by ‘outside investors’ (Grossman and Miller 1988) underlie short-term reversals in the market index return, consistent with the model of Campbell, Grossman, and Wang (1993). An increase in outside investors by 1% of the average daily trade volume is estimated to cause a 1-2% standard deviation increase in aggregate liquidity imbalance, a 0.2 percentage point increase the likelihood of a next-day market index return reversal, and a 0.018 decrease in the daily index return autocorrelation.

*Email: tcho@fas.harvard.edu. I am grateful to John Y. Campbell for his guidance and support throughout the project. I thank Christopher Anderson, Melissa Dell, Claudia Goldin, Richard Hornbeck, Yosub Jung, Dyvia Kirti, Andrei Shleifer, Jeremy C. Stein, seminar participants at Harvard University, and especially David Laibson and Owen Lamont for helpful conversations. I gratefully acknowledge the financial support from the Hirtle Callaghan grant to Harvard. Excellent research assistance was provided by Dada Huang, Dian Jiao, and Michelle Namkoong.

1 Introduction

Return reversal is a dominant feature of stock market indices' short-horizon behavior. From 1926 to 2014, the Standard & Poor's 500 (S&P 500) index has had an average daily return of 0.03% but an average absolute value of the return of 0.72%, implying large reversals that cannot be plausibly explained by daily information flow. What explains these reversals in the market index? Campbell, Grossman, and Wang (1993) (CGW) provide a model in which non-informational trades cause a price movement that subsequently reverses and predict that the magnitude of the reversal is larger if there is a greater imbalance between the intensity of non-informational trades and the availability of liquidity suppliers. This theory has been tested on *individual* stocks using exogenous shocks to non-informational trades; for example, non-informational trades on individual stocks due to an index redefinition are found to cause a price change that shortly reverses (Harris and Gurel 1986; Greenwood 2005).¹ Testing this theory as an explanation for daily movements in a *market index*, however, has been difficult because a causal inference requires an exogenous shock affecting the intensity of non-informational trades in all stocks.

Previous studies test this theory of market return reversal using aggregate trade volume as a proxy for non-informational trades. For example, CGW show that when NYSE trade volume is unusually high, S&P 500 returns indeed exhibit lower autocorrelation. Following this approach, LeBaron (1992) reports similar evidence on the Dow Jones Index. Llorente et al. (2002), however, raise the concern that a unusually high trading volume may reflect informed trading and show that the volume-autocorrelation relation in individual stocks depends on the extent to which the volume is caused by informed trading. Although informed trading is less likely to affect aggregate trade volume, a public signal produces high trading volume if investors interpret the signal differently (Kandel and Pearson 1995; Bamber, Barron, and Stober 1999). Furthermore, even if non-informational trades were correctly measured, the estimated effect of non-informational trades on index reversals may capture reverse causality, since unusual stock price movements are known to trigger non-informational trades (Seasholes and Wu 2007; Barber and Odean 2008).

In this study I use exogenous variation in daily summer temperatures in 19th century Manhattan to identify the effect of non-informational trades on New York Stock Exchange (NYSE) daily index returns. Prior to the advancement in telecommunications and air conditioning, trading stocks in the summer required substantial physical effort. This suggests that particularly hot weather in New York City, where a large fraction of trades originated from, would have reduced the NYSE trading

¹Further evidence of the liquidity effect exists on individual stocks. Reversal profitability is higher in illiquid stocks (Pastor and Stambaugh 2003; Avramov, Chordia, and Goyal 2006a) and stocks with high trading activity, a proxy for non-informational trades (Conrad, Hameed, and Niden 1994). Relatedly, studies have also predicted and shown that uninformed trades create non-fundamental volatility over a short horizon (Hellwig 1980; Wang 1994; Avramov, Chordia, and Goyal 2006b; Koudijs 2015).

volume. Indeed, in the summers of 1889-1902, an unseasonal rise in daily Manhattan temperature by 1 degree Celsius led to a 1.3% decline in NYSE trade volume on the same day. Using the broad categorization of investors into ‘outside investors’ and ‘market makers’ as in Grossman and Miller (1988), I attribute this decline in trade volume to a reduction in non-informational trades by outside investors who did not have immediate access to telephones or automatic ticker machines, were further away from Wall Street, and did not trade for a living. This allows me to use a unseasonal rise in daily summer temperature as a proxy for the reduction in trades by outside investors.

Using the temperature variation, I find that outside investor trades harm aggregate market liquidity and increase the likelihood of a daily reversal in the market index. Using unique data of daily NYSE individual stock prices during 1889-1902, I show that an increase in outside investors equivalent to 1% of the average daily trade volume leads to a 1 – 2% standard deviation increase in the expected return from aggregate liquidity provision. This increase in reversal profitability occurs through two channels: an increase in the number of stocks having a next-day reversal and a decrease in the absolute value of returns on stocks having a next-day continuation (by allowing more informed trades to occur undetected). Outside investors, on the other hand, do not seem to affect the absolute returns on stocks having a next-day reversal; this quantity seems to be determined not by liquidity demand factors but by liquidity supply factors such as market volatility. Looking at market index reversals, the same 1% reduction in outside investor trades leads to a 1.2% fall in the probability of a market return reversal on the following day and a 0.018 increase in the daily market return autocorrelation.

The results support the findings of CGW and LeBaron (1992) that time-varying demand for liquidity caused by outside investor trades is an important determinant of daily movements in a stock market index.² This effect of outside investor trades through liquidity demand is distinct from the effect of liquidity supply conditions, although the two effects would interact. Nagel (2012) shows that reversal profitability falls during high volatility periods, as liquidity suppliers become more capital constrained. This supply side effect on aggregate market liquidity, however, moves slowly and is likely to matter only during periods in which liquidity suppliers are constrained. Therefore, day-to-day changes in the aggregate liquidity imbalance and in the valuation of market indices are more likely to be governed by time-varying demand for liquidity due to non-informational trades.

One implication of this finding is that the demand curve for the aggregate market index slopes down, just as it does for individual stocks (Shleifer 1986; Avramov, Chordia, and Goyal 2006a). That is, even at the aggregate stock market level, where limited attention or idiosyncratic risk would play a minimal role, arbitrageur capital does not immediately respond to counteract liquidity traders. In addition, non-informational traders exert significant price pressure on the market index,

²The findings are also consistent with the evidence that stocks heavily bought by retail investors, the majority of whom would be classified as outside investors, exhibit return reversal over a week (Barber, Odean, and Zhu 2009).

providing empirical grounds for the important roles played by noise traders in various asset pricing models (e.g., Verrecchia 1982; De Long et al. 1990; Shleifer and Summers 1990; Shleifer and Vishny 1997).

This study is related to three other strands of literature. First, several studies document higher stock returns on sunny days and explain this through investor psychology (Saunders 1993; Hirshleifer and Shumway 2003; Kamstra, Kramer, and Levi 2003; Cao and Wei 2005; Bassi, Colacito, and Fulghieri 2013).³ I find that weather prior to 1903 also affected the stock market by changing the cost of carrying out a trade. This work is related to the observation that trade volume tends to fall in the summer (Gallant, Rossi, and Tauchen 1992; Bouman and Jacobsen 2002; Hong and Yu 2009). I find that this summer seasonality in volume has gradually disappeared, due in part to the improvement in temperature control and communications technologies. Several studies outside finance document negative effect of temperature on economic activities.⁴ My study complements these findings by presenting a high frequency (daily) relationship between temperature and stock trading activity. Besides these three lines of literature, this paper's use of a natural experiment from a historical stock market is similar to Koudijs (2015), who studies the effect of information on individual stock price volatility using variation in the arrival of news to the Amsterdam exchange through sailing boats.

This paper proceeds as follows. In section 2, I present historical background, choose the test sample period, and discuss the data and variables used. In section 3, I outline my empirical approach and discuss identification assumptions related to using unseasonal temperature as a proxy for the presence of outside investors. I present the main empirical results in section 4. There, I first estimate the daily effect of Manhattan temperature on NYSE trading volume during the test sample period. Using unseasonal temperature as a proxy, I then study the effect of outside investor trading on aggregate market liquidity and daily market return reversals. Section 5 concludes.

³Goetzmann and Zhu (2005), Jacobsen and Marquering (2008), and Jacobsen and Marquering (2009) provide counter-evidence.

⁴The recent literature review by Dell, Jones, and Olken (2014) provides an overview of this growing literature. To mention a few recent works, Dell, Jones, and Olken (2009), Jones and Olken (2010), and Dell, Jones, and Olken (2012) use annual data to show that temperature has adverse effects on output, export, and income. Deschenes and Greenstone (2007) find similar evidence from the agricultural sector. Relatedly, Bansal and Ochoa (2011, 2015) show that temperature change poses long-run risk on equities through its effect on economic activity, and that this is already taken into account by asset covariance (i.e., beta) and return. My findings suggest that temperature can also influence stock returns through its effect on trade volume and investor composition, if technology does not adequately moderate the effect of temperature on human activity.

2 Historical Background, Data, and Measurement

2.1 Stock trading prior to 1903

This study relates Manhattan weather with NYSE stock trading volume and returns prior to 1903 at a daily frequency. Hot summer weather was a plausible source of exogenous variation in the cost of trading stocks prior to 1903 and presumably had the strongest negative impact on outside investors, who had to make substantial physical effort to trade stocks.⁵ Telephones and automatic ticker machines were rare, so a typical investor wishing to make two-way communications or receive continuous updates on the market would have had to travel to a nearby telephone exchange or visit the customer's room at a brokerage house.^{6,7} Air conditioning was virtually nonexistent until 1903, when the world's first large-scale air conditioner was installed in the new NYSE building (Buchanan 2013). Thus, outside investors prior to 1903 experienced the full extent of the summer heat, making them likely compliers to the temperature effect.⁸

Besides the outside investors who traded through brokers, there were important investors who traded directly. 'Capitalists' had seats in the exchange and traded large shares of stocks without having to pay commission to brokers.⁹ 'Room traders' made a living by betting on short-term fluctuations. 'Specialists' worked as the auctioneer for brokers in addition to trading their own similar to room traders.¹⁰ These traders together represent market makers and other informed traders.¹¹ As these individuals traded stocks for a living, they would have been less affected by hot weather. In fact, their non-compliance is a key assumption for my identification; although it is not essential that all outside investors comply to hot weather, it is quite important that neither market makers nor informed traders responded to hot weather.

Focusing on the pre-1903 sample allows me to use temperature in Manhattan as the temperature felt by a large fraction of NYSE investors. During this period, the underdevelopment of long-distance communications impaired the competitiveness of trading from another city. Long-distance automatic ticker service was unavailable until 1905, when such a service was first set up between New York and Philadelphia. Before then, it took brokerage houses in other cities an additional 15

⁵Records suggest that they were important part of the daily stock market (Selden 1919).

⁶One in every 300 persons had a telephone in 1890 (U.S. Census Bureau 1975), and approximately 1000 automatic ticker machines existed in bank and brokerage offices in New York as of 1880 (Donnan 2011).

⁷This description of stock trading in the late 19th century is a reconstruction based on Selden (1919) and Beckert (2003).

⁸It took about 30 more years for important government buildings such as the White House and the House of Representatives used air conditioning ("The History of Air Conditioning", retrieved in September 2013 from http://air-conditioners-and-heaters.com/air_conditioning_history.htm).

⁹The annual cost of membership was over \$3,000 a year as of 1917, equivalent to \$55,000 in 2013.

¹⁰The last three sentences including the cost of membership in the exchange are based on Selden (1919).

¹¹In reality, a small number of professional traders may behave as outside investors, trading without information. However, these traders are likely to have been driven out quickly by the market mechanism.

minutes to receive market updates from New York, because the employees had to hand-copy the stock quotations received by Morse on manifold sheets (Tilghman 1961). Furthermore, major cities like Boston and Philadelphia had regional exchanges that local traders could use, also implying a tight link between Manhattan temperature and NYSE trading volume.¹² Indeed, in section 3, I show that the temperature in Cambridge MA, despite its proximity to Manhattan, has no predictive power on NYSE trading volume once Manhattan temperature has been controlled for.

The natural experiment I exploit here would not be valid in the 20th century. Once brokerage houses started installing air conditioning, hot weather would have no longer discouraged the customers from visiting their brokers. Telephones became commonplace, with the number of telephones growing at an annual rate of approximately 20% from 1900 to 1910.¹³ Long-distance telephone service became widespread, with a service between New York City and San Francisco set up in 1915.¹⁴ Major cities began offering long-distance ticker services, allowing investors from other cities to participate in an exchange. Transportations in New York City improved with the advent of the subway system in the city in 1904 (DuTemple 2002). As shown later, the link between temperature and stock trading first weakens then disappears in the four latter sample periods (1903-1930, 1931-1960, 1961-1990, and 1991-2014) as these changes take place.

2.2 Data and measurement

I use the NYSE individual stock price data collected from the Financial and Commercial Chronicle, NYSE index returns based on the Dow Jones Industrial Average (DJIA) and S&P 500, NYSE total trading volume, and temperature data collected at Manhattan's Central Park by the National Climatic Data Center (NCDC), all at the daily frequency. My robustness checks also use temperatures collected in 2 other US cities and 3 foreign cities.

Unique data of pre-1903 NYSE individual stock daily returns

Individual stock returns data are required to construct returns from aggregate liquidity provision, but are not available electronically prior to 1926. I therefore collect the daily minimum and maximum prices on all NYSE-listed common stocks in the summers (Jun 11-Sep 10) of 1889 to 1902 by digitizing the price information in the Financial and Commercial Chronicles.¹⁵ I compute individual stock returns as the percentage change in the midpoint price, excluding observations with

¹²It is also worth noting that banks and financial services have long had high concentration near NYSE on Wall Street (Longcore and Rees 1996).

¹³The calculation is made based on numbers provided in U.S. Census Bureau (1975).

¹⁴"Phone to Pacific From the Atlantic". New York Times, January 26, 1915.

¹⁵Trading volume is available weekly, and I do not use this information in the paper. The hard copy of the Chronicle is not available for June 1894, which is excluded from my analyses.

no record of either a buy or a sell transaction. I also exclude days with a dividend payment because the lack of information about the dividend amount prevents me from computing the gross-dividend total returns. The digitized data initially contained a number of typos that resulted in extreme daily returns. Although most typos have been detected and revised, to prevent extreme returns due to any remaining errors from affecting my results, I exclude observations with an absolute value of the daily return above 50%. The result is an unbalanced panel of 89,729 daily return observations (by company and date).

NYSE index daily returns

The NYSE index return is measured daily by DJIA return for years 1889-1925 and S&P 500 return for 1926-2014. I obtain DJIA returns over 5/26/1896-12/31/1925 from the Dow Jones website, and the prior years' cumulative-dividend returns are kindly provided by Schwert (1990).¹⁶ I obtain S&P 500 returns since 1926 from the Center for Research in Security Prices (CRSP). A notable difference between the two indices is that DJIA is price-weighted but S&P 500 is value-weighted, although the difference would be small if larger stocks had higher prices.

NYSE daily trading volume

NYSE daily trading volume since 1888 is obtained from the NYSE website.¹⁷ I compute detrended trading volume, $Tradvol_t$, as the log deviation of trade volume per trading hour from its trailing 1-year moving average:

$$Tradvol_t = \ln \left(\frac{\text{trading volume}_t}{\text{trading hours}_t} \right) - \frac{1}{N_t} \sum_{s=1}^{N_t} \ln \left(\frac{\text{trading volume}_{t-s}}{\text{trading hours}_{t-s}} \right) \quad (1)$$

where N_t is the number of trading days in the previous 365 days. This measure of trading volume is similar to those in CGW and Chen, Hong, and Stein (2001) except that it adjusts for the trading hours to make weekday and Saturday trade volumes comparable (Saturday had shorter trading hours) and that it uses trading volume instead of share turnover.¹⁸

The result is daily volume series from 1/2/1889 to 12/31/2014, plotted in Figure 1 for the entire sample, the sample of 1889-1902, and the summer of 1895 (representing the middle of the

¹⁶The Dow Jones data was retrieved in November 2015 from <http://www.djaverages.com/?go=industrial-index-data&report=performance>.

¹⁷The data was retrieved in October 2015 from <https://www.nyse.com/data/transactions-statistics-data-library>. The data contains 4 instances in which two volume observations have the same date, and I average the two numbers in such instances. None of the dates, however, are in the period of interest (1889-1902). Analysts from the NYSE kindly went over and removed all other instances of duplicates.

¹⁸The number of total shares outstanding at NYSE is not available going back to the 19th century, but using trading volume is equally valid in this study exploiting daily variations.

test sample).¹⁹

[Figure 1 here]

The detrended process is still fairly persistent with the average daily autocorrelation of 0.73 in the 14 summers of 1889-1902. Although this is not considered near unit root, I do take the possibility of a spurious regression bias seriously, conducting placebo tests with temperature and trading volume data in post-1903 samples as well as with temperatures in other cities within the pre-1903 sample.

Daily temperature and merged data

I obtain daily temperatures in Manhattan Central Park during 1879-2014 from the NCDC website.²⁰ For placebo tests, I also collect daily temperature of Cambridge MA, Jacksonville FL, Birmingham in the United Kingdom, Sydney in Australia, and Edmonton in Canada during 1879-2014 (except Cambridge MA which starts in 1885).²¹ The collected data contain maximum and minimum temperatures measured in degrees Celsius ($^{\circ}\text{C}$). I obtain the trading hour average temperature T each day using the maximum and minimum temperatures on that day. To do so, I assume that an intraday variation in temperature reaches the daily maximum at 15:00 and minimum at 5:00 (Lonnqvist 1962) and that the temperature change is linear in time.

Stock trade volume is known to be lower in the summer (Gallant, Rossi, and Tauchen 1992; Bouman and Jacobsen 2002; and Hong and Yu 2009), which may be a combination of temperature and seasonal effects. Table 1 shows that the summer seasonality in trading volume was strong in earlier sample periods but has steadily declined over the century.

[Table 1 here]

To isolate the temperature effect from the seasonal effect on trading volume, I decompose T into seasonal and unseasonal components, denoted \bar{T} and \tilde{T} respectively. Seasonal temperature \bar{T} is the average trading hour temperature on the same day for the previous 10 years, and unsea-

¹⁹The drop in the volatility of trade volume around 1950 roughly coincides with the end of Saturday trading. This does not affect my analysis, which focuses on the 1889-1902.

²⁰The data was retrieved in July 2013 from <http://www.ncdc.noaa.gov/data-access/land-based-station-data>.

²¹Although the NCDC interface does not allow me to choose five cities in a systematic manner, I tried to obtain the longest time-series data from a variety of geographical locations.

sonal temperature \tilde{T} is the temperature deviation from the seasonal temperature.²² This results in decomposed temperature series for years 1889-2014 (1890-2014 for Cambridge MA with \bar{T} in 1890-1894 computed using all previous years).

I conduct my analyses using only the summer season, defined here as June 11th to September 10th, the hottest 90 days in a year when average temperature is computed in 10 days' interval (Figure 2).

[Figure 2 here]

Restricting my analyses to the summer not only serves an additional control for the effect of seasonality, but also avoids the need to deal with the “inverse-U” effect of temperature on human physiology. Sepannen, Fisk, and Lei (2006) show that task performance in office environment peaks around 20-22°C and decreases as temperature increases or decreases from this point. Summer temperatures in Manhattan are almost always above 22°C, so a rise in temperature is expected to have only an adverse effect on trading activity during the summer season.

Merging all data gives individual stock and index returns, detrended trade volume, and decomposed temperatures over the period 1/2/1889 to 12/31/2014. I divide this sample into 5 sample periods (1889-1903, 1903-1930, 1931-1960, 1961-1990, and 1991-2014) and use the earliest period as the test sample.

3 Estimation Framework

3.1 Challenges in estimating the effect of outside investors on aggregate liquidity

This study aims to estimate the causal effect of outside investors on aggregate liquidity imbalance and ultimately index return reversal (“*Imbalance*” for the purpose of exposition in this section). The existing approach uses aggregate trading volume as a proxy variable for unobserved outside investor trades to estimate the following regression model:

$$Imbalance_t = \phi_0 + \phi_1 Tradvol_t + X_{L,t} \Phi + \nu_t \quad (2)$$

²²Given this definition, \tilde{T} corresponds not to unexpected temperature given the preceding few days' data, but to unexpected temperature as of the beginning of the year or the summer. This choice reflects the fact that the main concern underlying the decomposition is the effect of seasonality. Being able to predict the temperature a few days earlier does not violate the identification assumption that the ratio of individual investors in the daily stock market falls on hotter days. To ensure that the remaining persistence of \tilde{T} (0.57) does not produce spurious regression results, I repeat all analyses with lagged unseasonal temperature as an additional control and obtain similar results.

where $X_{L,t}$ represents the vector of control variables for return reversal and v_t is an error term. Two problems arise. First, trading volume may respond not only to innovations in outside investor trades ($Outside_t$), but also to trades based on new information ($Informed_t$). Second, innovations in $Imbalance_t$ unobserved by the econometrician but observed by the investors at t may trigger outside investor trades, creating reverse causality. I discuss these two problems and suggest using temperature as an exogenous proxy for outside investor trades.

The first problem is that the omitted variable $Informed_t$ may be correlated with $Tradvol_t$. To see this, suppose that trading volume is determined by both outside investors and informed traders:

$$Tradvol_t = \theta_0 + \theta_1 Outside_t + \theta_2 Informed_t + X_{TV,t} \Theta + v_t \quad (3)$$

where $\theta_1 > 0$ and $\theta_2 \geq 0$ and $X_{TV,t}$ represents the vector of control variables for trading volume. Also, write the true model of liquidity imbalance as

$$Imbalance_t = \tilde{\phi}_0 + \tilde{\phi}_1 Outside_t + X_{L,t} \tilde{\Phi} + \tilde{v}_t \quad (4)$$

where $\tilde{\phi}_1 > 0$. This implies that the specification in (2) is equivalent to

$$Imbalance_t = \underbrace{\tilde{\phi}_0 - \tilde{\phi}_1 \theta_0 \theta_1^{-1}}_{\phi_0} + \underbrace{\tilde{\phi}_1 \theta_1^{-1}}_{\phi_1} Tradvol_t + \tilde{\phi}_1 \theta_1^{-1} X_{TV,t} \Theta + X_{L,t} \tilde{\Phi} - \underbrace{\tilde{\phi}_1 \theta_2 \theta_1^{-1} Informed_t - \tilde{\phi}_1 \theta_1^{-1} v_t + \tilde{v}_t}_{v_{t+1}} \quad (5)$$

Contrasting this to equation (2), we see that $Tradvol_t$ and the error term v_t would be correlated. This correlation is positive, so the positive effect of outside investors on index reversal would be underestimated. One may also argue that informed trading improves liquidity imbalance so that $Informed_t$ should enter the right-hand side of (4) with a negative coefficient. This would increase the magnitude of the positive correlation between $Tradvol_t$ and the omitted variable $Informed_t$ in the regression, producing a larger bias.

The second and perhaps more serious problem arises from reverse causality: a large price movement associated with an increase in market liquidity imbalance orthogonal to the innovation in outside investor trades (e.g., reduction in liquidity supply) can trigger outside investor trades. This concern has empirical grounds. Studies find that retail investors—the majority of whom would be outside investors—tend to trade on days with large stock price movements (e.g., Seasholes and Wu 2007; Barber and Odean 2008). This implies that the factors outside investor

trades can be written as a function of liquidity imbalance:

$$Outside_t = \varphi_0 + \varphi_1 Imbalance_t + \omega_t \quad (6)$$

where $\varphi_1 > 0$. This means that even if outside investor trades can be perfectly observed, running the regression in (4) results in a simultaneity bias. This bias would be negative, meaning that a significantly positive coefficient obtained from estimating the model (4) may not imply outside investors inducing aggregate liquidity imbalance and index return reversal.

3.2 Temperature as an exogenous proxy variable for outside investors

Using exogenous variation in temperature as a proxy for outside investor trades addresses both concerns raised above. In the summers of 1889-1902, part of the innovation in NYSE outside investor trades came from hot Manhattan weather, allowing us to rewrite (6) as

$$Outside_t = \varphi_0 + \varphi_1 Imbalance_t + \varphi_2 \tilde{T}_t + \tilde{\omega}_t \quad (7)$$

where \tilde{T}_t is daily unseasonal Manhattan temperature, $\varphi_1 > 0$, and $\varphi_2 < 0$. Combining this with the model of aggregate liquidity imbalance in (4) and solving for liquidity imbalance as a function of control variables and innovation terms gives

$$Imbalance_t = \frac{\tilde{\varphi}_0 + \tilde{\varphi}_1 \varphi_0}{1 - \tilde{\varphi}_1 \varphi_1} + \tilde{\varphi}_1 \frac{\varphi_2}{1 - \tilde{\varphi}_1 \varphi_1} \tilde{T}_t + \frac{1}{1 - \tilde{\varphi}_1 \varphi_1} X_{L,t} \tilde{\Phi} + \frac{\tilde{\varphi}_1 \tilde{\omega}_t + \tilde{v}_t}{1 - \tilde{\varphi}_1 \varphi_1} \quad (8)$$

The coefficient on \tilde{T}_t is the combination of the direct negative effect temperature has on outside investor trades and the multiplier effect of smaller liquidity imbalance associated with lower temperature reducing outside investor trades (the multiplier $1 / (1 - \tilde{\varphi}_1 \varphi_1)$ represents the infinite geometric sum and is positive unless the coefficients are such that the sum diverges to infinity). Estimating this model and finding the coefficient on \tilde{T}_t to be negative implies a positive $\tilde{\varphi}_1$.

What happens if hot weather also discourages informed trading? If informed trading helps reduce liquidity imbalance, this would lead to an underestimation of the coefficient on \tilde{T}_t , so that a statistically negative coefficient does imply a positive $\tilde{\varphi}_1$.²³ In other words, a negative effect of

²³The notion that informed traders reduce liquidity imbalance does not contradict the notion that informed trading harms liquidity (e.g., Kyle 1985). Liquidity in the sense of Kyle (1985) determines price reversals over a sequence of executed trades and is proxied by the bid-ask spread. In contrast, liquidity imbalance here refers to the size of non-informational trades relative to the size of market makers and informed traders in the daily market willing to absorb the non-informational trades. It can be proxied by returns from reversal strategies (e.g., Nagel 2012), is related to the difference between the current stock price and its correct valuation (e.g., Brunnermeier and Pedersen 2009), and is thought to determine return reversals over a few days. Informed trading is likely to reduce the gap between the

temperature on informed trading would dampen the effectiveness of high temperature in reducing liquidity imbalance through lower outside investor activities.

Although using temperature as an exogenous proxy for outside investor trades allows us to identify the direction of the effect, the magnitude of the effect would be difficult to interpret. For this reason, I also use temperature as an instrument for aggregate trading volume with outside investor trades as the compliers of the local treatment effect and report the result along with results from the proxy regression.

To see how this is done, note that if temperature affects outside investors but not informed traders, combining equations (3), (4), and (7) implies the following expression for trading volume:

$$Tradvol_t = \frac{1}{1-\theta_1\phi_1\phi_1} [\theta_0 + \theta_1(\phi_0 + \phi_1\phi_0) + \theta_1\phi_2\tilde{T}_t + \theta_2 Informed_t + \theta_3 X_{TV,t} + \theta_1\phi_1 X_{L,t}\Phi + \theta_1(\phi_1 v_t + \tilde{\omega}_t) + v_t] \quad (9)$$

where $X_{L,t}$ is the vector of control variables in the liquidity imbalance regression. Again, the multiplier $(1 - \theta_1\phi_1\phi_1)^{-1}$ arises from the reverse causality of liquidity imbalance on outside investor trades and is positive under convergence of the geometric series (the multiplier is one if there is no reverse causality, i.e. $\phi_1 = 0$). To the extent that temperature does not affect informed trading $Informed_t$, innovation in temperature \tilde{T}_t would affect $Tradvol_t$ only through its effect on outside investors. Using predicted trade volume from this first-stage regression, I run the second-stage regression,

$$Imbalance_t = \beta_0 + \beta_1 \widehat{Tradvol}_t + X_t \mathbf{B} + \varepsilon_t \quad (10)$$

which allows me to interpret β_1 as the marginal effect of an increase in trading volume due to outside investors on liquidity imbalance and on index return reversal.

3.3 Identification assumptions

In the previous subsection, I made explicit and implicit assumptions required to obtain correct identification. An explicit assumption was that temperature affects outside investors but not informed traders. This assumption was discussed in section 3.1, where I argued that hot weather is unlikely to have discouraged informed traders and market makers from participating in the daily market. An implicit assumption was that Manhattan temperature does not affect NYSE aggregate market liquidity or the likelihood of NYSE index return reversal except through its negative effect on outside investor trades. I first discuss this implicit assumption and make a brief comment about using Manhattan temperature as an instrument for trading volume.

One source of violation in the implicit assumption is that Manhattan temperature conveys in-

prevailing market price and the “correct” price of a stock, and thus can improve this notion of liquidity imbalance.

formation because hot weather can significantly reduce the productivity of firms traded on the exchange. The weather effect on the agricultural industry, with its tiny representation in the exchange, is less of a concern. The railroad industry, however, was an important part of NYSE in the pre-1903 sample and may have had reduced productivity on hotter days. Railroad development in the late 19th century nonetheless focused on the recent settlements in the West and the transcontinental connection between the Pacific Coast and the eastern states.²⁴ Temperature in Manhattan and nearby areas, with the railroad system already in place, would not affect the railroad stocks. Consistent with this notion, a rise in unseasonal or seasonal temperature in the summers of 1889-1902 has no effect on NYSE index returns.

This study's instrumental variable approach may seem unusual, and this is because I use temperature not as an instrument for outside investor trades, which is unobserved, but as an instrument for the aggregate trading volume. One way to understand this approach is that hot summer weather is an instrument for aggregate trading volume with outside investors as the compliers of the treatment effect. In a typical study where identifying the population average effect is the objective, employing an instrumental variable with only the local average treatment effect is a problem (Angrist and Imbens 1995). In my case, however, the objective is precisely to identify the local effect coming from the outside investors. In any case, my main approach is to use temperature as a proxy variable, and the instrumental variable regressions can be understood as a way to obtain coefficients that measure marginal effects of an increase in outside investors' trade volume and thus are easier to interpret than the effects in units of temperature.

3.4 Model specifications

In section 4, I study the outside investors' role in shaping the daily aggregate stock market using various regressions. First, I estimate the trading volume model specified in (9) to see if a unseasonal rise in summer temperature in Manhattan significantly reduces (economically and statistically) NYSE trading volume, a necessary condition for my approach to be valid. Then, I estimate the effect of outside investors on NYSE aggregate liquidity imbalance, using temperature to proxy for outside investor presence. After confirming that outside investors indeed lower aggregate market liquidity, I then study how they affect the behavior of NYSE index returns. As we will see, I use similar control variables in all these regressions, as the possibility of reverse causality forces me to include control variables for liquidity imbalance in the volume regression.

²⁴“Railroads in the Late 19th Century” on the Library of Congress website (<http://www.loc.gov/teachers/classroommaterials/presentationsandactivities/presentations/timeline/riseind/railroad/>), retrieved on December 6, 2015.

Rewriting the model of trading volume in (9) with simpler notation gives

$$Tradvol_t = b_0 + b_1 \tilde{T}_t + X_{TV,t} \mathbf{b}_{TV} + X_{L,t} \mathbf{b}_L + \varepsilon_t \quad (11)$$

where T_t is Manhattan temperature, $X_{TV,t}$ is a vector of control variables for aggregate volume, $X_{L,t}$ is a vector of control variables for aggregate liquidity imbalance, and innovations in trade volume due to information ($Informed_t$) is included in the error term ε_t . For the vector of control variables for volume, I use seasonal temperature (\bar{T}), day of the week dummies (\mathbf{D}_t), and dummy variables indicating the number of trading holidays until the next trading day ($\mathbf{G}_t = 1$ by 3 vector of dummy variables representing 1 day's gap, 2 day's gap, and 3 or more days' gap).

Seasonal temperature is used because even within each summer, I find trading volume to be lower in hotter weeks. Day of the week dummies are used in Gallant, Rossi, and Tauchen (1992) to control for the predicted patterns in trading volume within a week, and a variant of trading days' gap dummies are used in the same paper to control for the effect of variation in the number of trading days between two consecutive trading days. While Gallant, Rossi, and Tauchen (1992) construct dummy variables based on the number of trading holidays since the previous trading day, I construct dummy variables based the number of trading holidays until the *next* trading day and find that it has greater explanatory power. Using the alternative specification of \mathbf{G}_t has little effect on results.

In the vector of control variables for liquidity imbalance, I include return volatility ($Volatility_t$) and an indicator for negative market ($\mathbf{1}(R_t < 0)$). Return volatility is an important determinant of liquidity imbalance because liquidity suppliers tend to be more capital-constrained in volatile times (Nagel 2012). I estimate volatility using a generalized autoregressive heteroskedasticity (GARCH) model with two moving average and two autoregressive terms²⁵ I include the negative return dummy because the volume-autocorrelation relation is known to be stronger for positive returns (LeBaron 1992), which suggests that the relative size of outside investors ($Outside_t$) and informed traders ($Informed_t$) may differ in positive and negative markets.

In summary, I estimate the following model of daily trading volume:

$$Tradvol_t = b_0 + b_1 \tilde{T}_t + b_2 \bar{T}_t + b_3 V_t + b_4 \mathbf{1}(R_t < 0) + \mathbf{D}_t b_5 + \mathbf{G}_t b_6 + b_{7,y} year_t + \varepsilon_t \quad (12)$$

where b_1 is the coefficient of interest with $b_1 < 0$ as the prediction and $year_t$ indicates the year fixed effect. I include year fixed effects to ensure that I only use temperature variation within each

²⁵Specifically, it is the conditional volatility obtained from the simple autocorrelation regression $R_{m,t+1} = \alpha + (\beta_0 + \beta_1 \mathbf{1}(R_{m,t} < 0) + \mathbf{G}_t \beta_2) R_{m,t} + \varepsilon_t$, where $\sigma_t^2 = a_0 + \sum_{s=1}^2 a_s \varepsilon_{t-s}^2 + b \sigma_{t-1}^2$ and \mathbf{G}_t is the vector of dummy variables indicating the number of trading holidays until the next trading day. This control variable is motivated by findings that trading volume represents a disagreement, which may widen during times of high volatility.

summer.

Next, I specify the model of NYSE aggregate liquidity imbalance and daily index reversal. Because theory suggests that short-run return reversals arise due to liquidity imbalances, I specify both regressions using the model of aggregate liquidity specified in (8), which I rewrite as

$$Imbalance_t = \beta_0 + \beta_1 \tilde{T}_t + X_{L,t} \mathbf{B}_L + \varepsilon_t \quad (13)$$

where $X_{L,t}$ is the vector of control variables. I specify $X_{L,t}$ using all control variables used in the volume regression and add aggregate trading volume ($Tradvol_t$) as an additional control. Trading volume is included to capture innovations in outside investor trades or informed trades not captured by other variables. Day of the week dummies are included because they are reported to be significant in the CGW specification of return autocorrelation regression. The trading days' gap dummies are included because the larger the gap until the next trading day, the less persistence in market index returns we would expect. The justification for volatility and the negative return indicator is the same as above.

In summary, I use the following specification to estimate the effect of outside investors on aggregate liquidity imbalance and index return reversal:

$$LHS_t = \beta_0 + \beta_1 \tilde{T}_t + \beta_2 \bar{T}_t + \beta_3 Tradvol_t + \beta_4 V_t + \beta_5 \mathbf{1}(R_t < 0) + \mathbf{D}_t \beta_6 + \mathbf{G}_t \beta_7 + \beta_{8,y} year_t + \varepsilon_t \quad (14)$$

where LHS_t indicates the dependent variable of interest. As I do in the volume regression, I use year fixed effects to ensure that variation in temperature across different years do not affect the regression results. I would expect the estimated β_1 to be positive.

4 Main Empirical Results

4.1 Effect of Manhattan temperature on NYSE trading volume

Lower trading activities in hotter summers

Although the causal link I need to establish is the daily temperature variation causing variation in trading volume on the same day, I begin by presenting evidence that the temperature-volume relationship existed at the yearly frequency. To study the effect, I compute NYSE summer trading as the average daily NYSE trading volume (before detrending) over the summer divided by its annual average, and regress it on average Manhattan summer temperature:

$$NYSE\ Summer\ Trading_t = b_0 + b_1 Manhattan\ Summer\ Temperature_t + \varepsilon_t \quad (15)$$

where t indexes a year.

As shown in Table 2, in the earliest sample period, years with hotter summers had lower trading activity in the summer (an effect statistically significant at 10%).

[Table 2 here]

A smaller yet significant effect exists in the 1903-1930 sample, but not in the latter sample periods. During 1889-1902, a 1 degree Celsius increase in the summer temperature led to a large 11 percentage point drop in NYSE's relative summer trading. This strengthens the evidence that hot weather discouraged trading activity in the late 19th century. Figure 3 visualizes this temperature-volume relationship in 5 sample periods.

[Figure 3 here]

When the regression for the first sample period is repeated with temperature in other cities, no significant result is obtained. Table 3 shows that the effect is negative only in Cambridge MA, although the effect is not significant at the 10% level, illustrating the importance of Manhattan temperature to the participants of NYSE.

[Table 3 here]

Daily effect of Manhattan temperature on NYSE total trading volume

The hypothesized relationship between temperature and trade volume in pre-1903 NYSE is stronger at a daily frequency. To estimate the daily temperature effect on volume, I regress log detrended trading volume ($Tradvol_t$) on unseasonal Manhattan temperature (\tilde{T}_t) and other controls. As shown earlier, summer temperature and average trading activity in the summer have a negative relationship at a yearly frequency. Although this relationship is also likely to be capture the temperature effect, I take the conservative approach of including year fixed effects to exploit only the within-year variation in trading volume and temperature. I also report the result where I included lagged unseasonal temperature (\tilde{T}_{t-1}) in the regression but do not find a significant effect once \tilde{T}_t has been controlled for. The result is the following model of daily trade volume discussed in the previous

section:

$$Tradvol_t = b_0 + b_1\tilde{T}_t + b_2\bar{T}_t + b_3V_t + b_4\mathbf{1}(R_t < 0) + \mathbf{D}_t b_5 + \mathbf{G}_t b_6 + b_{7,y}year_t + \varepsilon_t \quad (16)$$

I estimate this model in each of the five sample periods using the summer season, and report the estimated coefficients as well as Newey-West HAC standard errors with 20 lags in Table 4. The coefficient on unseasonal temperature \hat{b}_1 is the correct measure of temperature effect on trading volume, as the coefficient on seasonal temperature \bar{T} likely contains both the temperature effect and seasonal effects.

[Table 4 here]

The temperature effect is economically and statistically significant in the test sample, with a rise in daily Manhattan summer temperature by 1 degree Celsius (1.8 degrees Fahrenheit) reducing NYSE trading volume by 1.3% on average (dependent variable is in log). The effect is smaller but significant in the second sample but disappears in the last three sample periods, suggesting that the temperature effect on trading volume had gradually disappeared.²⁶ Seasonal temperature is highly significant in the most recent sample period, but given that unseasonal temperature has no effect, it should be interpreted as hotter times in the summer having higher trading activity for seasonal rather than temperature-related reasons. Trading volume significantly rises during volatile periods in all samples. Trading volume during a negative market tends to be lower in 1931-60 and 1961-90 samples but higher in the most recent sample.

A placebo test using temperatures in other cities

To rule out the possibility that the impact of Manhattan temperature on trading volume is a result of a spurious regression (e.g., Ferson, Sarkissian, and Simin 2003), I repeat the estimation in the first sample period using temperatures in other cities (Table 5). The correlations in the temperature (unseasonal and seasonal) of Manhattan and another city are reported at the bottom of the table.

[Table 5 here]

As expected, only the unseasonal temperature in Cambridge, MA has explanatory power, presumably due to its strong correlation with unseasonal temperature in Manhattan (0.75). Despite

²⁶Although not reported here, I also find that trading volume falls on snowy days in the first sample period and that low winter temperature does not affect trading volume.

the low statistical significance, however, unseasonal temperatures in both U.S. cities have similar economic magnitudes (-0.010) as Manhattan’s unseasonal temperature (-0.013).

The large magnitudes in the placebo tests disappear once I include Manhattan temperatures in the regressions. In Table 6, I run a horse race between unseasonal temperature in Manhattan and that in other cities by including Manhattan temperatures as additional explanatory variables in the placebo regressions.

[Table 6 here]

In all regressions, only the unseasonal temperature in Manhattan retains explanatory power. The magnitudes of other U.S. cities’ temperature effects are now smaller, confirming that the previously large magnitudes came from their correlation with Manhattan temperature.

Historical evidence discussed in section 2 suggests that the temperature effect on trading volume prevented here can be interpreted as hot weather discouraging outside investors from participating in the daily market. Thus, in the next three subsections, I use Manhattan unseasonal temperature either as a proxy variable for outside investor trades or as an instrumental variable for volume with outside investors as the compliers of the treatment effect.

4.2 Outside investor trades induce aggregate liquidity imbalance

How does the reduction in outside investors due to hot weather affect aggregate liquidity imbalance? Using the shift in temperature that affects the participation of outside investors as a whole, I explore whether a smaller presence of outside investors in the daily market leads to smaller liquidity imbalance, measured by lower than average returns from aggregate liquidity provision.

Returns from aggregate liquidity provision is measured in four ways, one using index returns and three using the cross-section of individual stock returns. In all of the four measures, the return on liquidity provision is defined as the one-day return from a portfolio of individual stocks that mimics the daily positions taken by market makers:

$$L_t^s = \sum_{i=1}^N \omega_{i,t}^s R_{i,t+1} \quad (17)$$

where $i = 1, \dots, N$ indexes the individual stocks, $s \in \{0, 1, 2, 3\}$ indexes the four strategies, and $\omega_{i,t}^s$ is the time- t position on stock i taken by strategy s . I focus on returns over a day because, as stated in Nagel (2012), the half-life of dealer inventory holding is estimated to be 1/2 to 2 days (Hansch, Naik, and Viswanathan 1998; Hendershott and Menkveld 2013).

For each measure of return on liquidity provision L_t^s , I study whether outside investors increase the expected return, using unseasonal temperature as the proxy variable. Both here and in the return reversal regressions in section 4.3, the control variables are the same as those used in volume regressions except for the addition of $Tradvol_t$ (see section 3.4 for model specifications). The regression specification is thus

$$L_t^s = \beta_0 + \beta_1 \tilde{T}_t + \beta_2 \bar{T}_t + \beta_3 Tradvol_t + \beta_4 V_t + \beta_5 \mathbf{1}(R_t < 0) + \mathbf{D}_t \beta_6 + \mathbf{G}_t \beta_7 + \beta_{8,y} year_t + \varepsilon_t \quad (18)$$

where V_t measures volatility, \mathbf{D}_t are the day of the week dummies, and \mathbf{G}_t are the trading days' gap forward dummies. Year fixed effects are used so that only the effects within each year are identified; this is to ensure that low-frequency variation in temperature over years does not drive the regression results.

Using returns from a simple reversal strategy on the market index

The first specification of returns from aggregate liquidity provision uses the market index return itself. The strategy goes long 1 dollar of S&P 500 ETF if the market went down today and short 1 dollar of S&P 500 ETF if the market went up today. Thus, this reversal strategy return mimics the one-day return earned by an investor who takes on the same dollar-exposure to a market index each day but in the direction opposite to the market movement on that day. The position on each individual stock taken by this first strategy is

$$\omega_{i,t}^0 = -\omega_{i,t}^m \mathbf{1}(R_{m,t} > 0) \quad (19)$$

Although a more accurate replication of positions taken by market makers requires the cross-section of individual stock returns, this measure of market liquidity has the advantage that it can be calculated in periods when individual stock price data are unavailable (1903-1925). Furthermore, because larger stocks have greater influence on returns from this strategy, analyzing this first reversal strategy addresses the concern that the three cross-sectional strategies implicitly assume equal sizes of all stocks and thus may not accurately describe market maker's positions in this regard.

Table 7 reports the regression results in 5 different sample periods based on the proxy variable approach as well as the instrumental variable regression in the test sample.

[Table 7 here]

Column (1) shows that in the first sample period, an increase in unseasonal temperature decreases

the return from this reversal strategy. Interpreting a rise in unseasonal temperature as a proxy for the reduction in outside investors, this implies that a smaller presence of outside investors in the market lowers the market maker's required return on accommodating non-informational trades. The instrumental-variable regression allows for an interpretation in units of volume. Column (2) shows that an increase in outside investors equivalent to 1% of the average daily trade volume leads to a 1.8 basis point increase in the return from the reversal strategy. The 1.8 basis point increase in the return represents a 1.8% standard deviation increase in the return, as the standard deviation of this strategy return is 1.02% in the summers of 1889-1902.

Volatility increases the return from liquidity provision in the first sample, consistent with the notion that higher volatility leads to a lower supply of liquidity so that each liquidity provider needs to be compensated at a higher rate. The sign of the volatility effect, however, flips in later periods. This is similar to the finding of CGW that once volume has been controlled for, volatility sometimes has a negative impact on return reversal. The statistically significant effect of temperature during 1961-1990 is puzzling. Interestingly, the effect is stronger during the period of heightened energy prices (1971-1985) but insignificant in other years of the sample period. This suggests that the differential temperature effect on outside investors and market makers at the NYSE may have existed during the energy crisis, although further examination is required to validate this conjecture. A similar effect shows up in the study of the likelihood of index return reversal and index return autocorrelation. For now, I continue to use the pre-1903 sample as the test sample.

Using returns from aggregate liquidity provision based on cross-section of daily returns

The previous measure of market maker position does not use the rich information contained in the cross-section of individual stock prices. Accordingly, I repeat the previous analysis using three measures of return on aggregate liquidity provision based on individual stock price data. The three cross-sectional reversal strategies ($s = 1, 2, 3$) are used in Nagel (2012) and take the following positions on each individual stock:

$$\omega_{i,t}^1 = -\frac{R_{i,t-1} - R_{m,t-1}^{EW}}{\frac{1}{2} \sum_{i=1}^N |R_{i,t-1} - R_{m,t-1}^{EW}|} \quad (20)$$

$$\omega_{i,t}^2 = -\frac{R_{i,t-1} - R_{m,t-1}^{EW}}{N} \quad (21)$$

$$\omega_{i,t}^3 = -\frac{R_{i,t-1} - R_{m,t-1}^{EW}}{\sum_{i=1}^N (R_{i,t-1} - R_{m,t-1}^{EW})^2} \quad (22)$$

where $R_{m,t-1}^{EW} = N^{-1} \sum_{i=1}^N R_{i,t-1}$ is the equal-weighted market index return between $t - 1$ and t .²⁷

The three strategies use the same relative allocation of funds to the cross-section of stocks but use different rules to allocate funds across time. The three strategies' identical numerators indicate that the relative allocation of funds across stocks depends only on the deviation of individual stock return from the market return—the lower the individual stock return compared to the market return, the higher the strategy's long exposure to the stock, and vice versa. Differences come from the denominator. In the first strategy (Lehman 1990), a dollar of either long or short position is assumed to require the same 50 cents, implying a constant margin of 1/2 across time and across all stocks (Nagel 2012). This implies that the market maker's asset position in the strategy is also constant over time, unaffected by time-varying volatility. The second strategy (Lo and MacKinlay 1990) assumes a more speculative market maker. Because the denominator is constant, the market maker's position in the entire strategy increases with the cross-sectional dispersion of stock returns. As one will see, this speculative nature of the strategy makes its expected return increase with market volatility. The third strategy assumes a market maker faced with a form of time-varying margin requirement (margin requirement falls with volatility). Because each exposure of $R_{i,t-1} - R_{m,t-1}^{EW}$ requires its squared value to be expended by the market maker, the strategy is less dependent on extreme deviations in individual stock return from the average return.

Nagel (2012) shows that the three strategies also exhibit different levels of sensitivity to the informativeness that the liquidity imbalance has about the asset value as well as the sensitivity to the change in volatility of public information. He finds that the second strategy is insensitive to the change in volatility of public information, whereas the third strategy is less sensitive to the variation in the liquidity imbalance's informativeness; the first strategy strikes a balance between the two sensitivities.

Regress each of the three strategy returns on unseasonal temperature and other controls, I find that a rise in unseasonal temperature lowers the expected returns from liquidity provision (Table 8).

[Table 8 here]

Based on the first strategy, a reduction in outside investors associated with an unseasonal rise in Manhattan temperature by 1 degree Celsius decreases the expected return from aggregate liquidity provision by 4.9 basis points, or 2.0% of the standard deviation.²⁸ This suggests that a smaller presence of outside investors reduces liquidity imbalance. Columns (2), (4), and (6) report results

²⁷To equalize the scales, $\omega_{i,t}^2$ is multiplied by 100, and $\omega_{i,t}^3$ is divided by 10 before forming portfolio returns.

²⁸The standard deviation of this return over the first sample is 2.36%.

from the regression that uses unseasonal temperature as an instrumental variable for trading volume with outside investors as the compliers. Column (2) shows that an increase in outside investor trades by 1% of the average daily trading volume increases the expected return from liquidity provision by 2.9 basis points, or 1.2% of the standard deviation.

The effect of seasonal temperature (\bar{T}) on returns from liquidity provision tends to be small or has a sign opposite to that in unseasonal temperature, suggesting that its seasonal effect may be offsetting any of its temperature effect. A high trade volume is associated with smaller liquidity imbalance, and this remains true (i.e. the coefficient stays negative) even when temperature variables are excluded, consistent with the observation that reversal profitability based on large NYSE and AMEX stocks falls with trading activity (Cooper 1999). This reflects that trading based on private information (Wang 1994) or public information (Kandel and Pearson 1995; Bamber, Barron, and Stober 1999) mitigates the price pressure created by non-informational trades. Volatility tends to decrease liquidity, consistent with the finding that the supply of liquidity falls during volatile periods (Nagel 2012). The volatility effect is largest in strategy 2 and smallest in strategy 3, consistent with the earlier discussion that strategies 2 and 3 represent the strategies of aggressive and conservative market makers, respectively.

Do outside investors affect reversal profitability through the number of stocks with a return reversal or the magnitude of returns?

The baseline regression reported in column (1) of Table 8 can be examined further. It shows that unseasonal temperature reduces reversal profitability, but through which channel does this occur? That is, do outside investors increase reversal profitability by (a) increasing the number of stocks having a reversal, (b) increasing the absolute return on stocks that reverse, or (c) decreasing the absolute return on stocks that continue? If outside investors flock into all stocks in general, the first and third would be true. If outside investors tend to chase after a few stocks having large price changes, the second would be true. Similarly, does trade volume reduce reversal profitability by reducing the number of stocks with a reversal on the next day, reducing the absolute return on stocks that reverse, or increasing that on stocks that continue?

To answer these questions, in Table 9, I regress the following three measures of reversal patterns in NYSE individual stocks on unseasonal temperature, trading volume, and other controls: fraction of stocks with a return reversal next day (Panel A); mean absolute return on stocks with a return reversal next day (Panel B); and mean absolute return on stocks with a return continuation next day (Panel C).

[Table 9 here]

The effect of unseasonal temperature is negative and significant, suggesting that a reduction in outside investors lowers the fraction of stocks having a reversal on the next day. The instrumental-variable regression suggests that when outside investor trades increase by 1% of the daily trading volume, the fraction of stocks reversing the next day falls by 0.2 percentage points. Neither volume nor volatility, however, has a significant effect on this fraction. The effect of volatility here can be linked to the effect of liquidity supply. As discussed in Nagel (2012), high volatility reduces the market makers' capacity to provide liquidity. That a higher volatility does not significantly reduce the number of stocks having a reversal is consistent with the intuition that liquidity suppliers do not determine whether the stock is subject to non-informational price pressure.

How do unseasonal temperature, volume, and volatility affect the equal-weighted average absolute return on stocks whose returns reverse on the next day (Panel B of Table 9)? The results contrast those in the previous regression explaining the fraction having reversals. Outside investors, proxied by unseasonal temperature, do not significantly increase the magnitude of reversals, although the negative sign indicates a weak positive effect (unseasonal temperature and outside investor trades are negatively related). Both trade volume and volatility, on the other hand, significantly increase the absolute returns on reversing stocks. If trading volume increases the absolute returns on stocks that reverse, why does it appear to reduce reversal profitability?

The answer lies in the effect of volume on absolute returns on stocks that continue the previous day's return. In Panel C of Table 9, the average absolute return on stocks that continue the previous day's returns is regressed on unseasonal temperature and other controls. Interestingly, the reduction in outside investors associated with a rise in unseasonal temperature increases the mean absolute returns on continuation stocks. This suggests that in the absence of outside investors, private information traders may find it difficult to trade without releasing information through their price pressure, leading them to conduct more trades and create higher price pressure in the same direction on the next day. This is consistent with the theory that price reflects information and having a large number of noise traders in the market helps informed investors trade without leaking much information (Glosten and Milgrom 1985; Kyle 1985). Furthermore, trading volume has a larger coefficient than in Panel B, suggesting that the negative coefficient found in Table 8 is caused by a higher trading volume increasing the continuation stocks' absolute returns more than reversing stocks'. As will be seen later, a higher trading volume has a positive effect on index reversal. This suggests that a higher trading volume does increase the absolute return on reversing stocks more than that of continuation stocks but that its affect on reversing stocks may be more concentrated in large stocks.

Using four specifications of market make's expected returns, I showed that lower demand for liquidity associated with higher temperature improves aggregate market liquidity. I also presented evidence that smaller presence of outside investors associated with higher temperature affects the

reversal profitability through two channels: a reduction in the number of stocks having a reversal on the next day and an increase in the absolute return on stocks having a return continuation on the next day. Also, the negative overall effect of trade volume on reversal profitability results from higher trade volume increasing the absolute return on continuation stocks more than that of reversal stocks. Next, I examine how such an improvement in liquidity affects the daily behavior of market index returns.

4.3 Outside investor trades induce market index return reversals

I study the short-horizon reversal phenomenon in market index returns in two ways. First, I study the likelihood of a next-day return reversal in the market index. Despite the information loss associated with discretizing the market return behavior into a binary event of reversal and non-reversal, this approach has the advantage that it is less subject to the influence of outlier returns. Next, I study the daily autocorrelation in the market index as in CGW. Here, I winsorize all daily returns at the 5% level to prevent outliers from dominating the regression outcome (e.g., Pinegar 2002). Before analyzing the two measures of reversal using the temperature proxy, I discuss their general patterns during the test sample and during the period in which S&P 500 returns are available (1926-2014).

Market return reversals and autocorrelations in 1889-1902 and 1926-2014

The probability of daily return reversal in the market index was 48% in 1889-1902 and 46% in 1926-2014, suggesting that continuation was slightly more common than reversal in the market index return over two consecutive trading days (Table 10).

[Table 10 here]

This is also reflected in the positive autocorrelation in the market index in the two sample periods. These positive daily autocorrelations in the market index are consistent with the observation that, despite the predominantly negative daily autocorrelation in individual stocks, the cross-autocorrelation among individual stocks bring the overall market autocorrelation to a positive value. Froot and Perold (1995), however, document that this autocorrelation in the market index has fallen since the 1970s, with the better dissemination of market-wide information and the use of futures contracts.

How does the market index behave differently during more volatile times? Table 10 reports the same statistics during above-median volatility times in each sample period. Both the prob-

ability of reversal and return autocorrelation indicate that the market index return is more likely to reverse in more volatile times. This may be due to a combination of higher trading activity of outside investors (i.e. higher demand for liquidity) and market makers requiring higher return on liquidity provision (i.e. lower supply of liquidity). I now study how the probability of index return reversal and index return autocorrelation are affected by aggregate liquidity demand, proxied by the exogenous variation in Manhattan's unseasonal temperatures.

Likelihood of a next-day return reversal in the market index

To study the outside investors' affect on reversal probability, I define $Reversal_t \equiv \mathbf{1}[R_{m,t}R_{m,t+1} < 0]$ as the indicator of return reversal so that $Reversal_t = 1$ indicates a return reversal and $Reversal_t = 0$ indicates no return reversal on the next trading day, where $R_{m,t}$ denotes the daily market index return. I then estimate the following logit model of the index reversal event:

$$\ln\left(\frac{\Pr[rev_t=1]}{\Pr[rev_t=0]}\right) = \beta_0 + \beta_1 \tilde{T}_t + \beta_2 \bar{T}_t + \beta_3 Tradvol_t + \beta_4 V_t + \beta_5 \mathbf{1}(R_t < 0) + \mathbf{D}_t \beta_6 + \mathbf{G}_t \beta_7 + \beta_{8,y} year_t + \varepsilon_t \quad (23)$$

The explanatory variables are unseasonal Manhattan temperature and other control variables used when studying liquidity imbalance, and I refer the reader to the discussions in 4.1 and 4.2 for the choice of controls. To interpret the effect of outside investors in units of trading volume, I also use unseasonal temperature as an instrument for trading volume and estimate the impact of predicted volume on reversal likelihood using a linear probability model. As pointed out in Wooldridge (2002), this is an acceptable approach in estimating a discrete response model with an endogenous explanatory variable and avoids making strong assumptions required to estimate a probit model with an endogenous explanatory variable.

Columns 1 through 3 of Table 11 report the determinants of the likelihood of a NYSE index return reversal during 1889-1902.

[Table 11 here]

The first two columns are estimated using the logit model, whereas the third column uses the two-stage least squares method that assumes the linear probability model. A reduction in outside investors associated with a 1 degree Celsius increase in Manhattan temperature reduced the probability of reversal by 1.6 percentage points (column 1). In units of trade volume, a reduction in outside investors by 1% of the NYSE trading volume increased the probability of a reversal by 1.1 percentage points (column 3).

The temperature effect on reversal probability does not exist in later periods except in 1961-1990, when the effect was significantly negative at the 10% level. Although positive in all periods, the effect of trading volume on the likelihood of a reversal is not consistently significant and is not significant in the test sample period. When the proxy for outside investors is excluded, an increase in trade volume is estimated to have a larger effect on the likelihood of a reversal (column 2). Volatility does not have a consistent direction of effect, as previously noted in CGW. In all sample periods, reversal likelihood is much higher during negative markets, which may reflect the short-term overreaction to negative news and underreaction to positive news (e.g., Braun, Nelson, and Sunier 1995; Veronesi 1999; Avramov, Chordia, and Goyal 2006b).

That non-informational trades by outside investors increase the likelihood of a market return reversal is consistent with the theoretical prediction and empirical findings of CGW. Similar evidence is found in autocorrelations.

Daily return autocorrelation in market index returns

To study how outside investors affect the persistence of market index returns, I regress the next trading day's index return on unseasonal temperature and control variables interacted with today's return:

$$R_{m,t+1} = \alpha_0 + \left(\beta_0 + \beta_1 \tilde{T}_t + \beta_2 \bar{T}_t + \beta_3 Tradvol_t + \beta_4 V_t + \beta_5 \mathbf{1}(R_t < 0) + \mathbf{D}_t \beta_6 + \mathbf{G}_t \beta_7 + \beta_{8,y} year_t \right) \times R_{m,t} + \varepsilon_t \quad (24)$$

This is the approach used in CGW and allows the explanatory variables to affect the daily autocorrelation coefficient. I report the regression results based on market returns winsorized at 5% level (Table 12).

[Table 12 here]

Column (1) shows that a rise in unseasonal temperature increases the persistence of daily index returns. The effect is large, with an increase in unseasonal temperature by 1 degree Celsius lowering the autocorrelation coefficient by 0.019 on average. Column (3) quantifies this effect in terms of outside investors' trading volume by instrumenting the term $TradVol_t \times R_t$ through $\tilde{T}_t \times R_t$.²⁹ It suggests that an increase in outside investors equivalent to 1% of the average daily trading volume decreases the autocorrelation coefficient by 0.018. This quantity can be compared to the effect

²⁹Although this implies that the moment condition will involve the term $R_t^2 \tilde{T}_t$, the approach is still valid if temperature does not affect squared return—or volatility—in a systemic way.

found in CGW, who finds that the same 1% rise in trading volume lowers the autocorrelation of S&P 500 daily returns by 0.004 over the sample period 7/3/62-9/30/87 (Table II of CGW). This suggests that information-driven aggregate trading volume may have a positive effect on daily autocorrelation of market index returns, counteracting non-informational trading volume's negative effect on the autocorrelation. A similar conclusion holds when column (3) is compared to column (2), which repeats CGW's autocorrelation regression in the test sample.

Looking at other sample periods, I find that unseasonal temperature has a large and significant effect on autocorrelation in 1961-1990, a pattern observed in the study of the simple reversal strategy on the NYSE index and in the study of the reversal probability. Trading volume has a negative effect and volatility has mixed effects on autocorrelation, consistent with CGW. Finally, autocorrelation falls significantly during a negative market.

To summarize, a reduction in outside investors affects the behavior of daily market index returns by lowering the likelihood of a market index return reversal and increasing the daily autocorrelation coefficient. The effect is economically and statistically significant, as shown in the interpretation through predicted trading volumes. Together with section 4.2, this establishes that outside investors' demand for liquidity is an important cause of short-run reversals in the daily market returns.

5 Conclusion

In this paper, I suggest a natural experiment in the late 19th century in which hot daily weather removed a sizable quantity of outside investors from the daily stock market. This experiment helps establish a causal link between outside investors' non-informational trades and short-horizon reversals in stock market indices through the channel of aggregate liquidity imbalance. The experiment also provides an insight on how noise trading and informed trading interact.

The paper's focus on the sample period 1889-1902 allows for a causal inference, but it comes at the cost of having to take extra care in drawing implications on today's market, which I briefly discuss. Perhaps the most important change from the test sample period is that the supply of liquidity is likely to be more flexible in today's market.³⁰ Hence, the same fluctuation in liquidity demand due to non-informational trades may lead to a smaller change in the aggregate liquidity imbalance or market return reversal today than what the estimated coefficients imply. In fact, this may explain why we now observe less short-horizon reversals in the market index than before (e.g., Froot and Perold 1995; also see Table 10).

³⁰This may be due to a combination of factors such as the information revolution, automated market making, and more integrated asset markets.

The suggested natural experiment in the 19th century can be used to explore other topics in asset pricing. For example, combining the individual stock price data based on daily high and low prices with the temperature proxy for outside investors, one can test theories of bid-ask spreads or volatilities. This would be possible using the range-based measures of bid-ask spreads (Corwin and Schultz 2012) and volatility (Alizadeh, Brandt, and Diebold 2002).

References

- [1] Angrist, Joshua D. and Guido W. Imbens, 1995. Identification and estimation of local average treatment effects. NBER Technical Working Paper, No. 118.
- [2] Alizadeh, Sassan, Michael W. Brandt, and Francis X. Diebold, 2002. Range-based estimation of stochastic volatility models. *Journal of Finance* 57(3), 1047-1091.
- [3] Avramov, Doron, Tarun Chordia, and Amit Goyal, 2006a. Liquidity and autocorrelations in individual stock returns. *Journal of Finance* 61(5), 2365-2394.
- [4] Avramov, Doron, Tarun Chordia, and Amit Goyal, 2006b. The impact of trades on daily volatility. *Review of Financial Studies* 19(4), 1241-1277.
- [5] Bamber, Linda Smith, Ori E. Barron, and Thomas L. Stober, 1999. Differential interpretations and trading volume. *Journal of Financial and Quantitative Analysis* 34(3), 369-386.
- [6] Bansal, Ravi, and Marcelo Ochoa, 2011. Welfare costs of long-run temperature shifts. Duke University, working paper.
- [7] Bansal, Ravi, Dana Kiku, and Marcelo Ochoa, 2015. Climatic change and growth risks. Duke University, working paper.
- [8] Barber, Brad M., and Terrance Odean, 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21(2), 785-818.
- [9] Barber, Brad M., Terrance Odean, and Ning Zhu, 2009. Do retail trades move markets?. *Review of Financial Studies* 22(1), 151-186.
- [10] Bassi, Anna, Riccardo Colacito, and Paolo Fulghieri, 2013. 'O sole mio: an experimental analysis of weather and risk attitudes in financial decisions. *Review of Financial Studies* 26(7), 1824-185.
- [11] Beckert, Sven, 2003. *The monied metropolis: New York City and the consolidation of the American bourgeoisie, 1850-1896*. Cambridge: Cambridge University Press.
- [12] Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2012. Hedge fund stock trading in the financial crisis of 2007–2009. *Review of Financial Studies* 25(1), 1-54.
- [13] Bouman, Sven, and Ben Jacobsen, 2002. The Halloween indicator, "Sell in May and go away": Another puzzle. *American Economic Review* 92(5), 1618-1635.

- [14] Buchanan, Matt, 2013. An apparatus for treating air: The modern air conditioner. *The New Yorker*, Jun. 5, 2013. <http://www.newyorker.com/tech/elements/an-apparatus-for-treating-air-the-modern-air-conditioner>.
- [15] Braun, Phillip A., Daniel B. Nelson, and Alain M. Sunier, 1995. Good news, bad news, volatility, and betas. *Journal of Finance* 50(5), 1575-1603.
- [16] Brunnermeier, Markus K. and Lasse Heje Pedersen, 2009. Market liquidity and funding liquidity. *Review of Financial studies* 22(6), 2201-2238.
- [17] Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993. Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics* 108(4), 905-939.
- [18] Cao, Melanie, and Jason Wei, 2005. Stock market returns: A note on temperature anomaly. *Journal of Banking and Finance* 29(6), 1559-1573.
- [19] Chen, Joseph, Harrison Hong, and Jeremy C. Stein, 2001. Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices. *Journal of Financial Economics* 61, 345-381.
- [20] Conrad, Jennifer S., Allaudeen Hameed, and Cathy Niden, 1994. Volume and autocovariances in short-horizon individual security returns. *Journal of Finance* 49(4), 1305-1329.
- [21] Corwin, Shane A., and Paul Schultz, 2012. A simple way to estimate bid-ask spreads from daily high and low prices. *Journal of Finance* 67(2), 719-760.
- [22] Dell, Melissa, Benjamin Jones, and Benjamin Olken, 2009. Temperature and income: reconciling new cross-sectional and panel estimates. *American Economic Review Papers and Proceedings* 99(2), 198-204.
- [23] Dell, Melissa, Benjamin Jones, and Benjamin Olken, 2012. Temperature shocks and economic growth: evidence from the last half century. *American Economic Journal: Macroeconomics* 4(3), 1866-95.
- [24] Dell, Melissa, Benjamin Jones, and Benjamin Olken, 2014. What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature* 52(3), 740-798.
- [25] De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990. Noise trader risk in financial markets. *Journal of political Economy* 98(4), 703-738.

- [26] Deschenes, Olivier, Michael Greenstone, 2007. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *American Economic Review*, 97(1), 354-385.
- [27] Donnan, Chris, 2011. A brief history of market data. Presentation at the European Financial Information Summit 2011. http://www.financialinformationsummit.com/digital_assets/6006/Chris_Donnan.pdf.
- [28] DuTemple, Lesley A., 2002. *The New York Subways*. Minneapolis: Lerner Publications Company.
- [29] Froot, Kenneth A., and André F. Perold, 1995. New trading practices and short-run market efficiency. *Journal of Futures Markets* 15(7), 731-765.
- [30] Gallant, A. Ronald, Peter E. Rossi, and George Tauchen, 1992. Stock prices and volume. *Review of Financial Studies* 5(2), 199-242.
- [31] Glosten, Lawrence R., and Paul R. Milgrom, 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14(1), 71-100.
- [32] Goetzmann, William N., and Ning Zhu, 2005. Rain or shine: where is the weather effect?. *European Financial Management* 11(5), 559-578.
- [33] Greenwood, Robin. Excess comovement of stock returns: Evidence from cross-sectional variation in Nikkei 225 weights. *Review of Financial Studies* 21(3), 1153-1186.
- [34] Grossman, Sanford J., and Merton H. Miller, 1988. Liquidity and market structure. *Journal of Finance* 43(3), 617-633.
- [35] Harris, Lawrence, and Eitan Gurel. Price and volume effects associated with changes in the S&P 500 list: New evidence for the existence of price pressures. *Journal of Finance* 41(4), 815-829.
- [36] Hellwig, Martin F., 1980. On the aggregation of information in competitive markets. *Journal of Economic Theory* 22(3), 477-498.
- [37] Hirshleifer, David, and Tyler Shumway, 2003. Good day sunshine: stock returns and the weather. *Journal of Finance* 58(3), 1009-1032.
- [38] Hong, Harrison, and Jialin Yu, 2009. Gone fishin': Seasonality in trading activity and asset prices. *Journal of Financial Markets* 12(4), 672-702.

- [39] Jacobsen, Ben, and Wessel Marquering, 2008. Is it the weather?. *Journal of Banking and Finance* 32(4), 526-540.
- [40] Jacobsen, Ben, and Wessel Marquering, 2009. Is it the weather? Response. *Journal of Banking and Finance* 33(3), 583-587.
- [41] Jones, Ben, and Ben Olken, 2010. Climate shocks an exports. *American Economic Review Papers and Proceedings* 100(2), 454-459.
- [42] Kyle, Albert S., 1985. Continuous auctions and insider trading. *Econometrica* 53(6), 1315-1335.
- [43] Longcore, Travis R., and Peter W. Rees, 1996. Information technology and downtown restructuring: The case of New York City's financial district. *Urban Geography* 17(4), 354-372.
- [44] Lonnqvist, Olov, 1962. On the diurnal variation of surface temperature. *Tellus* 14(1), 96-101.
- [45] Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi, 2003. Winter blues: A SAD stock market cycle. *American Economic Review* 93(1), 324-343.
- [46] Kandel, Eugene, and Neil D. Pearson, 1995. Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy* 103(4), 831-872.
- [47] Koudijs, Peter, 2015. The boats that did not sail: Asset price volatility in a natural experiment. *Journal of Finance*, accepted online manuscript.
- [48] LeBaron, Blake, 1992. Persistence of the Dow Jones Index on rising volume. University of Wisconsin, unpublished paper.
- [49] Llorente, Guillermo, Roni Michaely, Gideon Saar, and Jiang Wang, 2002. Dynamic volume return relation of individual stocks. *Review of Financial studies* 15(4): 1005-1047.
- [50] Longcore, Travis R., and Peter W. Rees, 1996. Information technology and downtown restructuring: The case of New York City's financial district. *Urban Geography* 17(4), 354-372.
- [51] Nagel, Stefan, 2012. Evaporating liquidity. *Review of Financial Studies* 25(7), 2005-2039.
- [52] Pastor, Lubos and Robert F. Stambaugh, 2003. Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3), 642-685.

- [53] Saunders, Edward M., 1993. Stock prices and wall street weather. *American Economic Review* 83(5), 1337-1345.
- [54] Seasholes, Mark S., and Guojun Wu, 2007. Predictable behavior, profits, and attention. *Journal of Empirical Finance* 14(5), 590-610.
- [55] Selden, George Charles, 1919. *The machinery of Wall street: why it exists, how it works and what it accomplishes*. New York: The Magazine of Wall street.
- [56] Seppanen, Olli, William J. Fisk, and Q. H. Lei, 2006. Effect of temperature on task performance in office environment. Lawrence Berkeley National Laboratory, working paper.
- [57] Schwert, G. William, 1990. Indexes of United States stock prices from 1802 to 1987. *Journal of Business* 63, 399-426. <http://schwert.ssb.rochester.edu/dstock.htm> (accessed July 2013).
- [58] Shleifer, Andrei, 1986. Do demand curves for stocks slope down?. *Journal of Finance* 41(3), 579-590.
- [59] Shleifer, Andrei, and Lawrence H. Summers, 1990. The noise trader approach to finance. *Journal of Economic Perspectives* 4(2), 19-33.
- [60] Shleifer, Andrei, and Robert Vishny, 1997. The limits of arbitrage. *Journal of Finance* 52(1), 35-55.
- [61] Tilghman, Charles R., 1961. Telegraph history: Some early days of Western Union's stock ticker service 1871-1910. *Western Union Technical Review* 15(2), 75-79.
- [62] U.S. Census Bureau, 1975. *Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition, Part 2*. Washington, D.C.: U.S. Government Printing Office. <http://www2.census.gov/prod2/statcomp/documents/CT1970p2-05.pdf>.
- [63] Veronesi, Pietro, 1999. Stock market overreactions to bad news in good times: a rational expectations equilibrium model. *Review of Financial Studies* 12(5), 975-1007.
- [64] Verrecchia, Robert E., 1982. Information acquisition in a noisy rational expectations economy. *Econometrica* 50(6), 1415-1430.
- [65] Wang, Jiang, 1994. A model of competitive stock trading volume. *Journal of Political Economy* 102(1), 127-168.

A Tables and Figures

Table 1: Seasonality in NYSE Trading Volume and Manhattan Temperature

NYSE trading activity and Manhattan temperature are computed in each season of the four seasons, for five different sample periods. Trading activity is computed as the average trading volume over the season divided by its annual average. The four seasons used here are spring (Mar-May), summer (Jun-Aug), fall (Sep-Nov), and winter (Dec-Feb).

	(1)	(2)	(3)	(4)	(5)
	1889-1902	1903-1930	1931-1960	1961-1990	1991-2014
<i>NYSE trading activity</i>					
Mar-May	1.05	1.01	0.97	0.99	1.00
Jun-Aug	0.81	0.86	0.91	0.95	0.96
Sep-Nov	1.07	1.12	1.04	1.02	1.03
Dec-Feb	1.07	1.02	1.08	1.04	1.01
<i>Average Manhattan temperature (degrees Celsius)</i>					
Mar-May	11.3	12.1	12.9	13.7	13.9
Jun-Aug	24.9	24.7	25.8	26.0	25.9
Sep-Nov	14.8	15.4	16.4	16.3	16.2
Dec-Feb	1.7	1.7	2.9	2.8	3.7

Table 2: Regression of NYSE Summer Trade Volume on Manhattan Summer Temperature

This table reports the OLS regression results for five distinct sample periods. Dependent variable is summer relative volume, computed as the average NYSE total trading volume over Jun 11-Sep 10 divided by its annual average. Independent variable is the summer average temperature, computed as the average trading hour temperature in Manhattan over Jun 11-Sep 10. Standard errors corrected for heteroskedasticity are reported in parentheses. *, **, and ***, indicates significance at 10, 5, and 1 percent, respectively.

	(1)	(2)	(3)	(4)	(5)
	1889-1902	1903-1930	1931-1960	1961-1990	1991-2014
Summer Temperature	-0.11*	-0.04*	0.02	-0.01	0.01
	(0.06)	(0.02)	(0.05)	(0.02)	(0.01)
Observations	14	28	30	30	24
R^2	0.144	0.057	0.005	0.008	0.014

An intercept included but not reported here.

Table 3: Regression of NYSE Summer Trade Volume on Summer Temperature (1889-1902)

A Placebo Test Using Temperature in Other Cities

This table reports the OLS regression results using temperatures in five different cities during pre-1903 sample. Dependent variable is summer relative volume, computed as the average NYSE total trading volume over Jun 11-Sep 10 divided by its annual average. Independent variable is the summer average temperature, computed as the average trading hour temperature in each city over Jun 11-Sep 10. Standard errors corrected for heteroskedasticity are reported in parentheses. *, **, and ***, indicates significance at 10, 5, and 1 percent, respectively.

	(1)	(2)	(3)	(4)	(5)
	Cambridge	Jacksonville FL	Birmingham UK	Sydney AU	Edmonton CD
Summer Temp.	-0.04 (0.05)	0.29 (0.19)	0.00 (0.06)	0.25 (0.17)	0.13 (0.08)
Observations	14	14	14	14	14

An intercept included but not reported here.

Table 4: Determinants of Daily NYSE Total Trade Volume

Dependent variable is daily detrended NYSE total trading volume, computed as the deviation of log total trading volume per hour from its 1-year moving average. \tilde{T}_t and \bar{T}_t are unseasonal and seasonal temperatures in Manhattan on day t , respectively. $\mathbf{1}(R < 0)$ is an indicator variable for a negative NYSE index return. OLS regression results for pre-1903 sample are reported in columns (1) and (2), while the results for the four other sample periods are reported columns (3)-(6). Newey-West HAC standard errors with 20 lags are reported in parentheses. *, **, and ***, indicates significance at 10, 5, and 1 percent, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	1889-1902		1903-1930	1931-1960	1961-1990	1991-2014
\tilde{T}	-0.013** (0.006)	-0.015*** (0.005)	-0.009** (0.004)	0.002 (0.003)	-0.000 (0.002)	0.002 (0.002)
$\tilde{T}(t-1)$		0.002 (0.004)				
\bar{T}	-0.085*** (0.017)	-0.085*** (0.017)	-0.022* (0.012)	0.005 (0.012)	0.002 (0.005)	0.013*** (0.004)
Volatility	0.716*** (0.224)	0.714*** (0.222)	1.990*** (0.300)	0.494*** (0.158)	1.671*** (0.223)	0.604*** (0.095)
$\mathbf{1}(R < 0)$	-0.037 (0.024)	-0.038 (0.024)	0.015 (0.018)	-0.042** (0.018)	-0.052*** (0.011)	0.018** (0.008)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,056	1,051	2,080	2,083	1,895	1,529

An intercept and dummy variables for day of the week and trading days' gap are included but not reported.

Table 5: Determinants of Daily NYSE Total Trade Volume (1889-1902)

A Placebo Test Using Temperatures in Other Cities

This table reports the OLS regression results using temperatures in other cities during pre-1903 sample. Dependent variable is daily detrended NYSE total trading volume, computed as the deviation of log total trading volume per hour from its 1-year moving average. \tilde{T}_t and \bar{T}_t are unseasonal and seasonal temperatures in the specified city on day t , respectively. $\mathbf{1}(R < 0)$ is an indicator variable for a negative NYSE index return. Newey-West HAC standard errors with 20 lags are reported in parentheses. *, **, and ***, indicates significance at 10, 5, and 1 percent, respectively.

	(1)	(2)	(3)	(4)	(5)
	Cambridge MA	Jacksonville FL	Birmingham UK	Sydney AU	Edmonton CD
\tilde{T}	-0.010* (0.005)	-0.010 (0.011)	-0.005 (0.010)	0.004 (0.008)	0.006 (0.006)
\bar{T}	-0.079*** (0.016)	-0.142*** (0.028)	-0.083*** (0.029)	0.086*** (0.022)	-0.047*** (0.011)
Volatility	0.701*** (0.236)	0.738*** (0.226)	0.691*** (0.237)	0.731*** (0.226)	0.714*** (0.222)
$\mathbf{1}(R < 0)$	-0.031 (0.024)	-0.050** (0.024)	-0.042* (0.024)	-0.031 (0.024)	-0.040* (0.024)
Year FE	Yes	Yes	Yes	Yes	Yes
$\rho(\tilde{T}, \tilde{T} NY)$.75	.08	.04	.02	-.14
$\rho(\bar{T}, \bar{T} NY)$.85	.54	.61	-.06	.59
Observations	905	1,056	1,056	1,056	1,056

An intercept and dummy variables for day of the week and trading days' gap are included but not reported.

Table 6: Determinants of Daily NYSE Total Trade Volume (1889-1902)**A Horse Race Between Temperature in Manhattan and Temperature in Other Cities**

This table reports the OLS regression results using temperatures in other cities during pre-1903 sample. Dependent variable is daily detrended NYSE total trading volume, computed as the deviation of log total trading volume per hour from its 1-year moving average. “NY \tilde{T}_t ” and “NY \bar{T}_t ” are unseasonal and seasonal temperatures in Manhattan on day t , respectively. \tilde{T}_t and \bar{T}_t are unseasonal and seasonal temperatures in the specified city on day t , respectively. $\mathbf{1}(R < 0)$ is an indicator variable for a negative NYSE index return. Newey-West HAC standard errors with 20 lags are reported in parentheses. *, **, and ***, indicates significance at 10, 5, and 1 percent, respectively.

	(1)	(2)	(3)	(4)	(5)
	Cambridge MA	Jacksonville FL	Birmingham UK	Sydney AU	Edmonton CD
NY \tilde{T}	-0.017** (0.008)	-0.011* (0.006)	-0.013** (0.006)	-0.012** (0.006)	-0.011** (0.006)
NY \bar{T}	-0.070*** (0.023)	-0.050*** (0.016)	-0.076*** (0.015)	-0.056*** (0.019)	-0.070*** (0.017)
\tilde{T}	0.004 (0.007)	-0.005 (0.012)	-0.000 (0.010)	-0.000 (0.008)	0.008 (0.006)
\bar{T}	-0.027 (0.022)	-0.093*** (0.029)	-0.037 (0.026)	0.053** (0.024)	-0.022* (0.012)
Volatility	0.695*** (0.233)	0.737*** (0.227)	0.715*** (0.231)	0.733*** (0.226)	0.720*** (0.220)
$\mathbf{1}(R < 0)$	-0.029 (0.024)	-0.044* (0.024)	-0.037 (0.024)	-0.031 (0.024)	-0.037 (0.024)
Year FE	Yes	Yes	Yes	Yes	Yes
$\rho(\tilde{T}, \tilde{T} NY)$.75	.08	.04	.02	-.14
$\rho(\bar{T}, \bar{T} NY)$.85	.54	.61	-.06	.59
Observations	905	1,056	1,056	1,056	1,056

An intercept and dummy variables for day of the week and trading days' gap are included but not reported.

Table 7: Manhattan Temperature and Returns from A Simple Reversal Strategy on NYSE Index (1889-1902)

Dependent variable L_t^0 measures aggregate market liquidity using expected return from the simple reversal strategy of going long (short) one dollar of NYSE index if today's index has a negative (positive) return. \tilde{T}_t and \bar{T}_t are unseasonal and seasonal temperatures in Manhattan on day t , respectively. $\mathbf{1}(R < 0)$ is an indicator variable for a negative NYSE index return. NYSE market index uses the DJIA during 1889-1925 and S&P 500 during 1926-2014. OLS regression results for five distinct sample periods are reported in (1) and (3)-(6), while 2SLS regression result using the pre-1903 sample is reported in column (2). Manhattan unseasonal temperature \tilde{T}_t is used as a proxy for the decrease in outside investor trades in (1) and as an instrumental variable for trading volume with outside investors as the compliers in (2). The return has been multiplied by 100. Newey-West HAC standard errors with 20 lags are reported in parentheses. *, **, and ***, indicates significance at 10, 5, and 1 percent, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	1889-1902		1903-30	1931-60	1961-90	1991-2014
\tilde{T}	-0.024*** (0.009)		-0.007 (0.008)	-0.006 (0.008)	-0.012** (0.006)	0.002 (0.008)
\bar{T}	-0.012 (0.028)	0.144* (0.082)	-0.017 (0.015)	-0.010 (0.015)	0.004 (0.011)	-0.005 (0.016)
Tradvol	0.071 (0.077)		0.121** (0.049)	0.018 (0.094)	0.136 (0.123)	0.597*** (0.225)
$\widehat{\text{Tradvol}}$		1.834* (0.996)				
Volatility	0.907*** (0.276)	-0.342 (0.814)	-1.671 (1.901)	-0.591*** (0.219)	-2.187*** (0.815)	-1.627*** (0.465)
$\mathbf{1}(R < 0)$	0.072 (0.070)	0.135 (0.087)	0.049 (0.054)	0.097 (0.068)	0.022 (0.043)	0.035 (0.051)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
IV Tradvol		Yes				
Observations	1,056	1,056	2,080	2,083	1,895	1,529

An intercept and dummy variables for day of the week and trading days' gap included but not reported.

Table 8: Manhattan Temperature and Returns from Three Reversal Strategies Based on the Cross-Section of Individual Stock Returns (1889-1902)

Each of the three different dependent variables (denoted L_t^1 , L_t^2 , and L_t^3) measures aggregate market liquidity using market maker's expected returns from aggregate liquidity provision. The three measures are calculated using individual stock price data, as discussed in section 4.2. \tilde{T}_t and \bar{T}_t are unseasonal and seasonal temperatures in Manhattan on day t , respectively. $\mathbf{1}(R < 0)$ is an indicator variable for a negative NYSE index return. OLS regression results for five distinct sample periods are reported in (1) and (3)-(6), while 2SLS regression result using the pre-1903 sample is reported in column (2). Manhattan unseasonal temperature \tilde{T}_t is used as a proxy for the decrease in outside investor trades in (1) and as an instrumental variable for trading volume with outside investors as the compliers in (2). The return has been multiplied by 100 to be interpreted in percentage unit. Newey-West HAC standard errors with 20 lags are reported in parentheses. *, **, and ***, indicates significance at 10, 5, and 1 percent, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Liquidity 1		Liquidity 2		Liquidity 3	
\tilde{T}	-0.047** (0.022)		-0.055* (0.030)		-0.051* (0.027)	
\bar{T}	-0.002 (0.042)	0.276* (0.152)	0.008 (0.039)	0.334* (0.202)	0.003 (0.060)	0.302 (0.192)
Tradvol	-0.287* (0.152)		-0.340** (0.140)		-0.296 (0.210)	
$\widehat{\text{Tradvol}}$		2.917 (1.778)		3.415 (2.278)		3.151 (2.172)
Volatility	0.023*** (0.003)	-0.000 (0.013)	0.036*** (0.004)	0.009 (0.017)	0.017*** (0.003)	-0.008 (0.016)
$\mathbf{1}(R < 0)$	0.164 (0.174)	0.262 (0.192)	0.187 (0.187)	0.302 (0.232)	0.232 (0.219)	0.338 (0.226)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
IV Tradvol		Yes		Yes		Yes
Observations	1,018	1,018	1,018	1,018	1,018	1,018

An intercept and dummy variables for day of the week and trading days' gap are included but not reported.

Table 9: Manhattan Temperature and the Daily Reversal Pattern of NYSE Individual Stocks (1889-1902)

Dependent variables are the following: fraction of NYSE individual stocks that experience a return reversal on the next trading day (Panel A); average absolute return on NYSE individual stocks that experience a return reversal on the next trading day (Panel B); and average absolute return on NYSE individual stocks that experience a return continuation on the next trading day (Panel C). \tilde{T}_t and \bar{T}_t are unseasonal and seasonal temperatures in Manhattan on day t , respectively. $\mathbf{1}(R < 0)$ is an indicator variable for a negative NYSE index return. OLS regression result is reported in odd numbered rows, while 2SLS regression results is reported in even numbered rows. Manhattan unseasonal temperature \tilde{T}_t is used as a proxy for the decrease in outside investor trades in odd numbered rows and as an instrumental variable for trading volume with outside investors as the compliers in even numbered rows. Standard errors corrected for heteroskedasticity are reported in parentheses. *, **, and ***, indicates significance at 10, 5, and 1 percent, respectively.

	\tilde{T}	\bar{T}	Tradvol	$\widehat{\text{Tradvol}}$	Volatility	$\mathbf{1}(R < 0)$	Year FE	IV
<i>Panel A: Fraction of stocks with a return reversal next day</i>								
(1)	-0.003** (0.001)	-0.004 (0.004)	-0.017 (0.012)		0.024 (0.038)	0.018* (0.010)	Yes	
(2)		0.015 (0.009)		0.193* (0.113)	-0.129 (0.094)	0.024** (0.012)	Yes	Yes
<i>Panel B: Mean absolute return on stocks with a return reversal next day</i>								
(3)	-0.007 (0.007)	-0.003 (0.025)	0.252*** (0.077)		0.870*** (0.168)	-0.026 (0.043)	Yes	
(4)		0.037 (0.039)		-0.204 (0.475)	1.200*** (0.393)	-0.040 (0.048)	Yes	Yes
<i>Panel C: Mean absolute return on stocks with a return continuation next day</i>								
(5)	0.014** (0.006)	0.011 (0.018)	0.324*** (0.058)		0.356*** (0.153)	-0.047 (0.038)	Yes	
(6)		-0.073 (0.044)		-0.644 (0.582)	1.057*** (0.427)	-0.076* (0.046)	Yes	Yes

An intercept and dummy variables for day of the week and trading days' gap are included but not reported. $N = 1,018$

Table 10: NYSE Market Index Return Reversal and Autocorrelation over 1889-1902 and 1926-2014

This is a summary table of probability of market index daily return reversal and return autocorrelation during the summers of 1889-1902 and 1926-2014. Volatile sample is defined as the dates with above-median volatility within each sample period.

<i>Probability of daily return reversal in the market index</i>	
1889-1902	
All sample	47.8%
Volatile sample	50.0%
1926-2014	
All sample	46.0%
Volatile sample	48.0%
<i>Daily return autocorrelation in the market index</i>	
1889-1902	
All sample	0.037
Volatile sample	0.013
1926-2014	
All sample	0.059
Volatile sample	0.041

Table 11: Manhattan Temperature and the Likelihood of NYSE Index Return Reversal

Dependent variable is the binary variable indicating the event that the NYSE index return switches the sign on the next trading day (1=Reversal, 0=No Reversal). \tilde{T}_t and \bar{T}_t are unseasonal and seasonal temperatures in Manhattan on day t , respectively. $\mathbf{1}(R < 0)$ is an indicator variable for a negative NYSE index return. NYSE market index uses the DJIA during 1889-1925 and S&P 500 during 1926-2014. Marginal effects from logit regressions for five distinct sample periods are reported in columns (1), (2), and (4)-(7), and 2SLS regression result based on LPM and the pre-1903 sample is reported in column (3). Manhattan unseasonal temperature \tilde{T}_t is used as a proxy for the decrease in outside investor trades in (1) and as an instrumental variable for trading volume with outside investors as the compliers in (3). Standard errors corrected for heteroskedasticity are reported in parentheses. *, **, and ***, indicates significance at 10, 5, and 1 percent, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	1889-1902			1903-30	1931-60	1961-90	1991-2014
\tilde{T}	-0.016*** (0.005)			-0.002 (0.003)	-0.005 (0.003)	-0.006* (0.004)	-0.000 (0.004)
\bar{T}	0.007 (0.013)		0.106*** (0.040)	-0.001 (0.008)	0.005 (0.007)	-0.002 (0.008)	0.006 (0.009)
Tradvol	0.015 (0.038)	0.017 (0.036)		0.050* (0.027)	0.008 (0.028)	0.041 (0.063)	0.174** (0.077)
$\widehat{\text{Tradvol}}$			1.149** (0.472)				
Volatility	0.200 (0.129)	0.182 (0.121)	-0.630* (0.369)	0.210 (0.189)	0.001 (0.047)	-0.406 (0.404)	-0.255 (0.178)
$\mathbf{1}(R < 0)$	0.085*** (0.032)	0.078** (0.032)	0.122*** (0.046)	0.103*** (0.023)	0.081*** (0.023)	0.044* (0.024)	0.078*** (0.026)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Model	Logit	Logit	LPM	Logit	Logit	Logit	Logit
IV Tradvol			Yes				
Observations	1,056	1,056	1,056	2,080	2,083	1,895	1,529

An intercept and dummy variables for day of the week and trading days' gap are included but not reported.

Table 12: Manhattan Temperature and the Daily Autocorrelation in NYSE Index Returns

Dependent variable is NYSE market index return on the next trading day, and independent variables are variables interacted with today's market index return. \tilde{T}_t and \bar{T}_t are unseasonal and seasonal temperatures in Manhattan on day t , respectively. $\mathbf{1}(R < 0)$ is an indicator variable for a negative NYSE index return. NYSE market index uses the DJIA during 1889-1925 and S&P 500 during 1926-2014. OLS regression results for five distinct sample periods are reported in (1), (2), and (4)-(7), while 2SLS regression result using pre-1903 sample is reported in column (3). All returns are winsorized at 5%. Manhattan unseasonal temperature \tilde{T}_t is used as a proxy for the decrease in outside investor trades in (1) and as an instrumental variable for trading volume with outside investors as the compliers in (3). Newey-West HAC standard errors with 20 lags are reported in parentheses. *, **, and ***, indicates significance at 10, 5, and 1 percent, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	1889-1902			1903-30	1931-60	1961-90	1991-2014
$\tilde{T} \times R$	0.019** (0.009)			0.007 (0.006)	0.003 (0.007)	0.018*** (0.007)	-0.001 (0.008)
$\bar{T} \times R$	0.004 (0.021)		-0.059 (0.086)	0.000 (0.012)	0.026* (0.014)	0.012 (0.014)	0.013 (0.018)
Tradvol $\times R$	-0.116** (0.058)	-0.125** (0.057)		-0.062 (0.052)	-0.036 (0.042)	-0.219* (0.118)	-0.400*** (0.139)
($\widehat{\text{Tradvol}} \times R$)			-1.759 (1.445)				
Volatility $\times R$	-0.188** (0.096)	-0.173* (0.102)	0.376 (0.524)	-0.335 (0.224)	0.001 (0.074)	0.687* (0.350)	0.655*** (0.246)
$\mathbf{1}(R < 0) \times R$	-0.238* (0.123)	-0.244** (0.123)	-0.391* (0.231)	-0.226*** (0.078)	-0.152** (0.073)	-0.037 (0.076)	-0.151* (0.086)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV Tradvol $\times R$			Yes				
Observations	1,056	1,056	1,056	2,080	2,083	1,895	1,529

An intercept, return, and return interacted dummies for day of the week and trading days' gap not reported.

Figure 1: Detrended NYSE Trade Volume

1889-2014 (top), 1889-1902 (middle), and the summer of 1895 (bottom)

These graphs show, for three different time horizons, daily detrended NYSE total trading volume, computed as the deviation of log total trading volume per hour from its 1-year moving average.

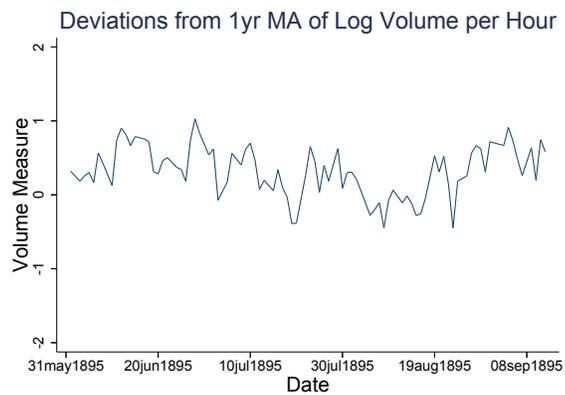
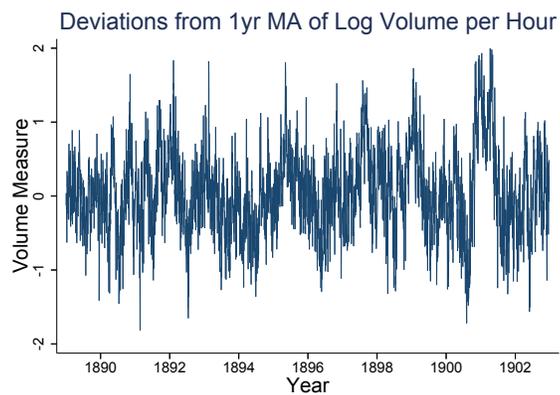
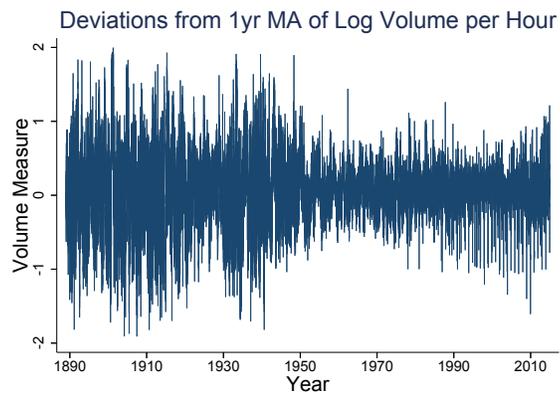


Figure 2: Manhattan Temperature by 10-Days' Interval (1889-1902)

The summer period used in the paper (Jun 11-Sep 10) is highlighted in dark gray. I obtain the trading hour average temperature T each day using the maximum and minimum temperatures on that day. To do so, I assume that an intraday variation in temperature reaches the daily maximum at 15:00 and minimum at 5:00 (Lonnqvist 1962) and that the temperature change is linear in time.

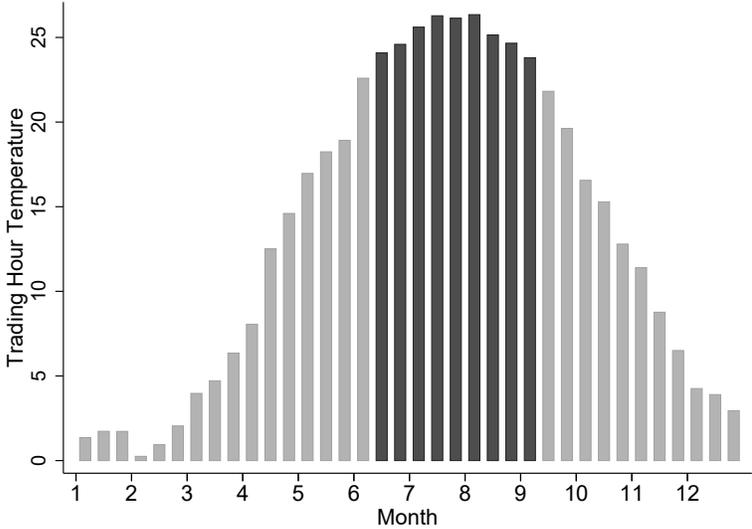


Figure 3: Summer NYSE Trade Volume and Summer Manhattan Temperature

Summer relative volume is the average NYSE total trading volume over Jun 11-Sep 10 divided by its annual average. Summer average temperature is the average trading hour temperature over Jun 11-Sep 10.

