Journal of Empirical Finance 70 (2023) 165–181

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

Acute illness symptoms among investment professionals and stock market dynamics: Evidence from New York City

Gabriele M. Lepori[1]

Department of Banking and Finance, Southampton Business School, University of Southampton, Highfield Campus, Southampton SO17 1BJ, United Kingdom

ARTICLE INFO

JEL classification:
G12
G14
G23
G41

Keywords:
Stock market professionals
Price discovery
Volatility
Trading activity
Acute illness

ABSTRACT

In the U.S., stock market professionals (e.g., traders, portfolio managers, and analysts) are clustered in New York City (NYC). In view of this, I exploit daily changes in the incidence of acute illness symptoms among 18–64 year old New Yorkers to identify exogenous variation in the rate of acute illness among market professionals and estimate its *causal* impact on key stock market outcomes. A detailed analysis of taxi trips from a sample of financial institutions to local hospitals provides support for my identification assumption. Other things equal, increased rates of acute physical illness (i.e., reduced productivity) among market professionals hamper price discovery and lower trading activity, volatility, and returns. A one-standard-deviation increase in my illness incidence proxy reduces by 18% (6.7%) the immediate response of stock prices to earnings surprises (changes in analysts' consensus recommendations) and increases by 29% (42%) their delayed response.

1. Introduction

The investor portrayed in neoclassical finance models tends to resemble a robot that works uninterruptedly and never experiences productivity fluctuations. Yet, real-world market participants are made of flesh and bones. Just like the rest of us, some of them, sometimes, may get a stomach bug, or get an asthma attack, or come down with a bad cold. What happens then when acute illness symptoms affect the productivity of a large number of stock market professionals at once? How does that affect market trading activity and volatility? How does it impact stock returns and the price discovery process?

To address these questions, in this paper I take advantage of the fact that, in the U.S., investment professionals are clustered in NYC. For example, statistics on users of the Bloomberg Terminal (a leading provider of financial data) show that, among those located in the U.S. and with a focus on equities, 46.7% of traders, 25.8% of portfolio managers, and 45.8% of analysts work in NYC. Based on this insight, I exploit daily changes in the incidence of acute illness symptoms (e.g., asthma, vomiting) among 18–64 year old New Yorkers to identify exogenous variation in the rate of acute illness among NYC's investment professionals. This allows me to estimate its *causal* impact on key market outcomes and price discovery.

Previous work on the links between human physiology and financial markets is scant. The few studies that have investigated this topic provide evidence that disruptions in sleep patterns (Kamstra et al., 2000), hormonal fluctuations (Bose et al., 2020), and allergies (Pantzalis and Ucar, 2018) may affect large groups of investors across the country and alter market dynamics. McTier et al. (2013) show that increases in weekly flu incidence (nationally and in the Middle Atlantic division containing the State of New York) are associated with reduced market trading activity, volatility, and stock returns.

E-mail address: g.m.lepori@soton.ac.uk.

[1] This work began when I was employed at Keele University. I thank the Editor, an anonymous reviewer, John Wald, Anthony Zhang, seminar participants at the University of Southampton, and participants at the Academy of Behavioral Finance & Economics annual meeting in Los Angeles for useful comments and suggestions. Any errors are my own. The study was approved by the Faculty Ethics Committee of the University of Southampton (ERGO number: 55996.A2).

https://doi.org/10.1016/j.jempfin.2022.12.003
Received 10 November 2021; Received in revised form 26 November 2022; Accepted 4 December 2022
Available online 6 December 2022
0927-5398/

My research design builds on McTier et al.'s (2013) work, but using novel data sets I extend their analyses and findings in several new directions. In particular, the present paper contributes to the literature by enriching our understanding of how acute illness symptoms among a substantial number of professional market participants may hinder the price discovery process and affect trading activity, volatility, and stock returns.

My key findings are as follows: By combining data on daily emergency department (ED) visits made by New Yorkers with data on taxi trips between financial institutions headquartered in NYC and local hospitals, I provide evidence that acute illness incidence among NYC's investment professionals is positively correlated with acute illness incidence among 18–64 year old New Yorkers. Using the latter variable as a proxy, I find that increased rates of acute physical illness among NYC's investment professionals are accompanied by reduced market trading activity, volatility, and returns. Specifically, based on a panel of NYSE stocks, my estimates indicate that a one-standard-deviation increase in acute illness incidence among NYC's stock market professionals leads to a contemporaneous 2% decrease in daily turnover, a 1.2% decrease in intraday volatility, and a 0.10% decrease in returns. At the same time, in line with my identification assumption, I observe no relation between changes in the incidence of acute illness symptoms in the 0–4, 5–17, and 65+ age groups and stock market dynamics.

Secondly, I provide evidence that acute illness symptoms among stock market professionals hamper price discovery by decreasing the speed with which prices impound new information. Specifically, I estimate that a one-standard-deviation increase in illness incidence among NYC's professionals decreases by 4.2% (10%) the amount of trading activity (intraday volatility) that occurs in response to the arrival of firm-specific news.

Thirdly, when I focus on two specific types of public information, namely earnings surprises and changes in analysts' consensus recommendations, I find that acute illness symptoms among investment professionals diminish the immediate reaction of prices and exacerbate the post announcement drift that has been well documented in the literature (Bernard and Thomas, 1989). Specifically, a one-standard-deviation increase in acute illness incidence among NYC's investment professionals reduces by 18% (6.7%) the immediate response of stock prices to earnings surprises (changes in analysts' consensus recommendations) and increases by 29% (42%) their delayed response.

Lastly, I show that the negative relation between my illness incidence proxy and stock returns can be explained neither by economic news nor by changes in turnover, liquidity, and investor risk aversion. Consistent with the interpretation that it is driven by illness-induced mood fluctuations among investment professionals (Finch et al., 2012; Goetzmann et al., 2015), I document that this effect is stronger among stocks that are more sensitive to investor sentiment and more difficult to arbitrage/value. Moreover, I find that the initial response of stock returns is followed by a complete reversal within five trading days.

The rest of the paper is organized as follows. Section 2 discusses the clustering of stock market professionals in NYC. Section 3 explains NYC's syndromic surveillance system. Section 4 describes the data. Section 5 presents the empirical analysis, and Section 6 concludes.

2. Stock market professionals in New York City

Data on stock ownership and trading show that investment professionals came to dominate the U.S. stock market by the last quarter of the 20th century (Ben-David et al., 2019). Recent work reveals that retail investors' trades currently account for only about 1%–2% of NYSE trading volume (Kadan et al., 2018; O'Hara et al., 2019).

With regard to the geographic distribution of professional market participants, research in urban geography documents that NYC has long been "the dominant investment centre" in the U.S. (Green et al., 2015). However, to provide more insights into the extent to which stock market professionals are clustered in NYC, I examine data from the Bloomberg Terminal and the SEC's EDGAR database.

It is well known that the Bloomberg Terminal is among the primary sources of information for financial market professionals (Ben-Rephael et al., 2017). As such, I regard the geographic distribution of Bloomberg users as a fitting proxy for the geographic distribution of stock market professionals. Panel A of Fig. 1 reveals that, at the time of this writing, among Bloomberg users located in the U.S. and with a focus on equities, 46.7% of traders, 25.8% of portfolio managers, and 45.8% of analysts are based in NYC.

Since historical statistics on Bloomberg users are not available, I also search the SEC's EDGAR database to collect historical data on the geographic distribution of large financial institutions that file Form 13F.[2] Specifically, I download and parse all the Form 13F-HR filings for the second quarter of each year between 2007 and 2017. The results, displayed in panel B of Fig. 1, indicate that, throughout the sample period, about 20% of the large financial institutions operating in the U.S. were headquartered in NYC. And between 2013 and 2017, NYC's institutions held about 20% of the 13(f) securities held by institutions.[3]

Overall, the statistics displayed in Fig. 1 provide hard evidence that the investment professionals based in the City play an outsized role among the professionals who participate in the U.S. stock market. The implication is that it is reasonable to assume that sizeable fluctuations in the incidence of acute illness symptoms among NYC's stock market professionals may affect price discovery and aggregate stock market dynamics.

[2] https://www.sec.gov/cgi-bin/srch-edgar. As explained on the SEC's website, "[a]n institutional investment manager that uses the U.S. mail [. . .] in the course of its business, and exercises investment discretion over $100 million or more in Section 13(f) securities [. . .] must report its holdings on Form 13F with the [. . .] SEC".

[3] With regard to institutional holdings, I limit my attention to the period from 2013 to 2017 because prior to the second quarter of 2013, when the SEC introduced a new technical specification, the contents of these filings did not follow a standard format, which makes their automatic extraction more susceptible to data errors. More details on these specification changes are available at https://www.sec.gov/info/edgar/specifications/form13fxml.1_d.htm.

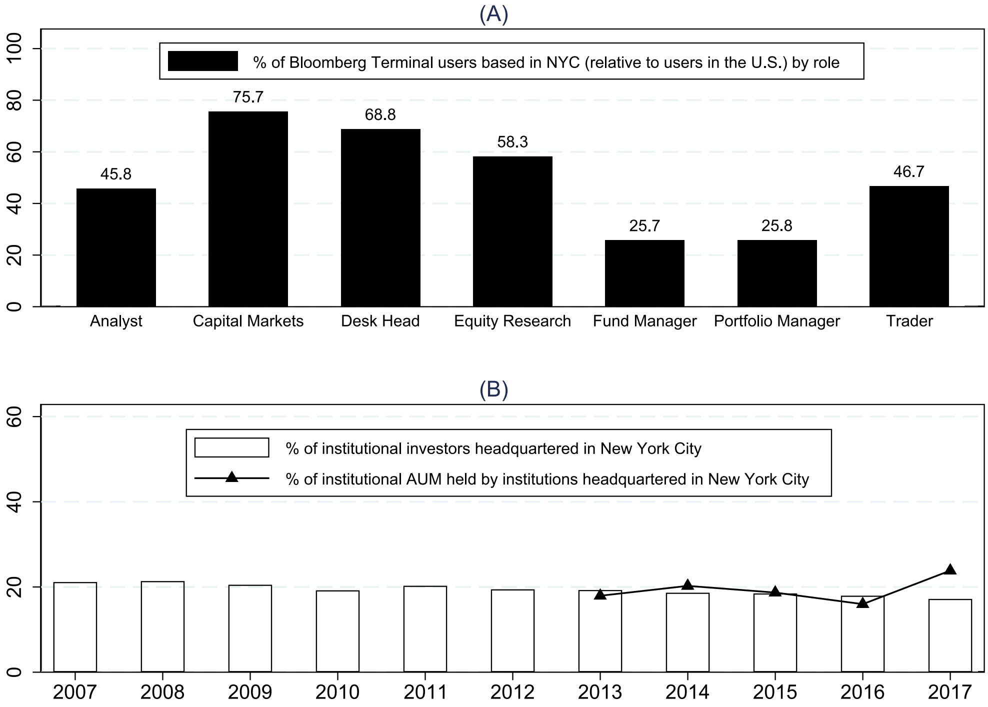

Fig. 1. Geographic distribution of stock market professionals: NYC vs. elsewhere. This figure shows some statistics on the geographic distribution of stock market professionals in the U.S. Panel A displays some statistics on users of the Bloomberg Terminal who, according to their public profile, are based in the U.S. and have a focus on equities/stocks; the data are broken down by role and were collected on July 13, 2022. Specifically, the vertical columns represent the percentage of users in a given role (e.g., trader) who are based in New York City relative to those based in the U.S. Panel B displays some statistics on a sub-group of stock market professionals, i.e., institutional investors, that report their holdings on Form 13F with the SEC. The vertical columns represent the percentage of institutional investors that are headquartered in New York City, relative to those headquartered in the U.S., between 2007 and 2017 (second quarter of each sample year). The triangles represent the percentage of institutional assets under management (AUM) held by institutions headquartered in the City relative to those held by institutions in the U.S. Author's own calculations based on data retrieved from the SEC's EDGAR database.

3. New York City's syndromic surveillance system

In 1995, NYC launched a computer-based surveillance system that relied on syndrome-specific data and whose initial purpose was to "detect outbreaks of waterborne illness" (Heffernan et al., 2004). In response to the September 11th terrorist attacks, the system began to monitor ED visits "to track the acute health effects of the attacks and to detect possible biologic terrorism" (Heffernan et al., 2004).[4]

On a daily basis, NYC's EDs transmit to the Department of Health and Mental Hygiene (DOHMH) electronic data about the previous day's number of visits, including information about chief complaint and "age, sex, and ZIP code" of each patient (Lall et al., 2017). Thereafter, based on the patient's chief complaint and the International Classification of Diseases (ICD-10), a computer algorithm determines whether the visit falls into one of the following five syndrome categories: asthma, diarrhea, influenza-like illness, respiratory, vomiting (Mathes et al., 2017).

It is worth noting that, since syndromic surveillance systems are not based on laboratory tests and only a fraction of individuals with given symptoms seek medical attention in EDs, syndromic data do not offer exact measures of illness among the population. Nevertheless, the literature has highlighted their validity and usefulness, as "[f]ew other data sources are comparable for understanding population-level health and health needs in near-real time" (Lall et al., 2017). For example, Olson et al. (2007) claim that "ED surveillance can provide a timely surrogate measure of morbidity", and Westheimer et al. (2012) document that "[c]itywide ED visits for [influenza-like illness] correlated well with influenza laboratory diagnoses".

4. Data

I gather the syndromic data from the website of NYC's DOHMH.[5] Specifically, I obtain the daily time series of the number of ED visits in the City broken down by syndrome category (asthma, diarrhea, influenza-like illness, respiratory, vomiting), age group (0–4, 5–17, 18–64, and 65+), and borough (Brooklyn, Bronx, Manhattan, Queens, and Staten Island). Visits are classified by borough on the basis of the ZIP code of the patient's residence.

[4] In 2004, 48 EDs participated in the system, which covered about 86% of the ED visits in the city (Heffernan et al., 2004). By 2005, this figure grew to 90% (Olson et al., 2007), and by 2010 to 95% (Mathes et al., 2011). Since May 1, 2016, the system has been collecting data from all EDs in NYC and capturing 100% of the ED visits in the city. See http://www1.nyc.gov/site/doh/data/data-sets/epi-syndromic-surveillance-data.page.

[5] https://nyc.gov/health/epiquery.

These data are available starting from January 1, 2006. However, since the NYSE gradually introduced a new hybrid system between October 2006 and January 2007 (Hendershott and Moulton, 2011), "trade data after [the switch] may not be comparable with earlier data" (McTier et al., 2013). Consequently, my sample period starts on February 1, 2007 and ends on December 31, 2017.

Following McTier et al. (2013), my primary explanatory variable is the daily percentage change in acute illness incidence, i.e., the change in the natural log of the number of ED visits. Specifically, I focus on ED visits by 18–64 year olds (*ΔLogIllness18_64*). In what follows, I posit that this variable proxies for changes in illness incidence among NYC's stock market professionals. This seems reasonable because, in 2011, according to the Bureau of Labor Statistics, 94.7% of the individuals employed in the finance and insurance industry in the fields of "securities, commodities, funds, trusts, and other financial investments" were between the age of 16 and 64 years.[6] Additionally, about 80% of the people who work in the City are NYC residents (NYC Planning, 2016). Sections 5.1–5.4 present a number of statistical tests and theoretical arguments that provide further support for my assumption.

In their empirical analysis, McTier et al. (2013) limit their attention to NYSE listed stocks. I follow the same approach because NYSE trading volume is largely driven by professional investors. Furthermore, NASDAQ reported volume "does not accurately measure the trading volume by public buyers and sellers" (Atkins and Dyl, 1997). The sample consists of an unbalanced panel of 1609 stocks, and its construction is described in Section A1 of the online Appendix.

The main dependent variables, which capture market trading activity, volatility, liquidity, and returns, are defined in Table 1, as are the control variables. To ensure that all dependent variables are stationary and comparable with changes in illness incidence, I follow McTier et al. (2013) and measure them as flows.

Panel A of Table 2 displays some summary statistics on the citywide number of ED visits by age group. For example, on the average day, about 844 NYC residents in the 18–64 age group visit an ED.[7] Panels B-D of Table 2 display summary statistics for the main dependent variables and some of the controls.

5. Empirical analysis

5.1. Taxi trips from financial institutions in NYC to local hospitals

Daily changes in the incidence of acute physical illness (e.g., asthma, flu-like symptoms) can be thought of as randomly assigned with respect to potential stock market outcomes. Consequently, the evidence presented in what follows lends itself to *causal* interpretation. However, there are two potential threats to my identification strategy. I address the first one here and the second one in Section 5.4.

The first threat concerns the identification assumption that variation in the number of ED visits made by 18–64 year old New Yorkers captures illness patterns among NYC's professional market participants. To investigate its validity, I analyze a data set containing information about all yellow taxi trips recorded in NYC between January 1, 2009 and June 30, 2016.[8] The rationale is that when NYC's investment professionals are at work and sickness strikes, they may choose to leave the office and visit an ED to seek medical attention. As such, if mine is a valid proxy, the daily number of taxi trips between financial institutions in NYC and local hospitals should be positively correlated with the number of ED visits made by 18–64 year old New Yorkers.

A map of the 51 EDs in the City is displayed in Fig. 2.[9] To identify a suitable sample of financial institutions, I focus on SEC Registered Investment Advisers with more than 250 employees performing investment advisory functions. A map of the headquarters of the 20 institutions that match these criteria is provided in Fig. 3. Though their names are not disclosed here due to the sensitivity of the data, it is safe to say that all of them are prominent players in the investment industry.[10]

To model the phenomenon of interest, I employ variants of the following specification:

$$EDtrips_{i,t} = exp\Bigg(\alpha + \beta_1 Illness18\_64_t + \varphi_1 ADSBusConditions_t + \mu_1 Rain_t + \mu_2 Snow_t + \vartheta_1 Holiday_t + \sum_{j=2}^{7}\gamma_j Day_{j,t} + \sum_{j=2}^{12}\delta_j Month_{j,t} + \sum_{j=2010}^{2016}\theta_j Year_{j,t} + Institution\,FE + Pick\text{–}up\,location\,FE\Bigg) \quad (1)$$

[6] Labor force statistics for 2011 are available from the website of the Bureau of Labor Statistics.

[7] It is worth stressing that only some affected individuals seek medical attention this way. Consequently, public health experts often resort to the so-called "syndromic multiplier" to estimate population measures of morbidity. For example, according to estimates by Metzger et al. (2004), each ED visit for influenza-like (diarrheal) illness represents about 76.5 (262.4) cases among 18–64 year old New Yorkers.

[8] The data are from the City of New York's website. The data set contains information about 1.25 billion taxi rides; for each ride, information is available about exact pick-up and drop-off locations (latitude and longitude), date/time, and number of passengers. The data are available from the start of 2009, but the fields capturing exact pick-up and drop-off locations were discontinued after June 30, 2016.

[9] I obtain the list of hospitals with an ED from the website of the New York State Department of Health. I then manually transform each address into latitude/longitude coordinates using OpenStreetMap. The list of hospitals and GPS coordinates is in Table A.3 in the online Appendix.

[10] The names of the 20 institutions were disclosed to the Editor and the reviewers during the review process. Data at monthly frequency about SEC Registered Investment Advisers are available from the SEC's website. Their addresses are available starting from November 2009. Therefore, I assume that there were no changes of address between January and November 2009, and I manually transform each address into latitude/longitude coordinates using OpenStreetMap. Note that Fig. 3 shows more than 20 locations because some of these 20 institutions moved their headquarters during the sample period.

Table 1
Variable definitions.

Variable	Definition	Notes
Illness incidence proxies		
IllnessX	Number of ED visits made by New Yorkers in the *X* age group, where *X* is one of the following: 0_4, 5_17, 18_64, or 65+	
ΔLogIllnessX	Daily percentage change in acute illness incidence among New Yorkers in the *X* age group. Specifically, this variable measures the daily change in the log of *IllnessX*. Standardized to have zero mean and unit variance	
Dependent variables		
Shares traded	Number of shares traded in millions	a
ΔLogShares	100 × daily change in the log of *Shares traded*	
Dollar volume	Trading volume in billions of U.S. dollars	a
ΔLogVolume	100 × daily change in the log of *Dollar volume*	
Turnover	100 × shares traded/shares outstanding	a
ΔLogTurnover	100 × daily change in the log of *Turnover*	
CLHO_volatility	Intraday volatility. Equal to 100 times the Garman and Klass (1980) volatility estimator adjusted for opening jumps (Molnár, 2012), which is based on the previous day's closing price and the current day's high, low, opening, and closing prices	a
ΔLogVolatility	100 × daily change in the log of *CLHO_volatility*	
Closing bid–ask spread	100 × (Ask − Bid)/((Ask + Bid)/2)	a
ΔSpread	Daily change in *Closing bid–ask spread*	
Return	100 × daily change in the log of the stock's total return index, which includes reinvested dividends	a
Firm-specific news		
AbNumStories	Continuous variable calculated by Bloomberg that captures abnormal daily news flow, ranging between 0 and 4. Specifically, it measures "the amount of stories currently being published on a company relative to the flow over the previous 45 days," with higher values indicating unusually high story flow (Bloomberg, 2020)	b
NewsSentiment	Continuous variable calculated by Bloomberg that measures daily news sentiment for a given company and ranges between −1 (mostly negative) and +1 (mostly positive)	b
SUE	Standardized unexpected earnings, i.e., the spread between actual and expected EPS on day t divided by the stock price on day $t-5$, as in DellaVigna and Pollet (2009)	c
Abs(SUE)	Absolute value of *SUE*	
RecChng	Change in analysts' consensus recommendation. Bloomberg first converts each recommendation into a numerical rating ranging between 1 (Sell) and 5 (Buy). Subsequently, the ratings are averaged across the analysts following the company	c
Abs(RecChng)	Absolute value of *RecChng*	
Market conditions		
ADSBusConditions	Aruoba et al.'s (2009) real-time daily business conditions index for the U.S. economy	d
MktUncertainty	Baker et al.'s (2018) U.S. equity market daily uncertainty index	d
BEXRiskAversion	Bekaert et al.'s (2020) daily risk aversion index for the U.S. market	d
Other variables		
Day (Month, Year)	Day-of-the-week (month, year) dummy variables	
PreH (PostH)	Pre- (post-) holiday dummy variable, where holidays = New Year's Day, Martin Luther King Day, Presidents' Day, Easter Sunday, Memorial Day, Independence Day, Labor Day, Thanksgiving Day, and Christmas Day	
SADOR	Proxy that Kamstra et al. (2015) use to measure seasonal depression in the U.S.	
Cloudy (Sunny)	Dummy variable that takes the value of 1 if the average percentage of sky cover in NYC between 6 am and 4 pm is greater than 90% (less than 10%), and 0 otherwise	e
HighPollution	Dummy variable that takes the value of 1 if the overall air quality index (AQI) value for New York county is greater than 100, and 0 otherwise	f
$\text{Ab\_Ret}^{[t+h,t+k]}$	Stock's cumulative abnormal return from day $t+h$ ($h=0$, 1) to $t+k$ ($k=0$, 3, 5, 7, 10, 20, 30, 40), where t is an earnings announcement day, and abnormal returns are computed using the Fama and French (1993) 3-factor model with an estimation window from $t-200$ to $t-21$	
HLtoH	Spread between the stock's daily high price and low price divided by the daily high price	
Return(t − 27,t − 6)	Stock's cumulative return from day $t-27$ to $t-6$	
AVG_Turnover(t − 5,t − 1)	Stock's mean turnover from day $t-5$ to $t-1$	

(*continued on next page*)

Table 1 (*continued*).

Variable	Definition	Notes
AVG_Spread(t − 5,t − 1)	Stock's mean closing bid–ask spread from day $t-5$ to $t-1$	
SDRET	Standard deviation of the stock's returns from day $t-27$ to $t-6$	
LnSize	Log of the stock's average market capitalization from day $t-27$ to $t-6$	
InstHold	Log of 1 plus the stock's percentage of institutional ownership	
LnNumEst	Log of 1 plus the number of analysts following the stock	

[a]To mitigate the risk of data errors and the effect of outliers, I assign a missing value to dollar volume (number of shares traded) when its value equals zero. I assign a missing value to turnover when its value is equal to 0 or greater than 50%. I assign a missing value to intraday volatility when its value is equal to 0 or greater than 50%. I assign a missing value to the closing percent quoted spread when its value is negative or greater than 50%. Lastly, I assign a missing value to a stock's return when the simple return is greater than +50% or less than −50%.

[b]*AbNumStories* is available from January 8, 2008, whereas *NewsSentiment* is available from the start of the sample period. However, both variables feature a large number of missing observations. I assign a missing value to *NewsSentiment* when its value is greater than +1 or less than −1, as these values represent data entry errors.

[c]When EPS surprises or changes in analysts' consensus recommendations occur on a non-trading day or on a day when the stock price is missing, I shift them to the next trading day with a non-missing stock price. To mitigate the risk of data errors and the impact of outliers, I assign a missing value to *SUE* when its value is greater than +20% or less than −20%.

[d]To ensure that these variables are comparable with the dependent variables, I enter them in the regressions as flows, i.e., in first difference (*ΔADSBusConditions*, *ΔMktUncertainty*, *ΔBEXRiskAversion*).

[e]The sky coverage data are from the weather station at LaGuardia Airport and are available from the Iowa Environmental Mesonet.

[f]The air pollution data for New York county are from the Environmental Protection Agency's website.

Table 2
Summary statistics.

	Count	Mean	std	min	max
Panel A: Illness incidence proxies					
Illness0_4	6,427,805	407.757	157.301	125.00	1131.00
ΔLogIllness0_4	6,427,805	0.000	0.127	−0.88	0.85
Illness5_17	6,427,805	294.035	192.815	71.00	3417.00
ΔLogIllness5_17	6,427,805	−0.000	0.192	−1.09	0.97
Illness18_64	6,427,805	844.217	218.686	355.00	2387.00
ΔLogIllness18_64	6,427,805	0.000	0.139	−0.67	0.58
Illness65+	6,427,805	241.544	60.278	111.00	561.00
ΔLogIllness65+	6,427,805	0.000	0.175	−1.00	0.73
Panel B: Dependent variables					
Shares traded (millions)	3,570,379	2.027	7.266	0.00	1226.79
ΔLogShares	3,567,131	−0.000	0.543	−10.49	11.51
Dollar volume ($ billions)	3,570,339	0.071	0.186	0.00	28.67
ΔLogVolume	3,567,128	−0.000	0.544	−10.37	17.11
Turnover (%)	3,564,211	1.107	1.454	0.00	49.94
ΔLogTurnover	3,560,441	−0.000	0.541	−10.49	11.51
CLHO_volatility (%)	3,568,264	2.209	1.956	0.00	49.89
ΔLogVolatility	3,564,125	−0.000	0.515	−9.66	9.63
Closing bid–ask spread (%)	3,516,001	0.243	1.311	0.00	50.00
ΔSpread	3,463,740	−0.002	1.131	−49.95	49.93
Return (%)	3,566,644	0.018	2.743	−69.18	40.55
Panel C: Firm-specific news					
AbNumStories	3,048,171	0.411	0.486	0.00	4.00
NewsSentiment	762,687	0.132	0.283	−1.00	1.00
SUE (%)	51,245	0.056	1.295	−19.88	19.94
Abs(SUE) (%)	51,245	0.476	1.206	0.00	19.94
RecChng	124,559	0.004	0.275	−4.00	5.00
Abs(RecChng)	124,559	0.142	0.236	0.00	5.00
Panel D: Market conditions					
ADSBusConditions	6,427,805	−0.389	0.842	−4.08	0.92
MktUncertainty	6,427,805	50.422	72.640	4.80	1117.23
BEXRiskAversion	4,398,651	2.906	1.424	2.27	29.65

Notes: This table displays descriptive statistics on the main dependent and explanatory variables used in the analysis. All variables are measured at a daily frequency, and the sample period runs from February 1, 2007 to December 31, 2017. In panel A, *IllnessX* (where *X* = *0_4*, *5_17*, *18_64*, or *65+*) measures the number of visits to local EDs made by New Yorkers in *X* age group. Only visits related to the following syndrome categories are included in the count: asthma, diarrhea, influenza-like, respiratory, vomiting. The remaining variables are as defined in Table 1.

where *EDtrips* represents the number of taxi trips on day t starting from financial institution i's headquarters and ending at a local ED. Specifically, a trip is included in the count if the pick-up time is between 8 am and 8 pm, the number of passengers is 1 or 2

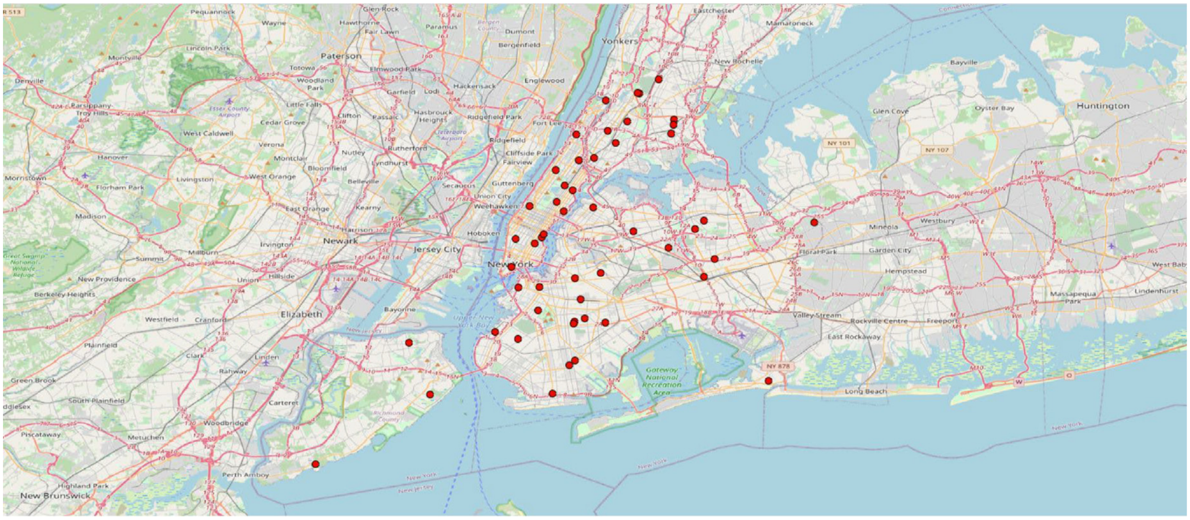

Fig. 2. Emergency departments in NYC. This figure shows the locations (dots) of the 51 hospitals in New York City that have an emergency department and are included in the taxi trip analysis discussed in Section 5.1. The list of hospitals, including their addresses, is from the New York State Department of Health. The address of each hospital is manually transformed into latitude/longitude coordinates using OpenStreetMap and subsequently plotted on the map using gpspointplotter.com.

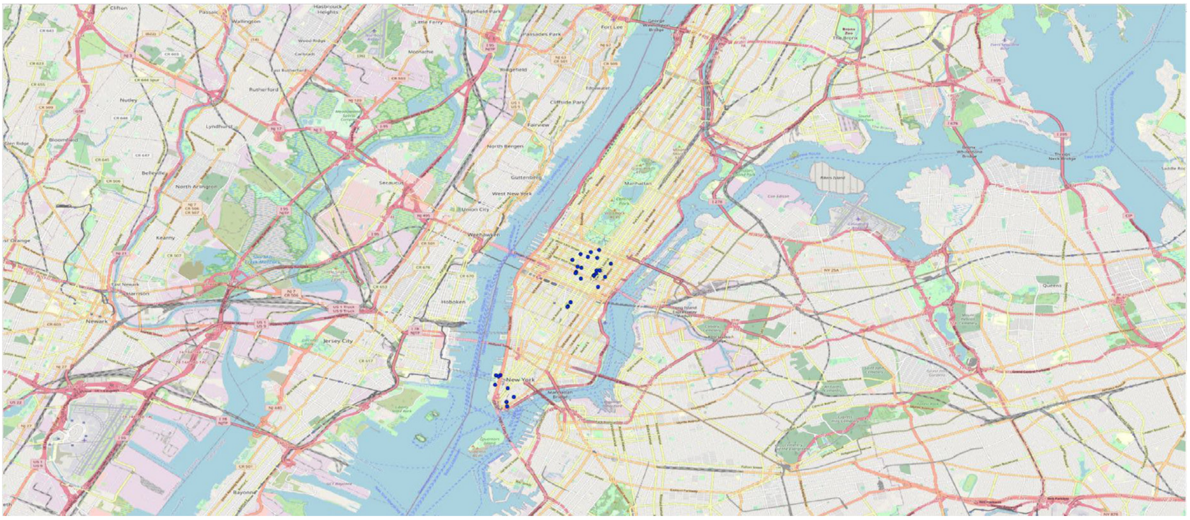

Fig. 3. Headquarters of a sample of financial institutions in NYC. This figure is a map showing the headquarters (dots) of the 20 financial institutions included in the taxi trip analysis discussed in Section 5.1. The sample consists of SEC Registered Investment Advisers headquartered in New York City and with more than 250 employees performing investment advisory functions. Note that the map shows more than 20 locations because some of these institutions moved their headquarters during the sample period, which runs from January 1, 2009 to June 30, 2016. Each address is manually transformed into latitude/longitude coordinates using OpenStreetMap and subsequently plotted on the map using gpspointplotter.com.

(allowing for the presence of a support person), the pick-up location is within a 140 m × 140 m square centered on institution *i*'s headquarters, and the drop-off location is within one of the 140 m × 140 m squares centered on the 51 EDs in the City.[11]

For comparability with the dependent variable, *Illness18_64* measures the number of ED visits made by 18–64 year old New Yorkers on day *t* and is standardized to have zero mean and unit variance. Eq. (1) also contains a number of covariates to control for general patterns in taxi use (Hochmair, 2016): *Rain* (*Snow*) is a dummy that takes the value of 1 when the amount of precipitation (snow) in NYC is greater than 0.1 inches (positive), and 0 otherwise. *Holiday* is a dummy that takes the value of 1 on the three

[11] The results are analogous when using 120 m × 120 m or 160 m × 160 m squares.

Table 3
Taxi trips from financial institutions headquartered in NYC to local hospitals.

Dependent variable	(1) # of taxi trips to hospital	(2) # of taxi trips to hospital	(3) # of taxi trips to hospital	(4) = 1 if # of taxi trips to hospital > 0	(5) = 1 if # of taxi trips to hospital > 0	(6) = 1 if # of taxi trips to hospital > 0
Illness18_64	1.105***	1.019**	1.005	1.171***	1.037**	1.010
	(0.000)	(0.011)	(0.744)	(0.000)	(0.028)	(0.725)
Large			1.007			0.994
			(0.901)			(0.942)
Medium			1.018			1.002
			(0.832)			(0.969)
Illness18_64 × Large			1.032			1.062
			(0.201)			(0.160)
Illness18_64 × Medium			1.010			1.034
			(0.657)			(0.448)
Institution FE	No	Yes	Yes	No	Yes	Yes
Pick-up location FE	No	Yes	Yes	No	Yes	Yes
Calendar FE	No	Yes	Yes	No	Yes	Yes
Weather Controls	No	Yes	Yes	No	Yes	Yes
Business Conditions	No	Yes	Yes	No	Yes	Yes
N	54760	54760	54760	54760	54760	54760
H_0: Illness18_64 + Illness18_64 × Large = 0			1.037**			1.073**
p-value			(0.026)			(0.028)
H_0: Illness18_64 + Illness18_64 × Medium = 0			1.015			1.045*
p-value			(0.275)			(0.095)

Notes: columns 1–3 of this table display the incidence-rate ratios generated by fitting an unconditional negative binomial model based on Eq. (1). The dependent variable measures the number of taxi trips on day t starting from financial institution i's headquarters and ending at a local ED. Specifically, a taxi trip is counted if the pick-up time is between 8 am and 8 pm, the number of passengers is 1 or 2 (allowing for the presence of a support person), the pick-up location is within a 140 m × 140 m square centered on institution i's headquarters, and the drop-off location is within one of the 140 m × 140 m squares centered on the 51 EDs in the City. The odds ratios displayed in columns 4–6 are generated by fitting an unconditional logit model based on Eq. (1), where the dependent variable takes the value of 1 if at least one taxi trip is observed on day t between financial institution i's headquarters and a local ED, and 0 otherwise. *Illness18_64* measures the number of ED visits made by 18–64 year old New Yorkers on day t, where a visit is included in the count if it is related to any of the following syndrome categories: asthma, diarrhea, influenza-like, respiratory, vomiting. *Illness18_64* is standardized to have zero mean and unit variance. *Large* (*Medium*) is a dummy that takes the value of 1 when financial institution i has more than 1000 (between 501 and 1000) employees performing investment advisory functions, and 0 otherwise. The calendar fixed effects consist of *Day*, *Month*, *Year*, and *Holiday*, which is a dummy taking the value of 1 on the three days around each major holiday, and 0 otherwise. The weather controls consist of *Rain* and *Snow*, where *Rain* (*Snow*) is a dummy that takes value of 1 when the amount of precipitation (snow) in NYC is greater than 0.1 inches (positive), and 0 otherwise. The business conditions control refers to Aruoba et al.'s (2009) real-time daily business conditions index for the U.S. economy. Each regression contains institution and pick-up location dummy variables. The p-values in parentheses are computed using the wild cluster bootstrap-t procedure (Roodman et al., 2019), where the standard errors are clustered by institution and day and bootstrapped on the institution dimension (null hypothesis imposed; 999 replications). *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

days around each major holiday, and 0 otherwise. *Institution FE* and *Pick-up location FE* represent two sets of institution and pick-up location dummies, respectively. The remaining variables are as defined in Table 1.[12]

Given the nature of the dependent variable, I estimate unconditional (mean-dispersion) negative binomial models.[13] Since the number of institutions in the sample is relatively small, I compute p-values using the wild cluster bootstrap-t procedure (Roodman et al., 2019), where the standard errors are clustered by institution and day and bootstrapped on the institution dimension (null hypothesis imposed; 999 replications).

Table 3 displays the estimated incidence-rate ratios in columns 1–3. In column 1 (2), where no (all) controls are included, the estimates indicate that a one-standard-deviation increase in *Illness18_64* is accompanied by a 10.5% (1.9%) increase in the number of taxi trips between a financial institution and a local hospital. In both cases, the coefficient of interest is statistically different from zero at least at the 5% level. As such, the results are consistent with my identification assumption that illness incidence among 18–64 year old New Yorkers is a valid proxy for illness incidence among NYC's investment professionals.

To exploit the variation in the size of the institutions in the sample, I also re-estimate model (1) with the inclusion of the interactions *Illness18_64* × *Large* and *Illness18_64* × *Medium*, where *Large* (*Medium*) is a dummy that takes the value of 1 when the institution in question has more than 1000 (between 501 and 1000) employees performing investment advisory functions, and 0 otherwise. The rationale is that, when illness strikes, the greater is the size of a financial institution, the greater is the number of its employees that may be expected to get sick and visit an ED. Indeed, the results in column 3 of Table 3 indicate that the point estimate of the coefficient on *Illness18_64* is increasing in the size of the institution: 1.037, 1.015, and 1.005 for large, medium,

[12] Precipitation and snow data are from the weather station at LaGuardia Airport and are available from the National Oceanic and Atmospheric Administration's website. Note that institution and pick-up location dummies do not fully overlap because, as mentioned earlier, some institutions changed address during the sample period.

[13] I use a negative binomial model instead of a Poisson model because Allison and Waterman (2002) show that the unconditional negative binomial model with a dummy for each cross-sectional unit "does not suffer from incidental parameter bias, and has much better sampling properties than the fixed-effects Poisson estimator". Additionally, the data reveal the presence of overdispersion, which renders the Poisson model unsuitable.

and small institutions, respectively. Additionally, only the coefficient for large institutions (*Illness18_64* + *Illness18_64* × *Large*) is statistically significant.

Lastly, to assess the robustness of these findings, I also estimate unconditional logit models where the regressors are the same as in model (1), but the dependent variable takes the value of 1 if *EDtrips* is greater than 0, and 0 otherwise.[14] The odds ratios displayed in columns 4–6 of Table 3 are very much in line with the output of the negative binomial model. For example, when all controls are included (column 5), a one-standard-deviation increase in *Illness18_64* increases by 3.7% the odds of observing at least one taxi trip between a financial institution and a local hospital. Furthermore, this effect is increasing in the size of the institution. As such, these patterns support my identification assumption.

5.2. Illness incidence among New Yorkers and stock market outcomes

To investigate the relation between changes in acute illness incidence in NYC and stock market outcomes, I estimate the following panel regressions by OLS:

$$\begin{aligned} DepVar_{i,t} = {} & \alpha + \beta_1 \Delta LogIllnessX_t + \pi_1 SUE_{i,t} \times I^{SUE}_{i,t} + \pi_2 RecChng_{i,t} \times I^{RecChng}_{i,t} \\ & + \varphi_1 \Delta ADSBusConditions_t + \varphi_2 \Delta MktUncertainty_t + \mu_1 SADOR_t \\ & + \mu_2 Cloudy_t + \mu_3 Sunny_t + \mu_4 HighPollution_t + \vartheta_1 PreH_t + \vartheta_2 PostH_t \\ & + \sum_{j=2}^{5} \gamma_j Day_{j,t} + \sum_{j=2}^{12} \delta_j Month_{j,t} + \sum_{j=2008}^{2017} \theta_j Year_{j,t} + FirmFE + \varepsilon_{i,t} \end{aligned} \tag{2}$$

where *DepVar* is one of the following six dependent variables, as defined in Table 1: *ΔLogShares*, *ΔLogVolume*, *ΔLogTurnover*, *ΔLogVolatility*, $\Delta Spread$, and *Return*.[15] I^{SUE} ($I^{RecChng}$) is an indicator variable that takes the value of 1 when an EPS announcement (a change in analysts' consensus recommendation) involving stock *i* occurs on day *t*, and 0 otherwise.

Since the literature has detected day-of-the-week, holiday, and seasonal patterns in syndromic data (Mathes et al., 2017) and stock market outcomes, Eq. (2) contains a number of calendar controls (e.g., *Day*). And since ED visits (Stieb et al., 2009) and stock market outcomes (Kamstra et al., 2003) are correlated with some environmental factors, Eq. (2) contains a set of environmental/behavioral controls (e.g., *SADOR*). Each regression also includes firm fixed effects (Correia, 2017).

In what follows, unless otherwise stated, I compute robust standard errors with two-way clustering by firm and day following Petersen (2009). Additionally, to facilitate interpretations and comparisons, all variables that proxy for changes in acute illness incidence are standardized to have zero mean and unit variance.

Table 4 displays the estimates generated by fitting model (2). For X = 0_4, 5_17, and 65+, the point estimates of the coefficients on *ΔLogIllnessX* are small in size, and the coefficients themselves are statistically indistinguishable from zero. This is consistent with my expectations, as individuals in the 0–4, 5–17, and 65+ age groups typically do not belong to the category of investment professionals.

Conversely, the coefficients on *ΔLogIllness18_64* indicate that there is an economically and statistically significant (at the 5% level) relation between changes in acute illness incidence among 18–64 year old New Yorkers and NYSE trading activity, volatility, and returns. A one-standard-deviation increase in *ΔLogIllness18_64* reduces daily turnover, dollar volume, and shares traded by about 2%, reduces intraday volatility by 1.2%, and decreases stock returns by 0.10%.[16] Overall, these estimates are in line with the findings of McTier et al. (2013), though their data are not broken down by age group.

The patterns in Table 4 also help address potential reverse causality concerns: If daily stock market dynamics caused ED visits for the five syndrome categories under observation (and to the best of my knowledge, there is no evidence in the literature that they do), then the coefficients on *ΔLogIllness18_64* and *ΔLogIllness65+* should display similar patterns of size and statistical significance. This is because both the 18–64 and 65+ age groups exhibit a high degree of participation in the stock market (Guiso et al., 2003; Bonaparte and Kumar, 2013), and consequently the wealth levels of both groups are similarly affected by market fluctuations. Yet, the numbers in Table 4 reveal that this is not the case.

5.3. Portfolios sorted by firm characteristics

To provide additional support for my identification assumption, I also take advantage of the evidence that different groups of market participants tend to trade stocks with different characteristics. For example, one of the groups that falls into the category of stock market professionals is institutional investors. According to the literature, this group tends to trade large-cap, high-price, liquid, high institutional-ownership, low idiosyncratic-volatility, and older stocks (Kumar and Lee, 2006; Kumar, 2009; Puckett and Yan, 2011).

[14] The unconditional logit model with a dummy for each cross-sectional unit, also known as logit dummy variable estimator, is biased for small T, but it is consistent for small N as T goes to infinity (Katz, 2001). Additionally, when T is very large, the conditional logit model "runs into serious numerical issues", whereas the unconditional logit model works well (Beck, 2018).

[15] Note that, when the dependent variable is other than *Return*, the control variables *SUE*, *RecChng*, *ΔADSBusConditions*, and *ΔMktUncertainty* enter the model in absolute value, as what matters is not their sign, but rather the amount of information that they convey.

[16] Further analyses where data on ED visits are broken down by borough and syndrome category (see Section A2 in the online Appendix) produce results that are consistent with those presented above. Similarly, including additional regressors to control for the arrival of new information, news sentiment, time-varying risk aversion, and further determinants of bid–ask spreads does not alter the conclusions (see Section A3 in the online Appendix).

Table 4
Incidence of acute illness symptoms among New Yorkers and stock market outcomes: Analysis by age group.

	(1) ΔLogShares	(2) ΔLogVolume	(3) ΔLogTurnover	(4) ΔLogVolatility	(5) ΔSpread	(6) Return
ΔLogIllness0_4	−0.689	−0.658	−0.685	0.298	0.001	−0.016
	(0.447)	(0.447)	(0.447)	(0.368)	(0.006)	(0.029)
ΔLogIllness5_17	0.160	0.197	0.159	−0.083	−0.002	−0.003
	(0.512)	(0.510)	(0.512)	(0.442)	(0.006)	(0.034)
ΔLogIllness18_64	−1.945**	−2.025**	−1.944**	−1.220**	−0.014	−0.100**
	(0.895)	(0.894)	(0.895)	(0.564)	(0.010)	(0.043)
ΔLogIllness65+	−0.389	−0.312	−0.389	−0.030	0.000	0.016
	(0.607)	(0.605)	(0.608)	(0.515)	(0.005)	(0.040)
SUE	No	No	No	No	No	Yes
Abs(SUE)	Yes	Yes	Yes	Yes	Yes	No
RecChng	No	No	No	No	No	Yes
Abs(RecChng)	Yes	Yes	Yes	Yes	Yes	No
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar FE	Yes	Yes	Yes	Yes	Yes	Yes
Behavioral controls	Yes	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes	Yes
N	3,558,469	3,558,469	3,552,538	3,556,131	3,453,250	3,557,760

Notes: this table displays the estimates generated by fitting model (2). Only the estimated coefficients on the illness incidence proxies are displayed. The intersection of each row and column represents a different regression: The dependent variable varies across columns, whereas the key explanatory variable, which is standardized to have zero mean and unit variance, varies across rows. The key explanatory variable, *ΔLogIllnessX*, measures the change in the log of the number of ED visits made by New Yorkers in *X* age group (*X* = *0_4*, *5_17*, *18_64*, *65+*), where a visit is included in the count if it is related to any of the following syndrome categories: asthma, diarrhea, influenza-like, respiratory, vomiting. The calendar fixed effects consist of *Day*, *Month*, *Year*, *PreH*, and *PostH*. The behavioral controls consist of *SADOR*, *Cloudy*, *Sunny*, and *HighPollution*. In column 6 (1–5), the economic controls consist of *ΔADSBusConditions* (*Abs(ΔADSBusConditions)*) and *ΔMktUncertainty* (*Abs(ΔMktUncertainty)*). All the variables are as defined in Table 1. Each regression contains firm fixed effects. The t-statistics in parentheses are computed based on standard errors clustered by firm and day. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Conversely, retail investors, who are not market professionals, prefer trading low-price, high idiosyncratic-volatility, low institutional-ownership, small-cap, and young stocks (Kumar and Lee, 2006; Kumar, 2009). This presents an opportunity for a falsification test: According to my identification assumption, the relation between *ΔLogIllness18_64* and stock market outcomes should be stronger among stocks favored by institutional investors than among stocks favored by retail investors.[17]

To conduct this test, every three months I sort stocks into quintiles based on the average value of one of their characteristics (idiosyncratic volatility, number of analysts following the firm, market cap, age, intraday volatility, price, bid–ask spread, institutional ownership, and Amihud illiquidity) during the previous month. Quintile 1 represents the lowest quintile and quintile 5 the highest one. For each sorting variable, I then re-estimate model (2) after including interactions between *ΔLogIllness18_64* and the quintile dummies.

Table 5 displays the estimates of the effect of *ΔLogIllness18_64* for quintile 1 (Q1) and quintile 5 (Q5) along with the difference between the two effects (Q5-Q1). Since stock returns are also affected by arbitrage forces, I focus on columns 1–5, where the dependent variables measure trading activity, volatility, and the bid–ask spread. The estimates reveal that the effects of *ΔLogIllness18_64* on market trading activity and volatility are present only (or, at the very least, are statistically significantly stronger) among the types of stocks favored by institutional investors. As for bid–ask spread, the coefficients of interest are mostly statistically insignificant. In sum, these results are consistent with my identification assumption that *ΔLogIllness18_64* captures fluctuations in acute illness incidence among NYC's stock market professionals.

5.4. Firms headquartered in NYC vs. firms headquartered elsewhere

The second threat concerns the identification assumption that *ΔLogIllness18_64* affects key stock market outcomes *only* via stock market professionals. The evidence discussed in the previous section is consistent with the view that the effect of *ΔLogIllness18_64* is not mediated by non-professionals such as retail investors. However, in principle, increased rates of acute illness symptoms among 18–64 year old New Yorkers might also affect the productivity of employees of listed companies headquartered in NYC and directly impact the market values of these companies. Potentially, this channel might affect market trading activity and volatility as well.

If the relation between *ΔLogIllness18_64* and stock market outcomes were driven by this group of employees, then it should exist only for companies headquartered in the City. To run a falsification test, I first identify all listed companies headquartered in NYC, which amount to about 7.3% of the sample. Next, I re-estimate model (2) after including among the regressors the interaction $\Delta LogIllness18\_64 \times NYC$, where *NYC* is a dummy that takes the value of 1 if the company in question is headquartered in the City, and 0 otherwise.

[17] I do not specifically analyze the role of other agents such as designated market makers and supplemental liquidity providers because, during the period under investigation, their market-making activities were "mostly carried out by high-frequency proprietary algorithms" (O'Hara et al., 2019), which are immune to human illnesses. As for the NYSE's trading floor, critics have argued that, though still populated by humans, it has turned into a mere façade for TV cameras (Levine, 2018).

Table 5
Acute illness incidence among 18–64 year old New Yorkers and stock market outcomes: Quintile portfolios sorted by firm characteristics.

	(1) ΔLogShares	(2) ΔLogVolume	(3) ΔLogTurnover	(4) ΔLogVolatility	(5) ΔSpread	(6) Return
Idiosyncratic volatility						
Q1	−2.447***	−2.523***	−2.443***	−1.917***	−0.013	−0.062
	(0.890)	(0.889)	(0.890)	(0.599)	(0.010)	(0.040)
Q5	−1.187	−1.277	−1.192	−0.199	−0.013	−0.151***
	(0.899)	(0.898)	(0.899)	(0.547)	(0.011)	(0.049)
Q5-Q1	1.260***	1.246***	1.251***	1.719***	−0.001	−0.089***
	(0.210)	(0.212)	(0.211)	(0.242)	(0.003)	(0.023)
Number of analysts						
Q1	−1.396	−1.459	−1.401	−0.675	−0.016	−0.096**
	(0.914)	(0.913)	(0.914)	(0.559)	(0.011)	(0.044)
Q5	−2.388***	−2.491***	−2.388***	−1.672***	−0.014	−0.105**
	(0.875)	(0.874)	(0.875)	(0.593)	(0.009)	(0.043)
Q5-Q1	−0.992***	−1.032***	−0.987***	−0.997***	0.002	−0.008
	(0.272)	(0.272)	(0.272)	(0.218)	(0.004)	(0.010)
Market cap						
Q1	−1.060	−1.128	−1.063	−0.188	−0.014	−0.112**
	(0.919)	(0.918)	(0.919)	(0.546)	(0.012)	(0.045)
Q5	−2.604***	−2.704***	−2.600***	−1.811***	−0.013	−0.086**
	(0.877)	(0.876)	(0.877)	(0.597)	(0.010)	(0.042)
Q5-Q1	−1.544***	−1.576***	−1.536***	−1.623***	0.001	0.025*
	(0.306)	(0.307)	(0.306)	(0.263)	(0.005)	(0.014)
Age						
Q1	−1.607*	−1.673*	−1.620*	−0.684	−0.011	−0.113***
	(0.908)	(0.907)	(0.908)	(0.549)	(0.011)	(0.043)
Q5	−2.228**	−2.304**	−2.224**	−1.471**	−0.017	−0.084*
	(0.901)	(0.900)	(0.901)	(0.590)	(0.011)	(0.043)
Q5-Q1	−0.621***	−0.630***	−0.604***	−0.788***	−0.006**	0.028***
	(0.201)	(0.201)	(0.201)	(0.180)	(0.003)	(0.008)
Intraday volatility						
Q1	−2.227**	−2.296***	−2.233**	−1.912***	−0.013	−0.049
	(0.889)	(0.888)	(0.889)	(0.591)	(0.010)	(0.040)
Q5	−1.338	−1.436	−1.345	−0.229	−0.013	−0.162***
	(0.900)	(0.899)	(0.900)	(0.552)	(0.011)	(0.051)
Q5-Q1	0.889***	0.860***	0.888***	1.683***	0.000	−0.114***
	(0.220)	(0.223)	(0.221)	(0.241)	(0.004)	(0.028)
Stock price						
Q1	−1.529*	−1.600*	−1.530*	−0.449	−0.015	−0.130***
	(0.903)	(0.902)	(0.903)	(0.545)	(0.012)	(0.047)
Q5	−2.130**	−2.217**	−2.126**	−1.544***	−0.012	−0.083**
	(0.890)	(0.889)	(0.891)	(0.592)	(0.010)	(0.042)
Q5-Q1	−0.601**	−0.617**	−0.596**	−1.095***	0.003	0.047***
	(0.239)	(0.240)	(0.239)	(0.224)	(0.005)	(0.017)
Bid–Ask spread						
Q1	−2.440***	−2.526***	−2.438***	−1.905***	−0.010	−0.077*
	(0.878)	(0.876)	(0.878)	(0.600)	(0.009)	(0.041)
Q5	−0.777	−0.838	−0.789	−0.147	−0.016	−0.114**
	(0.912)	(0.911)	(0.913)	(0.539)	(0.013)	(0.045)
Q5-Q1	1.663***	1.688***	1.648***	1.758***	−0.006	−0.037**
	(0.273)	(0.273)	(0.274)	(0.265)	(0.006)	(0.015)
Institutional ownership						
Q1	−0.557	−0.569	−0.567	−0.654	0.003	−0.069*
	(1.039)	(1.037)	(1.039)	(0.627)	(0.007)	(0.039)
Q5	−2.173**	−2.210**	−2.170**	−1.596**	0.003	−0.075*
	(1.040)	(1.039)	(1.040)	(0.642)	(0.005)	(0.042)
Q5-Q1	−1.616***	−1.641***	−1.602***	−0.943***	−0.001	−0.006
	(0.231)	(0.230)	(0.231)	(0.206)	(0.003)	(0.010)

(*continued on next page*)

Table 5 (*continued*).

	(1) ΔLogShares	(2) ΔLogVolume	(3) ΔLogTurnover	(4) ΔLogVolatility	(5) ΔSpread	(6) Return
Amihud illiquidity						
Q1	−2.573***	−2.673***	−2.570***	−1.874***	−0.013	−0.084**
	(0.875)	(0.874)	(0.875)	(0.594)	(0.009)	(0.042)
Q5	−1.054	−1.119	−1.070	−0.207	−0.012	−0.108**
	(0.923)	(0.922)	(0.923)	(0.546)	(0.013)	(0.045)
Q5-Q1	1.519***	1.554***	1.500***	1.667***	0.001	−0.024*
	(0.313)	(0.314)	(0.314)	(0.264)	(0.006)	(0.014)

Notes: this table displays the estimates generated by fitting model (2) with the inclusion of interactions between *ΔLogIllness18_64* and a set of quintile dummies. *ΔLogIllness18_64* measures the daily change in the log of the number of ED visits made by 18–64 year old New Yorkers, where a visit is included in the count if it is related to any of the following syndrome categories: asthma, diarrhea, influenza-like, respiratory, vomiting. *ΔLogIllness18_64* is standardized to have zero mean and unit variance. Every three months, I sort stocks into quintiles based on the average value of one of their characteristics (e.g., idiosyncratic volatility) during the previous month. Quintile 1 represents the lowest quintile and quintile 5 the highest quintile. For each column, each panel represents the output of a different regression where the quintile dummies are constructed based on a different sorting characteristic. To save space, the table displays only the estimated effects of *ΔLogIllness18_64* for quintile 1 (Q1) and quintile 5 (Q5) along with the differences between the two effects (Q5-Q1). The t-statistics in parentheses are computed based on standard errors clustered by firm and day. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table 6
Firms headquartered in NYC vs. firms headquartered elsewhere.

	(1) ΔLogShares	(2) ΔLogVolume	(3) ΔLogTurnover	(4) ΔLogVolatility	(5) ΔSpread	(6) Return
ΔLogIllness18_64	−1.949**	−2.030**	−1.950**	−1.212**	−0.013	−0.100**
	(0.896)	(0.895)	(0.896)	(0.564)	(0.010)	(0.043)
ΔLogIllness18_64 × NYC	0.059	0.071	0.086	−0.105	−0.006*	−0.001
	(0.233)	(0.235)	(0.233)	(0.167)	(0.003)	(0.007)
SUE & RecChng						Yes
Abs(SUE) & Abs(RecChng)	Yes	Yes	Yes	Yes	Yes	
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar FE	Yes	Yes	Yes	Yes	Yes	Yes
Behavioral controls	Yes	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes	Yes
N	3,558,469	3,558,469	3,552,538	3,556,131	3,453,250	3,557,760
H_0: ΔLogIllness18_64 + ΔLogIllness18_64 × NYC = 0	−1.890**	−1.959**	−1.864**	−1.317**	−0.019*	−0.101**
	(0.912)	(0.911)	(0.911)	(0.582)	(0.010)	(0.043)

Notes: this table displays the estimates generated by fitting model (2) with the inclusion of the interaction *ΔLogIllness18_64×NYC*. *ΔLogIllness18_64* measures the change in the log of the number of ED visits by 18–64 year old New Yorkers, where a visit is included in the count if it is related to any of the following syndrome categories: asthma, diarrhea, influenza-like, respiratory, vomiting. *ΔLogIllness18_64* is standardized to have zero mean and unit variance. *NYC* is a dummy variable that takes the value of 1 when the firm is headquartered in New York City, and 0 otherwise. The calendar fixed effects consist of *Day*, *Month*, *Year*, *PreH*, and *PostH*. The behavioral controls consist of *SADOR*, *Cloudy*, *Sunny*, and *HighPollution*. In column 6 (1–5), the economic controls consist of *ΔADSBusConditions* (*Abs(ΔADSBusConditions)*) and *ΔMktUncertainty* (*Abs(ΔMktUncertainty)*). All the variables are as defined in Table 1. Each regression contains firm fixed effects. The t-statistics in parentheses are computed based on standard errors clustered by firm and day. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table 6 displays the estimates. In columns 1–4 and 6, the coefficient on $\Delta LogIllness18\_64 \times NYC$ is trivial in size and statistically indistinguishable from zero. As such, there is no evidence that the impact of *ΔLogIllness18_64* on market trading activity, volatility, and returns is any different for companies headquartered in NYC than for companies headquartered elsewhere. These results are consistent with my identification assumption.

As for column 5, where the dependent variable is $\Delta Spread$, the coefficient on *ΔLogIllness18_64 × NYC* is negative and statistically significant. And the coefficient on *ΔLogIllness18_64* is statistically significantly negative *only* for stocks headquartered in the City. Specifically, a one-standard-deviation increase in *ΔLogIllness18_64* leads to a reduction in closing bid–ask spreads of 1.9 basis points (=$\Delta LogIllness18\_64 + \Delta LogIllness18\_64 \times NYC$) for NYC firms. It is worth emphasizing that this last result is not inconsistent with my identification assumption. A possible interpretation is that investment professionals have an informational advantage on local stocks (Bernile et al., 2019; Gaspar and Massa, 2007). If this is the case, then an exogenous reduction in the productivity of NYC's investment professionals (caused by acute illness symptoms) may lower adverse selection costs for other market participants, thereby raising liquidity and reducing bid–ask spreads for stocks headquartered in the City.

5.5. Illness incidence among market professionals and price discovery

I now turn to examining how fluctuations in the incidence of acute illness symptoms among investment professionals affect the incorporation of public information into stock prices. My conjecture is that firm-specific news is impounded less quickly into prices when more professional market participants are absent or less productive due to physical illness.

To shed light on this question, I re-estimate model (2) after including among the regressors *AbNumStories* and the interaction $\Delta LogIllness18\_64 \times AbNumStories$, where *AbNumStories* measures the arrival of news concerning company i on day t (see Table 1).

Table 7
Acute illness incidence among investment professionals and price discovery: Public news flow.

	(1) ΔLogShares	(2) ΔLogVolume	(3) ΔLogTurnover	(4) ΔLogVolatility
ΔLogIllness18_64	−1.707*	−1.717*	−1.708*	−0.858
	(0.927)	(0.925)	(0.927)	(0.619)
AbNumStories	14.719***	14.757***	14.602***	15.494***
	(0.291)	(0.290)	(0.289)	(0.316)
ΔLogIllness18_64 × AbNumStories	−0.628*	−0.630*	−0.620*	−1.554***
	(0.333)	(0.332)	(0.334)	(0.266)
Abs(SUE) & Abs(RecChng)	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Calendar FE	Yes	Yes	Yes	Yes
Behavioral controls	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes
N	2,909,457	2,909,457	2,905,802	2,908,347

Notes: this table displays the estimates generated by fitting model (2) with the inclusion of *AbNumStories* and the interaction *ΔLogIllness18_64×AbNumStories*. *ΔLogIllness18_64* measures the daily change in the log of the number of ED visits made by 18–64 year old New Yorkers, where a visit is included in the count if it is related to any of the following syndrome categories: asthma, diarrhea, influenza-like, respiratory, vomiting. *ΔLogIllness18_64* is standardized to have zero mean and unit variance. *AbNumStories* is a continuous variable, calculated by Bloomberg, that captures abnormal daily news flow and ranges from 0 to 4. Specifically, it measures "the amount of stories currently being published on a company relative to the flow over the previous 45 days", with higher values indicating unusually higher story flow (Bloomberg, 2020). The calendar fixed effects consist of *Day*, *Month*, *Year*, *PreH*, and *PostH*. The behavioral controls consist of *SADOR*, *Cloudy*, *Sunny*, and *HighPollution*. The economic controls consist of *Abs(ΔADSBusConditions)* and *Abs(ΔMktUncertainty)*. All the variables are as defined in Table 1. Each regression contains firm fixed effects. The t-statistics in parentheses are computed based on standard errors clustered by firm and day. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

The estimates in Table 7 show that the coefficients on *AbNumStories* in columns 1–4 are positive, which confirms that, as expected, market trading activity and volatility rise in response to the arrival of new firm-specific information. However, the coefficients on $\Delta LogIllness18\_64 \times AbNumStories$ are negative and statistically significant, which suggests that increased rates of acute illness among investment professionals reduce the impact of public news on the two objects of interest.

Specifically, a one-standard-deviation increase in *ΔLogIllness18_64* decreases by 4.2% (=0.62/14.602) the effect on turnover of a one-unit increase in *AbNumStories*, and it decreases by 10% (=1.554/15.494) its effect on volatility. In other words, a piece of news has a smaller impact on market trading activity and volatility when more investment professionals are absent or less productive due to acute illness. This is consistent with the view that market professionals play a key role in the impounding of public news into prices.

Next, I focus on two specific types of public information that have received considerable attention in the literature, namely earnings announcements and changes in analysts' consensus recommendations. If stock market professionals trade in response to news, then, when more of them are impaired by physical illness, one would expect to observe a smaller immediate reaction of stock prices to an earnings surprise (or a change in analysts' consensus recommendation) and a larger delayed response.

To model the phenomenon of interest, in the spirit of DellaVigna and Pollet (2009) and Ben-Rephael et al. (2017), I employ variants of the following specification:

$$\begin{aligned}Ab\_Ret_i^{[t+h,t+k]} &= \alpha + \pi_1 SUE_{i,t} + \beta_1 \Delta LogIllness18\_64_t + \gamma SUE_{i,t} \times \Delta LogIllness18\_64_t \\ &+ \mu_1 AbNumStories_{i,t} + \mu_2 HLtoH_{i,t} + \mu_3 Ret_i^{[t-27,t-6]} + \mu_4 AVG\_Turnover_{i,t} \\ &+ \mu_5 AVG\_Spread_{i,t} + \mu_6 SDRET_{i,t} + \mu_7 LnSize_{i,t} + \mu_8 InstHold_{i,t} \\ &+ \mu_9 LnNumEst_{i,t} + \sum_{j=2}^{5} \theta_j Day_{j,t} + Quarter\ FE + \varepsilon_{i,t}.\end{aligned} \tag{3}$$

$Ab\_Ret_i^{[t+h,t+k]}$ is stock i's cumulative abnormal return from day $t+h$ ($h = 0, 1$) to $t+k$ ($k = 0, 3, 5, 7, 10, 20, 30, 40$), where t is an earnings announcement day. The remaining variables are as defined in Table 1. Each specification includes quarter fixed effects, and the standard errors are clustered by firm and day.

Table 8 displays the relevant estimates. In column 1, where the dependent variable is $Ab\_Ret_i^{[t]}$, the coefficient on *SUE* (i.e., standardized unexpected earnings) is positive, as expected, and the coefficient on *SUE ×ΔLogIllness18_64* is negative. Both are statistically different from zero at least at the 5% level. The implication is that a positive earnings surprise raises contemporaneous stock returns. However, a one-standard-deviation increase in *ΔLogIllness18_64* reduces the earnings response coefficient by 18% (=0.085/0.47).

In column 7, where the dependent variable is $Ab\_Ret_i^{[t+1,t+30]}$, the coefficient on *SUE* is positive, which is consistent with the post earnings announcement drift (PEAD) documented in the literature (Bernard and Thomas, 1989). However, the coefficient on *SUE×ΔLogIllness18_64* is also positive and statistically significant, suggesting that a one-standard-deviation increase in

Table 8
Acute illness incidence among investment professionals and price discovery: Earnings announcements.

	(1) Abnormal Return(t)	(2) Abnormal Return (t + 1,t + 3)	(3) Abnormal Return (t + 1,t + 5)	(4) Abnormal Return (t + 1,t + 7)	(5) Abnormal Return (t + 1,t + 10)	(6) Abnormal Return (t + 1,t + 20)	(7) Abnormal Return (t + 1,t + 30)	(8) Abnormal Return (t + 1,t + 40)
SUE	0.470***	0.634***	0.674***	0.727***	0.717***	0.698***	0.648***	0.728***
	(0.048)	(0.067)	(0.073)	(0.071)	(0.075)	(0.092)	(0.111)	(0.124)
ΔLogIllness18_64	−0.072	0.006	0.062	0.041	0.078	0.011	−0.003	0.018
	(0.046)	(0.057)	(0.075)	(0.096)	(0.105)	(0.136)	(0.172)	(0.203)
SUE × ΔLogIllness18_64	−0.085**	0.094	0.139*	0.115*	0.169**	0.193**	0.187*	0.163
	(0.040)	(0.066)	(0.075)	(0.070)	(0.074)	(0.098)	(0.110)	(0.129)
AbNumStories	0.370***	0.089	0.147**	0.212***	0.309***	0.505***	0.759***	0.811***
	(0.059)	(0.064)	(0.069)	(0.078)	(0.082)	(0.100)	(0.127)	(0.143)
HLtoH	−23.419***	−1.341	0.063	1.210	−0.314	0.129	1.557	0.115
	(2.860)	(1.443)	(1.605)	(1.806)	(1.929)	(2.129)	(2.486)	(2.942)
Return(t − 27,t − 6)	−0.011***	−0.005	−0.014**	−0.021***	−0.025***	−0.040***	−0.038***	−0.050***
	(0.004)	(0.006)	(0.007)	(0.007)	(0.008)	(0.011)	(0.014)	(0.016)
AVG_Turnover(t − 5,t − 1)	0.024	−0.117**	−0.112*	−0.105	−0.106	−0.256***	−0.346***	−0.499***
	(0.041)	(0.055)	(0.064)	(0.075)	(0.076)	(0.087)	(0.114)	(0.130)
AVG_Spread(t − 5,t − 1)	0.027	0.019	−0.008	0.054	0.105	0.120	0.089	0.280
	(0.062)	(0.070)	(0.081)	(0.099)	(0.111)	(0.130)	(0.135)	(0.173)
SDRET	0.261***	0.185***	0.251***	0.310***	0.318***	0.371***	0.646***	0.819***
	(0.044)	(0.070)	(0.095)	(0.114)	(0.108)	(0.134)	(0.175)	(0.211)
LnSize	−0.190***	−0.155***	−0.210***	−0.239***	−0.315***	−0.518***	−0.813***	−1.006***
	(0.029)	(0.039)	(0.045)	(0.049)	(0.053)	(0.063)	(0.077)	(0.090)
InstHold	0.002*	0.002	0.002	0.003*	0.004**	0.004*	0.005*	0.001
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
LnNumEst	−0.048	0.234***	0.326***	0.357***	0.449***	0.700***	1.120***	1.509***
	(0.060)	(0.083)	(0.096)	(0.110)	(0.124)	(0.149)	(0.186)	(0.204)
N	36,917	36,914	36,913	36,912	36,912	36,906	36,879	36,835

Notes: this table displays the estimates generated by fitting model (3). The dependent variable is the cumulative abnormal return of the given stock from day $t + h$ to $t + k$ ($h = 0$, 1 and $k = 0$, 3, 5, 7, 10, 20, 30, 40), where t is an earnings announcement day, and abnormal returns are computed using the Fama and French (1993) 3-factor model with an estimation window from $t - 200$ to $t - 21$. *SUE* measures standardized unexpected earnings, that is the spread between actual and expected EPS on day t divided by the stock price on day $t - 5$, as in DellaVigna and Pollet (2009). *ΔLogIllness18_64* measures the change in the log of the number of ED visits made by 18–64 year old New Yorkers, where a visit is included in the count if it is related to any of the following syndrome categories: asthma, diarrhea, influenza-like, respiratory, vomiting. *ΔLogIllness18_64* is standardized to have zero mean and unit variance. The remaining variables are as defined in Table 1. Each regression includes day-of-the-week and quarter fixed effects, and the t-statistics in parentheses are computed based on standard errors clustered by firm and day. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

ΔLogIllness18_64 increases by 29% (=0.187/0.648) the delayed response of stock prices to an earnings surprise. Similar patterns appear in columns 3–6. Therefore, the data are consistent with the interpretation that increased rates of acute physical illness among investment professionals lead to a smaller immediate reaction of stock prices to an earnings surprise and to more drift in the following days.

To investigate whether the same holds true for changes in analysts' consensus recommendations, I repeat the analysis after replacing *SUE* with *RecChng* in model (3).[18] Table 9 displays the relevant estimates. In column 1, where the dependent variable is $Ab\_Ret_i^{[t]}$, the coefficient on *RecChng* ×Δ *LogIllness18_64* is negative, whereas in column 7, where the dependent variable is $Ab\_Ret_i^{[t+1,t+30]}$, the same coefficient is positive. Both are statistically significant at least at the 5% level. Specifically, a one-standard-deviation increase in *ΔLogIllness18_64* reduces by 6.7% (=0.473/7.061) the immediate response of stock prices to a change in analysts' consensus recommendation and increases by 42% (=0.465/1.102) their delayed reaction.

Overall, these results are consistent with the literature on investor attention, which shows that reduced investor attention decreases the immediate (increases the delayed) response of stock prices to earnings news and changes in analysts' consensus recommendations (DellaVigna and Pollet, 2009; Ben-Rephael et al., 2017).

In summary, increased rates of acute illness among stock market professionals hinder the price discovery process and lead to an underreaction of prices to earnings news and changes in analysts' recommendations.[19] Put another way, this is further evidence that professional market participants play a key role in the incorporation of public news into prices.

[18] To mitigate the effect of outliers, I exclude an observation if less than 4 analysts follow the stock in question or the change in analysts' consensus recommendation is greater than +2 or less than −2. I also exclude an observation if the change in analysts' consensus recommendation occurs on the same day as an earnings announcement.

[19] Since some earnings announcements and changes in analysts' consensus recommendations may occur after the end of the trading day, I follow DellaVigna and Pollet (2009) and repeat the analyses using a window from day t to $t + 1$ to compute the immediate response of stock prices. Untabulated results lead to qualitatively similar conclusions. However, in the case of earnings announcements, the impact of *ΔLogIllness18_64* on the immediate response of stock prices is not statistically significant at conventional levels.

Table 9
Illness incidence among investment professionals and price discovery: Changes in analysts' consensus recommendations.

	(1) Abnormal Return(t)	(2) Abnormal Return (t + 1,t + 3)	(3) Abnormal Return (t + 1,t + 5)	(4) Abnormal Return (t + 1,t + 7)	(5) Abnormal Return (t + 1,t + 10)	(6) Abnormal Return (t + 1,t + 20)	(7) Abnormal Return (t + 1,t + 30)	(8) Abnormal Return (t + 1,t + 40)
RecChng	7.061***	1.515***	1.595***	1.566***	1.531***	1.320***	1.102***	0.913***
	(0.203)	(0.108)	(0.132)	(0.147)	(0.165)	(0.213)	(0.254)	(0.291)
ΔLogIllness18_64	−0.007	−0.019	−0.027	−0.027	−0.037	−0.018	0.025	−0.013
	(0.023)	(0.028)	(0.037)	(0.042)	(0.050)	(0.075)	(0.089)	(0.104)
RecChng × ΔLogIllness18_64	−0.473***	0.194**	0.213*	0.283**	0.309**	0.356*	0.465**	0.521**
	(0.133)	(0.095)	(0.117)	(0.129)	(0.155)	(0.191)	(0.221)	(0.247)
AbNumStories	−0.055	−0.027	−0.028	−0.020	−0.010	−0.001	0.032	0.160**
	(0.035)	(0.025)	(0.033)	(0.040)	(0.048)	(0.060)	(0.071)	(0.079)
HLtoH	−25.803***	1.382	1.123	2.272	1.973	−4.042	2.044	0.074
	(2.828)	(2.023)	(2.569)	(2.728)	(3.300)	(3.344)	(3.506)	(3.863)
Return(t − 27,t − 6)	−0.003	−0.008***	−0.015***	−0.020***	−0.030***	−0.042***	−0.070***	−0.094***
	(0.002)	(0.003)	(0.003)	(0.004)	(0.005)	(0.008)	(0.009)	(0.010)
AVG_Turnover(t − 5,t − 1)	0.047***	−0.031*	−0.038*	−0.051**	−0.068**	−0.042	−0.097	−0.215***
	(0.014)	(0.017)	(0.021)	(0.025)	(0.033)	(0.054)	(0.066)	(0.075)
AVG_Spread(t − 5,t − 1)	0.108***	−0.012	0.003	0.041	0.099	0.094	0.142	0.138
	(0.039)	(0.054)	(0.062)	(0.070)	(0.077)	(0.127)	(0.154)	(0.160)
SDRET	0.182***	0.003	−0.004	−0.051	−0.094	−0.185**	−0.155*	0.009
	(0.028)	(0.031)	(0.041)	(0.049)	(0.058)	(0.075)	(0.086)	(0.097)
LnSize	−0.054***	0.003	0.004	−0.012	−0.024	−0.073	−0.055	−0.045
	(0.017)	(0.019)	(0.024)	(0.028)	(0.034)	(0.046)	(0.055)	(0.065)
InstHold	−0.000	0.000	−0.000	−0.000	0.000	0.002	0.002	0.002
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
LnNumEst	−0.013	0.056	0.000	0.078	0.087	0.223*	0.301**	0.312*
	(0.039)	(0.044)	(0.060)	(0.070)	(0.087)	(0.120)	(0.147)	(0.173)
N	78,644	78,651	78,647	78,645	78,643	78,634	78,410	77,959

Notes: this table displays the estimates generated by fitting model (3), where *SUE* is replaced by *RecChng*. The dependent variable is the cumulative abnormal return of stock i from day $t + h$ to $t + k$ ($h = 0$, 1 and $k = 0$, 3, 5, 7, 10, 20, 30, 40), where t is a day when a change in analysts' consensus recommendation occurs for stock i, and abnormal returns are computed using the Fama and French (1993) 3-factor model with an estimation window from t − 200 to t − 21. *RecChng* measures changes in analysts' consensus recommendations, where recommendations are converted by Bloomberg into numerical ratings that range between 1 (Sell) and 5 (Buy) and subsequently averaged across the analysts following the company. *ΔLogIllness18_64* measures the change in the log of the number of ED visits by 18–64 year old New Yorkers, where a visit is included in the count if it is related to any of the following syndrome categories: asthma, diarrhea, influenza-like, respiratory, vomiting. *ΔLogIllness18_64* is standardized to have zero mean and unit variance. The remaining variables are as defined in Table 1. Each regression includes day-of-the-week and quarter fixed effects, and the t-statistics in parentheses are computed based on standard errors clustered by firm and day. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

5.6. What drives the relation between acute illness and stock returns?

As shown earlier in Tables 4, 5, and 6, increased rates of acute illness among NYC's investment professionals reduce not only market trading activity and volatility but also stock returns. The impact on the former can be attributed to the effect that physical illness exerts on investment professionals' productivity, but what drives the impact on the latter?

It should be noted first that neither economic news nor changes in turnover, liquidity, and investor risk aversion can explain the effect of *ΔLogIllness18_64* on returns (see Section A3 in the online Appendix). A possible explanation is that this effect is driven by the negative mood that stock market professionals experience in response to acute illness symptoms. This conjecture is motivated by the documented links between acute illness and negative mood (Clark and Watson, 1988; Finch et al., 2012) and between negative investor mood and stock returns (Goetzmann et al., 2015).

To investigate this channel, I start by re-examining the results of the quintile analysis introduced in Section 5.3. If the effect were driven by mood fluctuations, one would expect to observe a stronger impact of *ΔLogIllness18_64* on the returns of stocks that are harder to value and more difficult to arbitrage, as these are the stocks whose pricing is more sensitive to investor mood (Baker and Wurgler, 2006).

Indeed, the estimates in column 6 of Table 5 reveal that the effect of *ΔLogIllness18_64* on returns is negative and statistically significantly larger (in absolute value) among high idiosyncratic-volatility, high intraday-volatility, low market-cap, younger, low-price, and high bid–ask-spread stocks. These are precisely the types of stocks that are harder to value and more difficult to arbitrage because they feature higher information uncertainty, higher arbitrage risk, and higher transaction costs (Lam and Wei, 2011; Zhang, 2006).

Secondly, if this were a mood-induced phenomenon, then the initial drop in stock returns should be followed by abnormally high returns as prices revert to fundamental values. To investigate whether this is the case, I re-estimate model (2) after replacing the original dependent variable with $Return_i^{[t+m,t+n]}$, where the latter measures the cumulative return on stock i from day $t + m$ ($m = 0$, 1) to $t + n$ ($n = 0$, 2, 3, 4, 5).

Indeed, the estimates in columns 1–5 of Table 10 show that, while the contemporaneous effect of *ΔLogIllness18_64* on returns is negative, its lagged effect is positive. Specifically, a one-standard-deviation increase in *ΔLogIllness18_64* on day t is accompanied by

Table 10
Acute illness incidence among investment professionals and stock returns: Immediate impact and delayed reversal.

	(1) Return(t)	(2) Return(t + 1, t + 2)	(3) Return(t + 1, t + 3)	(4) Return(t + 1, t + 4)	(5) Return(t + 1, t + 5)	(6) Return(t, t + 5)
ΔLogIllness18_64	−0.100**	0.085	0.096	0.141	0.174*	0.073
	(0.043)	(0.056)	(0.072)	(0.086)	(0.095)	(0.102)
SUE & RecChng	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar FE	Yes	Yes	Yes	Yes	Yes	Yes
Behavioral controls	Yes	Yes	Yes	Yes	Yes	Yes
Economic controls	Yes	Yes	Yes	Yes	Yes	Yes
N	3,557,760	3,556,069	3,554,654	3,553,417	3,552,300	3,551,304

Notes: this table displays the estimates generated by fitting model (2) after replacing the original dependent variable with $Return^{[t+m,t+n]}$, which measures the cumulative return on stock i from day $t + m$ to $t + n$, where $m = 0$, 1 and $n = 0$, 2, 3, 4, 5. *ΔLogIllness18_64* measures the change in the log of the number of ED visits made by 18–64 year old New Yorkers, where a visit is included in the count if it is related to any of the following syndrome categories: asthma, diarrhea, influenza-like, respiratory, vomiting. *ΔLogIllness18_64* is standardized to have zero mean and unit variance. The calendar fixed effects consist of *Day*, *Month*, Year, *PreH*, and *PostH*. The behavioral controls consist of *SADOR*, *Cloudy*, *Sunny*, and *HighPollution*. The economic controls consist of *ΔADSBusConditions* and *ΔMktUncertainty*. All variables are as defined in Table 1. Each regression contains firm fixed effects. The t-statistics in parentheses are computed based on standard errors clustered by firm and day. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

a contemporaneous reduction in returns of 0.10% and is followed by a cumulative increase in returns of 0.17% between day $t + 1$ and $t + 5$. The former (latter) effect is statistically significant at the 5% (10%) level. In addition, the estimates in column 6 show that the cumulative effect between day t and $t + 5$ is statistically indistinguishable from 0. In other words, one cannot reject the null hypothesis that the initial decrease in returns is completely reversed during the following five days.

In summary, the results discussed in this section are consistent with a behavioral interpretation of the relation between *ΔLogIllness18_64* and stock returns. And they are in line with the findings of the literature on investor mood, which argues that, for example, weather-induced mood fluctuations affect professional investors' trading decisions and stock returns (Goetzmann et al., 2015; Jiang et al., 2021).

6. Conclusion

While understanding the mechanisms that drive (or hinder) the productivity of labor is one of the core preoccupations of economics, the finance discipline has shown so far little interest in studying the factors that may affect the productivity of investment professionals, who are the key players in developed financial markets.

Needless to say, mathematical modeling of investor behavior requires some simplifications; in practice, this has often led to dehumanizing investors to disembodied and emotionless automata. Yet, despite their training and experience, even professional market participants are subject to the same behavioral forces and human ailments as the rest of us. Hence, it seems important to shed light on the links between body, investor productivity, and aggregate market outcomes, which represent an under-researched area of study.

This paper contributes to this discussion. Its central idea is that acute physical illness symptoms momentarily reduce the productive capacity of investment professionals and, consequently, their ability to perform their jobs.[20] My results provide evidence that short-term fluctuations in the incidence of acute illness in stock market professionals do matter. Namely, increased rates of acute illness hamper the price discovery process and impact market trading activity, volatility, and returns. The hope is that these findings will stimulate more research on the factors that may influence the productivity of investment professionals and lead to a better understanding of its impact on the mechanics of financial markets.

CRediT authorship contribution statement

Gabriele M. Lepori: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jempfin.2022.12.003.

[20] It is worth noting that my research design is not suitable for studying what would happen, for example, in a pandemic such as that caused by COVID-19. While it is true that this kind of contagious disease may affect the productivity of stock market professionals, it also has a direct impact on economic activity and, consequently, on stock market outcomes.

References

Allison, P.D., Waterman, R.P., 2002. Fixed–effects negative binomial regression models. Sociol. Methodol. 32 (1), 247–265.

Aruoba, S., Diebold, F., Scotti, C., 2009. Real-time measurement of business conditions. J. Bus. Econom. Statist. 27 (4), 417–427.

Atkins, A.B., Dyl, E.A., 1997. Market structure and reported trading volume: NASDAQ versus the NYSE. J. Financ. Res. 20 (3), 291–304.

Baker, S.R., Bloom, N., Davis, S.J., 2018. Policyuncertainty.com. [Online] Available at: https://www.policyuncertainty.com/equity_uncert.html [Accessed 30 3 2018].

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. J. Finance 61 (4), 1645–1680.

Beck, N., 2018. Estimating grouped data models with a binary dependent variable and fixed effects: What are the issues. [Online] Available at: https://arxiv.org/abs/1809.06505.

Bekaert, G., Engstrom, E.C., Xu, N.R., 2020. SSRN. [Online] Available at: https://ssrn.com/abstract=3069078 [Accessed 27 4 2020].

Ben-David, I., Franzoni, F., Moussawi, R., Sedunov, J., 2019. The Granular Nature of Large Institutional Investors, s.l.: (No. 13427). CEPR Discussion Papers.

Ben-Rephael, A., Da, Z., Israelsen, R.D., 2017. It depends on where you search: Institutional investor attention and underreaction to news. Rev. Financ. Stud. 30 (9), 3009–3047.

Bernard, V.L., Thomas, J.K., 1989. Post-earnings-announcement drift - Delayed price response or risk premium. J. Account. Res. 27, 1–36.

Bernile, G., Kumar, A., Sulaeman, J., Wang, Q., 2019. Has local informational advantage disappeared? Rev. Financ. Econ. 37 (1), 38–60.

Bloomberg, 2020. NEWS_HEAT_PUB_DAVG field description, Bloomberg Professional. [Online] Available at: Bloomberg Subscription Service [Accessed 26 04 2020].

Bonaparte, Y., Kumar, A., 2013. Political activism, information costs, and stock market participation. J. Financ. Econ. 107 (3), 760–786.

Bose, S., Ladley, D., Li, X., 2020. The role of hormones in financial markets. Int. Rev. Financ. Anal. 67.

Clark, L.A., Watson, D., 1988. Mood and the mundane: Relations between daily life events and self-reported mood. J. Personal. Soc. Psychol. 54 (2), 296–308.

Correia, S., 2017. Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator. Working paper, [Online] Available at: http://scorreia.com/research/hdfe.pdf.

DellaVigna, S., Pollet, J.M., 2009. Investor inattention and friday earnings announcements. J. Finance 64 (2), 709–749.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33 (1), 3–56.

Finch, J.F., Baranik, L.E., Liu, Y., West, S.G., 2012. Physical health, positive and negative affect, and personality: A longitudinal analysis. J. Res. Personal. 46 (5), 537–545.

Garman, M.B., Klass, M.J., 1980. On the estimation of security price volatilities from historical data. J. Bus. 53 (1), 67–78.

Gaspar, J., Massa, M., 2007. Local ownership as private information: Evidence on the monitoring-liquidity trade-off. J. Financ. Econ. 83 (3), 751–792.

Goetzmann, W.N., Kim, D., Kumar, A., Wang, Q., 2015. Weather-induced mood, institutional investors, and stock returns. Rev. Financ. Stud. 28 (1), 73–111.

Green, M., O'Hagan, S., Lefebvre, M., 2015. The economic geography of institutional investment in the United States, 2010. Urban Geogr. 36 (6), 798–822.

Guiso, L., Haliassos, M., Jappelli, T., 2003. Household stockholding in europe: where do we stand and where do we go? Econ. Policy 18 (36), 123–170.

Heffernan, R., et al., 2004. New york city syndromic surveillance systems. In: s.l.: Syndromic Surveillance: Reports from a National Conference, 2003. Morbidity and Mortality Weekly Report September 24, 2004/ Vol. 53.

Hendershott, T., Moulton, P.C., 2011. Automation, speed, and stock market quality: The NYSE's hybrid. J. Financial Mark. 14 (4), 568–604.

Hochmair, H.H., 2016. Spatiotemporal pattern analysis of taxi trips in New York City. Transp. Res. Rec. (2542), 45–56.

Jiang, D., Norris, D., Sun, L., 2021. Weather, institutional investors and earnings news. J. Corp. Finance 69.

Kadan, O., Michaely, R., Moulton, P., 2018. Trading in the presence of short-lived private information: Evidence from analyst recommendation changes. J. Financ. Quant. Anal. 53 (4), 1509–1546.

Kamstra, M.J., Kramer, L.A., Levi, M.D., 2000. Losing sleep at the market: the daylight saving anomaly. Amer. Econ. Rev. 90 (4), 1005–1011.

Kamstra, M.J., Kramer, L.A., Levi, M.D., 2003. Winter blues: A SAD stock market cycle. Amer. Econ. Rev. 93 (1), 324–343.

Kamstra, M.J., Kramer, L.A., Levi, M.D., 2015. Seasonal variation in treasury returns. Crit. Finance Rev. 4 (1), 45–115.

Katz, E., 2001. Bias in conditional and unconditional fixed effects logit estimation. Political Anal. 9 (4), 379–384.

Kumar, A., 2009. Who gambles in the stock market? J. Finance 64 (4), 1889–1933.

Kumar, A., Lee, C.M., 2006. Retail investor sentiment and return comovements. J. Finance 61 (5), 2451–2486.

Lall, R., et al., 2017. Advancing the use of emergency department syndromic surveillance data, New York City, 2012–2016. Public Health Rep. 132 (1_suppl), 23S–30S.

Lam, F.E.C., Wei, K.J., 2011. Limits-to-arbitrage, investment frictions, and the asset growth anomaly. J. Financ. Econ. 102 (1), 127–149.

Levine, M., 2018. NYSE needed some pretend traders. [Online] Available at: https://www.bloomberg.com/opinion/articles/2018-12-14/nyse-needed-some-pretend-traders [Accessed 10 June 2020].

Mathes, R.W., Ito, K., Matte, T., 2011. Assessing syndromic surveillance of cardiovascular outcomes from emergency department chief complaint data in New York City. PLoS One 6 (2), e14677.

Mathes, R., et al., 2017. Evaluating and implementing temporal, spatial, and spatio-temporal methods for outbreak detection in a local syndromic surveillance system. PLoS One 12 (9), e0184419.

McTier, B., Tse, Y., Walk, J., 2013. Do stock markets catch the flu? J. Financ. Quant. Anal. 48 (3), 979–1000.

Metzger, K.B., Hajat, A., Crawford, M., Mostashari, F., 2004. How many illnesses does one emergency department visit represent? Using a population-based telephone survey to estimate the syndromic multiplier. Morb. Mortal. Wkly Rep. 53, 106–111.

Molnár, P., 2012. Properties of range-based volatility estimators. Int. Rev. Financ. Anal. 23, 20–29.

NYC Planning, 2016. Employment Patterns in New York City, Trends in a Growing Economy. [Online] Available at: https://www1.nyc.gov/assets/planning/download/pdf/planning-level/housing-economy/employment-patterns-nyc.pdf.

O'Hara, M., Saar, G., Zhong, Z., 2019. Relative tick size and the trading environment. Rev. Asset Pricing Stud. 9 (1), 47–90.

Olson, D.R., et al., 2007. Monitoring the impact of influenza by age: emergency department fever and respiratory complaint surveillance in New York City. PLoS Med. 4 (8), e247.

Pantzalis, C., Ucar, E., 2018. Allergy onset and local investor distraction. J. Bank. Financ. 92, 115–129.

Petersen, M.A., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. Rev. Financ. Stud. 22 (1), 435–480.

Puckett, A., Yan, X., 2011. The interim trading skills of institutional investors. J. Finance 66 (2), 601–633.

Roodman, D., Nielsen, M.Ø., MacKinnon, J.G., Webb, M.D., 2019. Fast and wild: Bootstrap inference in Stata using boottest. Stata J. 19 (1), 4–60.

Stieb, D.M., Szyszkowicz, M., Rowe, B.H., Leech, J.A., 2009. Air pollution and emergency department visits for cardiac and respiratory conditions: a multi-city time-series analysis. Environ. Health 8 (1), 1–13.

Westheimer, E., et al., 2012. Evaluating the New York City emergency department syndromic surveillance for monitoring influenza activity during the 2009-10 influenza season. PLoS Curr. 4.

Zhang, X.F., 2006. Information uncertainty and stock returns. J. Finance 61 (1), 105–137.