CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 630

The Geography of Investor Attention

Stefano Mengoli, Marco Pagano, and Pierpaolo Pattitoni

November 2021

University of Naples Federico II **University of Salerno** **Bocconi University, Milan**

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS AND STATISTICS – UNIVERSITY OF NAPLES FEDERICO II
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

WORKING PAPER NO. 630

The Geography of Investor Attention

Stefano Mengoli[*], Marco Pagano[†], and Pierpaolo Pattitoni[‡]

Abstract
Retail investors pay over twice as much attention to local companies than non-local ones, based on Google searches. News volume and volatility amplify this attention gap. Attention appears causally related to perceived proximity: first, acquisition by a nonlocal company is associated with less attention by locals, and more by nonlocals close to the acquirer; second, COVID-19 travel restrictions correlate with a drop in relative attention to nonlocal companies, especially in locations with fewer ights after the outbreak. Finally, local attention predicts volatility, bid-ask spreads and nonlocal attention, not viceversa. These findings are consistent with local investors having an information-processing advantage.

JEL classification: D83, G11, G12, G14, G50, L86, R32.

Keywords: attention, retail investors, local investors, distance, news, liquidity, volatility.

Acknowledgements: We thank for helpful comments Alberta Di Giuli, Andrew Ellul, Thierry Foucault, Tullio Jappelli, Christian Julliard, Annalisa Scognamiglio, Lorenzo Pandolfi, Raman Uppal, and participants at the Corporate Finance Webinar, the Higher School of Economics and New Economic School webinar, CSEF and LUISS seminars, and the European Economic Association Annual Meetings. Marco Pagano gratefully acknowledges financial support from the Italian Ministry for University and Research (MUR) and the Einaudi Institute for Economics and Finance (EIEF).

[*] Università di Bologna. E-mail: stefano.mengoli@unibo.it

[†] Università di Napoli Federico II, CSEF, and EIEF. E-mail: pagano56@gmail.com

[‡] Università di Bologna. E-mail: pierpaolo.pattitoni@unibo.it

Table of contents

1. Introduction

2. Model

3. Data

4. Local and nonlocal attention

4.1 Local-nonlocal attention gap

4.2 Investor attention, proximity and news

5. Shocks to Proximity and Attention Reallocation

5.1 M&A shocks to perceived proximity

5.2 COVID-19 shocks to proximity

6. Attention, Volatility and Liquidity

7. Conclusion

References

Appendix

Figures and Tables

1 Introduction

Attention is a scarce resource (Kahneman, 1973). So investors can be expected to direct it preferentially to news they are better at understanding (Van Nieuwerburgh and Veldkamp, 2009), and skew their trading activity and portfolios towards the corresponding assets (Barber and Odean, 2007; Van Nieuwerburgh and Veldkamp, 2010). In turn, their tendency to pay more attention to these assets affects their prices, for instance amplifying their volatility (Andrei and Hasler, 2014).

In this study, this argument is applied to the geography of investor attention: it is shown to account not only for investors' attention being skewed towards local firms, but also for the response of this attention gap to news volume, volatility and investors' perceived distance from companies. Furthermore, increases in local attention are shown to predict increases in volatility and reductions in market liquidity, as well as increases in nonlocal attention. All these findings are consistent with a model where investors optimally choose their attention and are better at processing information about local stocks than nonlocal ones. In equilibrium, local investors choose to pay more attention to news about local stocks than nonlocal investors do, so as to exploit their information-processing advantage, and this attention gap grows when news precision and/or fundamental volatility rise. Moreover, the stronger the local-nonlocal attention gap, the more informative is the order flow and the wider is the bid-ask spread.

Our empirical analysis starts by constructing a measure of attention based on Google web searches of U.S. stocks, in line with Da, Engelberg, and Gao (2011), Engelberg and Parsons (2011) and Andrei and Hasler (2014).[1] This approach provides a fine-grained measure of retail investor attention, which enables us to track web searches of a given stock depending on the state where they are done: for each firm, we can discriminate between local web searches, made in the state where the firm is headquartered, and nonlocal web searches, made in the other states. To identify web searches made by investors, we use stock financial tickers rather than company names as Google search keywords, which is likely to be the choice of an investor rather than a generic individual. An advantage of a web-based measure of attention is that, differently from other media, the Google search engine enables virtually all investors to access with equal ease news about local and nonlocal firms, so that any systematic difference in search intensity is likely to reflect intentional choices rather than differential search costs.

Equipped with this measure of investors' attention, we find that on average investors devote twice as much attention to news about companies headquartered in their own

[1] A similar approach is also used by Ben-Rephael, Da, and Israelsen (2017), who however rely on Bloomberg as a tool to gather financial information, as they focus on the attention of institutional investors.

state rather than elsewhere. News volume and volatility amplify this gap: local investors react about 7 times more than nonlocal ones to increases in the number of news about local companies (whether positive or negative) and 12 times more to an increase in their return volatility. This attention gap and its response to news and volatility are consistent with greater efficiency in processing information about local companies, as posited in our model: retail investors are more exposed to local information, being able to access local media that typically prioritize local company news (Engelberg and Parsons, 2011) and word-of-mouth information about local entrepreneurs, and therefore can interpret local company news surfacing on the web more easily and more rapidly than nonlocal investors.

The evidence is consistent with investors' attention being causally related to their proximity to firm headquarters. First, a decrease in perceived proximity due to an acquisition by a nonlocal firm is associated with a decline in local attention and a rise in attention by nonlocal investors close to the acquiring firm. Second, the increase in perceived distance of nonlocal firms due to 2020 pandemic-related travel restrictions were associated with an increase in the local-nonlocal attention gap, with relative attention paid to nonlocal stocks dropping more for firms in locations connected by fewer flights after the COVID-19 outbreak.

Finally, an increase in our local attention measure is associated with a subsequent increase in return volatility, in the bid-ask spread and in nonlocal attention, but is not predicted by these variables. This pattern is consistent with attention by local investors raising their informational advantage, and thus inducing market makers to protect themselves with wider bid-ask spreads, while at the same time accelerating price discovery and increasing volatility. Empirically, there findings dovetail with those reported by Shive (2012), who examines the effect of drops in local investors' trading activity due to large power outages, and finds that during local blackouts bid–ask spreads narrow and firm-specific price volatility drops, consistently with the view that local investors are informed traders whose orders tend to reduce market liquidity and increase volatility.

Our evidence suggests that the well-known local bias of investors' portfolios, that is, their tendency to overweight stocks issued by nearby firms (Coval and Moskowitz, 1999; Grinblatt and Keloharju, 2001) stems from an information-processing advantage of local investors, as argued by Van Nieuwerburgh and Veldkamp (2009), rather than from behavioral biases, such as familiarity (Huberman, 2015), loyalty (Cohen, 2008) or patriotism (Morse and Shive, 2011). This is not a forgone conclusion, considering that according to Grinblatt and Keloharju (2001) the evidence about the local portfolio bias of Finnish households "seems to support the hypothesis that the degree of these effects is inversely related to investor sophistication". However, others provide evidence that proximity confers an informational advantage. Malloy (2005) shows that local analysts' forecasts better predict stock returns and earn abnormal returns on their local assets, Ivković and Weisbenner (2005) find the average household generates an additional annualized return of 3.2% from its local holdings relative to its nonlocal holdings, suggesting that local investors can exploit local knowledge, and Massa and Simonov (2006) document that Swedish households' tendency to concentrate holdings in stocks to which they are geographically or professionally close enables them to earn higher returns than they would otherwise earn. Hence all these studies agree that households' strong preference for local investments is information driven rather than due to a behavioral bias.

Characterizing local retail investors' as featuring superior information processing skills may seem surprising, considering that Barber and Odean (2000, 2001, 2007, 2013) present retail investors as noise traders, subject to all kinds of behavioral biases. However more recent research documents that retail investors play a substantial role in price discovery. Kaniel, Saar, and Titman (2008) find that individual investor sentiment predicts future returns, and its information content is distinct from that of past returns or past volume, and Kaniel, Liu, Saar, and Titman (2012) provide evidence of informed trading by individual investors around earnings announcements, showing that their aggregate trading behavior predicts large abnormal returns on and after announcements. Relatedly, Kelley and Tetlock (2013) show that orders by retail investors positively predict firms' monthly stock returns with no evidence of return reversal, and contribute to market efficiency, and Friedman and Zeng (2021) find that retail trader activity on the Robinhood platform is associated with prices that are more responsive to earnings surprises. Finally, Boehmer, Jones, Zhang, and Zhang (2020) document that retail trades on OTC markets are quite informed, and Boehmer and Song (2021) find that retail short sellers' trades predict negative stock returns, suggesting that they profitably exploit public information, especially when it is negative.[2]

Even more relevant to the present paper, recent work documents that the attention level of retail investors, and particularly local ones, predicts subsequent risk-adjusted returns. Based on brokerage account data, Gargano and Rossi (2018) measure retail investors' attention based on the time they spend looking at data available via their brokerage account, and find that they devote more attention to local stocks, and that their attention level predicts stocks' risk-adjusted returns, and is particularly profitable when trading stocks with high uncertainty for which a lot of public information is available. Relatedly, Cziraki, Mondria, and Wu (2021), who also measure attention based on Google-search data, find that stocks featuring an abnormally large gap between local and nonlocal investors' attention earn higher risk-adjusted returns. These findings

[2] Along the same lines, Guiso and Jappelli (2020) find that households' investment in information is associated with significantly higher returns to financial wealth, based on survey data for the customers of a leading Italian bank.

are consistent with ours, and with the idea captured by our model that retail investors' attention to public information confers them an informational advantage, and more so for local than for nonlocal investors. As one would expect, this point also applies to institutional investors: Dyer (2021) constructs a measure of attention by these investors based on their requests for financial information from the SEC, and documents not only that they acquire approximately 20% more financial information for local stocks than for nonlocal ones, but also that local investors are faster in acquiring public information and make better trading decisions when acquiring public information.

The rest of the paper proceeds as follows. Section 2 presents the model and the predictions that guide our empirical analysis. Section 3 describes the data. Section 4 presents the empirical analysis of the geography of investor attention and its correlation with news intensity and stock return volatility. Section 5 presents the evidence that investor attention is causally related to their perceived proximity to firm headquarters: specifically, Subsection 5.1 analyzes the response of local and nonlocal attention to changes in perceived firm proximity due to M&A activity, while Subsection 5.2 analyses their response to the increase in perceived distance of nonlocal firms due to COVID-19-related travel restrictions. Next, Section 6 analyzes the times series relationships among local and nonlocal attention, stock return volatility and bid-ask spreads. Section 7 concludes.

2 Model

We start by presenting a simple model of investor attention allocation, trading decisions and asset price determination, whose predictions will guide the empirical analysis of subsequent sections. The model extends the Glosten and Milgrom (1985) setting by letting informed traders optimally choose how much of their (costly) attention they wish to devote to processing news and understand their implications for asset prices. Importantly, local investors are taken to be better at processing information about local stocks than nonlocal investors. The analysis is presented with reference to a stock issued by a representative local firm, but of course it applies symmetrically to a stock issued by a nonlocal firm, which nonlocal investors are better at analyzing.

Stock prices are set by risk-neutral and competitive dealers, who receive buy and sell orders for each stock from two types of investors: information-based investors, who trade with probability π, and noise traders, who buy or sell the stock with equal probability, and trade with frequency $1 - \pi$. All market participants place orders for a fixed amount of the stock, which for simplicity is standardized to 1. If both local and nonlocal investors choose to devote attention to news, the pool of informed traders is composed by an equal number of these two subgroups, respectively indexed by $i = L, N$, and each of them manages to place an order with the same probability $\pi/2$. If only members of a single group choose to exert attention, they manage to place their orders with probability π. Whether either type of investor chooses to devote attention to news and thus become informed is to be determined in equilibrium.

The fundamental value of the stock is $v \in \{v_b, v_g\} = \{\overline{v} - \sigma, \overline{v} + \sigma\}$, where $\overline{v}$ is its unconditional value and $\sigma = (v_g - v_b)/2$ measures its volatility. Before trading, investors observe a signal $s \in \{v_b, v_g\}$ that is correlated with the value of the stock. The informativeness of the signal depends on the quality q of news release (e.g., their detail and timeliness) and on the level of attention $a_i \in [0, 1/q]$ that investor i chooses to devote to the signal before trading. The greater are news quality and attention, the higher the probability of correctly estimating the probability distribution of the asset's value: $\Pr(s = v|v, a_i, q) = (1 + q \cdot a_i)/2$. So by paying more attention, investors read the signal more accurately. An investors i who chooses $a_i = 0$ would learn nothing from the signal: $\Pr(s = v|v, a_i, q) = 1/2$, while for one exerting the maximal attention $a_i = 1/q$ the signal would become perfectly informative: $\Pr(s = v|v, a_i, q) = 1$.

However, greater precision comes at increasing cost, implying decreasing returns to attention: the cost of information processing is $C_i(a_i, \theta_i)$, with $\partial C_i/\partial a_i > 0$ and $\partial^2 C_i/\partial a_i^2 > 0$, where the parameter θ_i measures investor i's ability to glean news' price implications. Local investors are assumed to be better at processing information about local stocks ($\theta_L > \theta_N$), for any given attention level. For concreteness, we posit a quadratic cost function: $C_i(a_i, \theta) = a_i^2/2\theta_i$. The greater θ_i, the easier it is for investor i to gauge the asset value's response to news. Importantly, non-local investors are not naive: they are aware of being less skilled than local investors at processing information about local stocks, and take this rationally into account in their market behavior.[3] Like Tirole (2009), Van Nieuwerburgh and Veldkamp (2009) and Van Nieuwerburgh and Veldkamp (2010), we do not assume bounded rationality: both local and non-local investors decide rationally how much information they wish to process, in the awareness that a low level of attention may lead to mistakes in trading.

The expected stock value from an investor's standpoint depends on her choice of attention a and on the signal s:

$$\widehat{v}(a, s) \equiv \mathbb{E}_h[v|s] = \frac{1}{2} \times \begin{cases} (1 + qa)v_g + (1 - qa)v_b & \text{if} \quad s = v_g, \\ (1 - qa)v_g + (1 + qa)v_b & \text{if} \quad s = v_b. \end{cases} \tag{1}$$

where $(1 + qa)/2$ is the conditional probability of the asset being high-value when the

[3]We assume that good and bad news are equally costly to process. However, in some settings bad news may be costlier to process for investors than good news. In this case, in equilibrium both types of investors would devote more attention to good than to bad news–the so-called 'ostrich effect'.

investor chooses attention a and receives a high-value signal.[4] This probability is an increasing function of the investor's attention a: in the limiting case $a = 0$, the estimate $\widehat{v}$ would be the unconditional average $\overline{v}$, whereas in the polar opposite case $a = 1/q$, the estimate would be perfectly precise. Thus, in equilibrium each investor holds a belief $\widehat{v}(a_i, s)$ whose precision depends on the chosen attention level.

Dealers are assumed to devote no attention to the processing of news, as they are taken to specialize in trading rather than in the analysis of stock fundamentals, also because they can infer already processed information at no cost directly from the order flow. With no loss of generality, we focus on the case where investors receive good news $s = v_g$, so that they may wish to buy at the ask price p_A (the analysis being symmetric for the case of bad news $s = v_b$ and trading at the bid price p_B). Investor i chooses her attention level a_i so as to maximize her expected utility:

$$\max_{a_i \in [0,1]} \frac{\pi}{2} \left[\widehat{v}(a_i, s = v_g) - p\right] - \frac{a_i^2}{2\theta_i} = \frac{1}{2}\left[\pi(\overline{v} - p_A + q a_i \sigma) - \frac{a_i^2}{\theta_i}\right], \text{ for } i \in \{L, N\}, \quad (2)$$

where the first term is the investor's expected profit (obtained with probability $\pi/2$) and the second her attention costs. The objective function (2) is increasing and strictly concave in a_i, and its parameters are assumed to be such that even local investors choose an interior solution $a_L^* < 1/q$. Hence the optimal choice of attention is

$$a_i^* = \frac{\theta_i \pi q \sigma}{2}, \quad (3)$$

so that an internal solution $a_L^* < 1/q$ requires $\theta_L < 2/(\pi q^2 \sigma)$, i.e. imposes an upper bound on the information-processing ability of local investors. Clearly, in equilibrium local investors devote more attention to local stocks than nonlocal ones, i.e. there is a local attention gap:

$$a_L^* - a_N^* = \frac{\pi q \sigma}{2}(\theta_L - \theta_N) > 0. \quad (4)$$

This immediately yields the following comparative statics:

Proposition 1 (Optimal attention and local attention gap) *If the local investor's attention problem has an internal solution, the optimal attention level a_i^* of investor i is*

[4] To see why, recall that the conditional probability that the signal is correct given an attention level a is $\Pr(\sigma = v_j | a, v = v_j) = (1 + qa)/2$, for $j = b, g$, and that the prior probability of a high asset value is $\Pr(v = v_g) = 1/2$. Therefore, the conditional probability that the investor assigns to $v = v_g$ is

$$\begin{aligned}\Pr(v = v_g | a, \sigma = v_g) &= \frac{\Pr(\sigma = v_g | a, v = v_g)\Pr(v = v_g)}{\Pr(\sigma = v_g | a, v = v_g)\Pr(v = v_g) + (1 - \Pr(\sigma = v_g | a, v = v_g))(1 - \Pr(v = v_g))} \\ &= \frac{(1 + qa)/2}{(1 + qa)/2 + [1 - (1 + qa)/2]} = \frac{1 + qa}{2}.\end{aligned}$$

increasing in her information processing efficiency θ_i. The attention levels of local and nonlocal investors and the difference between them are increasing in information quality q, order execution probability $\pi/2$, and stock volatility σ.

These results are intuitive: more able investors devote greater attention to news as they face lower costs of processing it, especially if news convey precise information, fundamentals are volatile (as volatility makes information more valuable) and there is a higher chance of trading on information.

As market makers earn zero expected profits,[5] the ask-side half-spread is

$$p_A - \overline{v} = \frac{\pi}{2} q\sigma (a_L^* + a_N^*) = \frac{1}{2}\overline{\theta}(\pi q \sigma)^2, \tag{5}$$

since total investors' attention is $a_L^* + a_N^* = \overline{\theta}\pi q\sigma$, where $\overline{\theta} \equiv (\theta_L + \theta_N)/2$ is the average information processing ability. Hence:

Proposition 2 (Equilibrium bid-ask spread) *The equilibrium bid-ask spread is increasing in investors' average information processing efficiency $\overline{\theta}$, the quality of information q, the probability of informed trading π and fundamental volatility σ.*

Note that since by (4) changes in information quality q and volatility σ affect local investors' attention more than that of nonlocal ones, the response of the bid-ask spread to changes in q or σ predicted by equation (5) travels more via the change in the attention of local investors, a_L^*, than via that of non-local ones, a_N^*. Hence, empirically one should expect the bid-ask spread to be more strongly correlated with local than with nonlocal attention.

Finally, we must verify whether both types of investors wish to exert attention in equilibrium, as initially assumed. This is not a forgone conclusion, as the adverse selection generated by attention may raise the bid-ask spread so much as to make informed trading unprofitable. In equilibrium, investor i expects not to make losses upon exerting attention[6] if and only if

$$\theta_i \geq 2\pi\overline{\theta} = \pi(\theta_L + \theta_N). \tag{6}$$

[5]Market makers' zero profit condition on the ask side is

$$\mathrm{E}(\Pi_A) = \frac{\pi}{2}\left(\overline{v} + q\sigma a_L^* - p_A\right) + \frac{\pi}{2}\left(\overline{v} + q\sigma a_N^* - p_A\right) + (1-\pi)\left(\overline{v} - p_A\right) = 0.$$

[6]The relevant non-negativity condition of investor i's equilibrium expected profits is

$$\mathrm{E}(\Pi_i) = \frac{1}{2}\left[\pi(\overline{v} - p_A + q a_i^* \sigma) - \frac{a_i^{*2}}{\theta_i}\right] = \frac{\pi^2 q^2 \sigma^2}{4}\left(\frac{\theta_i}{2} - \pi\overline{\theta}\right) \geq 0.$$

This condition indicates that the frequency of informed trading π reduces both local and non-local investors' incentives to exert attention, by raising the bid-ask spread. Clearly, condition (6) is more stringent for nonlocal than for local investors, as $\theta_L > \theta_N$. Hence, both groups exert attention if informed trading is sufficiently infrequent; only local investors do so if informed trading occurs with intermediate frequency, and no investor does if attention were to prompt very frequent informed trading (see the Appendix for a proof):

Proposition 3 (Frequency of informed trading and investors' attention) *Both local and non-local investors exert attention if $0 < \pi < \theta_N/(\theta_N + \theta_L)$; only local investors do if $\theta_N/(\theta_N + \theta_L) \leq \pi \leq 1/2)$, and no investor does if $1/2 < \pi \leq 1$.*

Interestingly, the participation constraint of nonlocal investors ($\pi < \theta_N/(\theta_N + \theta_L)$) implied by (6) can also be interpreted as requiring a limit to the ability gap between local and nonlocal investors rather than to the probability of informed trading, as done in Proposition 3. Indeed, the participation constraint of nonlocal investors can be restated as $\theta_L/\theta_N < (1 - \pi)/\pi$: if the informational advantage of local investors θ_L/θ_N exceeds this threshold, non-local investors' attention drops to zero, and so does their market participation as informed traders.

If the advantage θ_L/θ_N of local investors increases with the nonlocal investors' distance from the asset issuer's location, the attention gap between the two types of investors increases in the distance between them both along the intensive and the extensive margin: as the distance between them increases, the local-nonlocal attention gap $a_L^* - a_N^*$ increases in proportion to θ_L/θ_N and, as this exceeds the threshold $(1 - \pi)/\pi$, nonlocal investors' attention vanishes, while local investors' attention doubles (from $\theta_L \pi q\sigma/2$ to $\theta_L \pi q\sigma$, because now they trade with probability π instead of $\pi/2$). In other words, as distance exceeds a critical level, non-local investors exit the market, their attention being 'crowded out' by that of local investors. The model easily generalizes to multiple classes (or even a continuum) of nonlocal investors, each located at different distance from the issuer's location, hence at increasing disadvantage relative to local investors: their attention will be declining in their distance from the issuing firm, down to zero for the marginal nonlocal investor.

The model can also be modified to encompass a time dimension, assuming for instance that informed investors can place their orders in one of two periods, and that greater attention enables them to process information about stocks faster and therefore to be more likely to trade upon news in the first than in the second period (the probability of early trading being an increasing function of attention). In this alternative formulation, the greater ability of local investors in processing local news would confer them a speed rather than a precision advantage: by exerting greater attention, they could expect to trade before nonlocal investors, and thereby reap greater informational rents. The main prediction of this version of the model would be qualitatively similar to those of the model presented above, except for the additional prediction that local attention and trading would be predictors of nonlocal attention and trading over time.

3 Data

We draw web search activity data from Google Trends (https://trends.google.com). Google does not provide detailed information about web searches by individual users, but provides transformed measures that allow comparisons between states for a given firm and week or over time for a specific firm and state, but not both simultaneously. These characteristics of the data constrain our measurement of investor attention and requires different data downloads to address the different issues under analysis.

The measure used in Section 4 to study how attention to companies varies across residents of different states, is constructed on the basis of all weekly web searches performed in 2017 for each stock in the S&P500 (as at the beginning of 2017) and whose issuer is headquartered in the United States. Our sample consists of 480 tickers, and includes a total of 1,244,800 firm-week observations. The index is based on the number of times users digit a specific word, which in our case is a company ticker. However, as mentioned, Google does not provide the absolute number of web searches, but a standardized index of searches: it aggregates web searches into weekly counts and then converts them into a score that equals 100 in the state where the ticker is most frequently searched and a value ranging from 0 to 100 in other states, whose number of searches is scaled by that of the state where the ticker is most searched. For instance, if AAPL (the Apple ticker) is searched the most in California in a given week, California gets a score of 100 for Apple in that week. A score of 20 for Apple in Texas would mean that the web searches in Texas are one fifth (20/100) of those in California. So, our main attention metric – that we label as `GSearch` – is the number of Google web searches for a given ticker, week and state relative to the number of web searches for that ticker in the state where the ticker was most intensively searched during that week. Hence, by construction, each week a ticker has a state where `GSearch` = 100.

This variable is used as a proxy for potential retail investor attention even though, strictly speaking, it measures the search activity by any web user whose attention is caught by a stock. However, web searches of stock tickers are so specific that they are most likely made by potential investors, the stock ticker being often unknown to laymen. Furthermore, investors who search stock tickers on Google are most likely nonprofessional (retail) investors, as financial professionals typically use other channels to gather information about stocks (e.g., Bloomberg, as in Ben-Rephael, Da, and Israelsen (2017)). In what follows, we will check the robustness of this assumption.

By construction, `GSearch` only allows time-series comparison of investor attention in relative terms, not absolute ones, being rebased to 100 each week. As the analysis of Section 4 focuses on the geography of retail investors' attention for a given company, this characteristic is not problematic, and actually reduces potential confounding effects of time-series patterns in the absolute scale of the variable. We distinguish between local and nonlocal attention via the dummy variable `SameState`, which equals 1 for searches made by residents of the state where the company is headquartered, and 0 otherwise, in keeping with previous research that identifies a company's location with that of its headquarters (Coval and Moskowitz, 1999; Grinblatt and Keloharju, 2001; Malloy, 2005; Ivković and Weisbenner, 2005).

In Section 5.1, instead, we wish to measure the dynamics of attention of the investors potentially affected by M&A-induced shocks in perceived firm proximity. To this purpose, we download the time-series of `GSearch` in the bidder's state and in target's state. This is possible because, as mentioned above, Google allows downloading the time series of web searches for one state at a time for a given firm, even though it does not allow to download such series for all states at the same time. Finally, in Section 6, where we need to construct a panel of local and nonlocal attention for each firm and state, we download the aggregate time series for each ticker, and combine this information with the measures employed in Section 4 to compute the time series of web searches made by local investors (i.e., residents of the state where the company is headquartered) and nonlocal ones (i.e. residents in all other states).

Figure 1 illustrates the relationship between `GSearch` and `SameState` for six firms in the sample. For all firms, the figure plots weekly data for `GSearch` in the state where the firm is headquartered, while nonlocal attention is measured by the average of `GSearch` in the other states. The first two plots refer to the largest firms by market capitalization in our sample, i.e. Apple (headquartered in California) and Microsoft (headquartered in Washington). The remaining plots refer to firms that can be expected to be more local in character, being the only ones headquartered in their respective states: IDEXX Laboratories in Maine, Micron Technology in Idaho, SCANA Corporation in South Carolina and Danaher Corporation in D.C. Despite the different characteristics of these companies, for all of them local web searches are more frequent than the respective nonlocal web searches—a feature that will be seen to be widespread in the sample.

In order to characterize more finely the geography of investor attention, besides `SameState`, we construct a continuous variable, `Distance`, which measures the geographic distance (in thousands of kilometers) between a firm's headquarters and the capital of the state where the relevant web searches are done.

To test how investor attention correlates with available public information, we construct four variables. First, we measure the volume of information that is nationwide available about a given company by the variable `News`, defined as the log of 1 plus the number of news published by Thomson Reuters regarding a given ticker and in a given week. We filter out researches, videos, and stories related to the firm and count only headlines reported in English. Insofar as more news should also carry more information, this variable can be seen as an empirical counterpart of the information quality q in the model. A second variable measures the information locally available to investors in each US State about a given company, `LocNews`, defined as the log of 1 plus the number of news published by the local newspaper in a given US State regarding a given ticker in a given week.[7] This variable will enable us to gauge whether the geography of attention is shaped by superior access of local investors to local public information or rather by their superior ability to process information about about local companies, stemming for instance from better understanding of local businesses or access to private information about them. Thirdly, we measure the volatility of price-relevant news in a given week, `Vol`, by the absolute value of the weekly stock return (drawn from Thomson Reuters Eikon). This variable can be regarded as the empirical counterpart of the range of variation of firm value σ in the model. To capture the possible asymmetric effect of volatility on investors' attention at times of good and bad news, we also compute the absolute value of returns separately for weeks of negative and positive returns (when the indicator variable `NegRet`=1 or 0, respectively). This enables us to test the so-called 'ostrich effect', according to which investors pay more attention to good news (Galai and Sade, 2006).

[7]We link each newspaper to the state where it is published, except for three national newspapers: US Today (Virginia), Wall Street Journal and The New York Times (both published in New York City). The local newspapers used are Alaska Dispatch News (Alaska), Arizona Republic (Arizona), Pine Bluff Commercial (Arkansas), The Sacramento Bee (California), Denver Post (Colorado), Hartford Courant (Connecticut), Cape Cod Times (Delaware), Tampa Bay Times (Florida), The Atlanta Journal - Constitution (Georgia), Honolulu Star (Hawaii), Idaho Statesman (Idaho), Chicago Tribune (Illinois), Indianapolis Star (Indiana), Des Moines Register (Iowa), Topeka Capital Journal (Kansas), Courier - Journal (Kentucky), Times - Picayune (Louisiana), Portland Press Herald (Maine), The Baltimore Sun (Maryland), Boston Globe (Massachusetts), Detroit News (Michigan), Minneapolis Star and Tribune (Minnesota), The Sun Herald (Mississippi), St. Louis Post - Dispatch (Missouri), The Billings Gazette (Montana), Lincoln Journal Star (Nebraska), Las Vegas Review - Journal (Nevada), Concord Monitor (New Hampshire), Press of Atlantic City (New Jersey), Albuquerque Journal (New Mexico), New York Post (New York), The Charlotte Observer (North Carolina), Bismarck Tribune (North Dakota), The Columbus Dispatch (Ohio), The Oklahoman (Oklahoma), The Oregonian (Oregon), Pittsburgh Post - Gazette (Pennsylvania), The Providence Journal (Rhode Island), The Post and Courier (South Carolina), The Dickinson Press (South Dakota), Chattanooga Times Free Press (Tennessee), The Houston Chronicle (Texas), The Salt Lake Tribune (Utah), The Burlington Free Press (Vermont), Virginian - Pilot (Virginia), Seattle Post - Intelligencer (Washington), Charleston Daily Mail (West Virginia), Milwaukee Journal Sentinel (Wisconsin), and Wyoming Tribune - Eagle (Wyoming).

Our data also include state-level characteristics used as controls in the regression analysis: the log of state population (`Population`), state-level fraction of inhabitants over age 65 ($\texttt{Age65}^{+}$) and with at least a college degree (`Edu`), and per-capita GDP (`Income`), drawn from the US Census Bureau database.

Table 1 provides summary statistics of variables in our sample. On average, the frequency with which stocks are searched in a state is about 20% of the frequency with which they are searched in the state featuring the highest number of searches: the average value of `GSearch` is 20.14, while it would be about 100 if web searches were equally distributed across states. This indicates that web searches tend to cluster in a few states. The average distance between the company headquarters and the web search location is about 1860 kilometers, and the average number of news per week regarding a firm is about 90.[8] The summary statistics about `LocalNews` indicate an average of about 16 news per company-year published on local newspapers, with 1,017 news being the maximum).[9] The mean of the absolute weekly return (`Vol`) is about 2%, and the tenor of the news is negative in 43% of the cases.

Figure 2 shows the geographic distribution of company headquarters and web searches. The number of company headquarters is highest in California (73 companies, about 15% of the total) and in New York (62 companies, about 13% of the total). Ten states host no S&P500 company headquarters: Alaska, Hawaii, Kansas, Mississippi, New Hampshire, New Mexico, North Dakota, South Dakota, Vermont, West Virginia, and Wyoming. The highest average attention is also recorded in the states of New York (27.39) and California (26.38) and the lowest in South Dakota and Wyoming, both below 15. Hence, the states hosting more company headquarters also feature greater web search activity, both being likely driven by state-level characteristics such as population and economic activity.

4 Local and nonlocal attention

In this section we start by documenting that investors pay more attention to local companies than to nonlocal ones, and then test whether the response of this attention gap to news availability and return volatility is in line with the model's predictions.

4.1 Local-nonlocal attention gap

In principle, the geography in investors' attention can be analyzed either from the standpoint of companies, testing whether they attract more attention from local than nonlocal investors, or from the standpoint of investors, asking whether they devote more attention

[8]The average of the number of news is calculated before taking logs.

[9]The average per week is 0.31 with a maximum of 83 news.

to local than nonlocal companies. Table 2 documents that, irrespective of which of these two perspectives is taken, proximity plays a sizeable and statistically significant role in the attention that investors devote to firms.

The table shows the distribution of company headquarters across states and the average level of web searches (as measured by the above-defined `GSearch` variable) by search states and by company headquarters' states in 2017. For each of the states listed in column 1, local investors are defined as the residents of that state, and nonlocal investors as those resident in other states; symmetrically, local firms are defined as the companies whose headquarters are located in that state, and nonlocal firms as those whose headquarters are located elsewhere. Columns 2 and 3 report the number and fraction of company headquarters located in the corresponding state indicated in column 1. Columns 4 and 5 report the average values of the total weekly web searches made by local investors, and of those directed to local firms (whenever the state hosts at least one company's headquarters), respectively. By construction, the two average to the same figure (20.14) for the whole of the United States.

The subsequent columns of the table quantify the local-nonlocal attention gap in each state. The total searches directed to local firms (column 5) are decomposed into those made by same-state residents (column 6) and those undertaken on average by residents of any other state (column 7).[10] Column 8 shows that the difference between the two is positive and significantly different from zero at the 1% level for all the states except for the District of Columbia (hosting a single company), Louisiana (3 companies) and Maryland (10 companies). Hence, seen from the perspective of the 480 firms in the sample, in most states they receive significantly more attention from local investors than nonlocal ones.

Symmetrically, the last two columns of the table document the allocation of attention from the standpoint of investors: this time the total searches made by local investors (column 4) are decomposed into those that these investors on average direct to local companies (again, column 6) and those that on average they direct to nonlocal companies (column 9). Column 10 shows that, also in this case, the difference between the former and the latter is positive and statistically significant at the 1% level, except for Louisiana and Maryland, where it is positive but not significantly different from zero. Hence, in most states the average investor devotes significantly more attention to local companies than to nonlocal ones.

Irrespective of whether the geography of attention is measured from the standpoint of firms or investors, the country-wide average local-nonlocal attention gap is 23.23, namely,

[10]Note that the total number of searches in column 5 is a weighted average of the numbers of local and nonlocal searches for local firms shown in columns 6 and 7. By construction, this average is close to the number of nonlocal searches, which are weighted by 49/50 (the fraction of other states over the total), while local searches are weighted by 1/50.

the difference between the average of local searches for local firms (42.90) and the average of nonlocal searches for such firms or, equivalently, the average of searches directed by local investors to nonlocal firms (19.68), the latter two being identical by construction. Hence Table 2 documents that firms' geographic proximity plays an important role in the allocation of attention by U.S. retail investors: local attention is about twice as much than non-local attention. Figure 3 conveys this result graphically, plotting the web searches directed to firms in each state by local investors against those by nonlocal investors: all states (except for the above-mentioned three) lie above the bisector. The same picture emerges at firm level: in Figure 4 most points, which represent combinations of local and nonlocal `GSearch` for each sample firm, lay above the bisector.

4.2 Investor attention, proximity and news

Recall that the model in Section 2 not only predicts that firms attract more attention from local than by nonlocal investors, but also provides several testable predictions about the relationships between geographic proximity and the response of investor attention to news and volatility. In this section, we test these predictions by estimating the following baseline model and several variants of it:

$$\texttt{GSearch}_{ist} = \boldsymbol{\beta}'\boldsymbol{x}_{ist} + \boldsymbol{\gamma}'\boldsymbol{c}_{ist} + \mu_i + \tau_t + \epsilon_{ist}, \tag{7}$$

where $\texttt{GSearch}_{ist}$ is the web search activity for ticker i, in week t and state s; $\boldsymbol{x}_{ist}$ is the vector of news-related variables (i.e., `News`, `LocalNews`, and `Vol`), geographic variables (`SameState` or `Distance`) and their interactions with news-related variables; $\boldsymbol{c}_{ist}$ is the vector of state-level controls, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are the vectors of coefficients to be estimated, the parameters μ_i and τ_t indicate firm and week fixed effects, and ϵ_{ist} is the error term. While model (7) suggests a multilevel structure of the data, this is not the case for our data, as each ticker i is searched in multiple search states s.[11] By the same token, our data do not have a pure panel structure either, as at each date t web searches for firm i vary across states s. We estimate equation (8) by OLS augmented with firm fixed effects and year effects, and make inference based on robust standard errors clustered at firm level.

In Table 3 we investigate if investor attention exhibits a geographic trait, and whether local and nonlocal investors' attention correlates differently with the intensity of news regarding a firm and the volatility of its returns, as predicted by the model presented in Section 2. All specifications include firm and week fixed effects, to control for unobserved heterogeneity in attention between stocks and for aggregate time patterns in attention;

[11] Multilevel data have a hierarchical structure, featuring multiple units of analysis, each nested within the other.

moreover, the specifications shown in columns 3 to 7 include controls for search-state demographic and economic characteristics.

The estimated coefficient of the `SameState` dummy shown in the first row of the table indicate that `GSearch` is between 15 and 20 points larger in the state where the firm is headquartered than in other states, confirming the visual evidence in Figures 3 and 4. Column 2 shows that the relationship between investors' attention and their distance from issuing firms goes beyond the distinction between in-state and out-of-state firms: replacing the dummy `SameState` with the continuous variable `Distance`, we find that a 1,000 km increase in distance between a firm's headquarters and the state of the web search is associated with a drop in web searches of 0.65 points, that is, 3% of the sample average. These findings highlight the key role of geographic proximity in investor attention, in line with our model. The estimates of the state-level controls indicate that investor attention is higher in states with a more numerous, older, richer and more educated population, consistently with the idea that retail investor attention correlates with greater sophistication, hence lower costs of information processing, as documented by household finance research (Guiso and Jappelli, 2005).

To explore whether the local-nonlocal attention gap also responds to news as predicted by the model, in columns 3 and 4 we expand the specification of the model to allow for the news-related variables (`News` and `LocalNews`) and volatility (`Vol`), as well as their interactions with the `SameState` dummy variable. The estimates in column 3 show that investors' attention is positively and significantly related to the number of financial news (the coefficient of `News` being 0.20) but more so in the state where the firm is headquartered (the coefficient of the interaction `SameState` $\times$ `News` being 1.28 and significantly different from zero), again in line with the predictions of the model. Since `GSearch` is a score ranging between 0 and 100, to appreciate the economic significance of these results, consider that a 1-standard-deviation increase in the number of news relative to its sample average (from 90 to 237 news items) is associated with a 0.1 increase in nonlocal (i.e., out-of-state) web searches and a 0.7 increase in local (in-state) web searches. Hence, local investors react about 7 times more than nonlocal investors to news releases.[12]

The positive and significant coefficient of the `LocalNews` variable indicates that investors' attention responds to information published on local newspapers. However, such information does not appear to elicit a stronger response from investors close to the company headquarters, considering that the interaction of this variable with `SameState` is not statistically different from zero. This suggests that local investors' attention does not

[12]The effect of the number of news on nonlocal attention is given by $0.20 \ln(147+90+1) - 0.20 \ln(90+1) \approx 0.1$, while that on local attention is given by $(0.20 \ln(247+90+1) + 1.28 \ln(247+90+1) - 0.20 \ln(90+1) - 1.28 \ln(90+1) \approx 0.7$. Notice that we add 1 to the argument of the logarithm because `News` is defined as the log of one plus the number of news.

lie in their superior access to local news but in their different ability to process publicly available information, in line with the model's assumptions.

Changes in volatility also elicit widely different responses of nonlocal and local attention: in column 3, the coefficient of the interaction between `SameState` and `Vol` is 57.24, so that an increase in weekly volatility by one standard deviation (2%) is associated with increases in nonlocal and local attention of 0.11 and 1.14 points, respectively.[13] In other words, an increase in volatility is associated with an increase in attention by local investors about 12 times as large as that by nonlocal investors.

A possible concern about the estimates shown in column 3 is that the relationship between our measure of investors' attention and return volatility may be bi-directional, as highlighted by our model: on the one hand, more volatile returns should elicit greater attention; on the other, greater attention can be expected to increase the informational content of the order flow, and thus increase stock price volatility, as found by Andrei and Hasler (2014). To address the potential endogeneity of volatility due to reverse causality, we instrument the `Vol` variable and its interaction with the `SameState` dummy with industry-level return volatility and its interaction with that dummy: the number of web searches for a specific company should affect industry-level volatility less than firm-level volatility, being likely to focus mostly on firm idiosyncratic information. Column 4 presents 2SLS estimates of the specification shown in column 3 of the table. The estimates indicate that the baseline level of local-nonlocal attention gap (captured by the coefficient of the `SameState` dummy) is very close to that obtained by OLS in column 4, but the response of local attention to volatility (the estimated coefficient of the interaction between `Vol` and the `SameState` dummy) is about twice as large: hence, the endogeneity of stock return volatility appears to lead – if anything – to a sizable underestimate of its impact on local investors' attention.

The last two columns of the table show that this differential impact of volatility on local and nonlocal attention applies both to good and bad news: when the specification of column 3 is estimated separately on the subsample with positive returns (column 5) and that with negative returns (column 6), both the coefficient of the `Vol` variable and that of its interaction with `SameState` are positive and statistically different from zero in both cases. Indeed, the relevant coefficients are larger in the presence of negative returns than positive ones, although the difference is not significantly different from zero. Hence, the evidence is inconsistent with the 'ostrich effect', i.e. investors paying less attention to their portfolios upon receiving bad news than good ones (Galai and Sade, 2006; Karlsson, Loewenstein, and Seppi, 2009), and is instead consistent with the finding by Boehmer and Song (2021) that retail investors profitably exploit negative public information by

[13]The two figures result respectively from 5.41×0.02 and $(5.41 + 57.24) \times 0.02$.

short-selling overvalued stocks.[14]

To sum up, the results from Table 3 are in line with the model's predictions: not only local attention systematically exceeds nonlocal attention, but the gap between the former and the latter is amplified by news volume and stock return volatility.

These results are quite robust, as shown by the robustness and validation tests reported in Table A1 of the Appendix. One possible concern is that our sample may be biased towards a few states where firms cluster their headquarters, 28% of the sample firms being headquartered in either California or New York. But, upon re-estimating the specification of column (1) of Table 3 without the observations for those two states, the estimates reveal an even stronger local-nonlocal attention gap (23.14) than when these observations are retained (20.46).

A second possible concern is that the ticker of ten stocks in our data is formed by a single letter,[15] so that for these companies the number of searches may reflect typos by web users rather than intentional web searches. But when observations for these stocks are dropped, the results are unaffected.[16]

We also perform a test to validate that the `SameState` variable captures the gap between local and nonlocal attention in Table 3, based on the idea that local attention should drop on state holidays, local investors being more distracted by leisurely activities than on working days. Hence, we generate the dummy variable `Holi`, which equals 1 when a state-specific holiday occurs in a given week, and 0 otherwise, and we add this variable and its interaction with `SameState` to the specification: as expected, while on other days local attention continues to exceed nonlocal one, during such holidays the gap drops significantly (by about 1/4) in the headquarter state.

Finally, we conduct a falsification test to provide further evidence that the results are unlikely to be driven by unobservable characteristics of the states where firms' headquarters are located. Specifically, we randomize the states of firms' headquarters, while preserving the state-level frequencies of headquarters' distribution in our sample, by bootstrapping without replacement and forming 500 samples of firms with randomly-assigned

[14] This result is instead in contrast with the evidence by Sicherman, Loewenstein, Seppi, and Utkus (2015). The difference in results may stem from the difference in the measure of attention: the measure used by Sicherman, Loewenstein, Seppi, and Utkus (2015) is based on investor online account logins, while ours is measure on behavior that is likely to precede account logins, reflecting the first attempt to acquire information by investors. For this reason, our proxy could be considered a more fine-grained measure of attention compared to account logins. Moreover, our measure is likely to refer to the attention of less self-selected individuals rather than a sample of individuals who already have an online account and login into it to place an order.

[15] This applies to Agilent Technologies, AT&T, Citigroup, Dominion Energy, Ford Motor, Kellogg, Loews, Macy's, Realty Income and VISA.

[16] The results are also robust to excluding observations for firms that have a name quite similar to the financial tickers (e.g. CBS), as unreported results show. A further unreported analysis shows that the words most often searched in conjunction with our tickers are 'stock' and 'equity'.

headquarters. Then, these data are used to re-estimate the specification in column 1 of Table 3 for each of the 500 samples. If any underlying characteristic of the states where local investors search more intensively for firms were to drive our results, we should expect to find similar results to those reported in Table 3 using randomized headquarter data. Reassuringly, with randomly assigned headquarters the coefficient of the `SameState` variable is on average -0.0057 and is not statistically larger than zero at the 5% significance level for 496 of the 500 regressions. Hence, if company headquarters were randomly assigned, local web searches would be very unlikely to exceed nonlocal ones.

5 Shocks to Proximity and Attention Reallocation

The evidence presented in the previous section documents that U.S. companies attract more attention by local than nonlocal investors, and the gap between the two is positively correlated with news volume and return volatility, consistently with the predictions of our model. However, this evidence does not enable us to isolate the effect of geographic proximity *per se* on investors' attention from the effect of stock ownership. If for instance, due to a familiarity bias, investors tend to overweight local stocks in their portfolios, then it would be natural for them to pay more attention to local stocks than to nonlocal ones. In this case, the local-nonlocal attention gap would not stem from a comparative advantage of local investors in collecting and/or processing information about local stocks, as in our model, but from a familiarity bias, via the implied home bias in stock ownership. Hence, a sharper test of the hypothesis that the local skew in attention arises from a proximity-induced informational advantage requires focusing on situations in which investors face a sudden change in their perceived proximity to a company, irrespective of their initial equity portfolio: for instance, in their eyes a formerly distant firm suddenly becomes closer to them, and as such easier to analyze and worthy of greater attention, although they may have initially had little or no ownership stake in such a firm.

To this purpose, in this section we consider how investors' attention changes in response to shocks to their perceived proximity to firms. We focus on two very different types of shocks: first, those induced by acquisitions of local firms by nonlocal ones or, conversely, acquisitions of nonlocal firms by local ones (Section 5.1); second, those triggered by the COVID-19 pandemic, which has notoriously hindered travel to distant locations, thus raising the relative cost of collecting first-hand information about nonlocal firms relative to local ones, and thus local investors' informational advantage for local stocks relative to nonlocal ones (Section 5.2).

5.1 M&A shocks to perceived proximity

In this section, we document the dynamics of local attention when the perception of a firm's proximity is modified by an acquisition. To do so, we collect data about all 142 completed takeovers' announcements in 2017 concerning a U.S. company present in the S&P500, either as a bidder or as a target. From these, we remove the 38 cases where the acquisition targeted a firm headquartered in the same state of the acquirer, as these cases are not associated with a change in the perceived distance of the relevant firms. Of the remaining 104 M&A transactions involving a bidder and a target located in different states (based on their respective headquarters' location), 82 are initiated by a S&P500 bidder, and 22 are directed to a S&P500 target.

To measure the dynamics of attention of the investors potentially affected by M&A-induced shocks in perceived firm proximity, we download the time-series of `GSearch` both in the bidder's state and in target's state. As mentioned in Section 3, while Google does not allow to download time series of web searches disaggregated by firm and state for all states at the same time, it allows downloading this time-series for one state at a time. Specifically, for the 82 transactions initiated by a bidder in our sample, we measure the web searches of the bidder's ticker both in its own state and in the state of the target. Symmetrically, for the 22 transactions directed to a target in our sample, we measure the web searches of the target's ticker both in its own state and in the bidder's state. On the whole, these cross-state M&A transactions are not concentrated in few states: bidders are present in 27 states and targets in 29 states, naturally with a prevalence of the states hosting more company headquarters according to the top-left panel in Figure 2.

We analyze separately the short-term and long-term response of investors' attention to changes in perceived distance triggered by M&As. This analysis is similar in spirit to that typically done in event studies, where a normal predicted value of the response variable is compared with its actual value to detect any event-induced anomaly. We estimate linear fixed-effects models with company-specific slopes to capture the potentially differential response of investors' attention towards the acquiring and the target firms, depending on their location (Wooldridge, 2010). Having specific intercepts and slopes for each company allows for meaningful comparisons, since pre- and post-event comparisons are firm-specific. Furthermore, the average effect shown as a result in the tables benefits from the fact that all the search variables have a predetermined 0-100 range, which lowers the risk of the results being driven by a few influential observations.

Table 4 shows how the M&A event affects investors' attention in the acquirer's and in the target's state, in a three-week window around the announcement ($t = -1$; $t = +1$). The estimates in column 1 indicate that the attention paid to the acquirer by residents in its own state (bidder-home) does not significantly change either in the week of the public announcement ($t = 0$) or in the previous week ($t = -1$). Relative to its 41.7 average (the estimated constant), `GSearch` increases significantly only in the week after the event (4.33 points more). Column 2, instead, shows the estimates regarding the attention paid to the acquirer by residents of the state where the target is headquartered. Unsurprisingly, given the local skew in attention allocation, these investors pay less attention to the bidder than those located in the bidder's own state (the relevant constant being 38.82). However, both in the week before and in the week of the announcement, attention paid to the bidder by these nonlocal investors rises by 6.47 points and 4.83 points, respectively. The last two columns of Table 4 concern the attention devoted to the target, both in the state where the firm is headquartered (column 3: target-home) and in the state where the bidder is headquartered column 4: target-other). Web searches regarding the target company rise significantly in the weeks around the announcement in both states: the event attracts a significant increase in attention by residents of both the bidder's and the target's state.

To summarize, in the short run the M&A event raises out-of-state investors' attention for both target and bidder: residents in the target company's state become more interested in information about the out-of-state bidder (bidder-other), and residents in the bidder's state become more interested in the out-of-state target company (target-other). Insofar as M&A events are unanticipated, this increase in attention is unlikely to be driven by former shareholdings in the out-of-state bidder or target, respectively, and may instead reflect an increase in the perceived closeness of investors to the bidder's and target's headquarters, respectively: the acquirer turns into a quasi-local company in the eyes of residents in the target company's state, and the target does the same in the eyes of residents in the acquirer's state.

However, the surge of interest for the target immediately after its acquisition documented by Table 4 interest is transitory. This is apparent in Table 5, which is based on data for a longer time window, spanning 52 weeks before to 52 weeks after the public announcement, where the coefficient of the dummy variable `PostEvent` enables to test whether there are persistent changes in investors' attention. The estimates in column 1 of this table show that in the post-acquisition year the attention devoted to the bidder by its local investors rises significantly by 2 points (`PostEvent`) relative to its 41.75 average. Column 2 indicates that this permanent increase in attention for the bidder is present also in the state where the target is headquartered (bidder-other). In contrast, the estimates in columns 3 and 4 reveal a generalized and persistent drop in attention for the target, especially by residents in the target's own state (target-home). Hence, on the whole the results point to a persistent shift in investors' interest away from the target and in favor of the acquirer, probably reflecting awareness that control over the target has now shifted in the acquirer's hands. However, it is striking that this reallocation in attention is especially large among the residents of the target's state, consistently with the idea that these investors now perceive the target as quasi-nonlocal, rather than truly local: its headquarters are still local, but decisions over its management are taken elsewhere.

This interpretation of the results of the regressions shown in columns 2 and 3 of Table 5 is strengthened by those shown in Table 6, where these regressions are re-estimated splitting the sample on the basis of the average size of the acquisition: columns 1 and 2 report the estimates for attention paid by the target's state residents to large and small bidders, respectively; column 3 and 4 show the estimates for attention paid by the bidder's state residents to large and small targets, respectively. The estimates in columns 1 and 2 show that the permanent increase in attention for the acquirer by residents in the target's state is present – and indeed quite strong – for large transactions only: after the announcement, for those events the coefficient indicates a 6.44 increase in web searches in the target's state, while for acquisitions that involve small acquirers, the effect vanishes. The estimates in columns 3 and 4 confirm the permanent drop in attention for the target company by local investors but indicate that the drop is particularly large for small targets: -8.11 versus -3.73 points, to be compared with similar average attention levels (34.42 and 36.23 respectively). These estimates suggest that investors located in the target company's state pay more attention to the acquirer only when it is large enough to dictate the policy of the target company, and lower their attention towards the target more significantly when the target is small, and therefore more likely to be governed by a nonlocal owner. in other words, they regard large out-of-state acquirers as more local, and small same-state targets are less local.

5.2 COVID-19 shocks to proximity

The COVID-19 outbreak provides another opportunity to identify the causal relationship between investors' perceived proximity from firms and their attention to stocks. One effect of the pandemic has been to restore the relevance of geographic proximity to economic choices, by hindering travel activity. As a result, the COVID-19 outbreak can be expected to have reduced investors' ability to collect first-hand information about nonlocal stocks, and thus to have made it harder for them to assess news found on the web about such companies, compared to news about local companies. By the same token, the pandemic should have improved their comparative advantage in collecting and processing information about local stocks. Hence, the prediction is that COVID-19 should have triggered an increase in the local skew of investors' attention, being a shock increase to the perceived distance of nonlocal companies.

To test this hypothesis, we download weekly data for the `GSearch` variable from 2

December 2019 to 17 May 2020 as done for the data used to estimate the models in Tables 3, and use them to estimate model specifications that allow for different levels and sensitivities to distance of investors' attention before and after the inception of the pandemic. Using data drawn from the New York Times, we identify the week starting on 2 March 2020 as that in which the pandemic became relevant in investors' eyes, as the first fatality due to COVID-19 was reported on Saturday 29 February, and deaths rose dramatically to 96 in the subsequent week (from 2 to 8 March 2020). Hence we create a `DCovid` dummy that equals 1 starting on that week, and 0 before.[17]

Table 7 presents the results. The estimates reported in column 1 refer to the same model specification presented in column 1 of Table 3, and confirm that the local-nonlocal attention gap is present also in this more recent sample. The results in column 2 show that, since the onset of the pandemic, retail investors' attention not only increased significantly, in line with the strong increase in retail trading,[18] but did so in particular for local companies: the estimated coefficient of the interaction `SameState` $\times$ `DCovid` implies that local searches during the pandemic are more than twice as large than nonlocal searches ($2.32 = 4.22/1.82$), and the attention gap rises by almost one fourth above its pre-COVID-19 period ($0.24 = 4.22/17.50$). Similar results are obtained by replacing the `SameState` dummy with the continuous `Distance` variable: the estimates shown in column 3 indicate that the rise in perceived distance triggered by the pandemic (`Distance` $\times$ `DCovid`) has reduced investors' attention for nonlocal stocks: the negative effect of distance on attention during the pandemic increases by one fourth relative to its previous value ($0.16/0.64$).

To further test the hypothesis that the drop in investors' attention for nonlocal stocks after the onset of COVID-19 stems from an increase of the perceived distance due to travel restrictions, we collect data about the number of direct flights available between U.S. cities in any given month, so as to construct a variable capturing the availability of air travel from any state to the state where the headquarters of sample companies are located (source: US Department of Transportation), both before and after the onset of COVID-19 (as defined above). Specifically, we merge the weekly number of searches made in any state for an out-of-state company with the total number of flights between the search state and the company headquarter state.

These data enable us to construct a new variable, `AdjDistance`, namely, the ratio between the `Distance` variable and the number of monthly flights from the search state (i.e., the state where the web search occurs) to the headquarters state (i.e., the state

[17] In that same week, the Google search for the word "COVID" increased more than 3 times relative to the previous week, the maximum growth rate reached by this proxy for the growth in attention.

[18] According to estimates by JPMorgan, since the pandemic retail investors have accounted for a larger fraction of total trading volumes, ranging between 20 and 30 per cent, as opposed to 10 to 15 percent before the pandemic (see The *Financial Times* "Is the army of lockdown traders here to stay?", 18 October 2021).

where the company's headquarters are located). We increase by 1 the number of flights in the denominator of this fraction to avoid dividing by zero if no flights are available – hence, no passengers travel – between the search state and the headquarters state in a given week. So by construction `AdjDistance` is decreasing in the number of flights between the two states, its maximum value being the physical distance between the two states (i.e., `Distance`) when the number of flights connecting them drops to zero.

During the COVID-19 period, the number of flights across states dropped dramatically, as illustrated by Figure 5: their median (mean) dropped from 133 (500) in February to 126 (470) in March and 28 (150) in April 2020. Flight cancellations were not uniform: they affected short-range flights more than long-range ones; moreover, relatively more flights were canceled on routes to and from some states, such as California and New York, because these states were more severely struck by the pandemic. Therefore, `AdjDistance` increased on average, but not uniformly for the whole country: it rose differently depending on how many flights were canceled across couples of them during the pandemic.

As a result, while `Distance` varies only across states, `AdjDistance` also varies over time, and does so differently across states, capturing different increases in the difficulty of covering the distance between states by air travel, and thus of collecting on-site information about distant companies: as such, this variable is a time-varying measure of the perceived distance between an investor and a company. Hence, we expect not only investors' attention to be decreasing in `AdjDistance`, but also to respond more to this variable during the pandemic, when air travel became generally harder or impossible. Hence, the incremental response of attention to `AdjDistance` during the pandemic should be larger than that estimated on the basis of physical distance, since it takes into account not only cross-sectional variation in distance, but also the differential time-series variation in the number of flights across states.

The estimates obtained when perceived distance is measured by `AdjDistance` are consistent with all three predictions: (i) its coefficient is negative in both columns 4 and 5; (ii) the coefficient of its interaction with the COVID-19 dummy is also negative, witnessing the incremental negative impact of COVID-19 on investor attention; (iii) this incremental impact is larger relative to the corresponding pre-COVID-19 effect in column 5 ($0.18/0.26 = 0.69$) than it is if estimated on the basis of physical distance in column 3 ($0.16/0.84 = 0.25$). Intuitively, for any given physical distance between states, those harder to reach (i.e., with fewer flights) are perceived as being farther away, and the differential impact of COVID-19 is better captured by its interaction with such perceived distance than with physical distance. The specification reported in column 6, where physical distance and number of flights are entered as separate variables (rather than as a ratio as in the `AdjDistance` variable) shows that each of the two variables has a distinct explanatory value, both before and after the onset of the pandemic, and in each case their coefficients have the expected sign and are significantly different from zero.

Finally, the specification shown in column 7 explores whether during the pandemic investors' attention reacted differently to local news compared to normal times. First, the estimates concerning the response of investors' attention to local news in normal times confirm the results found in Table 3: the coefficient of the `LocalNews` variable indicates that, when local newspapers report news about a company located in the same state, the company attracts more attention by local and nonlocal investors alike, as the coefficient of the interaction with `SameState` is not significantly different from zero. But during the pandemic news in the local press appeared to attract more attention than usual by investors, especially by local ones, whose incremental response to local news was almost four times as large as that of nonlocal investors: the coefficient of the triple interaction (`SameState` x `DCovid` x `LocalNews`) is 4.39, while that of the interaction `DCovid` x `LocalNews`) is 1.16. This witnesses the increased importance of local information at a time when investors were almost unable to travel and collect first-hand information in distant locations, and therefore is consistent with the idea that the pandemic was perceived by investors as a sudden increase in perceived distance, leading them to refocus on assets physically close to them.

6 Attention, Volatility and Liquidity

So far, the empirical analysis has focused on the geography of investor attention and its determinants, and on the extent to which they conform to the predictions of the model in Section 2. But that model also predicts that the geography of investor attention affects market outcomes, namely, local investors' attention should increase the price impact of orders – hence return volatility – and the bid-ask spread more than the attention of nonlocal investors. Moreover, insofar as local investors' attention also confers them a time advantage in information processing over other investors, local attention should lead nonlocal attention in time. In this section, we test these additional predictions of the model.

To estimate the dynamic relationships between the variables of interest, our 2017 weekly data must be converted into a panel. The measures used in Section 4.2 allow comparing local and nonlocal web searches within a week, not across weeks, since the data used so far `GSearch` is re-based each week. As already mentioned, Google does not provide time series of web searches disaggregated by company and state for all states at the same time. To circumvent this problem, for each week we multiply the fraction of web searches regarding a given company occurring in each state by the aggregate number of searches occurring for that company in the whole of the U.S. in the same week. This produces two weekly time-series for each company, namely, the number of web searches made in the state where the company is headquartered, `LocalSearch`, and those made in other states, `NonLocalSearch`. Moreover, for each week and company, we construct time-series of additional variables, i.e. return volatility, measured by the absolute value of weekly returns (`Vol`) and the average weekly spread (`Spread`).

Then, we estimate panel vector autoregressive (pVAR) models of the following type:

$$\boldsymbol{y}_{it} = \sum_{s=1}^{S} \boldsymbol{A}_s \boldsymbol{y}_{it-s} + \boldsymbol{\mu}_i + \boldsymbol{\epsilon}_{it} \tag{8}$$

where $\boldsymbol{y}_{it}$ is the vector of endogenous variables, $\boldsymbol{A}_s$ are matrices of parameters to be estimated, $\boldsymbol{\mu}_i$ is a vector of variable-specific firm fixed effects accounting for systematic cross-sectional heterogeneity and $\boldsymbol{\epsilon}_{it}$ is a vector of variable-specific idiosyncratic error terms. We estimate the system of equations (8) using the Generalized Method of Moments (GMM), as the fixed effects estimator would generate biased estimates due to the presence of lagged dependent variables. As Nickell (1981) notes, the problem arises because the demeaning process of the fixed effect estimators generates correlation between regressors and error terms, causing endogeneity. Instead, the GMM estimator provides consistent estimates. Since the estimation requires stationarity, we first perform Fisher-type unit-root tests based on Phillips-Perron tests (Choi, 2001). Those tests reveal that the variables used in the analysis are stationary.

Table 8 presents the estimates of three different specifications of the system of equations (8): in the first, shown in column 1, the vector $\boldsymbol{y}_{it}$ only includes `LocalSearch` and `NonLocalSearch`; in the second (column 2) and in the third (columns 3), the vector of endogenous variables also includes stock price volatility (`Vol`) and the weekly average bid-ask spread of the relevant company (`Spread`), respectively. The autoregressive order is one. We also estimate models with more lags, but we just report results with a single lag as we find that this lag structure is sufficient to account for the dynamic features of the variables. Since data for the `Spread` variable feature outliers, we trim the top and bottom 1% of its distribution.

In the model whose estimates are shown in column 1, both local and nonlocal attention feature autoregressive memory. But the most interesting result is that their time precedence relationship is asymmetric: while lagged `LocalSearch` has a strong predictive power for `NonLocalSearch`, the latter has little predictive power for the former. In other words, local attention anticipates nonlocal attention much more than the opposite. While we estimate simple temporal correlations, this result is consistent with the idea that local investors' attention confers them a time advantage over nonlocal investors.

The estimates shown in column 2 are obtained by adding return volatility (`Vol`) to local and nonlocal attention in the vector of endogenous variables. Beside confirming the time precedence relationship between the two attention variables shown in column 1, the estimates of column 2 show that increases in attention predict increases in volatility, with local attention being a stronger predictor than nonlocal attention – the difference between the respective coefficients (0.96 and 0.63) being statistically significant. This aligns with our model's prediction that orders placed by local investors have a stronger informational content than those placed by nonlocal ones. It is also consistent with other evidence that investors' attention amplifies asset volatility (Andrei and Hasler, 2014). Moreover, neither attention variable is predicted by absolute returns, as one would expect. It is worth relating these results to those presented in Table 3, which indicate that a rise in return volatility elicits a stronger contemporaneous increase in local investors' attention than in that of nonlocal ones, also when the potential feedback of attention on volatility is controlled for with IV estimation. The estimates in column 2 of Table 8 investigate precisely this feedback effect from attention to volatility, exploiting the dynamic structure of the data, and show that also in this case the relationship is stronger for local than nonlocal attention.

Column 3 presents estimates of the dynamic relationship between investors' attention and the bid-ask spread: both local and nonlocal web searches are positively and significantly correlated with subsequent values of the bid-ask spread. The difference between the two coefficients (0.26 and 0.34) is not statistically significant: increases in investors' attention predict a drop in market liquidity, consistently with the model, although there is no evidence of this dynamic relationship being stronger for local rather than nonlocal investors. Instead, the bid-ask spread has no predictive power for either investor group's attention.

7 Conclusion

In this paper, we document that retail investors pay more than twice as much attention to news about firms whose headquarters are located in their state as those located elsewhere; more generally, their attention is decreasing in their distance from the company's headquarters. News availability and volatility amplify this attention gap: local attention reacts about 7 times more than nonlocal one to local firm news (whether positive or negative) and 12 times more to an increase in stock volatility. These findings are consistent with a model where the local investors skew their attention allocation towards local asset because they are better at processing information about them, rather than because of behavioral biases. They are also consistent with the findings of other recent papers (Gargano and Rossi (2018), Cziraki, Mondria, and Wu (2021) and Dyer (2021)) documenting that investors' attention predicts stocks' risk-adjusted returns, and that this relationship is much stronger for local investors' attention.

Importantly, the evidence is consistent with investors' attention being causally related to their perceived proximity to firm headquarters. First, a decrease in perceived firm proximity due to an acquisition by a nonlocal firm is associated with lower local attention and greater attention by nonlocal investors located close to the acquiring firm. Second, the increase in perceived distance of nonlocal firms due to COVID-19-related travel restrictions have been associated with a greater local attention bias, and the decline in the relative attention to nonlocal stocks has been stronger for firms in locations connected by fewer flights after the outbreak. This evidence indicates that the local attention bias documented in this paper is not simply a mechanical consequence of the presence of the local ownership bias shown by earlier studies such as Coval and Moskowitz (1999) and Grinblatt and Keloharju (2001)), which in principle could stem from behavioral biases, such as familiarity.

Finally, increases in local attention are shown to predict subsequent increases in return volatility and bid-ask spreads, as well as in nonlocal attention, while they are not predicted by them. Also this empirical finding is consistent with an information-processing advantage of local investors: the greater informational content of their orders should trigger larger price movements, and lead market makers to widen their bid-ask spread, as also highlighted by our model.

References

Andrei, Daniel, and Michael Hasler, 2014, Investor attention and stock market volatility, *The Review of Financial Studies* 28, 33–72.

Barber, Brad M., and Terrance Odean, 2000, Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors, *The Journal of Finance* 55, 773–806.

——— , 2001, Boys will be Boys: Gender, Overconfidence, and Common Stock Investment, *The Quarterly Journal of Economics* 116, 261–292.

Barber, Brad M, and Terrance Odean, 2007, All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, *The Review of Financial Studies* 21, 785–818.

Barber, Brad M., and Terrance Odean, 2013, The Behavior of Individual Investors, in George M. Constantinides, Milton Harris, and Rene M. Stulz, ed.: *Handbook of the Economics of Finance*vol. 2 . chap. 22, pp. 1533–1570 (Elsevier).

Ben-Rephael, Azi, Zhi Da, and Ryan D Israelsen, 2017, It depends on where you search: Institutional investor attention and underreaction to news, *The Review of Financial Studies* 30, 3009–3047.

Boehmer, Ekkehart, Charles M. Jones, Xiaoyan Zhang, and Xinran Zhang, 2020, Tracking Retail Investor Activity, *Journal of Finance,* forthcoming.

Boehmer, Ekkehart, and Wanshan Song, 2021, Smart Retail Traders, Short Sellers, and Stock Returns, SSRN Working Paper No. 3723096.

Choi, In, 2001, Unit Root Tests for Panel Data, *Journal of International Money and Finance* 20, 249–272.

Cohen, Lauren, 2008, Loyalty-Based Portfolio Choice, *The Review of Financial Studies* 22, 1213–1245.

Coval, Joshua D, and Tobias J Moskowitz, 1999, Home Bias at Home: Local Equity Preference in Domestic Portfolios, *The Journal of Finance* 54, 2045–2073.

Cziraki, Peter, Jordi Mondria, and Thomas Wu, 2021, Asymmetric Attention and Stock Returns, *Management Science* 67, 48–71.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In Search of Attention, *The Journal of Finance* 66, 1461–1499.

Dyer, Travis A., 2021, The demand for public information by local and nonlocal investors: Evidence from investor-level data, *Journal of Accounting and Economics* 72, 101417.

Engelberg, Joseph E, and Christopher A Parsons, 2011, The Causal Impact of Media in Financial Markets, *The Journal of Finance* 66, 67–97.

Friedman, Henry L., and Zitong Zeng, 2021, Retail Investor Trading and Market Reactions to Earnings Announcements, SSRN Working Paper No. 3817979.

Galai, Dan, and Orly Sade, 2006, The "Ostrich Effect" and the Relationship between the Liquidity and the Yields of Financial Assets, *The Journal of Business* 79, 2741–2759.

Gargano, Antonio, and Alberto G Rossi, 2018, Does It Pay to Pay Attention?, *The Review of Financial Studies* 31, 4595–4649.

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, 71 – 100.

Grinblatt, Mark, and Matti Keloharju, 2001, How Distance, Language, and Culture Influence Stockholdings and Trades, *The Journal of Finance* 56, 1053–1073.

Guiso, Luigi, and Tullio Jappelli, 2005, Awareness and Stock Market Participation, *Review of Finance* 9, 537–567.

———, 2020, Investment in Financial Information and Portfolio Performance, *Economica* 87, 1133–1170.

Huberman, Gur, 2015, Familiarity Breeds Investment, *The Review of Financial Studies* 14, 659–680.

Ivković, Zoran, and Scott Weisbenner, 2005, Local does as local is: Information content of the geography of individual investors' common stock investments, *The Journal of Finance* 60, 267–306.

Kahneman, Daniel, 1973, *Attention and Effort* (Prentice-Hall: Englewood Cliffs, NJ).

Kaniel, Ron, Shuming Liu, Gideon Saar, and Sheridan Titman, 2012, Individual Investor Trading and Return Patterns around Earnings Announcements, *The Journal of Finance* 67, 639–680.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, 2008, Individual investor trading and stock returns, *The Journal of Finance* 63, 273–310.

Karlsson, Niklas, George Loewenstein, and Duane Seppi, 2009, The Ostrich Effect: Selective Attention to Information, *Journal of Risk and uncertainty* 38, 95–115.

Kelley, Eric K., and Paul C. Tetlock, 2013, How Wise Are Crowds? Insights from Retail Orders and Stock Returns, *The Journal of Finance* 68, 1229–1265.

Malloy, Christopher J., 2005, The Geography of Equity Analysis, *The Journal of Finance* 60, 719–755.

Massa, Massimo, and Andrei Simonov, 2006, Hedging, Familiarity and Portfolio Choice, *The Review of Financial Studies* 19, 633–685.

Morse, Adair, and Sophie Shive, 2011, Patriotism in Your Portfolio, *Journal of Financial Markets* 14, 411 – 440.

Nickell, Stephen, 1981, Biases in Dynamic Models with Fixed Effects, *Econometrica* pp. 1417–1426.

Shive, Sophie, 2012, Local investors, price discovery, and market efficiency, *Journal of Financial Economics* 104, 145 – 161.

Sicherman, Nachum, George Loewenstein, Duane J Seppi, and Stephen P Utkus, 2015, Financial Attention, *The Review of Financial Studies* 29, 863–897.

Tirole, Jean, 2009, Cognition and Incomplete Contracts, *The American Economic Review* 99, 265—-294.

Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2009, Information Immobility and the Home Bias Puzzle, *The Journal of Finance* 64, 1187–1215.

——— , 2010, Information Acquisition and Under-Diversification, *Review of Economic Studies* 77, 779—805.

Wooldridge, Jeffrey M, 2010, *Econometric Analysis of Cross Section and Panel Data* (MIT press: Cambridge) 2 edn.

Appendix

Proof of Proposition 3. Recalling that $\theta_L > \theta_N$, condition (6) is met for both investors if $\pi < \theta_N/(\theta_N + \theta_L) < 1/2$. In the region $\pi \in (\theta_N/(\theta_N + \theta_L), \theta_L/(\theta_N + \theta_L))$, there is no equilibrium where nonlocal investors are willing to exert attention, but to check that there is an equilibrium where local investors do, the equilibrium must be recalculated on the assumption that local investors are the only informed ones and trade with probability π rather than $\pi/2$. Absent nonlocal investors, local ones choose their attention level a_L to maximize their objective function:

$$\max_{a_i \in [0,1]} \pi(\overline{v} - p_A + q a_i \sigma) - \frac{a_L^2}{2\theta_L},$$

Hence, their optimal attention is

$$a_L^* = \theta_L \pi q \sigma, \tag{9}$$

the equilibrium bid-ask spread (deviation of the ask from the mid-price) is

$$p_A - \overline{v} = \pi q \sigma a_L^* = \theta_L (\pi q \sigma)^2,$$

and the local investors' expected profits are

$$\Pi_L = \pi(\overline{v} - p_A + q a_L^* \sigma) - \frac{a_L^{*2}}{2\theta_L} = \theta_L (\pi q \sigma)^2 \left(\frac{1}{2} - \pi\right), \tag{10}$$

which is positive for $\pi < 1/2$, i.e. the same condition for both types of investors to exert attention and trade when they are identical. Hence, for $\pi \geqslant 1/2$ even local investors exert no attention, so that there is no informed trading in equilibrium.

Table A1
Robustness checks

This table presents robustness checks. Model 1 shows OLS estimates from the regression of `GSearch` on `SameState` and control variables for the subsample that exclude firms headquartered in California and the State of New York. Model 2 shows OLS estimates from the regression of `GSearch` on `SameState` and control variables for the subsample of firms with a ticker longer than one letter. Model 3 shows OLS estimates from the regression of `GSearch` on `SameState` and control variables, including the interaction effects of `SameState` with the dummy variable `Holi`. All variables but `Holi` are described in Section 3. `Holi` is described in Section 4.2. All models include week fixed effects among regressors. Inference is based on unit-cluster standard errors. t-statistics are reported in brackets. Significance code: * p<0.10, ** p<0.05, *** p<0.01.

Variable	1	2	3
`SameState`	23.14***	20.88***	20.72***
	[15.16]	[16.99]	[16.90]
`Holi`			-0.04
			[-0.37]
`SameState` × `Holi`			-4.90***
			[-4.90]
`Population`	1.88***	2.02***	1.99***
	[14.83]	[17.85]	[17.86]
`Age65`$^+$	10.46**	16.41***	15.97***
	[2.16]	[3.87]	[3.83]
`Income`	3.11***	3.67***	3.35***
	[2.95]	[4.10]	[3.78]
`Edu`	13.08***	17.59***	17.89***
	[5.11]	[7.70]	[7.96]
constant	-46.98***	-57.81***	-53.11***
	[-4.28]	[-6.10]	[-5.65]
firm FEs	yes	yes	yes
week FEs	yes	yes	yes
N	893800	1218800	1244800
adj. R^2	0.48	0.455	0.487

Figure 1
Local `GSearch` and nonlocal `GSearch` for six firms in the sample

This figure shows local `GSearch` (when `SameState` = 1; blue solid line) and nonlocal `GSearch` (when `SameState` = 0; red dash line) for Apple (California), Microsoft (Washington), IDEXX Laboratories (Maine), Micron Technology (Idaho), SCANA Corporation (South Carolina) and Danaher Corporation (District of Columbia). The first two companies are the largest in the sample by market capitalization. The remaining four are the only companies headquartered in their respective states.

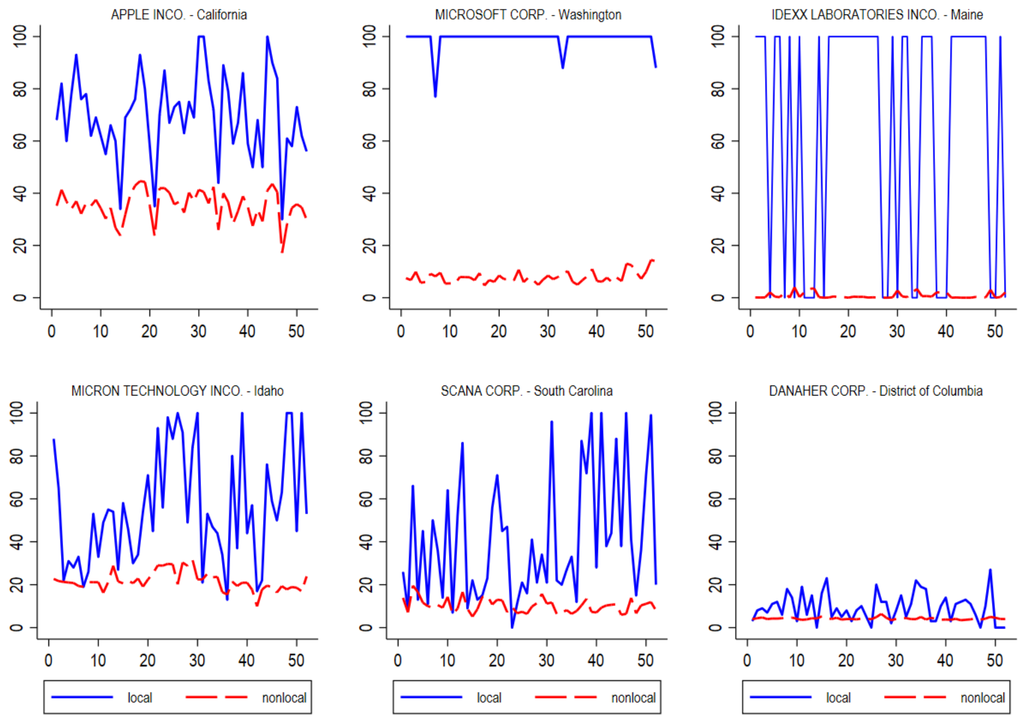


Figure 2
Number of companies by headquarters and `GSearch` split by US state

This figure reports four maps that show the (i) number of companies headquartered in a state (up-left map); (ii) `GSearch` in a state (up-right map); (iii) local `GSearch` of companies headquartered in a state (bottom-left); (iv) nonlocal `GSearch` of companies headquartered in other states (bottom-right). Darker colors indicate higher values. Alaska and Hawaii are removed to facilitate the graphical representation.

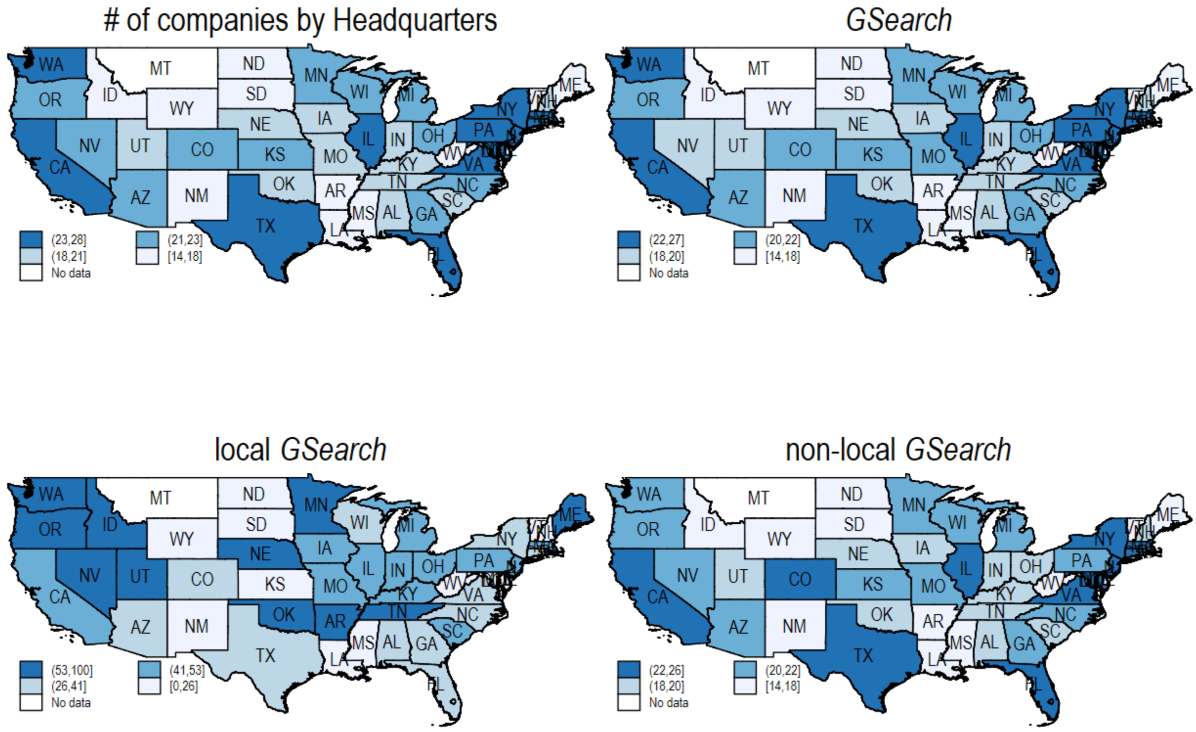


Figure 3
Scatterplot of average `GSearch` by state

This figure shows the scatterplots of local (y-axis) vs. nonlocal (x-axis) `GSearch` by state. The solid line is the bisector.

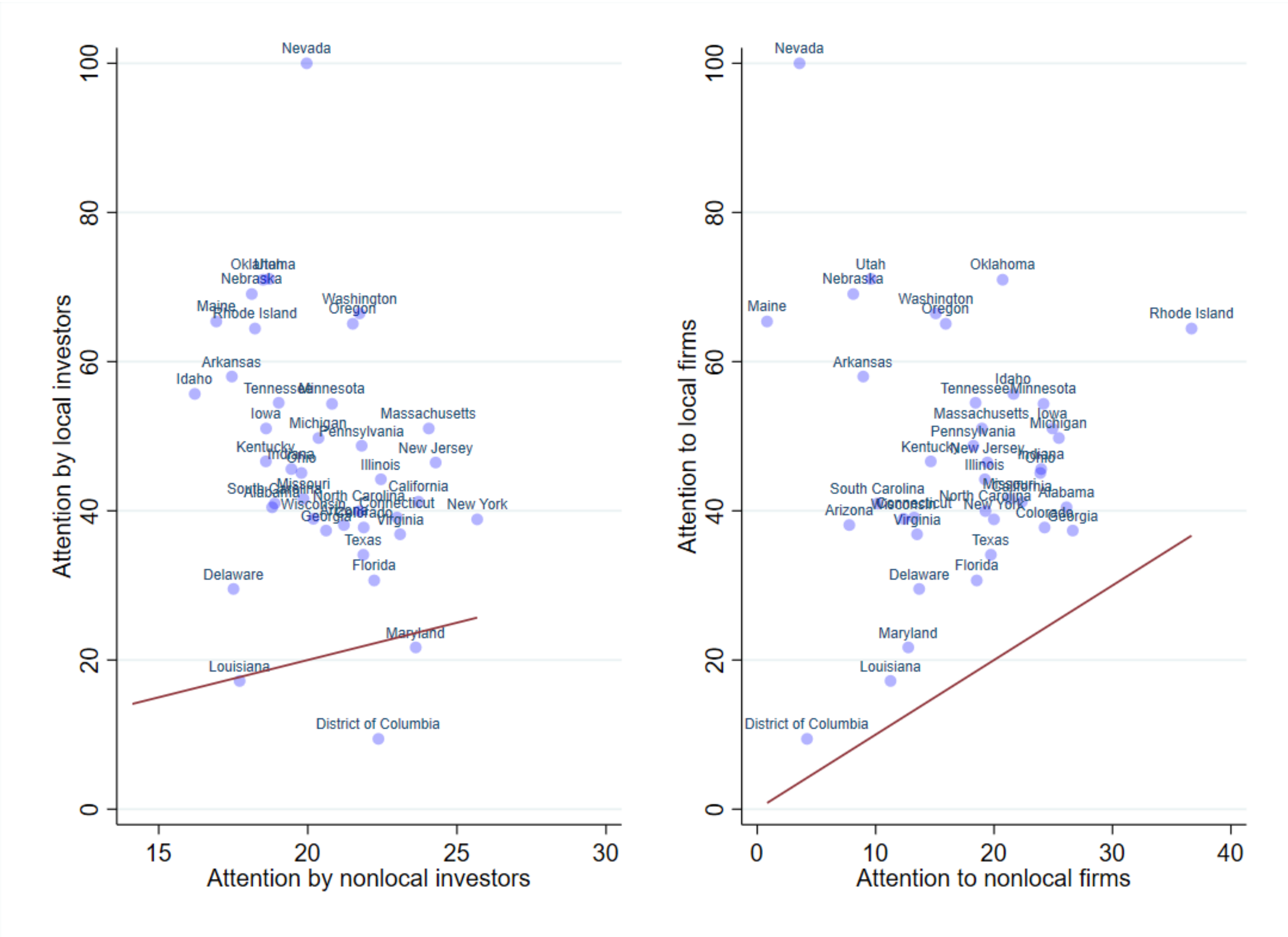


Figure 4
Scatterplot of average `GSearch` by firm

This figure shows the scatterplot of local (y-axis) vs. nonlocal (x-axis) `GSearch` by firms in our sample. The solid line is the bisector.

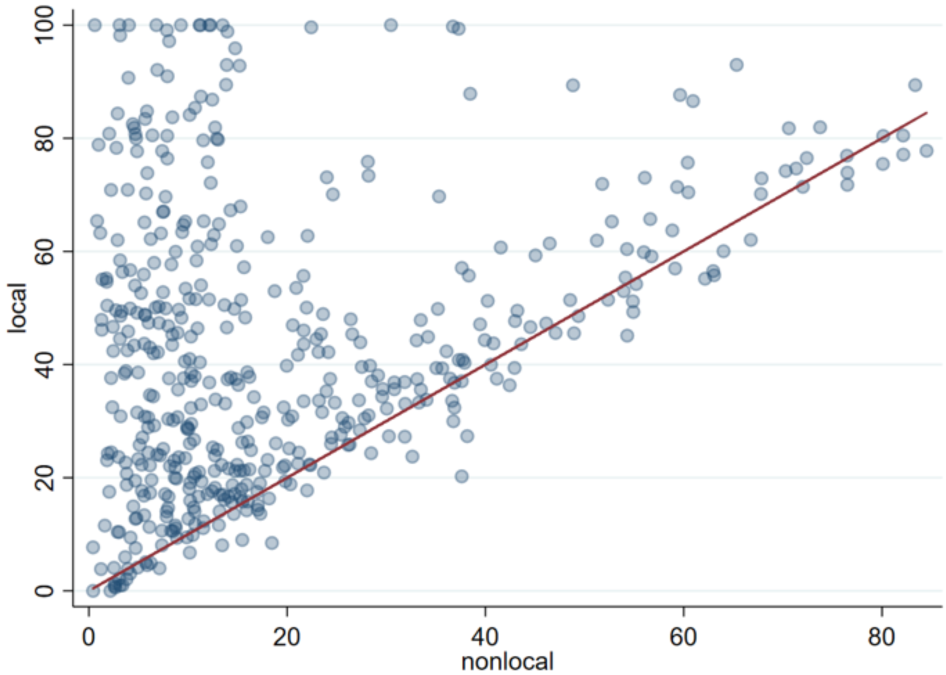
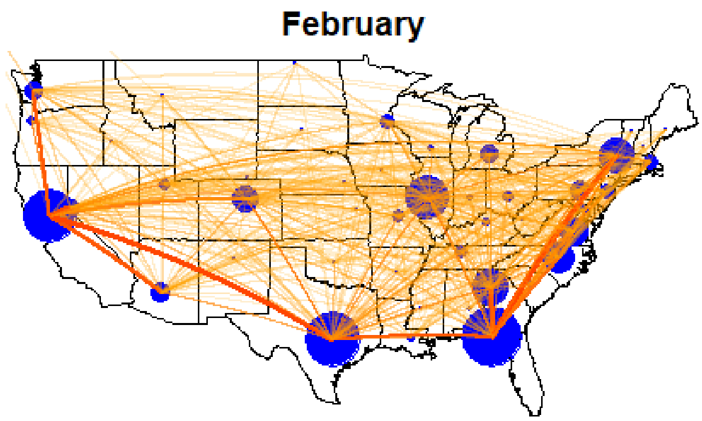


Figure 5
Flight connections between states: before vs. after COVID

The figure shows the number of flights connecting states in February and May 2020, respectively. The size of the blue circles measures the number of flights to the corresponding state. The darkness and thickness of the orange lines measure the number of active routes connecting the corresponding couple of states.



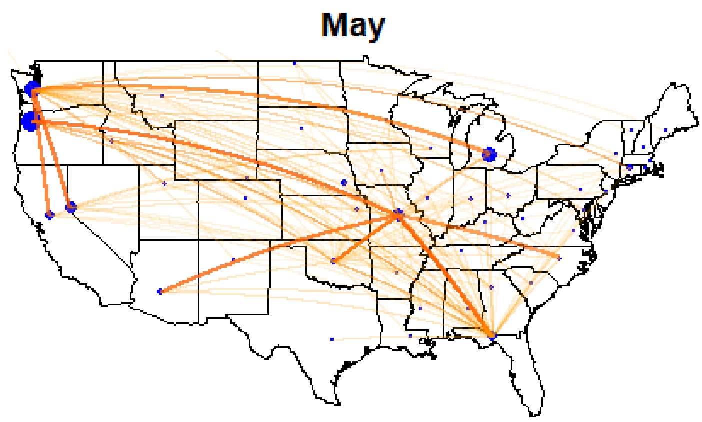


Table 1
Summary statistics

This table shows summary statistics for the variables used in the empirical analysis. All variables are described in Section 3. Reported statistics are means, standard deviations (SD), minima (Min) and maxima (Max). The total number of observations is 1,244,800.

	Mean	SD	Min	Max
`GSearch`	20.14	26.85	0.00	100.00
`SameState`	0.02	0.14	0.00	1.00
`Distance`	1.86	1.37	0.00	8.22
`News`	3.40	1.89	0.00	8.45
`LocNews`	0.00	0.06	0.00	3.53
`Vol`	0.02	0.02	0.00	0.72
`NegRet`	0.43	0.50	0.00	1.00
`Population`	15.17	1.03	13.26	17.48
$\texttt{Age65}^{+}$	0.16	0.02	0.11	0.19
`Income`	10.26	0.16	9.95	10.73
`Edu`	0.30	0.06	0.19	0.55

Table 2
Summary statistics by search state

This table shows the distribution of company headquarters across states and the mean of web searches (GSearch) by search states and company headquarters' states in 2017. For each of the states listed in column 1, local investors are defined as the residents of that state, and nonlocal investors as those resident in other states; local firms as the companies whose headquarters are located in that state, and nonlocal firms as those whose headquarters are located in other states. Columns 2 and 3 report the number and fraction of company headquarters located in the corresponding state indicated in column 1. Columns 4 and 5 show the mean total web searches made by local investors, and of those directed to local firms, respectively. Columns 6 and 7 show the mean web searches directed to local firms by local and nonlocal investors, respectively. Column 8 shows the difference between mean web searches directed to local firms by local and nonlocal investors. Columns 9 shows the mean web searches directed to nonlocal firms by local investors. Column 10 reports the difference between mean web searches directed to local and to nonlocal firms by local investors. Significance code: * p<0.10, ** p<0.05, *** p<0.01.

(1) State name	(2) No. of firms	(3) % of firms	(4) Total searches by local investors	(5) Total searches for local firms	(6) Local searches for local firms	(7) Nonlocal searches for local firms	(8) Local-nonlocal gap: (6) − (7)	(9) Local searches for nonlocal firms	(10) Local-nonlocal gap: (6) − (9)
Alabama	2	0.42	18.89	26.41	40.47	26.12	14.35***	18.80	21.67***
Alaska			15.81					15.81	
Arizona	4	0.83	21.35	8.38	38.10	7.78	30.32***	21.21	16.88***
Arkansas	3	0.63	17.71	9.93	58.00	8.95	49.05***	17.45	40.55***
California	73	15.21	26.38	22.73	41.22	22.35	18.87***	23.71	17.51***
Colorado	10	2.08	22.20	24.54	37.76	24.27	13.5***	21.87	15.89***
Connecticut	14	2.92	23.46	13.77	39.10	13.25	25.85***	22.99	16.11***
Delaware	2	0.42	17.56	13.99	29.53	13.68	15.85***	17.51	12.02***
District of Columbia	1	0.21	22.34	4.31	9.42	4.20	5.22**	22.37	-12.94***
Florida	14	2.92	22.47	18.78	30.66	18.53	12.13***	22.22	8.44***
Georgia	16	3.33	21.17	26.86	37.35	26.64	10.71***	20.61	16.74***
Hawaii			19.17					19.17	
Idaho	1	0.21	16.29	22.33	55.67	21.65	34.03***	16.21	39.46***
Illinois	32	6.67	23.91	19.73	44.23	19.23	25.00***	22.45	21.78***
Indiana	7	1.46	19.83	24.41	45.60	23.97	21.63***	19.45	26.15***
Iowa	2	0.42	18.74	25.45	51.04	24.93	26.11***	18.60	32.44***
Kansas			21.12					21.12	
Kentucky	3	0.63	18.77	15.29	46.63	14.65	31.98***	18.59	28.03***
Louisiana	3	0.63	17.71	11.38	17.19	11.26	5.94***	17.71	-0.52
Maine	1	0.21	17.03	2.12	65.38	0.83	64.55***	16.93	48.45***
Maryland	10	2.08	23.58	12.93	21.69	12.75	8.93***	23.62	-1.93
Massachusetts	21	4.38	25.24	19.62	51.04	18.97	32.07***	24.06	26.98***
Michigan	10	2.08	20.93	25.96	49.76	25.47	24.29***	20.36	29.4***
Minnesota	13	2.71	21.72	24.78	54.34	24.18	30.17***	20.81	33.53***
Mississippi			16.66					16.66	
Missouri	10	2.08	20.31	21.71	41.60	21.30	20.29***	19.85	21.74***
Nebraska	2	0.42	18.33	9.33	69.07	8.11	60.96***	18.12	50.95***

(continues on next page)

Table 2
Summary statistics by search state (continued)

This table shows the distribution of company headquarters across states and the mean of web searches (`GSearch`) by search states and company headquarters' states in 2017. For each of the states listed in column 1, local investors are defined as the residents of that state, and nonlocal investors as those resident in other states; local firms as the companies whose headquarters are located in that state, and nonlocal firms as those whose headquarters are located in other states. Columns 2 and 3 report the number and fraction of company headquarters located in the corresponding state indicated in column 1. Columns 4 and 5 show the mean total web searches made by local investors, and of those directed to local firms, respectively. Columns 6 and 7 show the mean web searches directed to local firms by local and nonlocal investors, respectively. Column 8 shows the difference between mean web searches directed to local firms by local and nonlocal investors. Columns 9 shows the mean web searches directed to nonlocal firms by local investors. Column 10 reports the difference between mean web searches directed to local and to nonlocal firms by local investors. Significance code: * p<0.10, ** p<0.05, *** p<0.01.

(1) State name	(2) No. of firms	(3) % of firms	(4) Total searches by local investors	(5) Total searches for local firms	(6) Local searches for local firms	(7) Nonlocal searches for local firms	(8) Local-nonlocal gap: (6) − (7)	(9) Local searches for nonlocal firms	(10) Local-nonlocal gap: (6) − (9)
Nevada	2	0.42	20.30	5.50	100.00	3.57	96.43***	19.96	80.04***
New Hampshire			19.36					19.36	
New Jersey	18	3.75	25.12	19.97	46.48	19.43	27.04***	24.29	22.19***
New Mexico			17.54					17.54	
New York	62	12.92	27.39	20.36	38.86	19.99	18.87***	25.68	13.18***
North Carolina	13	2.71	22.19	19.69	39.96	19.28	20.69***	21.72	18.25***
North Dakota			14.85					14.85	
Ohio	21	4.38	20.89	24.32	45.07	23.90	21.17***	19.79	25.29***
Oklahoma	4	0.83	18.94	21.73	70.99	20.72	50.26***	18.50	52.48***
Oregon	2	0.42	21.69	16.90	65.08	15.92	49.16***	21.51	43.57***
Pennsylvania	16	3.33	22.71	18.86	48.72	18.25	30.48***	21.81	26.92***
Rhode Island	4	0.83	18.61	37.23	64.44	36.67	27.76***	18.23	46.21***
South Carolina	1	0.21	18.94	10.78	40.98	10.16	30.82***	18.89	22.09***
South Dakota			14.86					14.86	
Tennessee	9	1.88	19.69	19.16	54.48	18.44	36.05***	19.02	35.46***
Texas	36	7.50	22.78	20.02	34.10	19.73	14.37***	21.86	12.24***
Utah	2	0.42	18.93	10.84	71.09	9.61	61.48***	18.71	52.38***
Vermont			15.89					15.89	
Virginia	16	3.33	23.56	13.96	36.86	13.50	23.36***	23.10	13.76***
Washington	12	2.50	22.86	16.11	66.47	15.08	51.39***	21.74	44.73***
West Virginia			16.63					16.63	
Wisconsin	8	1.67	20.50	12.89	38.93	12.36	26.57***	20.19	18.74***
Wyoming			14.11					14.11	
Total	480	100	20.14	20.14	42.90	19.68	23.23***	19.68	23.23***

Table 3
Investor attention, geographic proximity and news

This table shows estimates from the regression of GSearch on geographic variables (SameState or Distance) and control variables. All the regressions in the table are estimated by OLS except for those in column 4, where the specification of column 4 is re-estimated with 2SLS, instrumenting Vol and the interaction SameState×Vol by a volatility measure based on industry index returns and by its interaction with the SameState dummy. In columns 5 and 6 the regression is re-estimated separately for the subsamples with positive and negative stock returns, respectively (based on the dummy NegRet). All variables are described in Section 3. All specifications include week and firm fixed effects among the explanatory variables. Inference is based on unit-cluster standard errors. t-statistics are reported in brackets. Significance code: * p<0.10, ** p<0.05, *** p<0.01.

	(1)	(2)	(3)	(4) IV	(5) NegRet = 0	(6) NegRet = 1
SameState	20.46***		14.89***	14.67***	15.44***	14.17***
	[16.89]		[7.06]	[14.81]	[7.16]	[6.61]
Distance		-0.65***				
		[-7.10]				
News			0.20***	0.21***	0.19**	0.23**
			[2.61]	[5.03]	[2.14]	[2.22]
SameState × News			1.28**	1.09***	1.22**	1.38**
			[2.29]	[9.08]	[2.14]	[2.44]
LocNews			3.29***	3.11***	2.85**	3.83***
			[2.79]	[8.89]	[2.06]	[3.74]
SameState × LocNews			-3.14	-1.99	-4.94	-0.69
			[-0.74]	[-1.06]	[-1.27]	[-0.13]
Vol			5.41***	7.32	3.61*	8.04***
			[3.39]	[1.62]	[1.89]	[3.11]
SameState × Vol			57.24***	104.03**	54.72***	59.11**
			[2.86]	[2.25]	[2.64]	[2.41]
Population	1.99***	2.32***	1.99***	2.04***	1.99***	1.98***
	[17.84]	[20.10]	[17.82]	[111.73]	[17.57]	[17.19]
Age65^{+}	16.05***	8.49***	16.15***	16.61***	15.51***	16.97***
	[3.84]	[2.00]	[3.88]	[14.86]	[3.62]	[3.87]
Income	3.36***	5.93***	3.36***	2.65***	3.17***	3.60***
	[3.79]	[6.27]	[3.79]	[10.65]	[3.56]	[3.78]
Edu	17.86***	14.30***	17.80***	19.84***	17.84***	17.74***
	[7.94]	[6.02]	[7.91]	[29.56]	[7.75]	[7.38]
constant	-53.23***	-80.66***	-53.62***	16.25***	11.90	5.91***
	[-5.66]	[-8.03]	[-5.70]	[6.71]	[1.26]	[377.46]
firm FEs	yes	yes	yes	yes	yes	yes
week FEs	yes	yes	yes	yes	yes	yes
N	1244800	1244800	1244800	1124350	703500	541300
$\bar{R}^2$	0.49	0.48	0.49	0.50	0.49	0.49

Table 4
Short-term changes in attention around out-of-state acquisitions

This table presents the estimated coefficients of regressions in which the dependent variable is the number of web searches regarding the bidder company in its own state (column 1: bidder-home) or in the target company's state (column 2: bidder-other), and those regarding the target company in its own state (column 3: target-home) or in the bidder's state (column 4: target-other) in a 3-weeks event window bracketed around an out-of-state company acquisition in 2017. The $t = -1$, $t = 0$ and $t = +1$ dummy variables take value 1 in the pre-acquisition week, in the acquisition week and in the post-acquisition week, respectively, for the relevant company and set of investors. Their respective coefficients are the estimated changes in attention by the relevant group of investors for the bidder company (in columns 1 and 2) or the target company (in columns 3 and 4) in the pre-acquisition week, in acquisition week and in the post-acquisition week, respectively. The constant measures the estimated average level of attention by the relevant group of investors for the bidder company (in columns 1 and 2) or the target company (in columns 3 and 4). Inference is based on unit-cluster standard errors. t-statistics are reported in brackets. Significance code: * p<0.10, ** p<0.05, *** p<0.01.

	(1)	(2)	(3)	(4)
	bidder-home	bidder-other	target-home	target-other
$t = -1$	2.20	6.47*	21.16***	12.09***
	[0.93]	[1.94]	[4.44]	[3.18]
$t = 0$	2.93	4.83*	8.93***	7.22**
	[1.30]	[1.73]	[2.73]	[2.26]
$t = +1$	4.33*	0.39	9.43***	7.86*
	[1.93]	[0.16]	[2.65]	[1.81]
constant	41.70***	38.82***	35.02***	29.91***
	[15.45]	[12.89]	[21.11]	[15.33]
N	4182	4182	1122	1122

Table 5
Long-term changes in attention around out-of-state acquisitions

This table presents the estimated coefficients of regressions in which the dependent variable is the number of web searches regarding the bidder company in its own state (column 1: bidder-home) or in the target company's state (column 2: bidder-other), and those regarding the target company in its own state (column 3: target-home) or in the bidder's state (column 4: target-other) in a 105-weeks event window bracketed around an out-of-state company acquisition in 2017. The `PostEvent` dummy variable takes value 1 in the year following the acquisition for the relevant company and set of investors. Its coefficient is the estimated change in attention by the relevant group of investors for the bidder company (in columns 1 and 2) or the target company (in columns 3 and 4) in the post-acquisition year, respectively. The constant measures the estimated average level of attention by the relevant group of investors for the bidder company (in columns 1 and 2) or the target company (in columns 3 and 4). Inference is based on unit-cluster standard errors. t-statistics are reported in brackets. Significance code: * p<0.10, ** p<0.05, *** p<0.01.

	(1) bidder-home	(2) bidder-other	(3) target-home	(4) target-other
`PostEvent`	2.02***	1.26*	-6.52***	-1.94***
	[2.90]	[1.79]	[-8.23]	[-4.13]
constant	41.75***	38.95***	35.57***	30.24***
	[21.71]	[18.30]	[31.14]	[22.33]
N	8200	8200	2200	2200

Table 6
Long-term changes in attention around out-of-state acquisitions, by size

This table presents the estimated coefficients of regressions in which the dependent variable is the number of web searches regarding the bidder company in the target company's state, broken down between acquisitions by above-average bidders (column 1: big-bidder-other) and by below-average bidders (column 2: small-bidder-other), and those regarding the target company in its own state, broken down between acquisitions of above-average targets (column 3: big-target-home) and of below-average targets (column 4: small-target-home) in a 52-weeks event window after an out-of-state company acquisition in 2017. The `PostEvent` dummy variable takes value 1 in the year following the acquisition for the relevant company and set of investors. Its coefficient is the estimated changes in attention by the relevant group of investors for the bidder company (in columns 1 and 2) or the target company (in columns 3 and 4) in the post-acquisition year, respectively. The constant measures the estimated average level of attention by the relevant group of investors for the bidder company (in columns 1 and 2) or the target company (in columns 3 and 4). Inference is based on unit-cluster standard errors. t-statistics are reported in brackets. Significance code: * p<0.10, ** p<0.05, *** p<0.01.

	(1) big-bidder-other	(2) small-bidder-other	(3) big-target-home	(4) small-target-home
`PostEvent`	6.44***	0.78	-3.73***	-8.11***
	[5.99]	[1.22]	[-5.99]	[-9.73]
constant	44.23***	38.45***	34.42***	36.23***
	[21.08]	[18.10]	[36.45]	[29.28]
N	700	7500	800	1400

Table 7
COVID-induced shocks to retail investor attention

This table shows estimates from the regression of `GSearch` on distance measures (`SameState`, `Distance` or `AdjDistance`), the dummy `DCovid`, the the corresponding interaction terms and control variables. All the regressions in the table are estimated by OLS. All variables are described in Section 3 and 5.2. All specifications include firm fixed effects among the explanatory variables. Inference is based on unit-cluster standard errors. t-statistics are reported in brackets. Significance code: * p<0.10, ** p<0.05, *** p<0.01.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
`SameState`	19.43***	17.50***					16.73***
	[16.34]	[14.96]					[14.04]
`DCovid`		1.82***	2.20***		2.10***	2.12***	1.82***
		[9.70]	[9.96]		[10.46]	[9.33]	[9.71]
`SameState` × `DCovid`		4.22***					3.40***
		[8.62]					[6.22]
`Distance`			-0.64***			-0.57***	
			[-6.90]			[-5.81]	
`Distance` × `DCovid`			-0.16**			-0.13*	
			[-2.40]			[-1.91]	
`AdjDistance`				-0.30***	-0.26***		
				[-4.09]	[-3.16]		
`AdjDistance` × `DCovid`					-0.18***		
					[-3.33]		
`NFlights`						0.25***	
						[3.47]	
`NFlights` × `DCovid`						0.38***	
						[4.74]	
`LocNews`							1.54***
							[3.20]
`SameState` × `LocNews`							3.17
							[1.58]
`DCovid` × `LocNews`							1.16*
							[1.87]
`SameState` × `DCovid` ×`LocNews`							4.39**
							[2.44]
`Population`	0.60***	0.60***	0.89***	0.87***	0.84***	0.73***	0.56***
	[3.90]	[3.90]	[5.71]	[5.56]	[5.40]	[4.66]	[3.66]
`Age65`$^+$	16.05***	16.05***	16.61***	17.55***	17.66***	17.53***	14.55***
	[3.58]	[3.58]	[3.58]	[3.73]	[3.76]	[3.77]	[3.27]
`Income`	0.33	0.33	1.71***	1.36**	1.36**	1.37**	0.17
	[0.61]	[0.61]	[3.09]	[2.45]	[2.45]	[2.48]	[0.32]
`Edu`	24.33***	24.33***	22.13***	22.41***	22.31***	22.70***	24.37***
	[10.57]	[10.57]	[9.50]	[9.59]	[9.55]	[9.68]	[10.62]
Constant	-1.56	-2.4	-19.90***	-16.06**	-16.61**	-14.39**	-0.01
	[-0.24]	[-0.37]	[-2.93]	[-2.34]	[-2.43]	[-2.13]	[-0.00]
Firm Fes	yes	yes	yes	yes	yes	yes	yes
N	560541	560541	560541	560541	560541	560541	560541
$\bar{R}^2$	0.49	0.49	0.48	0.48	0.48	0.48	0.49

Table 8
Local and nonlocal attention, return volatility and bid-ask spread

This table shows GMM estimates of panel VAR models whose endogenous variables are the number of weekly web searches by residents in the relevant company's state (`LocalSearch`) and by residents in other states (`NonLocalSearch`), the absolute value of the relevant stock return (`Vol`) and its average relative bid-ask spread (`Spread`) over the relevant week, using 2017 data. Inference is based on robust standard errors. z-statistics are reported in brackets. Significance code: * p<0.10, ** p<0.05, *** p<0.01.

	(1)	(2)	(3)
`LocalSearch` equation			
$\texttt{LocalSearch}_{t-1}$	0.11***	0.11***	0.12***
	[3.84]	[3.85]	[3.90]
$\texttt{NonLocalSearch}_{t-1}$	0.02***	0.02***	0.02***
	[2.68]	[2.66]	[2.78]
$\texttt{Vol}_{t-1}$		0.00	
		[0.06]	
$\texttt{Spread}_{t-1}$			0.00
			[0.28]
`NonLocalSearch` equation			
$\texttt{LocalSearch}_{t-1}$	0.23***	0.23***	0.23***
	[5.00]	[5.02]	[4.62]
$\texttt{NonLocalSearch}_{t-1}$	0.40***	0.40***	0.41***
	[9.48]	[9.47]	[8.61]
$\texttt{Vol}_{t-1}$		-0.00	
		[-0.78]	
$\texttt{Spread}_{t-1}$			0.00
			[0.36]
`Vol` equation			
$\texttt{LocalSearch}_{t-1}$		0.96***	
		[5.10]	
$\texttt{NonLocalSearch}_{t-1}$		0.63***	
		[5.32]	
$\texttt{Vol}_{t-1}$		-0.00	
		[-0.30]	
`Spread` equation			
$\texttt{LocalSearch}_{t-1}$			0.26***
			[4.98]
$\texttt{NonLocalSearch}_{t-1}$			0.34***
			[5.19]
$\texttt{Spread}_{t-1}$			0.36***
			[6.71]
N	23531	23531	18646