The disclosure function of the U.S. patent system: evidence from the PTDL program and extreme snowfall

Tim Martens[1]

Abstract

Are retail investors using uncurated disclosures in form of patents for their investment decisions? This study uses the investment decisions of retail investors and variation in the local availability of patent information to answer this question. The variation comes from changes in the locations of U.S. Patent and Trademark Depository Libraries over time. I find a strong increase in the local trading volume of stocks after the release of a patent in counties that have easier access to patent information. In addition, trades made by retail investors with easier access to this information yield higher returns, compared to trades made by other investors. These results indicate that disclosures of the U.S. Patent and Trademark Office facilitate the dissemination of patent information to retail investors. Furthermore, these results suggest that retail investors complement traditional curated disclosures with uncurated disclosures in form of patents.

> Instead of concentrating on the metrics in financial statements, [Herbert] Wertheim is devoted to reading patents and spends two six-hour blocks each week poring over technical tomes. 'What's more important to me is, what is your intellectual capital to be able to grow?' Thanks to his engineering background, the technical nature of optometry and his experience as an inventor, the patent library is Wertheim's comfort zone. Stocks he invested in based on their impressive patent portfolios include IBM, 3M and Intel.
>
> Forbes (2019)

✉ Tim Martens
  tim.martens@unibocconi.it

1   Bocconi University, Milan, Italy

1 Introduction

Are retail investors using uncurated disclosures in form of patents for their investment decisions? While the literature examines retail investors' use of curated disclosures, such as financial statements, their use of uncurated disclosures remains largely unstudied. Whether retail investors use uncurated disclosures matters for regulators who aim to maintain equal access to public information. It also matters for managers, who may opt to use uncurated disclosures when communicating with investors. This study uses variation in the local availability of patent information and local investors' investment decisions to address how retail investors use patent disclosures.

Patents are uncurated disclosures published by the US Patent and Trademark Office (USPTO), and the information in them is not readily available from financial statements or other curated firm disclosures. However, patents are important in that they contain value-relevant information about the firm that is used by sophisticated market participants, such as institutional investors and financial analysts (e.g., Hirschey and Richardson 2004; Hochberg et al. 2018; Martens and Sextroh 2021).

Despite this value-relevance, it is unclear whether retail investors also use this information for their investment decisions. They face obstacles in doing so. The processing costs of patents, due to their technical nature, are higher than those of other sorts of information. On top of this, information that is harder to interpret is generally more likely to be ignored by less sophisticated investors (Bloomfield 2002). Hence retail investors may view patent information as too costly to be worth their time and effort (Blankespoor et al. 2019; Christensen et al. 2017; Hirshleifer and Teoh 2003; Huberman and Regev 2001; Merton 1987). As a consequence, they might not collect and use the information, despite its potential relevance for their investment decisions.

To examine whether retail investors do use patent information, I exploit features of the pre-Internet patent system in the United States from 1991 to 1996. In 1870, the USPTO started distributing copies of patent documents to libraries across the United States to broadly disseminate the information to the public. These Patent and Trademark Depository Libraries (PTDL) offer access to all resources necessary to conduct a full patent and trademark search. In 1977, the USPTO began to expand the PTDL program by designating at least one PTDL in every state. I use PTDL openings over time as a shock to the local availability of patent information.

To rule out that other determinants drive my results, I examine a second source of variation in local access to patent information that cannot plausibly relate to other determinants of trading: extreme snowfall, which inhibits access to the PTDL. During my sample period, local investors depended on physical access to the patent information. So I collect data on the daily weather in each county in my sample and identify snowfalls severe enough to disrupt or delay access to patent information. A large snowfall does not relate to either the USPTO's decision to open a PTDL or any other local characteristic affecting local trading volume, but it would have inhibited access to patent libraries.

I find that, all else equal, the existence of a PTDL increases the trading volume of local retail investors after the release of a patent by 4.6%, the number of local

trades by 0.4%, and the probability of a local trade by 0.5%. This positive association is stronger for high-value patent information, and it breaks down on extremely snowy days. Overall, these results strongly suggest that retail investors use patent information to inform their trading.

I also find that the retail investor reaction is stronger for firms with strong information environments and low trading costs but weaker on days with lower investor attention. These results suggest that retail investors use uncurated disclosures in form of patents to complement curated financial statements and analyst reports.

Finally, I show that the access to patent information increases the returns of trades made by retail investors between 0.7% and 7.9%, compared to trades of other retail investors. This increase indicates that retail investors can understand the quality of the patents and use the information accordingly.

This study contributes to the literature examining the investment decisions of retail investors by documenting that they can collect and understand uncurated disclosures. Retail investors typically have fewer resources available for this sort of collection and analysis than do other stakeholders. As a result, any stakeholder with more resources can be inferred to likewise be able to use these uncurated disclosures (e.g., Glaeser 2018; Glaeser and Landsman 2020; Hegde et al. 2018; Kim and Valentine 2020; Saidi and Žaldokas 2020). Consequently, my results are relevant for regulators concerned with maintaining equal access to public information and who are charged with protecting retail investors.[1]

This study also contributes to the literature on the characteristics of alternative disclosures. By doing so, it addresses the call for research by Leuz and Wysocki (2016) to examine nontraditional disclosures. In this literature, few studies have examined the U.S. patent system (e.g., Glaeser and Landsman 2020; Glaeser et al. 2020). This study extends this literature by showing that firms can communicate with retail investors through the U.S. patent system.

The paper proceeds as follows. Section 2 describes the data and sample construction of this study. Section 3 contains the model and the results. Section 4 discusses of the generalizability of the results. Section 5 concludes.

2 Data and sample construction

I use the Patent and Trademark Depository Library program as a source of variation in the availability of patent information. Until 1870, patent documents were only located at the USPTO in Washington, D.C. In 1870, the USPTO started distributing copies of its documents to libraries across the United States for the sake of public dissemination. These Patent and Trademark Depository Libraries (PTDL) offer access to all resources necessary to conduct a full patent and trademark search. To ensure the efficient use of its resources, the USPTO targeted areas with large populations and with high patent and trademark activity (Sneed 2000). The information about the

---

[1]A former SEC chairperson, Mary Jo White, has highlighted that the SEC should focus on protecting retail investors: https://www.sec.gov/news/speech/mjw-speech-032114-protecting-retail-investor.

selection process allows me to control for the selection mechanism of the USPTO and use the PTDL openings as a shock to local patent information.[2]

The standards governing the PTDL program allowed it to foster innovation via information transfers between inventors (Furman et al. 2020). Additionally, these standards continue to ensure that investors and financial analysts have access to patent information (USPTO 2003). For example, Brown and Arshem (1993) show that 6% of the PTDL visitors used the patent information for economic research, while 17% used it for legal, product, and market research.

The PTDL system also allows retail investors to use other patents to assess the value of individual patents or patent portfolios. For example, the citation patterns included in the patent document can be used to determine the value of the innovation (Fitzgerald et al. 2019). Therefore the opening of a PTDL is not identical to the simple disclosure of current patent information by firms or information intermediaries.

Since patents and trademarks are legal instruments and applying for either is a matter of law, the PTDL staff, as opposed to attorneys, are not allowed to provide any legal advice concerning patent issues. They can only explain the application process, show how to search for a patent or trademark, and answer questions about searching. They are not allowed to execute the patent or trademark search, submit or process a patent application, or give advice on patent searches or on applications. This restriction ensures that the PTDL program does not affect any other local characteristics directly. Hence I use the openings of PTDL across time to study direct information transfer between the USPTO and investors.

The primary data source of this study is the discount brokerage database of Barber and Odean (2000), which contains a local trading database and a local portfolio database with data from 1991 to 1996. For my analysis, I use the local trading database, which contains the trades of almost 78,000 retail investors as well as their geographic and demographic information. I remove all accounts that do not have a valid zip code, which is needed to identify the investor locations. I aggregate the local trading database to the county-date level to calculate the local trading volume, which I use as the dependent variable. I include all counties that have shown any trading in firm j over the sample period as observations at time t in my sample. Hence county-date observations can have zero local trading volume.

I use the patent database of Kogan et al. (2017), which includes all patents filed with the USPTO from 1926 to 2010. I remove all observations that do not have information on the patents' technology class, citations, or economic value. The patent database has to be aggregated to the firm-date level, since a firm can receive several patents on the same date, which I refer to as the release of a patent.[3]

I merge the local-trading and patent databases with data from Compustat and Center for Research in Security Prices (CRSP). I supplement this data with the PTDL opening dates and locations from 1870 to 1996. The opening dates and locations are

---

[2]Please refer to Sneed (2000) for more information on the PTDL program.

[3]The release date/publication date and the grant date/issue date are identical over my sample period. Before the introduction of AIPA in 2000, patent information was only published at the grant date (Graham and Hegde 2015).

collected from Oliver (2002).[4] Table 2 shows the opening dates and locations of the PTDL in my sample.

I further supplement this data with population data from the Census State and County Intercensal database and snowfall data from the National Climatic Data Center GHCN-D database. The GHCN-D database contains weather data, including high and low temperatures, inches of snowfall, and precipitation. I calculate the average snowfall for all weather stations that are within a 10-mile radius of the counties in my sample. I add the number analyst earnings forecasts from the IBES summary database, the institutional ownership from the WRDS Thomson Reuters Institutional Holdings database, and M&A news from the Capital IQ Key Developments database.

The sample consists of 1,472 unique firms that experience 39,419 unique patent events (release of a patent). It further consists of 1,378 unique counties that account for 4,470,927 unique observations. Limiting the sample period to 1991 to 1996 ensures robust inferences, since data after 1996 is subject to significant changes to institutions (e.g., American Inventors Protection Act) and technology (e.g., online databases). The sample period is largely pre-Internet, implying that investors must rely on the PTDL to access patent documents.[5]

3 Methodology and estimation

To examine the information transfer between the USPTO and retail investors, I estimate the following regression:

$$Log(1 + Local\ trading_{i,j,t})|Patent_{j,t} = \beta_0 + \beta_1 Libraries_{i,t}$$
$$+ \sum Controls + \sum County_i + \sum Firm_j\ x\ Date_t \qquad (1)$$

where $Local\ trading$ is the $Local\ trading\ volume$ of retail investors in county i to firm j's release of a patent at time t. The variable is constructed over a three-day window; that is, patents are released on Tuesdays in my sample period and local trading volume accounts for trading from Tuesday to Thursday. All patents in my sample period are released on Tuesdays. $Libraries$ represents an indicator variable that takes the value of one if there is a PTDL in county i at time t and zero otherwise. The coefficient on $Libraries$ represents the trading volume after the release of a patent in counties with a PTDL, in contrast to those without one. I assume that investors who receive the patent information react similarly to the release of a patent across locations, since the information is identical in all locations. In other words,

---

[4]I adjust the opening date of the New Haven Free Public Library to August 11, 1986, since it replaced the Science Park Library, which opened at that date.

[5]The PTDL system was not the only source of patent information for investors. The USPTO started the AIDS online database, which contained AIDS-related patent information, on December 1, 1994, and the PatBib online database, which contained patent information from 1986 to 1995, on November 9, 1995. However, the PatBib database started to include the latest patents only on February 1, 1996. Furthermore, the sample period is not entirely pre-Internet. Barber and Odean (2002) show that 1,607 households in the local trading database switched from phone-based to Internet-based trading in the course of the sample period.

investors in Santa Clara County and New York County would react similarly. I also
assume investors near Boston are more likely to use the resources at the Boston Pub-
lic Library than those at other libraries. This assumption comports with the findings
of Brown and Arshem (1993), who show that most visitors originate from the area
near the PTDL. In addition, the setting requires that only a subset of local investors
use the PTDL. These investors then disseminate their information, ensuring local
transfer (Ivković and Weisbenner 2007; Hong et al. 2005).

I include four county characteristics to control for the decision to open a PTDL in
a county. (1) I include $Log(1 + County\ patents)$, which is defined as the natural
logarithm of one plus the the number of patents in thousands released in county i in
the year of the event date. (2) I include $Log(1 + County\ scientific\ value)$, which
is defined as the natural logarithm of one plus the adjusted number of citations in
thousands of patents released in county i in the year of the event date. (3) I include
$Log(1 + County\ economic\ value)$, which is defined as the natural logarithm of one
plus the economic value in billions of patents released in county i in the year of the
event date. (4) I include $Log(County\ population)$, which is defined as the natural
logarithm of the population in thousands in county i in the year of the event date. I
also include $Local\ firm$, which is an indicator variable that takes the value of one
if the headquarters of firm j is located in county i and zero otherwise. Please refer to
Appendix A for descriptions of all variables.

I include county fixed effects to control for county time-invariant characteristics
(e.g., larger counties might cause stronger reactions than smaller ones). I include firm
x date fixed effects to control for all fundamentals of the event (e.g., the scientific and
economic value of the patent) and any simultaneous actions (e.g., positive or negative
media coverage). By clustering residuals by county, I compute standard errors that
allow each county to have its own unobserved effect on trading volume (Petersen
2009).

3.1 Descriptive statistics

Table 3 reports the descriptive statistics for the variables in this analysis. 867,034
(19.4%) observations are in counties with a PTDL, while 3,603,893 (80.6%) are in
counties without one. In other words, for each observation with a PTDL, I observe
on average four observations without one, which allows me to vary the patent infor-
mation availability for different groups of investors for the same patent release. The
mean of the $Local\ trading\ volume$ is substantially higher in counties with a PTDL
than in counties without one ($550.25 versus $205.07). Furthermore, counties with a
PTDL have on average a larger population and generate more patents. These patents
have higher scientific and economic values.

3.2 PTDL and local trading volume

In Table 4, I examine the association between patent information availability and
local trading volume. In columns (1) to (3), I show the estimates of Eq. 1 with

the natural logarithm of one plus the absolute dollar trading volume in firm j's stock in county i within three days of the patent release as the dependent variable. Column (1) considers only the treatment variable *Libraries* and standard errors clustered by county. The result of column (1) suggests that investors in counties with a PTDL increase their local trading volume by 12.9%, in contrast to investors in counties without a PTDL. To ensure that these results are not driven by the USPTO selection mechanism, I include the variables $Log(1 + County\ patents)$, $Log(1 + County\ scientific\ value)$, $Log(1 + County\ economic\ value)$, and $Log(County\ population)$ as control variables in column (2) (Sneed 2000). These variables mimic the USPTO selection mechanism and ensure an unbiased coefficient estimate for the treatment variable *Libraries*. Furthermore, to ensure that the results are not driven by any simultaneous action that materializes in counties with a PTDL, due to the location choice of firms, I include the variable *Local firm*. Firms might select counties with a PTDL as a location for their headquarters to benefit from resources at the library. If any simultaneous action is correlated with the firm's location choice, the coefficient estimate for the treatment variable *Libraries* would be biased. The variable also accounts for any home bias that might affect local trading (Ivković and Weisbenner 2005; Seasholes and Zhu 2010). And column (2) includes county fixed effects to control for county time-invariant characteristics. The results of column (2) indicate that the existence of a PTDL increases local trading volume by 5.5%, versus counties without a PTDL. However, this specification ignores that some events might be associated with higher trading volumes for reasons unrelated to the PTDL. Hence, to control for these factors, column (3) adds firm x date fixed effects to the model. The inclusion of these fixed effects decreases the coefficient estimates for *Libraries* to 4.6%. The coefficient estimates for *Libraries* for these specifications are positive and significant at the $p \leq 0.1$ level. The adjusted R^2 increases from 0.2% in column (1) to 3.7% in column (3), highlighting the importance of these characteristics in this analysis.

To ensure that the results are not driven by outliers, columns (4) and (5) present the coefficient estimates for *Libraries* for two alternative dependent variables that treat each observation equally. Column (4) uses the natural logarithm of one plus the number of trades in firm j's stock in county i within three days of patent release. Column (5) uses a local trading indicator variable that takes the value one if there was trading in firm j's stock in county i within three days of patent release and zero otherwise. I use a linear probability model to account for the many fixed effects in this analysis. Both dependent variables minimize the influence of large or small trades on the estimation. The existence of a PTDL increases the number of local trades by 0.4% and the probability of a local trade by 0.5%. In addition, the coefficient estimates for the treatment variable *Libraries* are significant at the $p \leq 0.1$ level, indicating the results of the baseline specification are not driven by outliers.

The level of significance is due to the fact that the coefficient estimates represent the average effect across low-value and high-value patents. However, the retail investor reaction is driven by the value of the patent information and not by the PTDL itself. In Table 5, I show greater significance, due to the interaction with the scientific value in the analysis.

A frequently stated concern with difference-in-difference research designs is the violation of the parallel trends assumption. To address this concern, I define leads ($t - 4$ to $t - 2$) and lags ($t = 0$ to $t + 2$) of the treatment variable *Libraries* as half-year increments and use them as treatment variables.[6] For example, *Libraries* $t - 2$ represents an indicator variable that takes the value of one a year before the opening of a PTDL and zero otherwise. Similarly, *Libraries* $t + 2$ is an indicator variable that takes the value of one a year after the opening of a PTDL and zero otherwise. The baseline period in this analysis is the period $t - 1$. In Fig. 1, I examine the coefficient estimates before and after the opening of a PTDL to provide evidence of the validity of the parallel trends assumption. The coefficient estimates for the leads of the treatment variables *Libraries* are not statistically significant at the $p \leq 0.1$ level, highlighting that there are no systematic differences in local trading volume between both groups of counties before the opening of a PTDL.

The analysis only considers short-term market reactions after the release of a patent to ensure a clean identification of the information transfer. As a result, the size of the coefficient estimates for *Libraries* is not equivalent to the total economic effect of the information transfer. In fact, the results of this analysis likely constitute only a fraction of the total economic effect, since patent information is also used in licensing agreements and debt contracts, which might be affected by the information transfer (Mann 2018). In addition, these information transfers might influence the efficiency of stock price discovery (Hegde et al. 2018). In sum, the results of my analysis underline the existence of an information transfer between the USPTO and retail investors. These results also highlight that retail investors collect patent information and use it in their investment decisions.

3.3 Pre-release local trading volume, scientific value, and snow

To ensure that the results of the previous analysis are not driven by factors unrelated to an information transfer between the USPTO and retail investors, I examine the financial market reaction before the release of a patent. The specification of the previous analysis would be questionable if the existence of a PTDL were associated with the retail investor reaction before the release of a patent. The dependent variables in columns (1) and (2) of Table 5 are considered on days -3, -2, and -1 before the patent release date. The coefficient estimates for *Libraries* are not statistically significant at the $p \leq 0.1$ level for both specifications, suggesting that any increase in the local trading volume in the previous analysis is not driven by any factors unrelated to the release of a patent and hence unrelated to the information transfer.

In addition, if the results of the previous analysis are driven by an information transfer between the USPTO and retail investors, I expect the value of the information to drive the retail investor reaction. Consequently, columns (3) and (4) of Table 5 show the coefficient estimates for the treatment variable *Libraries* and its

---

[6]I also include the lead $t < -4$ and the lag $t > 2$, which are not shown in the graph since they have no meaningful interpretation.

interaction with the scientific value $Log(1 + Scientific\ value)$. I define $Log(1 + Scientific\ value)$ the natural logarithm of one plus the number of citations of the patent event. I follow prior research and deduct from the citation count the mean number of citations for all patents in the same technology class in the same year. This adjustment accounts for systematic differences between patents in different technology classes in different years (Hall et al. 2001). To ensure nonnegative values, I set the adjusted citations to zero for patents that have adjusted citations below zero.[7] The coefficient estimates for $Libraries$ are not statistically significant at the $p \leq 0.1$ level; however, the coefficient estimates for the interaction are significant at the $p \leq 0.01$ level for both specifications. The result of column (3) indicates that a 1% increase in the scientific value of a patent increases the local trading volume by 3.1% if a PTDL exists in county i at time t. A similar association is observable in column (4) for the number of local trades. These results suggest that the investors are not reacting to the opening of a PTDL; they are reacting to the release of a patent and the value of the information.

To rule out that the results of the previous analysis are driven by the USPTO selection mechanism, I follow the approach of Engelberg and Parsons (2011) and use extreme snowfall as a exogenous shock to patent information availability. I exploit the fact that the sample period is largely pre-Internet and investors must physically access patent information if they want to use it. Extreme snowfall is unrelated to the USPTO selection mechanism; however, it can prevent patent information from reaching investors by, for example, blocking roads. Columns (5) to (6) of Table 5 include the coefficient estimates for the variable $Snow$ and its interaction with the treatment variable $Libraries$. The variable is an indicator variable that takes the value of one if the snowfall in county i within three days after patent release exceeds 20 inches and zero otherwise.

In this setting, the coefficient estimates for $Libraries$ represent the effect of a PTDL on local trading volume on days without extreme snowfall. The inclusion of the interaction of $Libraries$ and $Snow$ does not substantially change the coefficient estimate for the treatment variable $Libraries$, compared to the baseline regression. Furthermore, the lack of statistical significance of the coefficient estimate for $Snow$ indicates that extreme snowfall itself does not influence local trading volume significantly. The variable of interest is the interaction of $Libraries$ and $Snow$, which is negative and significant at the $p \leq 0.05$ level. These coefficient estimates indicate that information transfer breaks down on days with extreme snowfall, suggesting the existence of a transfer between the USPTO and retail investors. It is noteworthy that the interaction variable offsets the positive effect of the PTDL (i.e., the sum of these coefficients is not statistically different from zero at the $p \leq 0.1$ level), which means that the information transfer is severely disrupted.[8] Overall, these results ensure that the results of the previous analysis are not driven by the USPTO selection mechanism but by an information transfer between the USPTO and retail investors.

---

[7]The economic value of the patent, measured through the aggregate market reaction, cannot be included as a control variable because the aggregate market reaction is mechanically affected by the local market reaction and hence would introduce endogeneity into the model.

[8]The results are not sensitive to various alternative definitions of the variable $Snow$ (e.g., 22 inches).

3.4 Characteristics of the information transfer

In Table 6, I examine the characteristics of the information transfer. In columns (1) to (5), I use the market reactions on the individual days after the release of a patent to measure the speed of the transfer. The coefficient estimates indicate that the transfer is relatively fast, since the strongest market reactions occur on day 0 (release date) and day 1 after the release ($p \leq 0.1$). The existence of a PTDL increases local trading volume by 1.6% on day 0 and by 2.5% on day 1. Market reactions from day 2 to day 4 are weaker, suggesting that the retail investors incorporate the patent information quickly into the stock price. This finding comports with the results of Kogan et al. (2017), who demonstrate that the strongest aggregate financial market reactions occur from day 0 to day 2. The speed of the market reactions also mitigates the concern that the results of the previous analysis are driven by local news coverage, since this would require patent information to be disseminated on release day, which is unlikely, given the existence of print deadlines.

The literature has shown that distance is an important determinant of information transfers (e.g., Belenzon and Schankerman 2013; Agrawal et al. 2017). I follow this literature and examine how distance affects the information transfer in my setting. In column (6) of Table 6, I include the variable $Log(Distance)$, which is defined as the natural logarithm of the distance of investors in county i to the closest PTDL at time t in miles.[9] The coefficient estimates for $Log(Distance)$ are negative and statistically significant at the $p \leq 0.05$ level. The result in column (6) indicates that a 1% increase in the distance to the closest PTDL decreases the local trading volume by 1.3%, suggesting that distance does affect the information transfer significantly. A possible explanation for this result is that the patent information needs time to reach investors farther from the PTDL.

3.5 Investor characteristics

In Table 7, I focus on the characteristics of different groups of retail investors. The results of this analysis should be viewed as descriptive and interpreted with caution. They do not rule out alternative explanations, and a large share of observations lack necessary data.

In columns (1) to (4), I examine whether investors with different jobs differ in their sensitivity to the availability of patent information. Columns (1) and (2) include the coefficient estimates for $Libraries$ for investors in management jobs and technical jobs, respectively. Both coefficient estimates are not statistically significant

---

[9]The distance to the closest PTDL is strictly larger than zero because the measure is based on PTDL zip codes, which differ from county zip codes. Hence even investors in a county with a PTDL exhibit a very small but positive distance.

at the $p \leq 0.1$ level. Columns (3) and (4) include the coefficient estimates for *Libraries* for retired investors and the remaining investors, respectively. The coefficient estimate for the retired investors is significant at the $p \leq 0.01$ level, while it is not statistically significant at the $p \leq 0.1$ level for the remaining investors. A possible explanation for these results is that retail investors have sufficient knowledge to process patent information, independent of their jobs, but that collecting and understanding patent information requires adequate time. However, the insignificant coefficient estimates should be interpreted with caution, as they are marginally significant and might be the result of a lack of statistical power.

In columns (5) and (6), I examine whether females and males differ in their use of patent information. The coefficient estimate for *Libraries* for female investors is not statistically significant at the $p \leq 0.01$ level, while it is at the $p \leq 0.05$ level for male investors. A possible explanation for the difference is differing levels of confidence in the precision of the information (Barber and Odean 2001).

3.6 Information environment, trading costs, and attention

In columns (1) and (2) of Table 8, I examine whether the information environment affects the retail investors and their use of patent information. I measure the information environment using two variables that are positively associated with a strong information environment: analyst following and institutional ownership (Bhushan 1989; Bushee and Noe 2000; Shores 1990). I define $Log(1 + Analyst\ following)$ as the natural logarithm of one plus the number of analyst earnings forecasts for firm i before the release of the patent. I define *Institutional ownership* as the percentage of firms j's shares held by institutional investors before the release of the patent. I expect that retail investors react more strongly when the information environment is stronger and hence search costs are lower. Columns (1) and (2) include the coefficient estimates for the interactions of *Libraries* with $Log(1 + Analyst\ following)$ and *Institutional ownership*, respectively. Both coefficients are positive and statistically significant at the $p \leq 0.01$ level, consistent with the argument that retail investors complement traditional curated disclosures, such as financial statements and analyst reports, with uncurated patent disclosures.

In columns (3) and (4), I examine whether trading costs affect the retail investors. I am using two measures of liquidity that are associated with trading costs: turnover and bid-ask spread. I define *Turnover* as the ratio of daily volume and shares outstanding the day before the release of the patent. Similarly, I define $Bid - ask$ $spread$ as the bid-ask spread the day before the release of the patent. The bid-ask spread is calculated using the Corwin and Schultz (2012) estimator. I expect that retail investors react more strongly to patent information when liquidity is high and trading costs are low. Columns (3) and (4) include the coefficient estimates for the interactions of *Libraries* with *Turnover* and $Bid - ask\ spread$, respectively. Both coefficients are statistically significant at the $p \leq 0.05$ level, consistent with the argument that low trading costs make trades more attractive for retail investors.

Finally, in columns (5) and (6), I examine whether the intensity of attention affects retail investors and their use of patent information. I use the amount of competing information and a salient corporate event as proxies for attention. I define $Log(Daily\ patents)$ as the natural logarithm of the number of patents released on day t. I define $M\&A\ news$ as an indicator variable that takes the value of one if there is any M&A news within the seven days before time t and zero otherwise. Since stocks in the news grab the attention of retail investors, I expect these investors will react more to patent information if there is less information competing for their attention and if there is M&A news, increasing attention on the pertinent firm (Barber and Odean 2008). Columns (5) and (6) include the coefficient estimates for the interactions of $Libraries$ with $Log(Daily\ patents)$ and $M\&A\ news$, respectively. The negative coefficient $Log(Daily\ patents)$ ($p \leq 0.05$) and the positive coefficient on $M\&A\ news$ ($p \leq 0.01$) are consistent with argument that attention influences the response of retail investors.

3.7 PTDL and profitability

In Table 9, I examine whether trades made by retail investors with access to patent documents are more profitable than those made by retail investors without access. I use a sample of individual trading days with buying activity, which allows me to accurately calculate the returns. I calculate the cumulative abnormal returns (CAR) and buy-and-hold abnormal returns (BHAR) for 20, 60, and 240 trading days. I use a market-adjusted model, which uses abnormal returns, defined as returns in excess of the CRSP value-weighted market return, to calculate the abnormal returns. I estimate a regression with the CAR and BHAR as dependent variables to determine the influence of patent information on profitability. Columns (1) to (6) include coefficient estimates for $Libraries$ across event windows. The coefficient on $Libraries$ represents the difference in returns for trades made by retail investors in counties with a PTDL, compared to trades made by retail investors in counties without a PTDL. The coefficient estimates for $Libraries$ are significant at the $p \leq 0.05$ level across all specifications. The return differentials range from 0.7% to 7.9%, depending on the model and the event window. These return differentials translate into returns of 115 and 1,293 dollars, respectively. This result suggests that retail investors collect and use patent information, since availability of this information is associated with significantly higher stock returns.

4 Discussion

To understand the generalizability of my results, it is necessary to answer the question whether these results transfer to the present. While the introduction of the Internet has made the PTDL program obsolete, the features of my setting still allow a discussion of generalizability.

The PTDL program as a disclosure channel is straightforward: it only offers access to patent information, and the information itself is not affected by the PTDL program. Hence any other disclosure channel that would offer access to current and historical patent documents to retail investors would most likely affect retail investors in a similar way. In addition, the format of patent documents did not change significantly over time. This simplicity and consistency suggest that today retail investors can collect and use patent information provided through other channels.

While investors today are confronted with a large body of information, it is unclear whether information overload has worsened over time, since firms try to address the issue by adjusting their disclosures (e.g., Chapman et al. 2019; EY 2014). Today retail investors have many tools to facilitate their collection and analysis of patent information (e.g., Google Patents and PatentsView). They also have many means of sharing information with each others (e.g., social media such as Reddit and Twitter). Furthermore, individual retail investors share their information with other retail investors (Ivković and Weisbenner 2007; Shiller and Pound 1989). Similarly, the transfer of technological information is strongly geographically localized (Belenzon and Schankerman 2013; Jaffe et al. 1993). These features suggest that the results of this study do transfer to today.

In addition, if retail investors can understand patent information, institutional investors, which are generally considered more sophisticated, should also be able to (Ayers et al. 2011). This conclusion is consistent with the growing stream of literature that highlights the value of patent information for investors (e.g., Hirshleifer et al. 2013, 2018; Hochberg et al. 2018).

5 Conclusion

This study uses the investment decisions of retail investors and variation in the local availability of patent information to determine whether retail investors use uncurated disclosures in form of patents. I document a positive association between the local availability of patent information and local trading volume after the release of a patent. All else equal, the existence of a PTDL increases the trading volume of local retail investors after the release of a patent by 4.6%, the number of local trades by 0.4%, and the probability of a local trade by 0.5%. The effect is stronger for high value information and for investors closer to the patent libraries. This association breaks down on days with extreme snowfall, indicating an information transfer between the USPTO and retail investors. Furthermore, I show that the response of retail investors is stronger for firms with strong information environments and low trading costs but weaker on days with lower investor attention. Finally, I show that the existence of a PTDL increases the return of trades made by retail investors between 0.7 and 7.9%. These results indicate that the disclosures of the USPTO facilitate the dissemination of patent information to retail investors. Furthermore, these results suggest that retail investors complement curated disclosures with uncurated ones in form of patents.

Appendix A: Variable descriptions

Table 1 Variable definitions

Trading-related variables	
Log(1 + Local trading volume	The natural logarithm of one plus the local dollar trading volume of retail investors in county i to firm j's release of a patent at time t. The variable is constructed over a three day window; i.e., patents are released on Tuesdays in my sample period, and local trading volume accounts for trading volume from Tuesday to Thursday.
Log(1 + Number of local trades	The natural logarithm of one plus the number of trades of retail investors in county i to firm j's release of a patent at time t. The variable is constructed over a three day window; i.e., patents are released on Tuesdays in my sample period, and number of local trades accounts for trades from Tuesday to Thursday.
Local trading indicator	An indicator variable that takes the value one if there was trading in firm j's stock in county i to firm j's release of a patent at time t and zero otherwise. The variable is constructed over a three day window; i.e., patents are released on Tuesdays in my sample period, and the local trading indicator accounts for trading activity from Tuesday to Thursday.
Log(1 + Local management trading volume	The natural logarithm of one plus the local dollar trading volume of retail investors in management jobs in county i to firm j's release of a patent at time t. The variable is constructed over a three day window; i.e., patents are released on Tuesdays in my sample period, and local trading volume accounts for trading volume from Tuesday to Thursday.
Log(1 + Local technical trading volume	The natural logarithm of one plus the local dollar trading volume of retail investors in technical jobs in county i to firm j's release of a patent at time t. The variable is constructed over a three day window; i.e., patents are released on Tuesdays in my sample period, and local trading volume accounts for trading volume from Tuesday to Thursday.
Log(1 + Local retiredtrading volume	The natural logarithm of one plus the local dollar trading volume of retired retail investors in county i to firm j's release of a patent at time t. The variable is constructed over a three day window; i.e., patents are released on Tuesdays in my sample period, and local trading volume accounts for trading volume from Tuesday to Thursday.
Log(1 + Local other trading volume	The natural logarithm of one plus the local dollar trading volume of retail investors who are not in management jobs or in technical jobs and who are not retired in county i to firm j's release of a patent at time t. The variable is constructed over a three day window; i.e., patents are released on Tuesdays in my sample period, and local trading volume accounts for trading volume from Tuesday to Thursday.

Table 1 (continued)

CAR	The cumulative abnormal return over the event window to firm j's release of a patent at time t. The cumulative abnormal return is calculated using a market-adjusted model that uses abnormal returns defined in excess of the CRSP value-weighted market return assuming a market beta of 1.
BHAR	The buy-and-hold abnormal return over the event window to firm j's release of a patent at time t. The buy-and-hold abnormal return is calculated using a market-adjusted model that uses abnormal returns defined in excess of the CRSP value- weighted market return assuming a market beta of 1.
County-specific variables	
Libraries	An indicator variable that takes the value of one if there is a PTDL in county i at time t and zero otherwise.
Log(1 + County patents)	The natural logarithm of one plus the the number of patents in thousands released in county i in the year of the event date i.
Log(1 + County scientific value)	The natural logarithm of one plus the adjusted number of citations in thousands of patents released in county i in the year of the event date. The mean number of citations for all patents in the same technology class in the same year is deducted from the citation count. The adjusted citations are set to zero for patents that have adjusted citations below zero to ensure nonnegative values.
Log(1 + County economic value)	The natural logarithm of one plus the economic value in millions of patents released in county i in the year of the event date. The adjusted economic value is the estimated economic value from Kogan et al. (2017) adjusted by the mean economic value for all patents in the same technology class in the same year.
Log(County population)	The natural logarithm of the population in thousands in county i in the year of the event date.
Log(Distance)	The natural logarithm of the distance of investors in county i to the closest PTDL at time t in miles.
Snow	An indicator variable that takes the value of one if the snowfall in county i within three days after patent release exceeds 20 inches and zero otherwise.
Firm-specific and patent event-specific variables	
Log(1 + Scientific value)	The natural logarithm of one plus the adjusted number of citations of the patent event. The mean number of citations for all patents in the same technology class in the same year is deducted from the citation count. The adjusted citations are set to zero for patents that have adjusted citations below zero to ensure nonnegative values.

Table 1 (continued)

Local firm	An indicator variable that takes the value of one if the headquarters of firm j is located in county i and zero otherwise.
Log(1 + Analyst following)	The natural logarithm of one plus the number of analyst earnings forecasts for firm i before the release of the patent.
Institutional ownership	as the percentage of stocks held by institutional investors before the release of the patent.
Turnover	The ratio of daily volume and shares outstanding the day before the release of the patent.
Bid-ask spread	The bid ask spread the day before the release of the patent. The bid ask spread is calculated using the Corwin and Schultz (2012) estimator.
Log(Daily patents)	The natural logarithm of the number of patents released on day t.
M&A news	An indicator variable that takes the value of one if there is any M&A news within the seven days prior to time t and zero otherwise.

Appendix B: Results
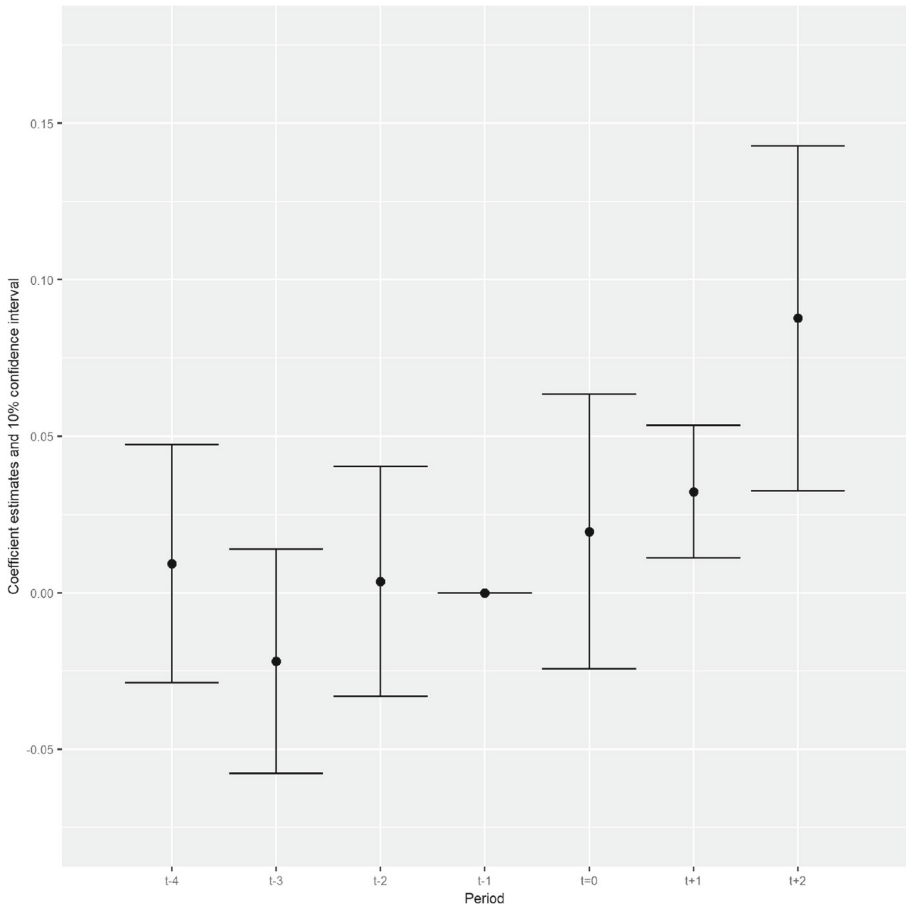

Fig. 1 Regression results: dynamic specification. *Notes*: The graph shows the coefficient estimates and 10% confidence intervals for the treatment variable before and after the opening of a PTDL. The leads and lags are defined as half years increments, relative to the opening of a PTDL. Standard errors are clustered by county. Please refer to Appendix A for descriptions of all variables

Table 2 PTDL locations and designation dates

State	County	City	Designation date
MA	Suffolk County	Boston	1870
NY	New York County	New York	1870
NY	Albany County	Albany	1870
OH	Franklin County	Columbus	1870
MO	St. Louis City	St. Louis	1870
CA	Los Angeles County	Los Angeles	1870
NY	Erie County	Buffalo	1871
OH	Hamilton County	Cincinnati	1871
MI	Wayne County	Detroit	1871
IL	Cook County	Chicago	1876
NJ	Essex County	Newark	1880
OH	Cuyahoga County	Cleveland	1890
RI	Providence County	Providence	1901
PA	Allegheny County	Pittsburgh	1902
OH	Lucas County	Toledo	1934
GA	Fulton County	Atlanta	1946
MO	Jackson County	Kansas City	1946
WI	Milwaukee County	Milwaukee	1949
OK	Payne County	Stillwater	1956
CA	Santa Clara County	Sunnyvale	1963
WI	Dane County	Madison	1976-04-26
TX	Harris County	Houston	1977-07-28
AL	Jefferson County	Birmingham	1977-08-08
WA	King County	Seattle	1977-08-08
NC	Wake County	Raleigh	1977-09-26
CO	Denver County	Denver	1977-11-11
TX	Dallas County	Dallas	1977-11-30
NE	Lancaster County	Lincoln	1978-01-05
TN	Shelby County	Memphis	1979-01-18
CA	Sacramento County	Sacramento	1979-04-13
PA	Centre County	University Park	1979-04-23
MN	Hennepin County	Minneapolis	1980-09-12
DE	New Castle County	Newark	1980-11-26
AZ	Maricopa County	Tempe	1981-06-15
LA	East Baton Rouge Parish	Baton Rouge	1981-06-24
NV	Washoe County	Reno	1983-05-16
TX	Travis County	Austin	1983-06-16
IN	Marion County	Indianapolis	1983-09-07
AL	Lee County	Auburn	1983-10-13
ID	Latah County	Moscow	1983-11-03

Table 2 (continued)

State	County	City	Designation date
NM	Bernalillo County	Albuquerque	1983-11-03
MI	Washtenaw County	Ann Arbor	1983-11-04
TX	Brazos County	College Station	1983-12-22
IL	Sangamon County	Springfield	1984-01-17
MD	Prince George's County	College Park	1984-01-25
CA	San Diego County	San Diego	1984-01-26
MT	Silver Bow County	Butte	1984-03-01
UT	Salt Lake County	Salt Lake City	1984-04-06
FL	Miami-Dade County	Miami	1984-07-25
FL	Broward County	Fort Lauderdale	1984-11-01
MA	Hampshire County	Amherst	1984-11-14
AK	Anchorage Municipality	Anchorage	1984-12-04
AR	Pulaski County	Little Rock	1985-01-18
TN	Davidson County	Nashville	1985-07-26
VA	Richmond city	Richmond	1985-09-26
PA	Philadelphia County	Philadelphia	1986-05-05
CT	New Haven County	New Haven	1986-08-11
DC	District of Columbia	Washington	1986-12-08
KY	Jefferson County	Louisville	1988-03-28
IA	Polk County	Des Moines	1988-12-09
FL	Orange County	Orlando	1988-12-30
NJ	Middlesex County	Piscataway	1989-05-09
HI	Honolulu County	Honolulu	1989-12-18
ND	Grand Forks County	Grand Forks	1990-01-26
FL	Hillsborough County	Tampa	1990-03-20
MS	Hinds County	Jackson	1990-06-19
KS	Sedgwick County	Wichita	1991-01-31
IN	Tippecanoe County	West Lafayette	1991-08-09
MI	Mecosta County	Big Rapids	1991-08-22
WV	Monongalia County	Morgantown	1991-12-02
SC	Pickens County	Clemson	1992-06-03
ME	Penobscot County	Orono	1993-12-23
CA	San Francisco County	San Francisco	1994-01-19
SD	Pennington County	Rapid City	1994-02-10
PR	Mayaguez Municipio	Mayaguez	1995-03-10
OR	Multnomah County	Portland	1995-07-27
OH	Summit County	Akron	1995-09-18
TX	Lubbock County	Lubbock	1995-10-31
NH	Merrimack County	Concord	1996-02-21

Table 3 Descriptive results

Statistic	N	Min	Pctl(25)	Median	Mean	Pctl(75)	Max	St. Dev.
Local trading volume	4,470,927	0.00	0.00	0.00	272.01	0.00	4,144,162.50	7,551.72
Number of local trades	4,470,927	0.00	0.00	0.00	0.02	0.00	34.00	0.16
Libraries	4,470,927	0.00	0.00	0.00	0.19	0.00	1.00	0.40
County patents	4,470,927	0.00	0.00	0.00	0.08	0.03	2.60	0.25
County scientific value	4,470,927	0.00	0.00	0.00	0.81	0.21	36.95	2.87
County economic value	4,470,927	0.00	0.00	0.00	0.99	0.12	42.45	3.29
County population	4,470,927	1.94	114.14	317.55	594.84	717.43	9,127.04	978.44
Local firm	4,470,927	0.00	0.00	0.00	0.01	0.00	1.00	0.08
Scientific value	4,470,927	0.00	0.00	5.22	41.47	37.05	1,276.04	97.51
Daily patents	4,470,927	1,457.00	1,858.00	1,982.00	1,937.34	2,048.00	2,256.00	187.02
Analyst following	4,470,927	0.00	0.00	8.00	8.85	15.00	35.00	8.36
Institutional ownership	4,470,927	0.00	0.00	0.44	0.35	0.58	5.24	0.29
Turnover	4,470,927	0.00	0.00	0.00	0.01	0.01	0.84	0.01
Bid-ask spread	4,470,927	0.00	0.00	0.00	0.01	0.01	0.95	0.01

Please refer to Appendix A for descriptions of all variables

Table 4 Regression results: PTDL and local trading volume

	Log(1 + Local trading volume) (Post-release)			Log(1 + Number of local trades) (Post-rel.)	Local trading indicator (Post-rel.)
	(1)	(2)	(3)	(4)	(5)
Libraries	0.121***	0.054***	0.045*	0.004*	0.005*
	(3.583)	(2.778)	(1.688)	(1.792)	(1.950)
Control variables					
Log(1 + County patents)		0.101	0.136	0.016	0.012
		(0.794)	(0.996)	(1.098)	(0.905)
Log(1 + County scientific value)		−0.018	−0.005	−0.001	−0.000
		(−0.825)	(−0.241)	(−0.406)	(−0.134)
Log(1 + County economic value)		0.112***	0.087***	0.009***	0.008***
		(5.530)	(4.435)	(4.176)	(4.282)
Log(County population)		0.021	−0.059	−0.006	−0.006
		(0.775)	(−1.281)	(−1.386)	(−1.282)
Local firm		0.691**	0.736**	0.076*	0.075**
		(1.990)	(2.149)	(1.950)	(2.257)
Constant	Yes	No	No	No	No
County fixed effects	No	Yes	Yes	Yes	Yes
Firm x date fixed effects	No	No	Yes	Yes	Yes
N	4,470,927	4,470,927	4,470,927	4,470,927	4,470,927
Adjusted R^2	0.002	0.016	0.037	0.043	0.036

Standard errors are clustered by county. T-statistics are in parentheses. ∗, ∗∗, and ∗ ∗ ∗ represent significance at the 10%, 5%, and 1% levels, respectively. Please refer to Appendix A for descriptions of all variables

Table 5 Regression results: pre-release local trading volume, scientific value and snow

	Log(1 + Local trading volume) (Pre-rel.) (1)	Log(1 + Number of local trades) (Pre-rel.) (2)	Log(1 + Local trading volume) (Post-rel.) (3)	Log(1 + Number of local trades) (Post-rel.) (4)	Log(1 + Local trading volume) (Post-rel.) (5)	Log(1 + Number of local trades) (Post-rel.) (6)
Libraries	0.015	0.001	−0.016	−0.002	0.045*	0.004*
	(0.765)	(0.738)	(−0.484)	(−0.775)	(1.689)	(1.792)
Libraries x Log(1 + Scientific value)			0.031***	0.003***		
			(3.964)	(3.591)		
Snow					−0.005	−0.001
					(−0.216)	(−0.702)
Libraries x Snow					−0.072**	−0.005**
					(−2.182)	(−1.963)
Control variables						
Log(1 + County patents)	0.109	0.012	0.138	0.016	0.136	0.016
	(0.833)	(0.934)	(1.015)	(1.116)	(0.996)	(1.098)
Log(1 + County scientific value)	−0.011	−0.001	−0.006	−0.001	−0.005	−0.001
	(−0.546)	(−0.595)	(−0.269)	(−0.431)	(−0.242)	(−0.406)
Log(1 + County economic value)	0.099***	0.009***	0.085***	0.008***	0.087***	0.009***
	(4.404)	(4.043)	(4.355)	(4.115)	(4.434)	(4.175)
Log(County population)	−0.054	−0.005	−0.052	−0.005	−0.059	−0.006
	(−1.134)	(−1.268)	(−1.130)	(−1.219)	(−1.281)	(−1.385)
Local firm	0.738**	0.075**	0.735**	0.076*	0.736**	0.076*
	(2.180)	(1.975)	(2.154)	(1.954)	(2.149)	(1.950)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm x date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	4,470,927	4,470,927	4,470,927	4,470,927	4,470,927	4,470,927
Adjusted R²	0.036	0.041	0.038	0.043	0.037	0.043

Standard errors are clustered by county. T-statistics are in parentheses. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. Please refer to Appendix A for descriptions of all variables

Table 6 Regression results: day 0 to day 4 local trading volume and distance

	Log(1 + Local trading volume) (Day 0) (1)	Log(1 + Local trading volume) (Day 1) (2)	Log(1 + Local trading volume) (Day 2) (3)	Log(1 + Local trading volume) (Day 3) (4)	Log(1 + Local trading volume) (Day 4) (5)	Log(1 + Local trading volume) (Post-rel.) (6)
Libraries	0.016*	0.025*	0.010	0.009	0.011	
	(1.746)	(1.688)	(1.143)	(0.820)	(1.355)	
Log(Distance)						−0.013**
						(−2.033)
Control variables						
Log(1 + County patents)	0.084	0.090	0.066	0.076	0.045	0.136
	(1.145)	(1.283)	(0.900)	(1.234)	(0.713)	(0.989)
Log(1 + County scientific value)	−0.004	−0.006	−0.002	−0.004	−0.013	−0.004
	(−0.384)	(−0.535)	(−0.201)	(−0.385)	(−1.286)	(−0.212)
Log(1 + County economic value)	0.044***	0.039***	0.035***	0.042***	0.046***	0.088***
	(4.620)	(3.506)	(3.794)	(4.129)	(3.788)	(4.461)
Log(County population)	−0.034	−0.017	−0.025	−0.025	−0.009	−0.062
	(−1.554)	(−0.758)	(−1.248)	(−1.213)	(−0.382)	(−1.347)
Local firm	0.341*	0.326**	0.339*	0.322*	0.328**	0.736**
	(1.833)	(1.975)	(1.943)	(1.875)	(1.966)	(2.149)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm x date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	4,470,927	4,470,927	4,470,927	4,470,927	4,470,927	4,470,927
Adjusted R^2	0.016	0.017	0.016	0.015	0.015	0.037

Standard errors are clustered by county. T-statistics are in parentheses. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. Please refer to Appendix A for descriptions of all variables

Table 7 Regression results: investor specification

	Log(1 + Local management trading volume) (Post-rel.)	Log(1 + Local technical trading volume) (Post-rel.)	Log(1 + Local retired trading volume) (Post-rel.)	Log(1 + Local other trading volume) (Post-rel.)	Log(1 + Local female trading volume) (Post-rel.)	Log(1 + Local male trading volume) (Post-rel.)
	(1)	(2)	(3)	(4)	(5)	(6)
Libraries	0.014	0.003	0.015***	0.017	−0.001	0.045**
	(1.473)	(1.585)	(3.642)	(0.989)	(−0.329)	(2.068)
Control variables						
Log(1 + County patents)	−0.005	0.005	−0.025	0.174	0.034	0.067
	(−0.211)	(0.132)	(−1.449)	(1.190)	(1.384)	(0.682)
Log(1 + County scientific value)	0.009	−0.000	0.004	−0.021	−0.002	0.009
	(1.192)	(−0.061)	(1.195)	(−1.044)	(−0.417)	(0.505)
Log(1 + County economic value)	0.020**	0.020***	0.001	0.068***	0.011**	0.075***
	(2.192)	(2.702)	(0.345)	(3.832)	(2.466)	(4.260)
Log(County population)	−0.037*	−0.022	0.024*	−0.035	−0.005	−0.054
	(−1.877)	(−1.351)	(1.883)	(−0.887)	(−0.348)	(−1.344)
Local firm	0.210**	0.097**	0.029***	0.568*	0.095*	0.626**
	(2.182)	(2.396)	(3.802)	(1.810)	(1.856)	(2.188)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm x date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
[5pt] N	4,470,927	4,470,927	4,470,927	4,470,927	4,470,927	4,470,927
Adjusted R^2	0.009	0.003	0.001	0.027	0.002	0.031

Standard errors are clustered by county. T-statistics are in parentheses. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively. Please refer to Appendix A for descriptions of all variables

Table 8 Regression results: firm and event characteristics

	Log(1 + Local trading volume) (Post-release)					
	(1)	(2)	(3)	(4)	(5)	(6)
Libraries	−0.027	0.029	0.025	0.049*	0.536***	0.044
	(−0.817)	(1.087)	(0.882)	(1.817)	(2.710)	(1.638)
Libraries x Log(1 + Analyst following)	0.045***					
	(4.268)					
Libraries x Institutional ownership		0.046***				
		(3.654)				
Libraries x Turnover			3.492**			
			(2.477)			
Libraries x Bid-ask spread				−0.451**		
				(−2.144)		
Libraries x Log(Daily patents)					−0.065**	
					(−2.411)	
Libraries x M&A news						0.120***
						(3.436)
Control variables						
Log(1 + County patents)	0.139	0.137	0.140	0.137	0.135	0.137
	(1.021)	(1.006)	(1.041)	(0.998)	(0.973)	(0.999)
Log(1 + County scientific value)	−0.006	−0.005	−0.006	−0.005	−0.004	−0.005
	(−0.278)	(−0.256)	(−0.306)	(−0.243)	(−0.214)	(−0.244)
Log(1 + County economic value)	0.086***	0.087***	0.084***	0.087***	0.089***	0.087***
	(4.357)	(4.404)	(4.339)	(4.433)	(4.481)	(4.422)
Log(County population)	−0.052	−0.057	−0.052	−0.059	−0.065	−0.058
	(−1.130)	(−1.244)	(−1.121)	(−1.276)	(−1.389)	(−1.266)
Local firm	0.733**	0.735**	0.735**	0.736**	0.736**	0.735**
	(2.143)	(2.148)	(2.162)	(2.148)	(2.148)	(2.148)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm x date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	4,470,927	4,470,927	4,470,927	4,470,927	4,470,927	4,470,927
Adjusted R^2	0.038	0.037	0.037	0.037	0.037	0.037

Standard errors are clustered by county. T-statistics are in parentheses. ∗, ∗∗, and ∗ ∗ ∗ represent significance at the 10%, 5%, and 1% levels, respectively. Please refer to Appendix A for descriptions of all variables

Table 9 Regression results: return specification

	CAR (20 days) (1)	CAR (60 days) (2)	CAR (240 days) (3)	BHAR (20 days) (4)	BHAR (60 days) (5)	BHAR (240 days) (6)
Libraries	0.008**	0.032***	0.049***	0.007**	0.031***	0.079***
	(2.164)	(6.059)	(4.046)	(2.065)	(5.372)	(4.927)
Control variables						
Log(1 + County patents)	−0.003	−0.042**	−0.096	−0.003	−0.050**	−0.082
	(−0.277)	(−1.980)	(−1.007)	(−0.213)	(−2.254)	(−0.792)
Log(1 + County scientific value)	0.010	0.017	0.039	0.010	0.017	0.060
	(1.439)	(1.105)	(0.692)	(1.449)	(1.185)	(0.905)
Log(1 + County economic value)	−0.005	0.001	−0.004	−0.005	0.000	−0.025
	(−0.736)	(0.043)	(−0.089)	(−0.797)	(0.013)	(−0.412)
Log(County population)	0.014	0.044	0.823***	0.007	0.026	0.994***
	(0.685)	(1.278)	(7.730)	(0.326)	(0.673)	(7.515)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	35,512	35,512	35,512	35,512	35,512	35,512
Adjusted R^2	0.002	0.001	0.016	0.002	0.001	0.016

Standard errors are clustered by county. T-statistics are in parentheses. ∗, ∗∗, and ∗ ∗ ∗ represent significance at the 10%, 5%, and 1% levels, respectively. Please refer to Appendix A for descriptions of all variables

References

Agrawal, Ajay, Alberto Galasso, and Alexander Oettl. 2017. Roads and innovation. *Review of Economics and Statistics* 99: 417–434.

Ayers, Benjamin C., Oliver Z. Li, and P. E. Yeung. 2011. Investor trading and the post-earnings-announcement drift. *The Accounting Review* 86: 385–416.

Barber, Brad M., and Terrance Odean. 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* 55: 773–806.

Barber, Brad M., and Terrance Odean. 2001. Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics* 116: 261–292.

Barber, Brad M., and Terrance Odean. 2002. Online investors: Do the slow die first? *Review of Financial Studies* 15: 455–488.

Barber, Brad M., and Terrance Odean. 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21: 785–818.

Belenzon, Sharon, and Mark Schankerman. 2013. Spreading the word: Geography, policy, and knowledge spillovers. *Review of Economics and Statistics* 95: 884–903.

Bhushan, Ravi. 1989. Collection of information about publicly traded firms: Theory and evidence. *Journal of Accounting and Economics* 11: 183–206.

Blankespoor, Elizabeth, Ed deHaan, John Wertz, and Christina Zhu. 2019. Why do individual investors disregard accounting information? The roles of information awareness and acquisition costs. *Journal of Accounting Research* 57: 53–84.

Bloomfield, Robert. 2002. The "incomplete revelation hypothesis" and financial reporting. *Accounting Horizons* 16: 233–243.

Brown, William H., and James A. Arshem. 1993. Survey of patent and trademark depository libraries 1991/1992, U.S. Patent and Trademark Office Report.

Bushee, Brian J., and Christopher F. Noe. 2000. Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research* 38: 171–202.

Chapman, Kimball L., Nayana Reiter, Hal. D. White, and Christopher D. Williams. 2019. Information overload and disclosure smoothing. *Review of Accounting Studies* 24: 1486–1522.

Christensen, Hans B., Eric Floyd, Lisa Yao Liu, and Mark Maffett. 2017. The real effects of mandated information on social responsibility in financial reports: Evidence from mine-safety records. *Journal of Accounting and Economics* 64: 284–304.

Corwin, Shane A., and Paul Schultz. 2012. A simple way to estimate bid-ask spreads from daily high and low prices. *Journal of Finance* 67: 719–760.

Engelberg, Joseph E., and Christopher A. Parsons. 2011. The causal impact of media in financial markets. *Journal of Finance* 66: 67–97.

EY. 2014. Disclosure effectiveness companies embrace the call to action.

Fitzgerald, Tristan, Benjamin Balsmeier, Lee Fleming, and Gustavo Manso. 2019. Innovation search strategy and predictable returns. *Management Science*, Forthcoming.

Forbes. 2019. The greatest investor you've never heard of: An optometrist who beat the odds to become a billionaire.

Furman, Jeffrey L., Markus Nagler, and Martin Watzinger. 2020. Disclosure and subsequent innovation: Evidence from the patent depository library program, *American Economic Review: Economic Policy*, Forthcoming.

Glaeser, Stephen. 2018. The effects of proprietary information on corporate disclosure and transparency: Evidence from trade secrets. *Journal of Accounting and Economics* 66: 163–193.

Glaeser, Stephen, and Wayne R. Landsman. 2020, Deterrent disclosure. *The Accounting Review*, Forthcoming.

Glaeser, Stephen, Jeremy Michels, and Robert E. Verrecchia. 2020. Discretionary disclosure and manager horizon: Evidence from patenting. *Review of Accounting Studies* 25: 597–635.

Graham, Stuart, and Deepak Hegde. 2015. Disclosing patents' secrets. *Science* 347: 236–237.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. 2001. The nber patent citation data file: Lessons, insights and methodological tools, National Bureau of Economic Research Report.

Hegde, Deepak, Baruch Lev, and Chenqi Zhu. 2018. Patent disclosure and price discovery. Working Paper.

Hirschey, Mark, and Vernon J. Richardson. 2004. Are scientific indicators of patent quality useful to investors? *Journal of Empirical Finance* 11: 91–107.

Hirshleifer, David, Po-Hsuan Hsu, and Dongmei Li. 2013. Innovative efficiency and stock returns. *Journal of Financial Economics* 107: 632–654.

Hirshleifer, David, Po-Hsuan Hsu, and Dongmei Li. 2018. Innovative originality, profitability, and stock returns. *Review of Financial Studies* 31: 2553–2605.

Hirshleifer, David A., and Siew H. Teoh. 2003. Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics* 36: 337–386.

Hochberg, Yael V., Carlos J. Serrano, and Rosemarie H. Ziedonis. 2018. Patent collateral, investor commitment, and the market for venture lending. *Journal of Financial Economics* 130: 74–94.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein. 2005. Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. *Journal of Finance* 60: 2801–2824.

Huberman, Gur, and Tomer Regev. 2001. Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *Journal of Finance* 56: 387–396.

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

Ivković, Zoran, and Scott Weisbenner. 2007. Information diffusion effects in individual investors' common stock purchases: Covet thy neighbors' investment choices. *Review of Financial Studies* 20: 1327–1357.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108: 577–598.

Kim, Jinhwan, and Kristen Valentine. 2020. The innovation consequences of mandatory patent disclosures. *Journal of Accounting and Economics*, Forthcoming.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics* 132: 665–712.

Leuz, Christian, and Peter D. Wysocki. 2016. The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research. *Journal of Accounting Research* 54: 525–622.

Mann, William. 2018. Creditor rights and innovation: Evidence from patent collateral. *Journal of Financial Economics* 130: 25–47.

Martens, Tim, and Christoph J Sextroh. 2021. Analyst coverage overlaps and interfirm information spillovers. *Journal of Accounting Research*, Forthcoming.

Merton, Robert C. 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42: 483–510.

Oliver, Jeanne. 2002. Celebrating 25 years of libraries and uspto cooperation, U.S. Patent and Trademark Office Report.

Petersen, Mitchell A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22: 435–480.

Saidi, Farzad, and Alminas Žaldokas. 2020. How does firms' innovation disclosure affect their banking relationships? *Management Science*, Forthcoming.

Seasholes, Mark S., and Ning Zhu. 2010. Individual investors and local bias. *Journal of Finance* 65: 1987–2010.

Shiller, Robert J., and John Pound. 1989. Survey evidence on diffusion of interest and information among investors. *Journal of Economic Behavior & Organization* 12: 47–66.

Shores, David. 1990. The association between interim information and security returns surrounding earnings announcements. *Journal of Accounting Research* 28: 164–181.

Sneed, Martha C. 2000. Fully disclosed yet merely descriptive: Intricacies of training the patent and trademark information professional. *Journal of Library Administration* 29: 59–78.

USPTO. 2003. Patent and trademark depository library program 2002 customer satisfaction survey, U.S. Patent and Trademark Office Report.