Passive Ownership and Price Informativeness

MARCO SAMMON *

September, 2022

ABSTRACT

I show that passive ownership negatively affects the degree to which stock prices anticipate earnings announcements. Estimates across several research designs imply that the rise in passive ownership over the last 30 years has caused the amount of information incorporated into prices ahead of earnings announcements to decline by roughly 16%. This effect occurs in part because passive owners collect less firm-specific information ahead of earnings announcements and limits to arbitrage prevent non-passive investors from fully offsetting this behavior.

Keywords: Passive ownership, Price informativeness.
JEL classification: G12, G14.

1   Introduction

Passive investing through index mutual funds and ETFs plays an increasingly large role in U.S. capital markets. From 1990 to 2018, the share of U.S. equities held by passive investors rose from less than 1% to almost 15%. There is still considerable debate about the costs and benefits of passive investment vehicles. Proponents of these instruments argue that they provide investors with access to a range of diversified portfolios at low costs due to a combination of lower fees, decreased turnover and greater tax efficiency (Wurgler (2010), Madhavan (2014), Madhavan (2016)). On the other hand, the growth of passive investing has raised concerns that capital market prices have become less informative, thereby distorting capital allocations (Brogaard et al., 2019). The general argument for this view is that passive investors pay less or no attention to the underlying securities and therefore their prices do not reflect all available information.

Research on the relationship between passive ownership and price informativeness has drawn mixed conclusions, in part due to differences in how price informativeness is measured. For instance, Kacperczyk et al. (2018a) find a positive relationship when measuring price informativeness using the ability of current prices to forecast future fundamentals. Their approach is based on the noisy rational expectations models of Grossman and Stiglitz (1980) and Bai et al. (2016). In contrast, Bennett et al. (2020b) build on Roll (1988) and find a negative relationship when measuring price informativeness based on a regression of individual security returns on market-wide returns.

In this paper, I bring new evidence to bear on this debate by studying how passive ownership affects the incorporation of information into prices in narrow windows around earnings announcements. My approach is motivated by early studies of market efficiency (Ball and Brown, 1968) documenting that a substantial portion of earnings news is incorporated into prices prior to the actual announcement. The usual interpretation of these findings is that private information is collected by market participants ahead of earnings announcements. I use this logic to test the following hypothesis: do stocks with more passive ownership have less of their earnings information incorporated into their prices ahead of earnings announcements?

I measure the amount of information incorporated into prices prior to announcements in two ways. In both cases, I examine prices in the month leading up to and including a

firm's earnings announcement date. First, I compute the fraction of the total gross return (the pre-earnings drift magnitude, DM) that occurs prior to the earnings release.[1] Second, I compute the fraction of total volatility (the quadratic variation share, QVS) that occurs prior to the earnings release. I interpret lower values of both as an indication that less private information was collected ahead of the earnings announcement.

Leveraging these measures, I establish several new facts about passive ownership and price informativeness before earnings announcements using the cross-section of U.S. equities from 1990 to 2018. My first main finding is that average price informativeness declined steadily over the past 30 years, mirroring the aggregate rise of passive ownership. In 1990, 92.1% of return volatility in the month leading up to and including an earnings announcement occurred prior to the release. By 2018, this number had declined to 72.4%. The average share of the total gross return occurring over the same window (DM) displays a similar downward trend. Both measures indicate that on average, less information is being incorporated into prices ahead of earnings announcements.

These aggregate patterns are mirrored in the cross-section of U.S. stocks. Through a series of panel regressions, I establish a robust negative relationship between pre-earnings announcement price informativeness and the fraction of individual firms' shares outstanding held by passive investors. My preferred regression estimates imply that a stock in the 90th percentile of passive ownership in 2018 has 9.1 pp less of its return volatility occur ahead of earnings announcements relative to a stock in the 10th percentile of passive ownership in 2018. For reference, the difference in passive ownership share between these two percentiles is 23%, slightly larger than the value-weighted average increase in passive ownership over my entire sample. Once again, the results for DM corroborate these findings. As additional supporting evidence for these results, I show that options markets internalize the relationship between passive ownership and decreased pre-earnings announcement price informativeness.

These reduced form cross-sectional correlations do not, however, conclusively establish a causal link between passive ownership and pre-earnings announcement price informativeness. An alternative interpretation is that causality runs the other way. For instance, passive vehicles may be more likely to own firms with larger market capitalizations. Larger firms are also more complex, so perhaps less information is incorporated into their prices ahead

---

[1]Weller (2018) uses a similar measure to quantify whether algorithmic trading activity deters information acquisition ahead of earnings announcements.

of earnings announcements.[2] Consider, for instance, a firm like Apple. To profitably trade ahead of Apple's earnings announcements, an investor would need to collect information spanning multiple business segments and geographies. To the extent that this effort is costly, Apple might have lower pre-earnings announcement price informativeness that is unrelated to the composition of its owners.

To establish a tighter causal link between passive ownership and pre-earnings announcement price informativeness, I build instruments for my baseline panel regressions using changes in passive ownership due to Russell 1000/2000 rebalancing (Appel et al. (2016), Ben-David et al. (2018), Gloßner (2018), Coles et al. (2022)) and S&P 500 index additions (Qin and Singal (2015), Bennett et al. (2020b)). My underlying assumption is that index rebalancing only affects price informativeness ahead of earnings announcements through its mechanical effect on passive ownership. Following Coles et al. (2022), I attempt to enforce this assumption by choosing an appropriate set of similar control firms that did not switch indices. For stocks switching to the Russell 2000, I choose a set of control firms which stayed in the Russell 1000 but were near the size cutoff used to determine index membership. I apply a similar logic for firms added to the S&P 500.

The IV estimates using both Russell and S&P 500 rebalancing reinforce a negative causal effect of passive ownership on pre-earnings announcement price informativeness. For example, when using the Russell rebalancing, I find that moving from the 90th percentile of passive ownership to the 10th percentile of passive ownership in 2018 decreases return volatility ahead of earnings announcements by 22.83 percentage points. Importantly, in the presence of the reverse causality described above, we would expect the OLS estimates to be biased upward in magnitude. The fact that the IV estimates are larger than those from the OLS suggest that the latter are not materially biased by these endogeneity concerns.

My preferred interpretation of these findings is that passive ownership decreases pre-earnings announcement price informativeness because passive investors gather less firm-specific information. This is motivated by recent work which has highlighted a trade off faced by passive investors, specifically that they may tend to increase the incorporation of systematic information at the expensive of security-specific information (Cong et al. (2020),

---

[2]In the cross-sectional regressions, I account for firm size by directly controlling for lagged market capitalization (Ben-David et al., 2018) and by value-weighting observations. In the Appendix, I show my results are robust to instead including fixed-effects for deciles of market capitalization, formed each quarter.

Glosten et al. (2021)). There are several pieces of evidence that lead me to favor this interpretation.

First, to show that the negative relationship between passive ownership and pre-earnings announcement price informativeness is being driven by firm-specific information, I examine how stock prices respond to fundamental news of a given size. The intuition is that if investors are not gathering private information before earnings announcements, they will have less precise beliefs. When the news arrives, therefore, they will update their beliefs significantly, which leads to larger average price changes (Ganuza and Penalva (2010), Ganuza and Penalva (2010)). If passive investors are neglecting firm-specific information, this effect should be especially strong for the idiosyncratic component of earnings news.

To quantify how responsive stock prices are to earnings news, I run regressions in the spirit of Kothari and Sloan (1992), with earnings-day-returns on the left-hand side and standardized unexpected earnings (SUE) on the right-hand side. My regression estimates imply that a stock in the 90th percentile of passive ownership in 2018 responds nearly 3 times as much to earnings news as a stock in the 10th percentile of passive ownership. To refine this test, following Glosten et al. (2021), I decompose earnings news into systematic and idiosyncratic components. Consistent with my proposed mechanism, high passive stocks' increased responsiveness to earnings information is concentrated in the idiosyncratic component of news.

Next, I show direct evidence of differences in information collection between high and low passive stocks. In the cross-section, passive ownership is correlated with fewer Bloomberg terminal searches, evidence of less attention by institutional investors (Ben-Rephael et al., 2017). Specifically, moving from the 10th to the 90th percentile of passive ownership in 2018 is associated with 34% less abnormal institutional investor attention, relative to its whole-sample mean. Passive ownership is also correlated with fewer non-robot downloads of SEC filings, evidence of less fundamental research Loughran and McDonald (2017).[3]

I then discuss the effect of passive ownership on the supply of information, which I measure using data on sell-side analysts (Martineau and Zoican, 2021). The logic is that if information is costly to produce, sell-side analysts may respond to passive owners' decreased demand for information by supplying less or lower quality forecasts. Consistent with this

---

[3]These results are similar to those in Israeli et al. (2017) and Coles et al. (2022), who also provide evidence that less information is gathered about stocks with more passive ownership.

hypothesis, passive ownership is correlated with decreased coverage, increased dispersion of analysts' estimates, decreased forecast accuracy and fewer updates.

Next, I examine trading volume before earnings announcements. This is useful to distinguish between theories that relate private information gathering to trading volume. Private information is a source of disagreement, which tends to increase trading volume (Wang, 1994), but it is also a source of information asymmetry, which can decrease trading volume through fear of adverse selection (Foster and Viswanathan, 1990). Over the past 30 years, average abnormal pre-earnings volume has declined by about 5% relative to its whole-sample mean. In the cross-section, passive ownership is negatively correlated with pre-earnings trading volume, suggesting that the effect of reduced private information on disagreement dominates its effect on adverse selection.

Finally, I turn to the question of why the remaining non-passive investors don't fully offset the behavior of passive investors by gathering more information (Coles et al., 2022). My first proposed explanation is based on the finding in Ben-David et al. (2018) that ETFs (but not index mutual funds) increase non-fundamental volatility. I argue that this may create noise-trader risk (De Long et al., 1990), which could deter informed investors from learning about high passive stocks. Consistent with this, I show the effect of passive ownership on pre-earnings announcement price informativeness is stronger for ETFs than index mutual funds. My second explanation is that passive ownership may reduce pre-earnings liquidity, which is corroborated by my results on its relationship to pre-earnings turnover.

Overall, my analysis contributes to several strands of research on passive ownership and price informativeness. By focusing on earnings days, I show that there has been a trend toward decreased pre-announcement price informativeness over the past 30 years. Through cross-sectional regressions and two instrumental variables designs, I show passive ownership causes pre-earnings price informativeness to decline. In terms of magnitudes, averaging the point estimates from the OLS and both IVs implies that a 15% increase in passive ownership decreases QVS by 14.87, an approximately 16% decline relative to its mean in 1990 ($14.87/92.1 \approx 0.16$). Finally, I provide evidence for why passive ownership decreases price informativeness, specifically that passive investors gather less firm-specific information.

Literature Review. My paper contributes to a growing literature that studies the relationship between passive ownership and price informativeness. The conclusions from this research are mixed. Some studies find a positive link (Buss and Sundaresan (2020),

6

Ernst (2020), Malikov (2020), Lee (2020), Kacperczyk et al. (2018a)), while others find a negative (Qin and Singal (2015), DeLisle et al. (2017), (Bond and Garcia, 2018), Garleanu and Pedersen (2018), Kacperczyk et al. (2018b), Breugem and Buss (2019), Brogaard et al. (2019), Bennett et al. (2020a), Bennett et al. (2020b)) or non-existent link (Coles et al., 2022). Part of the reason for this disagreement is that the papers differ in how they measure price informativeness. Another reason is that passive investors collect different types of information. For example, passive ownership may increase informativeness about systematic information while decreasing the incorporation of idiosyncratic information (Bhattacharya and O'Hara (2018), Cong et al. (2020), Antoniou et al. (2020), Glosten et al. (2021)).

The main difference between my approach and previous work is how I measure price informativeness. I focus specifically on the narrow window ahead of earnings announcements and measure how much of the news is incorporated into prices ahead of time. This allows me to abstract away from any particular model of price informativeness and only requires the assumption that prices reflect all of the information contained in the announcement shortly after its release. My main finding is that passive ownership causes less of the earnings information to be incorporated into prices ahead of the announcement. Moreover, I find that the effect occurs because passive owners gather less firm-specific information, consistent with theoretical work by Cong et al. (2020).

A related literature asks how price informativeness has evolved through time (Bai et al. (2016), Dávila and Parlatore (2018)). These studies measure price informativeness by looking at time trends in the relationship between current prices and future fundamentals. The main conclusion from this work is that the link between prices and future fundamentals has become stronger over time, likely due to improvements in financial and information technology (Farboodi and Veldkamp (2020), Farboodi et al. (2020)). My analysis focuses on a different question, namely when information is incorporated into prices. I find that over time, there has been a trend toward a larger share of earnings information being incorporated

into prices after the news is released.[4]

2   Measurement & data

This section motivates my two measures of pre-earnings announcement price informativeness. I then describe the data I use to compute these measures and the firm-level passive ownership share. Finally, I present facts on the time-series decline in average pre-earnings announcement price informativeness and increase in passive ownership from 1990 to 2018.

2.1   Measurement

Ball and Brown (1968) show that prices incorporate a substantial portion of earnings news before it is actually made public. In the Appendix, Figure A.1 replicates their main finding, showing that prices increase before the release of good earnings news and drift down before the release of bad earnings news. A natural measure of pre-earnings announcement price informativeness, therefore, is the percentage of the total information which was incorporated into prices ahead of time.

I leverage this intuition to create two measures of pre-earnings announcement price informativeness. The first, which directly builds on the logic of Ball and Brown (1968), captures the share of total returns in the month leading up to and including an earnings announcement that occurs before the actual release. To this end, I define the pre-earnings drift magnitude (DM) for firm i with an earnings announcement at time t as:

---

[4]My results are not inconsistent with Bai et al. (2016), who show that current valuation ratios have become better predictors of long-horizon future cashflows. By the logic of Campbell and Shiller (1988), this pattern must be driven by the fact that valuation ratios covary less with future returns (Cohen et al., 2003). My results speak to something different, namely that there has been a change in *when* return volatility occurs. The time-series trends in DM and QVS show that more return volatility occurs after the release of earnings information. These trends say nothing about total return volatility per se and thus the covariance of valuation ratios with long-run future returns. So, it can both be true that valuation ratios have become better forecasts of long-run future earnings but in the short-run, prices anticipate earnings announcement news less. This could be the case, for example, if improvements in financial and information technology have led prices to better reflect earnings information after the news is released, as Bai et al. (2016) are using prices after the announcement of December calendar quarter earnings (i.e., prices from the end of March) to forecast future fundamentals. The same logic applies as to why my results are not inconsistent with those in Dávila and Parlatore (2018).

$$DM_{i,t} = 100 \times \begin{cases} \frac{R_{i,(t-22,t-1)}}{R_{i,(t-22,t)}} & \text{if } r_{i,t} > 0 \\ \frac{R_{i,(t-22,t)}}{R_{i,(t-22,t-1)}} & \text{if } r_{i,t} < 0 \end{cases} \tag{1}$$

where $R_{i,(t-k,t+j)}$ denotes a cumulative *gross* market-adjusted return from $t - k$ to $t + j$ and $r_{i,t}$ denotes a *net* market-adjusted return, defined as the difference between firm i's return and the market factor from Ken French's data library (Campbell et al., 2001). The choice of 22 trading-days (roughly a calendar month) before the announcement is in line with previous literature on pre-earnings price informativeness (Weller, 2018).[5] When $r_{i,t}$ is positive, DM captures the percentage of the total gross return from $t - 22$ to t which is earned before the announcement itself. If $r_{i,t}$ is negative, this relationship would be reversed, which is why the measure is inverted when $r_{i,t}$ is less than zero.

DM is designed to capture the share of earnings information that is incorporated into prices before the announcement. The assumption underlying this interpretation is that the earnings announcement is fully incorporated into prices quickly after its release (within a day). In this case, the return over the month before and through the announcement can be used to proxy for the total amount of information contained in the announcement. This interpretation of DM yields the first empirical prediction I use to measure the effect of passive ownership on price informativeness.

Prediction 1: *If passive ownership decreases pre-earnings announcement price informativeness, it should cause DM to decline*

Using DM to measure informativeness does, however, have several limitations. For example, consider two firms with cumulative returns of 0% over the pre-earnings announcement month. One of them has an earnings day return of 5%, while the other has an earnings day return of -5%. Intuition suggests that these two firms have equally informative pre-earnings announcement prices, but they will have slightly different values of $DM_{i,t}$ (95.24 vs. 95.00).

The second issue is that DM is sensitive to the level of volatility. To fix ideas, consider two stocks with different volatilities: Leading up to an earnings announcement, Stock A has alternating returns of $\pm$ 1% while stock B has alternating returns of $\pm$ 5%. On the announcement day, stock A has a return of 1% while stock B has a return of 5%. It seems natural that both stocks have equally informative pre-earnings announcement prices, as the

---

[5] In the Appendix, I show that my results are robust to including various post-earnings announcement windows in both my measures of pre-earnings announcement price informativeness.

earnings day returns are the same magnitude as those over the prior month. These stocks will, however, have significantly different values of DM (99.01 vs. 95.24).

To address these limitations, I build on the model of information in Ganuza and Penalva (2010) and create a measure based on the share of total volatility which occurs before the earnings announcement. Specifically, I define the quadratic variation share (QVS) for firm i around earnings announcement t as:

$$QVS_{i,t} = 100 \times \sum_{\tau=-22}^{-1} r_{i,t+\tau}^2 / \sum_{\tau=-22}^{0} r_{i,t+\tau}^2 \tag{2}$$

where $r_{i,t}$ denotes a market-adjusted daily return.[6]

QVS measures the contribution of the earnings day return to volatility in the month leading up to and including the earnings announcement. Like DM, if more of the information contained in a given earnings announcement is being incorporated into prices ahead of the release, the magnitude of the earnings day return should be smaller, and so too will QVS. This interpretation of QVS yields the second empirical prediction I use to measure the relationship between passive ownership and price informativeness.

Prediction 2: *If passive ownership decreases pre-earnings announcement price infor-mativeness, it should cause $QVS_{i,t}$ to decline*

There are many other ways to measure price informativeness. For example, Weller (2018) quantifies pre-earnings announcement price informativeness using a measure similar to DM, defined using net returns instead of gross returns. I discuss these alternatives and show my conclusions are robust to them in Appendix C.

2.2 Data

My sample starts with all ordinary common shares (share codes 10-11) traded on major exchanges (exchange codes 1-3) that can be matched between CRSP and IBES between 1990 and 2018. For each stock, around each earnings announcement, I need to construct DM, QVS and the level of passive ownership.

---

[6]All my results are robust to defining QVS at the annual level i.e., defining the numerator to be the sum of squared non-earnings returns in year t, while the denominator is the sum of squared returns for all days in year t. This is evidence that my results are not driven by the choice of a 22-day pre-earnings window.

To construct the measures of price informativeness, I need to identify the first time investors could have traded on earnings information during normal market hours. I identify these days using the earnings release date and time in IBES. If earnings are released before 4:00 PM eastern time between Monday and Friday, that day will be labeled as the effective earnings date. If earnings are released on or after 4:00 PM eastern time between Monday and Friday, over the weekend, or on a trading holiday, the next trading date in CRSP is labeled as the effective earnings date. To be included in the final sample, a firm must have non-missing returns in CRSP each day from $t-22$ to t around the earnings announcement. I use these returns to construct DM and QVS.

The last object I need to construct for each observation is passive ownership, which I define as the fraction of a stock's shares outstanding which are held by passive funds. Following Appel et al. (2016), I identify passive funds using the CRSP mutual fund database, selecting all index funds, all ETFs and all funds with names that identify them as index funds. To calculate how many shares of each stock passive funds hold, I use the WRDS MFLINKS database to match the identified funds to Thompson S12, which contains data on funds' holdings. The passive ownership share is the sum of all shares held by passive funds, divided by shares outstanding in CRSP. In the Appendix, I show that my results are quantitatively unchanged by dropping all observations with zero passive ownership (Dannhauser, 2017).

2.3   Basic properties

To visualize the time-series and cross-sectional properties of the three key variables in my analysis, Figure 1 plots the 25th percentile, median, 75th percentile and value-weighted average of QVS, DM and passive ownership. The top left panel shows that QVS decreased steadily over my sample. Average QVS decreased from 92.1% in 1990 to 72.4% in 2018. This 19.7 percentage point decline is about the same size as QVS's whole-sample standard deviation of 21.1%.[7] There has also been a trend toward increased cross-sectional spread in QVS, with the interquartile range increasing from 10% to over 40%.

The decline in average QVS accelerates around 2001, which coincides with two changes to the amount of information released before earnings announcements. The first is Regulation Fair Disclosure (Reg FD), passed in August 2000, which reduced early selective disclosure

---

[7]The Appendix shows that the decrease in QVS was due to a simultaneous increase earnings-day volatility and a decrease in non-earnings-day volatility.
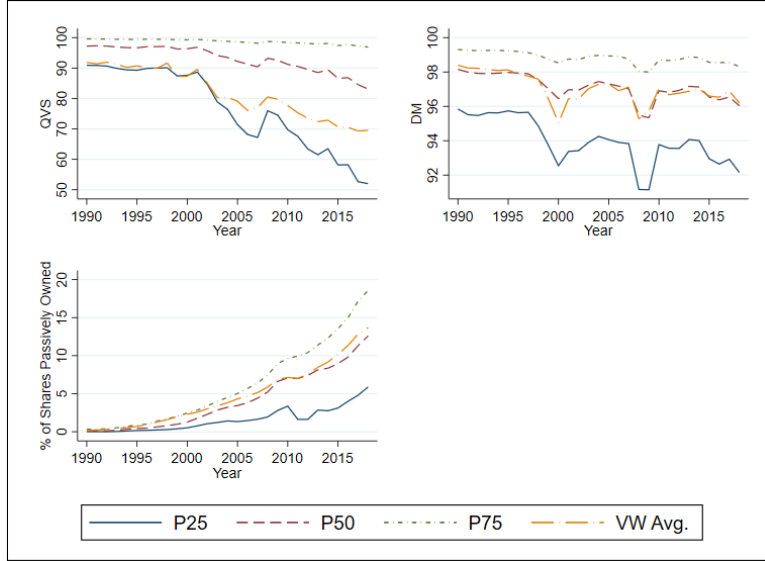
Figure 1. Trends in QVS, DM and passive ownership, 1990-2018. To compute the value-weighted average (*VW Avg.*), within each quarter, observations are weighted in proportion to their market capitalization at the end of the previous quarter. DM and QVS are defined in Equations 1 and 2. Passive ownership is defined as the fraction of a stock's shares which are held by all index funds, all ETFs and all mutual funds with names that identify them as index funds.

of earnings information. The second is the increased enforcement of insider trading laws (Coffee, 2007).

The top right panel of Figure 1 shows that, consistent with the trend in QVS, average DM decreased by about 2 between 1990 and 2018. This drop is roughly 40% of DM's whole sample standard deviation of 5.4. There are notable drops in average DM in the early 2000s and again in the late 2000s. As with QVS, the level shift down in the early 2000s may be the result of Reg FD and decreased insider trading.

Another explanation for the drop in DM is that these years correspond to the dot-com boom and the Global Financial Crisis. These were periods with higher overall volatility, leading to larger absolute earnings-day returns and lower values of DM on average. QVS may not have experienced a correspondingly large drop in the dot-com boom because it explicitly accounts for the level of volatility in the month leading up to the earnings announcement.

The bottom left panel of Figure 1 shows that passive ownership steadily increased over my sample. From 1990 to 2018, average passive ownership went from nearly zero to owning almost 15% of the US stock market. These numbers closely mirror those in the ICI factbook.

12

Like QVS and DM, the difference between high and low passive ownership stocks also grew over my sample, with the interquartile range increasing from 0% in 1990 to about 15% by 2018. Table 1 contains summary statistics on the measures of pre-earnings announcement price informativeness measures, as well as passive ownership.

		25%	50%	Mean	75%	St. Dev.
QVS		89.82	96.97	91.35	99.46	14.06
DM	1990-1999	95.36	97.84	96.41	99.17	4.48
$Passive$		0.06	0.38	0.69	1.03	0.82
QVS		61.23	88.37	76.68	97.91	26.19
DM	2010-2018	93.32	96.77	95.04	98.68	5.48
$Passive$		3.21	8.37	8.75	12.83	6.55
QVS		80.64	94.87	85.17	99.06	21.11
DM	All Years	94.08	97.26	95.50	98.92	5.36
$Passive$		0.32	1.66	3.69	5.18	4.90

Table 1 Summary Statistics. Cross-sectional equal-weighted means, standard deviations and distributions of price informativeness and passive ownership.

3 Passive ownership and pre-earnings announcement price informativeness in the cross-section

This section documents the relationship between passive ownership and pre-earnings announcement price informativeness. It starts with cross-sectional regressions of QVS and DM on passive ownership. Across both measures, the regressions show that higher passive ownership is correlated with decreased pre-earnings announcement price informativeness. I then provide evidence that options markets internalize the relationship between passive ownership and earnings-day volatility. Finally, I perform robustness checks to show that Regulation Fair Disclosure and the rise of algorithmic trading are not driving my OLS regression estimates.

3.1 Baseline analysis

I run the following regression to measure the relationship between pre-earnings announcement price informativeness and passive ownership:

$$\text{Price informativeness}_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t} \qquad (3)$$

where Price informativeness$_{i,t}$ is either $QVS_{i,t}$ or $DM_{i,t}$. Controls in $X_{i,t}$ include time since listing (age), one-month lagged market capitalization, returns from month $t-12$ to $t-2$, one-month lagged book-to-market ratio and the institutional ownership ratio.[8] $X_{i,t}$ also includes CAPM beta, CAPM R-squared, total volatility and idiosyncratic volatility, all computed over the previous 252 trading days. The controls in $X_{i,t}$ are selected to capture firm characteristics known to be correlated with passive ownership (Glosten et al., 2021). The Appendix contains details on the construction of all control variables.

Equation 3 also includes firm and year-quarter fixed effects. The firm fixed effects account for differences in average price informativeness e.g., investors may pay more attention to Apple's earnings announcements than to those of Dominion Energy. The year-quarter fixed effects account for the time trends in pre-earnings announcement price informativeness and the seasonality in earnings news. Standard errors are double-clustered at the firm and year-quarter level.

The regression results are in Table 2. Consistent with prediction 2, Column 1 shows that there is a negative relationship between passive ownership and QVS. The point estimate implies that a firm in the 90th percentile of passive ownership in 2018 (25%) has 9.1 pp lower QVS than a firm in the 10th percentile of passive ownership in 2018 (2%). For reference, 9.1 is roughly $2/5^{ths}$ of QVS's whole sample standard deviation. To allay concerns that small firms are driving my results, Column 2 weights observations by each firm's share of total market capitalization at the end of the previous quarter. Using value weights shrinks the estimated coefficient, but it remains statistically significant at the 1% level.

Column 3 shows that, consistent with prediction 1, there is also a negative correlation between DM and passive ownership. The point estimate implies that a firm in the 90th

---

[8]It's possible that firm size has a non-linear effect on price informativeness. In the Appendix, I show the cross-sectional regression results are quantitatively unchanged by including fixed efffxects for deciles of market capitalization, formed each quarter.

percentile of passive ownership in 2018 has 1.05 lower DM than a firm in the 10th percentile of passive ownership in 2018. For reference, 1.05 is approximately $1/5^{th}$ of DM's whole sample standard deviation. Column 4 shows that the relationship between passive ownership and DM is quantitatively unchanged by value weighting observations.

	QVS		DM	
	(1)	(2)	(3)	(4)
Passive Ownership	-39.48***	-26.23***	-4.78***	-4.96***
	(3.06)	(9.72)	(0.61)	(1.25)
Observations	430,489	430,489	430,489	430,489
R-Squared	0.23	0.24	0.22	0.28
Firm + Year/Quarter FE	✓	✓	✓	✓
Matched to Controls	✓	✓	✓	✓
Firm-Level Controls	✓	✓	✓	✓
Weight	Equal	Value	Equal	Value

Table 2 Cross-sectional regression of price informativeness on passive ownership. Table with estimates of β from:

$$\text{Price informativeness}_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$$

where Price informativeness$_{i,t}$ is either $QVS_{i,t}$ or $DM_{i,t}$. Controls in $X_{i,t}$ include age, one-month lagged market capitalization, returns from t-12 to t-2, one-month lagged book-to-market ratio, total institutional ownership, CAPM beta, CAPM R-squared, total volatility and idiosyncratic volatility. Passive ownership is expressed as a decimal, so $0.01 = 1\%$ of firm i's shares are owned by passive funds. Standard errors double clustered at the firm and year-quarter level in parenthesis.

3.2 Evidence from options

Columns 1 and 2 of Table 2 show that stocks with more passive ownership have relatively more volatility on earnings announcement days. If options markets internalize this relationship, then we would expect options exposed to earnings announcement risk to be relatively more expensive for high passive stocks. To quantify this effect, I adapt Kelly et al. (2016)'s Implied Volatility Difference (IVD) to measure how much higher implied volatility is for options that span earnings announcements, relative to options that expire the month before and after the announcement.

Specifically, letting τ denote an earnings announcement, I identify regular monthly expiration dates a, b and c, such that $a < \tau < b < c$. Then, I calculate the average implied volatility (in percentage points) $\overline{IV}_i$ for at the money options on stock i with these expiration dates. The final variable of interest, the implied volatility difference, is defined as:

$$IVD_{i,\tau} = \overline{IV}_{i,b} - \frac{1}{2}\left(\overline{IV}_{i,a} + \overline{IV}_{i,c}\right) \tag{4}$$

where higher values of $IVD_{i,\tau}$ imply that options which span earnings announcements are relatively more expensive than those not exposed to earnings announcement risk.[9] The sample for IVD is shorter than for QVS and DM because it relies on OptionMetrics, which begins in 1996.

My preferred interpretation of IVD is built on the same logic as QVS. As fewer investors gather information ahead of earnings announcements, volatility on the announcement day itself should increase (Ganuza and Penalva (2010), Åstebro and Penalva (2022)). If options markets internalize the negative relationship between passive ownership and pre-earnings announcement information gathering, the associated effect on earnings-day volatility should be reflected in higher option prices. This interpretation yields a testable prediction for the relationship between IVD and passive ownership.

Prediction 3: *If passive ownership decreases pre-earnings announcement information gathering, it should cause IVD to increase*

In terms of basic properties, the Appendix shows that average IVD is positive and has increased by about 5 percentage points over the past 25 years. This is consistent with the decline of QVS and DM, suggesting the trends towards decreased pre-earnings announcement price informativeness in Figure 1 are reflected in option prices.

To test prediction 3, I run a regression of IVD on passive ownership and the same controls and fixed effects as Equation 3. Table 3 contains the results. Column 1 shows that, consistent with prediction 3, IVD is positively correlated with passive ownership. The estimated coefficient implies that a firm in the 90th percentile of passive ownership in 2018

---

[9]See the Appendix for step-by-step details on how I construct IVD. One concern with this definition of IVD is that subtracting the average of $\overline{IV}_{i,a}$ and $\overline{IV}_{i,c}$ from $\overline{IV}_{i,b}$ accounts for firm-specific time trends in implied volatility, but not level differences in implied volatility across firms. This concern is partially alleviated by the inclusion of firm fixed effects. In addition, all the results are qualitatively unchanged instead defining the implied volatility difference as a ratio: $\widetilde{IVD}_{i,\tau} = \overline{IV}_{i,b} / \frac{1}{2}\left(\overline{IV}_{i,a} + \overline{IV}_{i,c}\right)$.

has a 2.16 percentage point higher average IVD than a firm in the 10th percentile of passive ownership in 2018. For reference, the whole sample mean of IVD is 5.07, and its standard deviation is 9.37. Column 2 shows the relationship between passive ownership and IVD is even stronger when using value weights instead of equal weights. These results imply that options markets reflect the relationship between passive ownership and pre-earnings announcement price informativeness, corroborating the findings in Table 2.

	IVD		Post Reg FD		AT Controls	
			QVS	DM	QVS	DM
	(1)	(2)	(3)	(4)	(5)	(6)
Passive Ownership	9.81***	17.27***	-33.59***	-4.36***	-28.25***	-4.918***
	(2.65)	(3.26)	(3.52)	(0.71)	(5.05)	(1.13)
Observations	111,415	111,415	250,068	250,068	80,990	80,990
R-Squared	0.35	0.45	0.22	0.22	0.25	0.27
Firm + Year/Quarter FE	✓	✓	✓	✓	✓	✓
Matched to Controls	✓	✓	✓	✓	✓	✓
Firm-Level Controls	✓	✓	✓	✓	✓	✓
Weight	Equal	Value	Equal	Equal	Equal	Equal

Table 3 Corroborating evidence for effect of passive ownership on pre-earnings announcement price informativeness.
Estimates of β from:

$$Outcome_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$$

In Columns 1-2, the left-hand side variable is IVD, while in Columns 3 and 5 it is QVS and in Columns 4 and 6 it is DM. Controls in $X_{i,t}$ include age, one-month lagged market capitalization, returns from t-12 to t-2, one-month lagged book-to-market ratio, total institutional ownership, CAPM beta, CAPM R-squared, total volatility and idiosyncratic volatility. All Columns contain year-quarter fixed effects and firm fixed effects. Columns 3-4 restrict to observations between 2001 and 2018. Columns 5-6 restrict to observations that can be matched to the SEC MIDAS data, and also include controls for the AT measures in Weller (2018). Standard errors double clustered at the firm and year-quarter level in parenthesis.

3.3 Additional robustness

One threat to my OLS regression results is Regulation Fair Disclosure (Reg FD), passed in August 2000, which reduced the early release of earnings information. Even though all the specifications in Table 2 have time fixed effects, this threat remains because Reg FD may have differently affected stocks with more passive ownership. Columns 3 and 4 of

Table 3 show that the OLS estimates are qualitatively unchanged when using only earnings announcements between 2001 and 2018, evidence that Reg FD is not driving my results.

Another threat to my OLS regressions is the rise of algorithmic trading (AT), which can reduce the returns to informed trading (Weller, 2018). This could threaten my results if e.g., high passive stocks also have high AT activity due to ETF arbitrage. Columns 5 and 6 replicate the baseline regressions, but explicitly control for the AT measures in Weller (2018).[10] The OLS estimates are not significantly changed by including these controls, evidence that a correlation between AT activity and passive ownership is not driving my results.

4   Causal evidence

One limitation of the regressions in Table 2 is that passive ownership is not randomly assigned in the cross-section of stocks. It's possible, therefore, that passive ownership increased the most in stocks with low pre-earnings announcement price informativeness and causality runs the other way. For example, Figure D.1 in the Appendix shows that passive ownership has a strong positive correlation with market capitalization. Large firms may be harder to value, because e.g., they are made up of multiple business segments (Cohen and Lou, 2012). In this case, we might expect large firms to have lower pre-earnings announcement price informativeness for reasons unrelated to their larger passive ownership share.

In my setting, reverse causality seems unlikely because a significant amount of passive ownership is determined by mechanical rules e.g., being one of the 100 lowest volatility stocks in the S&P 500 (Invesco's S&P 500 Low Volatility ETF, SPLV) or having one of the 1000 largest float-adjusted market capitalizations in the Russell 3000 (iShares' Russell 1000 ETF, IWB). Ex-ante, it's not obvious why the intersection of these rules would select stocks with low pre-earnings announcement price informativeness.

Even so, the cross-sectional correlations do not conclusively establish a causal link between passive ownership and pre-earnings announcement price informativeness. To establish causality, I construct two instruments for passive ownership using changes in index mem-

---

[10]These AT measures are constructed from the SEC's MIDAS data, which starts in 2012. This lack of a long historical time series is why I do not include these as controls in my baseline cross-sectional OLS regressions. See the Appendix for a detailed description of how the AT measures are constructed.

bership due to Russell 1000/2000 rebalancing and S&P 500 additions. Both IV designs are built on the logic of difference-in-differences. To this end, I identify a group of treated firms that experience a mechanical increase in passive ownership due to an index change. Then, to alleviate concerns of selection bias, I identify a corresponding group of similar control firms that do not. Finally, I instrument for passive ownership using the expected change in passive ownership from switching indices. My IV estimates confirm a negative causal relationship between passive ownership and pre-earnings announcement price informativeness.

4.1   Identifying treated & control firms

Until 2006, at the end of each May, FTSE Russell selected the 1000 largest stocks by float-adjusted market capitalization to be members of the Russell 1000, and selected the next 2000 largest stocks to be members of the Russell 2000. To reduce turnover between the two indices, in 2007, Russell switched to a banding rule. Now, as long as a potential switcher's market capitalization is within $\pm$ 2.5% of the Russell 3000E's total market capitalization, relative the 1000th ranked stock (the upper and lower bands), it will remain in the same index as the previous year.

Moving from the 1000 to the 2000 mechanically increases the fraction of a firm's shares that need to be held by passive funds. One reason is that switchers go from being one of the smallest stocks in a value-weighted index of big stocks, to one of the biggest stocks in a value-weighted index of small stocks, significantly boosting their index weight (Appel et al., 2016). Another reason is that the Russell 2000 has a higher average passive ownership share than the Russell 1000 (Pavlova and Sikorskaya, 2022).

In this setting, the ideal difference-in-differences design would compare potential switchers to those that actually switched. Identifying possible switchers is not straightforward, however, as the data that Russell uses to compute May market capitalizations is not made available to researchers. To compute a proxy for the Russell May market capitalizations, I follow the method in Coles et al. (2022).[11] Using their May market capitalization proxy, I correctly predict Russell 1000/2000 index membership for 98.63% of Russell 3000 stocks in my sample.

I also follow Coles et al. (2022) to identify groups of treated and control firms. Each

---

[11]I would like to thank the authors for sharing their replication code with me. The Appendix contains a step-by-step explanation of how I compute the May market capitalization proxy.

May, I create a cohort of possible switchers that were in the Russell 1000 the previous year. From 1990-2006, this is firms within ± 100 ranks around the 1000th ranked stock, while from 2007-2018, this is firms within ± 100 ranks of the lower band. The treated firms are those that ended up switching, while the control firms are those that stayed in the 1000.[12] A firm can be treated more than once if it switches to the 2000, goes back to the 1000 and then switches back to the 2000 at some future date. Control firms can appear more than once if they are near the index assignment threshold in multiple years, but don't switch.[13] These filters yield about 700 treated firms and 600 control firms.

My second set of treated and control firms are built using additions to the S&P 500. For a firm to be added to the index, it has to meet criteria set out by S&P, including a sufficiently large market capitalization, being representative of the US economy and financial health. Once a firm is added to the S&P 500, it experiences an increase in passive ownership, as the index mutual funds and ETFs tracking the index need to buy the stock.[14]

One concern with defining treatment as being added to the S&P 500 is that these changes are determined by a committee, rather than a mechanical rule. Therefore, it's possible that the increase in passive ownership is not fully exogenous to firm fundamentals. To ameliorate this concern, I follow the logic in the previous subsection and carefully choose a set of comparable control firms.

I start by obtaining daily S&P 500 index constituents from Compustat between 1990 and 2017. Motivated by the size and representativeness selection criteria, I identify a group of control firms that reasonably could have been added to the index at the same time as the treated firms. To this end, at the time of index addition, I sort firms into three-digit SIC industries and within each industry, form quintiles of market capitalization. For each added

---

[12]Another natural set of treated firms are those that switch from the Russell 2000 to the Russell 1000 because they experience a decrease in passive ownership. In the Appendix, I show that within one year of switching, this decrease is totally offset by the time trend toward increased passive ownership.

[13]One concern with defining treatment as switching to the 2000 instead of switching to *and staying in* the 2000 is that firms may change their index status in the post-treatment period. One could instead require treated firms to be out of the 2000 for the whole pre-treatment period and in the 2000 for the entire post-treatment period. This, however, is not my preferred specification, as whether or not a firm stays in/out of a particular index is endogenous and future index status is not known at the time of index addition.

[14]A natural extension is to examine firms that are dropped from the S&P 500 index, which experience a decrease in passive ownership. As I discuss in the Appendix, this is a less ideal setting than index addition, as firms are usually dropped from the index for (1) poor performance or lack of liquidity, which is related to firm fundamentals or (2) being acquired by or merged with another firm in which case there will be no post-index-deletion observations.

firm, the first set of control firms are those in the same three-digit SIC industry and same quintile of industry market capitalization which are outside the S&P 500 index. I also form a second control group of firms in the same 3-digit SIC industry and market capitalization quintile, but that are already in the S&P 500 index. Cohorts are defined as all matched treated and control firms in the same industry and size bucket in a given month.

As with the Russell 1000/2000 switchers, control firms can appear in more than one cohort. For example, the same firm outside the index can be a control for multiple firms added to the index at different points in time. These filters yield about 500 treated firms, 600 control firms in the index and 2,000 control firms out of the index.

4.2 Effect of treatment on passive ownership

The next step in building the IV is quantifying the effect of being treated on passive ownership (the first stage). To visualize this, the top left panel of Figure 2 compares the level of passive ownership around the index rebalancing month between Russell switchers and stayers. Within each cohort, I subtract the average level of passive ownership to ease comparison across years. Reassuringly, pre-addition changes and levels of passive ownership are similar between the treated and control groups. The treated firms, however, experience an increase in passive ownership at $t = 0$ and remain at a higher level of passive ownership over the next 12 months.[15]

The top right panel of Figure 2 shows the level of passive ownership for S&P 500 additions and matched control firms around the month of index rebalancing. Again, within each cohort, I subtract the average level of passive ownership to facilitate the comparison across industry-size buckets and across time. All three groups of firms have similar average pre-addition changes in passive ownership, although the firms already in the index have a higher average level of passive ownership. After index addition, the added firms experience an increase in passive ownership, essentially going from the level of the control firms outside

---

[15]Russell reconstitutions always coincide exactly with the end of a calendar quarter, so Figure 2 only plots data points for months with S12 filings (the last month of each calendar quarter).
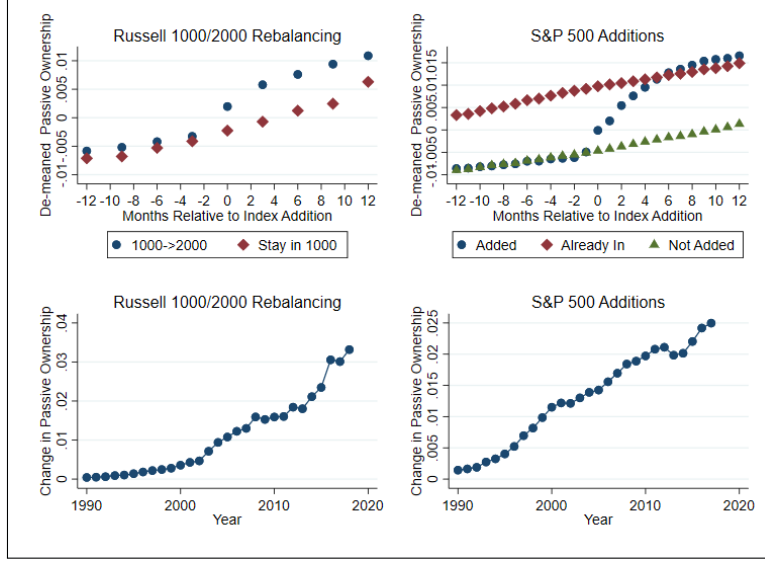
the index to the level of control firms inside the index.[16]

Figure 2. Effect of treatment on passive ownership. Top left panel: Average level of passive ownership for firms that stay in the Russell 1000 ("Stay in 1000") and firms that switched from the Russell 1000 to the Russell 2000 ("1000 $\rightarrow$ 2000"). Top right panel: Average level of passive ownership for control firms out of the index ("Not Added"), control firms in the index ("Already In") and added firms ("Added"). For both top panels, passive ownership is demeaned within each group of matched treated and control firms. Bottom left panel: 5-year moving average change in passive ownership for Russell 1000 to 2000 switchers from month $t = -3$ to $t = 3$ around the reconstitution date by year. Bottom right panel: 5-year moving average change in passive ownership for S&P 500 additions from month $t = -3$ to $t = 3$ around the index rebalancing date by year.

As shown by Figure 1, aggregate passive ownership has been increasing over time. One consequence of this trend is that the increase in passive ownership associated with switching from the Russell 1000 to the Russell 2000 and being added to the S&P 500 has grown over my sample. The two bottom panels of Figure 2 show the average change in passive ownership for treated firms between month $t = -3$ and month $t = 3$ relative to the index reconstitution. For Russell 2000 switchers, the increase grew from almost nothing in 1990 to about 3.5% by

---

[16]S&P 500 index additions do not always coincide with the end of a calendar quarter. Given that the S12 data I use to quantify passive ownership is quarterly, I do not always know the level of passive ownership exactly 3 months before, in the month of and 3 months after index addition for all treated and control firms. In constructing Figure 2, between quarter ends, I fix passive ownership at its last reported level each month. This is why passive ownership appears to increase slowly around the month of index addition, as I am averaging across observations with differences in time until the first set of post-index-addition S12 filings are released.

2018. The change in passive ownership accelerated after 2000, the year IWM (the largest Russell 2000 ETF) was launched. The change in passive ownership from being added to the S&P 500 exhibits a similar trend.

Given the trends in the bottom two panels of Figure 2, my IV design needs to account for the time series variation in passive ownership associated with index changes. To this end, I create a proxy for the expected increase in passive ownership from being treated, which I call Passive Gap$_{i,t}$. For the Russell switchers, it is defined as the difference in passive ownership between firms in the Russell 1000 and the Russell 2000 within $\pm$ 100 ranks of the 1000th ranked firm in March (the last S12 filing date before index rebalancing). For the S&P 500 additions, Passive Gap$_{i,t}$ is the difference in passive ownership between the matched control firms in the index and out of the index, three months before the treated firm is added to the index. If at the time of index addition there are not matched control firms *both* in and out of the index, I use the average Passive Gap$_{i,t}$ for all other added firms that year.

4.3   Instrumental variables design

The logic behind my IV is to use being treated, the post-treatment period and Passive Gap$_{i,t}$ to instrument for passive ownership. The three key pieces of the IV are therefore: (1) the instrumented change in passive ownership (2) the IV specification and (3) the reduced form:

$$Passive_{i,t} = \alpha + \beta_1 Post_{i,t} + \beta_2 \text{Passive Gap}_{i,t} \times Treated_{i,t} \times Post_{i,t} + FE + \epsilon_{i,t} \quad (5)$$

$$Outcome_{i,t} = \alpha + \beta_3 \widehat{Passive}_{i,t} + FE + \epsilon_{i,t} \quad (6)$$

$$Outcome_{i,t} = \alpha + \beta_4 Post_{i,t} + \beta_5 \text{Passive Gap}_{i,t} \times Treated_{i,t} \times Post_{i,t} + FE + \epsilon_{i,t} \quad (7)$$

where $Outcome_{i,t}$ is QVS or DM and $Post_{i,t}$ is an indicator for observations after the index change. Following Coles et al. (2022), all three equations include firm-by-cohort fixed effects. I restrict to data within three years before or after index addition, but exclude three months immediately before or after the event to avoid index inclusion effects (Morck and Yang (2001), Madhavan (2003)). Passive Gap$_{i,t} \times Treated_{i,t}$ is not included in the first stage or reduced form because it is constant within each firm-cohort and therefore is fully explained by the fixed effects. Standard errors are double clustered at the firm and quarter level.

Panel A of Table 4 shows the results of the IV built on Russell rebalancing and Column

1 shows the first stage. The associated F-statistic is large, which is not surprising given the increase in passive ownership pictured in Figure 2.[17] The coefficient on $Post \times Treated \times PassiveGap$ is larger than 1, implying that $PassiveGap$ tends to understate the actual change in passive ownership associated with switching to the Russell 2000. One reason for this is that there are three years of post-rebalancing observations for the treated firms and the trend toward increased passive ownership has been steeper for Russell 2000 firms than Russell 1000 firms.

Column 2 is the instrumental variables (IV) specification with QVS on the left hand side. The effect of passive ownership on QVS is negative, consistent with the cross-sectional regression results. The IV estimate of -99.28 is about 2.5 times the OLS estimate of -39.48. In Column 4, the analogue to Column 2 for DM, the IV estimate of -13.01 is also negative and about 2.5 times the OLS regression coefficient of -4.78. Importantly, in the presence of the reverse causality described at the start of this section, we would expect the OLS estimates to be biased upward in magnitude. The fact that the IV estimates are more negative than the OLS estimates suggest that the latter are not materially biased by this endogeneity concern.

Although the reduced-form estimates for QVS and DM are the same sign as the OLS estimates, they are statistically insignificant. It is not obvious, however, that the reduced form estimates should be directly comparable with the OLS results. One reason is that the cross-sectional regressions use the *level* of passive ownership, while the reduced form uses the expected *change* in passive ownership from index changes (i.e., Passive $\text{Gap}_{i,t}$), which may only be informative about the sign of the treatment effect. I present a more detailed discussion of the differences between the IV and RF specifications in the next subsection and in the Appendix.

Panel B of Table 4 is the analogue of Panel A using S&P 500 additions. Consistent with Panel A, the first stage regression in Column 1 has a large F-statistic. Columns 2 and 4 are the IV regressions, which both show a negative and statistically significant relationship between passive ownership and pre-earnings announcement price informativeness. Like Panel A, these point estimates are larger in magnitude than the cross-sectional regression estimates

---

[17]Because I am using both $Post$ and $Post \times Treated \times$ Passive $\text{Gap}_{i,t}$ as instruments for passive ownership, the time trend and the change in passive ownership associated with being treated in Figure 2 are jointly driving the large magnitude of the F-statistic in Table 4. In the first stage, both $Post$ and $Post \times Treated \times$ Passive $\text{Gap}_{i,t}$ are individually statistically significant at the 1% level.

| | | Panel A: Russell Rebalancing | | | |
| | | QVS | | DM | |
	First Stage	IV	RF	IV	RF
Post x Treated	1.819***		-53.81		-17.64
x Passive Gap	(0.174)		(59.16)		(11.40)
Passive Ownership		-99.28***		-13.01***	
		(19.92)		(3.41)	
Observations	31,030	31,030	31,030	31,030	31,030
F-statistic	202				
		Panel B: S&P 500 Additions			
		QVS		DM	
	First Stage	IV	RF	IV	RF
Post x Treated	0.547***		-24.48		-3.32
x Passive Gap	(0.044)		(18.11)		(4.33)
Passive Ownership		-158.17***		-18.96***	
		(17.36)		(5.37)	
Observations	185,494	185,494	185,494	185,494	185,494
F-statistic	388				
Cross-sectional regression estimate		-39.48	-39.48	-4.78	-4.78

Table 4 IV estimates for effect of passive ownership on pre-earnings announcement price informativeness. Estimates from:

$$Passive_{i,t} = \alpha + \beta_1 Post_{i,t} + \beta_2 \text{Passive Gap}_{i,t} \times Treated_{i,t} \times Post_{i,t} + FE + \epsilon_{i,t}$$

$$Outcome_{i,t} = \alpha + \beta_3 \widehat{Passive}_{i,t} + FE + \epsilon_{i,t}$$

$$Outcome_{i,t} = \alpha + \beta_4 Post_{i,t} + \beta_5 \text{Passive Gap}_{i,t} \times Treated_{i,t} \times Post_{i,t} + FE + \epsilon_{i,t}$$

where $Outcome_{i,t}$ is QVS or DM and $Post_{i,t}$ is an indicator for observations after the index change. Passive Gap$_{i,t}$ is the expected change in passive ownership from being treated. Column 1 in each panel is a first-stage regression. Columns 2 and 4 are instrumental variables regressions. Columns 3 and 5 are reduced-form regressions. Panel A contains observations from Russell rebalancing, while Panel B contains observations from S&P 500 additions. FE are fixed effects for each cohort. Standard errors, double clustered at the firm and quarter level, are in parenthesis.

by a factor of about 3. One possible reason for this is that my measure of passive ownership understates the true level of passive ownership firms experience after being added to the

S&P 500 index.[18]

4.4 Discussion

The assumption underlying my IV strategy is that index addition only affects price informativeness through its associated effect on passive ownership. One threat to this is that index switching/addition may be associated with an increase in total institutional ownership (Boone and White, 2015). Gloßner (2019) shows, however, that although there is an increase in passive ownership following Russell index reconstitution events, there is little change in overall institutional ownership.[19] To further alleviate the concern that institutional ownership is driving my results, in the Appendix, I show the IV results are quantitatively unchanged by including the institutional ownership ratio on the right-hand side.

A second issue, as discussed above, is that in both panels of Table 4, the IV is always significant, while the reduced form is insignificant. The concern is that, as discussed in Chernozhukov and Hansen (2008), a significant IV with an insignificant reduced form potentially indicates weak instruments. This is likely not a problem in my setting, as the first stage is very strong ($F > 200$ for the Russell 1000 to 2000 switchers and $F > 300$ for the S&P 500 additions). In the Appendix, following Lochner and Moretti (2004), I show why we might expect from a purely econometric perspective the reduced form to be less significant than the IV. I also discuss how the presence of shadow indexing (Mauboussin, 2019) could explain why the IV is significant but the reduced form is not.

An additional concern with the results in Table 4 is that many previous studies have used switching between from the Russell 1000 to the Russell 2000 and additions to the S&P 500 as natural experiments when studying the effects of passive ownership on a variety of outcomes e.g,. corporate governance, disclosure and investment. As discussed in Heath et al. (2020), this re-use of natural experiments can lead to false positives in later studies. The particular

---

[18]Suppose that what truly matters for price informativeness is the total amount of passive ownership. My measure, $Passive_{i,t}$, only captures funds that are explicitly passive, and misses e.g., shadow index funds (Mauboussin, 2019), as well as institutions that do index replication internally. If firms added to the S&P 500 experience an increase in these types of non-explicit passive ownership as well, we might expect their price informativeness to decline more than would be explained by index fund holdings alone.

[19]A related concern, raised in Appel et al. (2020), is that for the Russell switchers, the treatment is correlated with firm size. Given that my results are similar using both switching from the Russell 1000 to the Russell 2000, which applies to shrinking firms and S&P 500 index addition, which applies to growing firms, I find it unlikely that a pure size effect is driving my results.

issue is that my results could be driven by the effects of passive ownership on previously documented outcomes, rather than passive ownership per se.

The solution proposed by Heath et al. (2020) is to use t-statistics which explicitly account for how many times the natural experiment has been re-used. Table 4 shows that almost all of my IV t-statistics are over 3.62. This implies that even if previous research had looked at the effect of these index changes on over 300 other distinct outcomes, my results are unlikely to be spurious. Further, the Russell switcher IV yields similar point estimates to the S&P addition IV, even though these index changes have different implications for other known outcomes (e.g., firm size), again allaying concerns that my results are driven by factors other than passive ownership.

Finally, to confirm the robustness of my causal estimates, in the Appendix I build on the logic in Bernstein (2015) to develop an alternative IV strategy. Specifically, I use the interaction between a firm's CAPM beta at the end of March and cumulative market return from the start of April to the Russell ranking date in May to instrument for passive ownership over the year starting in July. Crucially, the IV regression includes dummy variables for deciles of firm size, formed at the end of March, interacted with year dummies. With these fixed effects, the instrument is leveraging the fact that firms which are similar in size in March, but have differential exposure to market returns from April to late May (based on their CAPM beta) will end up in different indices for index families that rebalance around the end of June (e.g., Russell and S&P). This alternative instrumentation approach is useful because it does not condition on future index membership and because it exploits a different source of variation than the two IVs above (cross-sectional vs. time series).

For this instrument, the exclusion restriction is that a firm's CAPM beta times the market return from April to May is exogenous to price informativeness in year following July. This assumption would be less plausible if stocks with high beta had high idiosyncratic volatility. To partially address this concern, I explicitly control for idiosyncratic volatility over the period used to compute CAPM beta. Reassuringly, in this alternative IV, the first stage and reduced form are both strong, and the causal estimates are comparable in magnitude to those found in this section.

5 Mechanisms

My preferred explanation for why passive ownership decreases pre-earnings announcement price informativeness is that passive investors gather less firm-specific information. To support this claim, I start by showing that the negative relationship between passive ownership and pre-earnings announcement price informativeness is coming from the firm-specific component of information. Then, I present both direct and indirect evidence which suggests that passive owners demand less information about firm specific news. Next, I leverage pre-earnings announcement trading volume to distinguish between theories that relate private information gathering to trade in financial markets. Finally, I discuss why the equilibrium response of non-passive investors doesn't fully offset the effects of passive ownership.

5.1 Systematic vs. idiosyncratic news

Recent work on the effect of passive ownership on information gathering has highlighted a trade off that passive investors face in terms of the information they collect (Cong et al. (2020), Glosten et al. (2021)). Because passive investors have diversified portfolios, systematic news is more important to them, so they will optimally collect more systematic information and less firm-specific news. To test these theories, I examine the response of stock prices to news of a specific size. The intuition is that if investors have less precise beliefs before an announcement, they will update significantly afterwards, leading to a larger price change (Ganuza and Penalva, 2010). Given that passive investors are less likely to collect stock-specific information, this effect should be especially strong for the firm-specific component of news. This yields a testable prediction for the effect of passive ownership on earnings responses.

Prediction 4: Stocks with more passive ownership should respond more to earnings news of a given size. This effect should be especially strong for firm-specific news

To test prediction 4, I use the following earnings-response regression (Kothari and Sloan, 1992) to quantify the market's reaction to a standardized measure of earnings news:

$$r_{i,t} = \alpha + \beta SUE_{i,t} + \theta Passive_{i,t} + \delta \left(SUE_{i,t} \times Passive_{i,t} \right) + \gamma X_{i,t} + \phi_t + \psi_i + \epsilon_{i,t} \qquad (8)$$

where $r_{i,t}$ denotes the market-adjusted return on the first day investors could trade on earnings information (in percentage points), Winsorized at the 1% and 99% level by year.[20] $SUE_{i,t} = \frac{E_{i,t} - E_{i,t-4}}{\sigma_{(t-1,t-8)}(E_{i,t} - E_{i,t-4})}$ where $E_{i,t}$ is earnings-per-share from the IBES unadjusted detail file (i.e., "street" earnings). In words, the numerator is the year-over-year (YOY) earnings growth, while the denominator is the standard deviation of YOY earnings growth over the past 8 quarters.[21] In the Appendix, I show that the market reaction to earnings news of a given size is about 3 times as large now as it was in the early 1990s.

I also run versions of Equation 8 breaking SUE into positive and negative components and decomposing the earnings news into a systematic and idiosyncratic component using the method in Glosten et al. (2021). This is done by regressing firm-level SUE on market-wide SUE and SIC-2 industry-wide SUE in five year rolling windows. The systematic component of earnings is the predicted value from this regression, while the idiosyncratic component is the residual.

Table 5 contains the regression results. Column 1 shows that, consistent with prediction 4, δ is positive and economically large, meaning that high passive ownership stocks are more responsive to earnings news. Column 2 shows that this effect is stronger for negative news than positive news. Column 3 shows that the increased responsiveness of high passive stocks to earnings news is concentrated in the firm-specific component, also consistent with prediction 4. Columns 4-6 confirm these results are robust to value weighting observations.

5.2 Information gathering

A natural explanation for a decrease in price informativeness is a decline in the share of informed investors or the precision of investors' signals (Grossman and Stiglitz (1980), Kyle (1985)). Passive managers, as well as investors in passive funds, lack strong incentives to gather and consume firm-specific information because these funds trade on mechanical rules, such as S&P 500 index membership (State Street's S&P 500 ETF Trust, SPY), or having

---

[20]In the Appendix, I show that the results in this subsection are similar when instead using cumulative returns in windows of up to 5 days after the earnings-announcement.

[21]I compute SUE this way, following Novy-Marx (2015), because it avoids (1) using prices as an input, whose average informativeness has changed over time and (2) using analyst estimates of earnings as an input, whose average accuracy has also changed over time. Using this method, the average absolute value of $SUE_{i,t}$ is roughly constant over my sample, except for large spikes during the tech boom/bust as well as during the Global Financial Crisis.

	(1)	(2)	(3)	(4)	(5)	(6)
SUE $\times Passive$	2.64*** (0.24)			2.13*** (0.26)		
$SUE \times 1_{SUE>0}$ $\times Passive$		1.15*** (0.37)			1.60*** (0.42)	
$|SUE| \times 1_{SUE\leq0}$ $\times Passive$		-3.63*** (0.34)			-2.12*** (0.63)	
$Sys.SUE \times 1_{Sys.SUE>0}$ $\times Passive$			1.40 (1.00)			1.00 (2.40)
$|Sys.SUE| \times 1_{Sys.SUE\leq0}$ $\times Passive$			0.67 (2.87)			1.5 (4.67)
$Idio.SUE \times 1_{Idio.SUE>0}$ $\times Passive$			1.04** (0.43)			1.47*** (0.46)
$|Idio.SUE| \times 1_{Idio.SUE\leq0}$ $\times Passive$			-3.89*** (0.32)			-2.81*** (0.48)
Observations	333,814	333,814	333,814	333,814	333,814	333,814
R-squared	0.06	0.06	0.06	0.03	0.03	0.03
Firm + Year/Quarter FE	✓	✓	✓	✓	✓	✓
Matched to Controls	✓	✓	✓	✓	✓	✓
Firm-Level Controls	✓	✓	✓	✓	✓	✓
Weight	Equal	Equal	Equal	Value	Value	Value

Table 5 Passive ownership and earnings responses. Estimates from:

$$r_{i,t} = \alpha + \beta SUE_{i,t} + \theta Passive_{i,t} + \delta \left(SUE_{i,t} \times Passive_{i,t} \right) + \gamma X_{i,t} + \phi_t + \psi_i + \epsilon_{i,t}$$

where $r_{i,t}$ is the market-adjusted return (in percentage points) on the effective earnings announcement date. Controls in $X_{i,t}$ include age, one-month lagged market capitalization, returns from t-12 to t-2, one-month lagged book-to-market ratio, total institutional ownership, CAPM beta, CAPM R-squared, total volatility and idiosyncratic volatility. All Columns contain year-quarter fixed effects and firm fixed effects. Standard errors double clustered at the firm and year-quarter level in parenthesis.

10 years of increasing dividend payments (Vanguard's Dividend Appreciation ETF, VIG). Further, because these funds are well diversified, even if they are traded by information-motivated investors, they are more likely to be used for bets on systematic, rather than firm-specific, information. This logic yields an empirical prediction for the relationship between passive ownership and information gathering.

Prediction 5: *Passive ownership should cause firm-specific information gathering to decline*

One way to quantify information gathering is with Bloomberg terminal searches for specific tickers. As discussed by Ben-Rephael et al. (2017), these searches capture attention

by institutional investors, who are the main users of Bloomberg's products. The timing of when investors will search for information relative to earnings announcements, however, is not obvious. Attentive investors may search (1) right before earnings are released to e.g., make a bet ahead of the announcement (2) on the earnings announcement date to e.g., bet on the announcement news or (3) some time after earnings are released to e.g., bet on a re-interpretation the announcement news.

Rather than trying to distinguish between these channels, I perform a more general test. At the stock/month level, I ask whether stocks with more passive ownership have fewer Bloomberg terminal searches than stocks with less passive ownership. To this end, I run a regression of the continuous abnormal institutional attention measure from Ben-Rephael et al. (2017) ($AIAC$) on passive ownership.[22] The sample is stock/month observations between 2010 and 2018 that can be linked between Bloomberg and CRSP on ticker. All the controls and fixed effects are identical to Equation 3.

Columns 1 and 2 of Table 6 contains the results. Consistent with prediction 5, passive ownership is correlated with fewer Bloomberg searches. In terms of magnitudes, a 15% higher level of passive ownership implies -0.23 lower $AIAC$, which is about 24% of its whole sample mean of 0.97. If the mechanism behind this decline was just that passive investors gather no information, this estimate is roughly in line with the 15% decrease in information gathering we would expect ex-ante. Institutional investors (13F filers), which is what $AIAC$ is designed to capture, however, only hold about 70% of the US stock market. So, if the rise of passive ownership was a re-allocation among institutional investors, we would expect to see a decline of 15%/70% $\approx$ 21%, which is almost exactly what we see in Table 6.

As an alternative way to measure investors' learning behavior, I examine downloads of SEC filings, with fewer downloads implying decreased gathering of fundamental information (Loughran and McDonald, 2017). Specifically, I define $Downloads_{i,t}$ as one plus the natural logarithm of the number of non-robot downloads, measured using the method in Loughran and McDonald (2017) and obtained from their website. The sample runs from 2003-2015, excluding the data lost/damaged by the SEC from 9/2005-5/2006, and I match the downloads to CRSP/Compustat merged on CIK. As with the regressions using Bloomberg ticker searches, the unit of observation is firm-month.

---

[22]These results are robust to instead using the other measures from Ben-Rephael et al. (2017) e.g., abnormal institutional attention (AIA) or the raw Bloomberg search intensity data.

	Bloomberg		Downloads		CAT	
	(1)	(2)	(3)	(4)	(5)	(6)
Passive Ownership	-1.52***	-2.29**	-1.14***	-1.41**	-11.49***	-10.25***
	(0.37)	(0.89)	(0.20)	(0.65)	(3.21)	(2.77)
Observations	58,629	58,629	533,099	533,099	407,283	407,283
R-Squared	0.54	0.50	0.81	0.89	0.08	0.15
Firm + Year/Quarter FE	✓	✓	✓	✓	✓	✓
Matched to Controls	✓	✓	✓	✓	✓	✓
Firm-Level Controls	✓	✓	✓	✓	✓	✓
Weight	Equal	Value	Equal	Value	Equal	Value

Table 6 Mechanisms. Estimates of β from:

$$Outcome_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$$

For Columns 1-2, the left-hand side variable is $AIAC$, the measure of continuous abnormal institutional attention from Ben-Rephael et al. (2017) built on Bloomberg searches. For Columns 3-4 the left-hand side variable is one plus the natural logarithm of the number of non-robot downloads from Loughran and McDonald (2017). For Columns 5-6, the left-hand side is cumulative abnormal pre-earnings turnover. Controls in $X_{i,t}$ include age, one-month lagged market capitalization, returns from t-12 to t-2, one-month lagged book-to-market ratio, total institutional ownership, CAPM beta, CAPM R-squared, total volatility and idiosyncratic volatility. All Columns contain year-quarter fixed effects, ϕ_t, and firm fixed effects ψ_i. Standard errors double clustered at the firm and year-quarter level in parenthesis.

Columns 3 and 4 of Table 6 are the analogue of Columns 1 and 2, but have $Downloads_{i,t}$ on the left-hand side. Also consistent with prediction 5, the estimated coefficient is negative and statistically significant, evidence that passive ownership is correlated with less investor attention.[23] In terms of magnitudes, a 15% higher level of passive ownership implies a decrease in $Downloads_{i,t}$ of -0.17, which is modest relative to its whole sample standard deviation of roughly 1.3. This magnitude, however, is harder to interpret than the results on Bloomberg searches, as we don't know who is downloading these SEC filings and whether or not they themselves are investors.

5.3 Sell-side analyst coverage

The relationship between passive ownership and information demand is intuitive. Investors buying an S&P 500 ETF probably care less about firm-specific fundamentals than

---

[23]This result is consistent with Israeli et al. (2017) and Coles et al. (2022), who also show that passive ownership is negatively correlated with downloads of SEC filings.

people buying the underlying stocks. In addition to this direct effect on information gathering, however, a change in demand may have corresponding equilibrium effects on the supply of information, which can be measured using sell-side analyst coverage (Martineau and Zoican, 2021).

To fix ideas, suppose the results in the previous subsection imply that passive ownership shifts the demand curve for information inward. Given that information is not costless to produce, we expect it to have an upward sloping supply curve. All else equal, therefore, this decrease in demand should also lower the equilibrium supply of information.[24] This logic yields a testable prediction for the relationship between passive ownership and information supply, as measured by sell-side analyst coverage.

Prediction 6: *Passive ownership should be correlated with decreased quantity and quality of sell-side analyst coverage*

To test prediction 6, I run versions of my baseline OLS regression (Equation 3) with measures of information production by sell-side analysts on the left-hand side. The sample is all quarterly earnings announcements in IBES, further restricting to observations that can be (1) matched to CRSP (2) have at least 3 estimates of earnings-per-share (3) have a non-missing value for realized earnings per share and (4) have a non-missing closing price on the last trading day before the earnings announcement in CRSP. Within each forecast period, I take the last statistical period (i.e., the last set of estimates before the earnings information is released).

Table 7 contains the results. Column 1 shows that higher passive ownership is correlated with lower analyst coverage, consistent with prediction 6. This mirrors Israeli et al. (2017) and Coles et al. (2022), who also show that ETF ownership is negatively correlated with the number of analyst estimates. Column 2 shows that passive ownership is correlated with a larger standard deviation of analyst estimates. Increased forecast dispersion is evidence of more uncertainty about the fundamental value of these firms (Diether et al. (2002), Zhang (2006)), which is also consistent with prediction 6.

One concern with these results, however, is that the increased standard deviation of forecasts is a mechanical function of the decrease in coverage documented in Column 1. To

---

[24]It's possible that the supply curve for information also shifts in response to rising passive ownership. Isolating this effect is difficult, however, without a way to measure the price of information e.g., the cost of analyst reports. My results, therefore, can only speak to the net effect of passive ownership on the supply of and demand for information.

address this, I construct a measure of analyst inaccuracy which explicitly accounts for the increase in dispersion. Specifically, I define inaccuracy as the absolute difference between realized earnings and the mean estimate of earnings, divided by the standard deviation of analysts' estimates. If analysts are producing lower quality information about high passive stocks, we would expect their forecasts to be less accurate, even when accounting for the increase in average uncertainty. Column 3 shows that this prediction holds empirically.

Columns 5 and 6 restrict to the subset of announcements which are covered by analysts who update their forecasts at least once between when they initiate coverage for a fiscal period and when earnings information is released. Columns 5 shows that analysts update their estimates of earnings less frequently for stocks with more passive ownership. In a similar vein, Column 6 shows that the average time between updates is higher for stocks with more passive ownership. Both Columns 5 and 6 imply analysts are expending less effort gathering information on stocks with more passive ownership, which also corroborates prediction 6.

	Num. Est (1)	SD(Est.) (2)	Dist./SD(Est.) (3)	Updates (4)	Time (5)
Passive Ownership	-11.64***	0.72***	1.97***	-0.45***	0.35***
	(1.38)	(0.17)	(0.45)	(0.10)	(0.11)
Observations	216,805	216,805	216,805	133,176	133,176
R-squared	0.79	0.64	0.13	0.26	0.55
Mean	8.62	0.09	2.23	2.23	3.76
St. Dev.	5.94	0.41	2.97	0.45	0.84

Table 7 Passive ownership and coverage by sell-side analysts. Estimates of β from:

$$Outcome_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$$

Num. Est. is the number of analyst estimates, SD(Est.) is the standard deviation of analyst estimates, Dist. is the absolute distance between realized earnings per share and the mean estimate of earnings per share, Updates is the average number of analyst updates within each forecasting period and Time is the average number of days between analyst updates within each forecasting period. Controls in $X_{i,t}$ include age, one-month lagged market capitalization, returns from t-12 to t-2, one-month lagged book-to-market ratio, total institutional ownership, CAPM beta, CAPM R-squared, total volatility and idiosyncratic volatility. All Columns contain year-quarter fixed effects and firm fixed effects. Standard errors double clustered at the firm and year-quarter level in parenthesis. The last two rows of the table present the means and standard deviations of the left-hand side variables.

The findings in Table 7 seem at odds with results in Section 4 because when firms are added to the S&P 500, they receive increased analyst coverage. Decreased incentives to gather or produce firm-specific information could still, however, explain those results. For example, suppose analysts know that after a firm is added to the S&P 500 index, a larger share of its investors are holding it as a part of a well-diversified portfolio. They may, therefore, choose not to expend the effort required to produce an equally accurate measure of firm fundamentals as they would if their clients were taking isolated bets on the stock. Consistent with this hypothesis, in the Appendix I show that even though S&P 500 index addition leads to increased analyst coverage, it also leads to increased dispersion in analyst forecasts and decreased analyst accuracy. Moving from the Russell 1000 to the 2000 causes a drop in analyst coverage and accuracy, but this may be because these firms are shrinking in size.

5.4 Pre-earnings abnormal turnover

Next, I study the relationship between passive ownership and pre-earnings announcement trading volume. This is useful for distinguishing between theories that relate private information collection to trade in financial markets. While the empirical quantity of volume is difficult to rationalize (Cochrane, 2004), a mechanism common to many models is that information is a key motivation for trade. In some models, private information can increase trading volume as it is a source of heterogeneity among investors, and such disagreement makes them willing to trade (Wang, 1994). Too much private information can, however, decrease volume as fears of adverse selection deter uninformed investors from trading (Foster and Viswanathan (1990), Foster and Viswanathan (1993)).

One challenge with bringing these theories to the data is that private information is hard to quantify. My results show, however, that passive ownership decreases pre-earnings announcement information gathering. Therefore, the relationship between passive ownership and pre-earnings announcement trading volume can speak to the relative strength of different channels proposed in the literature. This exercise is similar in spirit to Manela (2014), who examines trading and returns around a different set of public information release events (FDA drug approvals) to distinguish between the competing effects of the speed which with information diffuses through financial markets.

To quantify pre-earnings trading volume, let t denote an effective earnings announcement date. Define turnover T as total daily volume for stock i divided by shares outstanding. Then, define abnormal turnover for firm i, from event time $\tau = -22$ to $\tau = 22$ as:

$$AT_{i,t+\tau} = \frac{T_{i,t+\tau}}{\overline{T_{i,t-22}}} = \frac{T_{i,t+\tau}}{\sum_{k=1}^{252} T_{i,t-22-k}/252} \qquad (9)$$

Where abnormal turnover, $AT_{i,t+\tau}$, is turnover divided by the historical average turnover for that stock over the past year. I use abnormal turnover to account for differences across stocks and within stocks across time. Historical average turnover, $\overline{T_{i,t-22}}$, is fixed at the beginning of the 22-day window before earnings are announced to avoid mechanically amplifying or dampening changes in trading.

In the Appendix, I show that there has been a drop in trading volume throughout the month before earnings announcement over the past 3 decades. To summarize this decline, I define cumulative abnormal pre-earnings turnover as:

$$CAT_{i,t} = \sum_{\tau=-22}^{-1} AT_{i,t+\tau} \qquad (10)$$

In words, CAT is the sum of abnormal turnover from $t-22$ to $t-1$ for firm i around earnings date t. To reduce the influence of outliers, I Winsorize CAT at the 1% and 99% level by year. Between the 1990s and 2010s, average $CAT_{i,t}$ declined by about 1, which can be interpreted as a loss of 1 trading-day's worth of volume over the 22-day window before earnings announcements. The magnitude of this decrease is about 5% of CAT's whole-sample average of 22.

I run a regression of CAT on passive ownership with the same controls and fixed effects as Equation 3.[25] The results are in Columns 5 and 6 of Table 6, which show a strong negative relationship between passive ownership and pre-earnings announcement abnormal turnover. In terms of magnitudes, a firm in the 90th percentile of passive ownership in 2018 has a -2.6 lower CAT than a firm in the 10th percentile of passive ownership in 2018.

---

[25]CAT is similar to Manela (2014)'s measure of cumulative abnormal turnover ($CATO$) before FDA drug approvals. In the Appendix I show I obtain similar results using $CATO$ instead of CAT and I provide a more detailed explanation of CAT's advantages in my setting.

Returning to the theories discussed above, given that passive ownership decreases information gathering, we would expect it to decrease investor heterogeneity and therefore decrease trading volume. On the other hand, decreased information gathering should decrease adverse selection, which would tend to increase trading volume. The negative empirical relationship between passive ownership and pre-earnings trading volume suggests that in my setting, the effect of decreased disagreement tends to dominate the effect of decreased adverse selection.

One concern is that these regression results are mechanical functions of passive ownerships' effect on average trading volume. Passive investors may trade less and therefore more passive ownership leads to less trading overall. The way CAT is defined, however, should prevent passive ownership from causing a mechanical decrease in pre-earnings announcement trading volume, because passive ownership's effect of lowering average trading would be incorporated into past turnover (i.e., the denominator of Equation 9). By focusing on abnormal turnover, these regression results suggest there is a decline in trading volume before earnings announcements relative to firm-level average turnover, allaying this concern.

5.5   Equilibrium response of non-passive investors

Suppose passive investors gather no stock-specific information. Then, as passive ownership increases, we would mechanically expect total information gathering to decrease, holding fixed the behavior of the remaining non-passive investors. If pre-earnings announcement prices have become less informative, however, the returns to becoming informed should have increased. So, a question remains as to why the remaining non-passive investors don't increase their information production to capitalize on this, as occurs in the model of Coles et al. (2022).

A natural reason why non-passive investors wouldn't fully compensate for the decline in information production is that passive ownership's presence makes it harder to profit from private information. This might apply in my setting because as discussed in Ben-David et al. (2018), ETFs (but not non-ETF index funds) increase non-fundamental volatility in the underlying stocks. This could deter informed investors from gathering information, as there is some chance that before the end of their investment horizon, they are hit with a large volatility shock, which forces them to sell at a loss (De Long et al., 1990). An implication of

37

this is that the effects I document in Section 3 should be stronger for ETFs than non-ETF passive funds.

To test this hypothesis, I re-run the baseline OLS regressions (Equation 3), but break passive ownership into ETFs and all passive funds that are not ETFs (i.e., index mutual funds). Panel A of Table 8 shows that ETFs have a larger effect on $DM_{i,t}$ and $QVS_{i,t}$ than non-ETF passive funds. These results are consistent with ETFs increasing the limits to arbitrage by boosting non-fundamental volatility, which lowers the equilibrium response of non-passive investors.

One concern with the results in Panel A of Table 8 is that, because the coefficient on non-ETF passive ownership is insignificant, ETFs explain all the effects of passive ownership on pre-earnings announcement price informativeness. Separating the effects of index mutual funds from ETFs is difficult, however, as they have a correlation coefficient of almost 0.7. Therefore, the coefficient on non-ETF passive ownership could be insignificant because of collinearity with ETF ownership. As an additional check, in Panel B of Table 8, I replicate Panel A, but only include non-ETF passive ownership on the right-hand side. This restores the statistical significance of passive mutual fund ownership, evidence which suggests that my results are not entirely driven by a feature specific to ETFs.

Another possible channel is that passive ownership decreases pre-earnings liquidity. Consistent with this, as shown in Table 6, stocks with more passive ownership have relatively less pre-earnings trading volume. One explanation for decreased liquidity is that the nature of passive ownership makes it harder to hide informed orders. In models like Kyle (1985), the market maker cannot tell whether demand is coming from insiders or noise traders. Unlike this, at the end of every day, investors can observe the exact change in the number of shares held by ETFs. So, if more volume is coming from ETFs, it might be harder to profit from private information in the pre-earnings period, as other investors will be able to detect someone trading on information and push prices against them.

Risk and liquidity are just two of many possible explanations for why non-passive in-

| | Panel A: ETFs and Non-ETF Indexers | | | |
| | QVS | | DM | |
	(1)	(2)	(3)	(4)
ETF	-61.62***	-29.06**	-6.68***	-6.00***
	(5.13)	(13.97)	(0.86)	(1.63)
Non-ETF Passive	1.40	-17.47	-1.21	-2.10
	(5.79)	(20.48)	(1.25)	(2.50)
Observations	429,672	429,672	429,672	429,672
R-Squared	0.232	0.244	0.222	0.276
Firm + Year/Quarter FE	✓	✓	✓	✓
Matched to Controls	✓	✓	✓	✓
Firm-Level Controls	✓	✓	✓	✓
Weight	Equal	Value	Equal	Value
	Panel B: Only Non-ETF Indexers			
	QVS		DM	
	(1)	(2)	(3)	(4)
Non-ETF Passive	-36.37***	-45.66**	-5.30***	-7.91***
	(6.59)	(20.84)	(1.33)	(2.70)
Observations	429,672	429,672	429,672	429,672
R-Squared	0.231	0.244	0.221	0.276
Firm + Year/Quarter FE	✓	✓	✓	✓
Matched to Controls	✓	✓	✓	✓
Firm-Level Controls	✓	✓	✓	✓
Weight	Equal	Value	Equal	Value

Table 8 Breakdown of QVS and DM regression results by type of passive ownership. Table with estimates of b_is from:

$$PriceInformativeness_{i,t} = \alpha + b_1 ETF_{i,t} + b_2\text{Non-ETF Passive}_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$$

where $PriceInformativeness_{i,t}$ is either DM or QVS, $ETF_{i,t}$ is ETF ownership and Non-ETF Passive$_{i,t}$ is ownership by non-ETF passive funds. Controls in $X_{i,t}$ include age, one-month lagged market capitalization, returns from t-12 to t-2, one-month lagged book-to-market ratio, total institutional ownership, CAPM beta, CAPM R-squared, total volatility and idiosyncratic volatility. All ownership measures are expressed as decimals, so 0.01 = 1% of firm i's shares are owned by ETFs. Standard errors double clustered at the firm and year-quarter level in parenthesis.

vestors don't fully compensate for the lack of information gathering by passive investors.[26] These general equilibrium effects, however, are hard to measure. The evidence in Table 6 only speaks to net changes in information demand and, more broadly, the regressions in Tables 2 and 4 only speak to the net effect of passive ownership on pre-earnings announcement price informativeness. So, it is possible that non-passive investors respond by gathering more information, just not enough to fully offset the decrease coming from passive investors. Without being able to see individual investors' attention, however, it is difficult to quantify such effects.

6   Conclusion

In this paper, I propose two ways to measure the fraction of earnings information incorporated into prices before the announcement itself. I show that over the past 30-years, pre-earnings announcement price informativeness has been steadily declining. Passive ownership played an important role in this trend, as taking the average of the point estimates from the OLS and both IVs implies that a 15% increase in passive ownership decreases QVS by 14.87 , a roughly 16% decline relative to its 1990 mean.

My proposed mechanism is that passive investors gather less firm-specific information. To support this claim, I first use earnings response regressions to show that the decline in pre-earnings announcement price informativeness came from firm-specific news. Then, I show direct evidence of decreased information gathering for high passive stocks through Bloomberg terminal searches, downloads of SEC filings and sell-side analyst coverage. Finally, I argue why passive ownership may increase the limits to arbitrage, which prevents non-passive investors from fully offsetting these effects.

Relative to total institutional ownership, passive ownership is still small, owning only about 15% of the US stock market. Even at this level, passive ownership has led to significant changes in how stock prices anticipate the information contained in earnings announcements.

---

[26]Analogously, Haddad et al. (2021) find empirically that non-passive investors do not fully offset passive ownership's tendency to decrease demand elasticity. They propose several explanations for this, including costly information acquisition, risk, institutional mandates, bounded rationality and strategic interactions among investors (e.g., price impact and herding). In addition, Bond and Garcia (2018) develop a model where non-passive investors do not fully offset passive ownership's effect on price informativeness owing to participation costs and strategic complementary in participation decisions.

As passive ownership continues to grow, these effects may be amplified, further changing the way equity markets reflect firm-specific information.

References

Antoniou, C., Li, F. W., Liu, X., Subrahmanyam, A., and Sun, C. (2020). The real effects of exchange-traded funds. *Available at SSRN 3129369*.

Appel, I., Gormley, T. A., and Keim, D. B. (2020). Identification using russell 1000/2000 index assignments: A discussion of methodologies. *Critical Finance Review, Forthcoming*.

Appel, I. R., Gormley, T. A., and Keim, D. B. (2016). Passive investors, not passive owners. *Journal of Financial Economics*, 121(1).

Åstebro, T. and Penalva, J. (2022). Two-sided platform governance: Are founders manipulating the crowd in crowdfunding? *Available at SSRN 3991200*.

Bai, J., Philippon, T., and Savov, A. (2016). Have financial markets become more informative? *Journal of Financial Economics*, 122(3).

Ball, R. and Brown, P. (1968). An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research*, 6(2).

Ben-David, I., Franzoni, F., and Moussawi, R. (2018). Do ETFs Increase Volatility? *Journal of Finance*, 73(6).

Ben-Rephael, A., Da, Z., and Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, 30(9):3009–3047.

Bennett, B., Stulz, R., and Wang, Z. (2020a). Does the stock market make firms more productive? *Journal of Financial Economics*, 136(2):281–306.

Bennett, B., Stulz, R. M., and Wang, Z. (2020b). Does joining the s&p 500 index hurt firms? Technical report, National Bureau of Economic Research.

Bernstein, S. (2015). Does going public affect innovation? *The Journal of finance*, 70(4):1365–1403.

Bhattacharya, A. and O'Hara, M. (2018). Can ETFs Increase Market Fragility? Effect of Information Linkages in ETF Markets. *SSRN Electronic Journal*.

Bond, P. and Garcia, D. (2018). The equilibrium consequences of indexing. *Review of Financial Studies, forthcoming.*

Boone, A. L. and White, J. T. (2015). The effect of institutional ownership on firm transparency and information production. *Journal of Financial Economics*, 117(3).

Breugem, M. and Buss, A. (2019). Institutional investors and information acquisition: Implications for asset prices and informational efficiency. *Review of Financial Studies*, 32(6).

Brogaard, J., Ringgenberg, M. C., and Sovich, D. (2019). The economic impact of index investing. *The Review of Financial Studies*, 32(9):3461–3499.

Buss, A. and Sundaresan, S. (2020). More Risk, More Information: How Passive Ownership Can Improve Informational Efficiency. *SSRN Electronic Journal.*

Campbell, J. Y., Lettau, M., Malkiel, B. G., and Xu, Y. (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance*, 56(1).

Campbell, J. Y. and Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. *the Journal of Finance*, 43(3):661–676.

Chernozhukov, V. and Hansen, C. (2008). Instrumental variable quantile regression: A robust inference approach. *Journal of Econometrics*, 142(1):379–398.

Cochrane, J. H. (2004). Asset pricing: Liquidity, trading, and asset prices. *NBER Reporter Online*, (Winter 2004/05):1–12.

Coffee, J. C. (2007). Law and the market: The impact of enforcement.

Cohen, L. and Lou, D. (2012). Complicated firms. *Journal of financial economics*, 104(2):383–400.

Cohen, R. B., Polk, C., and Vuolteenaho, T. (2003). The value spread. *The Journal of Finance*, 58(2):609–641.

Coles, J. L., Heath, D., and Ringgenberg, M. C. (2022). On index investing. *Journal of Financial Economics*, 145(3):665–683.

Cong, L. W., Huang, S., and Xu, D. (2020). Rise of factor investing: security design and asset pricing implications. Technical report, Working Paper.

Dannhauser, C. D. (2017). The impact of innovation: Evidence from corporate bond exchange-traded funds (ETFs). *Journal of Financial Economics*, 125(3).

Dávila, E. and Parlatore, C. (2018). Identifying Price Informativeness. *National Bureau of Economic Research Working Paper Series*.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of political Economy*, 98(4):703–738.

DeLisle, R. J., French, D. W., and Schutte, M. G. (2017). Passive institutional ownership, r2 trends, and price informativeness. *Financial Review*, 52(4):627–659.

Diether, K. B., Malloy, C. J., and Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance*, 57(5):2113–2141.

Ernst, T. (2020). Stock-Specific Price Discovery From ETFs. *Working Paper*, 1(October 2017).

Farboodi, M., Matray, A., Veldkamp, L., and Venkateswaran, V. (2020). Where has all the data gone? Technical report, National Bureau of Economic Research.

Farboodi, M. and Veldkamp, L. (2020). Long-Run growth of financial data technology. *American Economic Review*, 110(8).

Foster, F. D. and Viswanathan, S. (1990). A theory of the interday variations in volume, variance, and trading costs in securities markets. *The Review of Financial Studies*, 3(4):593–624.

Foster, F. D. and Viswanathan, S. (1993). Variations in trading volume, return volatility, and trading costs: Evidence on recent price formation models. *The Journal of Finance*, 48(1):187–211.

Ganuza, J. J. and Penalva, J. S. (2010). Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica*, 78(3):1007–1030.

Garleanu, N. B. and Pedersen, L. H. (2018). Active and Passive Investing. *SSRN Electronic Journal*.

Gloßner, S. (2018). The Effects of Institutional Investors on Firm Outcomes: Empirical Pitfalls of Quasi-Experiments Using Russell 1000/2000 Index Reconstitutions. *SSRN Electronic Journal*.

Gloßner, S. (2019). Investor horizons, long-term blockholders, and corporate social responsibility. *Journal of Banking and Finance*, 103.

Glosten, L., Nallareddy, S., and Zou, Y. (2021). ETF activity and informational efficiency of underlying securities. *Management Science*, 67(1).

Grossman, S. J. and Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets On the Impossibility of Informationally Efficient Markets. *Source: The American Economic Review*, 70(3).

Haddad, V., Huebner, P., and Loualiche, E. (2021). How Competitive is the Stock Market? Theory, Evidence from Portfolios, and Implications for the Rise of Passive Investing. *SSRN Electronic Journal*.

Heath, D., Ringgenberg, M. C., Samadi, M., and Werner, I. M. (2020). Reusing natural experiments. *Available at SSRN 3991200*.

Israeli, D., Lee, C. M., and Sridharan, S. A. (2017). Is there a dark side to exchange traded funds? An information perspective. *Review of Accounting Studies*, 22(3).

Kacperczyk, M., Sundaresan, S., and Wang, T. (2018a). Do foreign investors improve market efficiency? Technical report, National Bureau of Economic Research.

Kacperczyk, M. T., Nosal, J. B., and Sundaresan, S. (2018b). Market Power and Price Informativeness. *SSRN Electronic Journal*.

Kelly, B., Pástor, u., and Veronesi, P. (2016). The Price of Political Uncertainty: Theory and Evidence from the Option Market. *Journal of Finance*, 71(5).

Kothari, S. P. and Sloan, R. G. (1992). Information in prices about future earnings. Implications for earnings response coefficients. *Journal of Accounting and Economics*, 15(2-3).

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica*, 53(6).

Lee, J. (2020). Passive Investing and Price Efficiency. *SSRN Electronic Journal.*

Lochner, L. and Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American economic review*, 94(1):155–189.

Loughran, T. and McDonald, B. (2017). The use of edgar filings by investors. *Journal of Behavioral Finance*, 18(2):231–248.

Madhavan, A. (2003). The russell reconstitution effect. *Financial Analysts Journal*, 59(4):51–64.

Madhavan, A. (2014). Exchange-traded funds: an overview of institutions, trading, and impacts. *Annu. Rev. Financ. Econ.*, 6(1):311–341.

Madhavan, A. N. (2016). *Exchange-traded funds and the new dynamics of investing.* Oxford University Press.

Malikov, G. (2020). Information, Participation, and Passive Investing. *SSRN Electronic Journal.*

Manela, A. (2014). The value of diffusing information. *Journal of Financial Economics*, 111(1):181–199.

Martineau, C. and Zoican, M. (2021). Investor attention and the cross-section of analyst coverage. *Available at SSRN 3376162.*

Mauboussin, M. (2019). Looking for Easy Games in Bonds. *Bluemountain Investment Research.*

Morck, R. and Yang, F. (2001). The mysterious growing value of s&p 500 membership. Technical report, National Bureau of Economic Research.

Novy-Marx, R. (2015). Fundamentally, Momentum is Fundamental Momentum. *National Bureau of Economic Research.*

Pavlova, A. and Sikorskaya, T. (2022). Benchmarking intensity.

Qin, N. and Singal, V. (2015). Indexing and stock price efficiency. *Financial management*, 44(4):875–904.

Roll, R. (1988). R-squared. *Journal of finance*, 43(2):541–566.

Wang, J. (1994). A Model of Competitive Stock Trading Volume. *Journal of Political Economy*, 102(1).

Weller, B. M. (2018). Does algorithmic trading reduce information acquisition? *Review of Financial Studies*, 31(6).

Wurgler, J. (2010). On the economic consequences of index-linked investing. Technical report, National Bureau of Economic Research.

Zhang, X. F. (2006). Information uncertainty and stock returns. *The Journal of Finance*, 61(1):105–137.