Passive Ownership and Price Informativeness

MARCO SAMMON *

June, 2022

ABSTRACT

How does passive ownership affect the incorporation of information into stock prices? Focusing on the period before earnings announcements, I find price informativeness declined over the past 30 years and passive ownership is negatively correlated with price informativeness. To establish causality, I show price informativeness decreases after quasi-exogenous increases in passive ownership arising from index additions/rebalancing. My proposed mechanism is that passive ownership decreases learning about stock-specific information. Consistent with this, high passive ownership stocks receive less attention from analysts and institutional investors, have more ex-ante earnings uncertainty, are traded less before earnings announcements and react more to fundamental news.

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

*Harvard Business School. Online appendix. I would like to thank my dissertation committee: Scott Baker, Alex Chinco, Ravi Jagannathan, Robert Korajczyk and Dimitris Papanikolaou. I would also like to thank Jennie Bai (discussant), Nick Bloom, Ian Dew-Becker, Simon Gloßner, Robin Greenwood, Sam Hanson, Mina Lee (discussant), Matthew Pritsker, Adi Sunderam, Martin Thyrsgaard, and Luis Viceira for helpful comments.

1 Introduction

The rise of passive ownership is one of the most significant changes in asset markets over the past 30 years. Passive funds grew from owning less than 1% of the US stock market in the early 1990s to owing nearly 15% in 2018. As passive ownership continues to grow it is increasingly important to understand: How does passive ownership affect the incorporation of information into stock prices?

There is no consensus answer to this question in the theoretical or applied literature. For example: Cong et al. (2020) and Glosten et al. (2021) argue that passive ownership can increase the incorporation of systematic news. On the other hand, Ben-David et al. (2018) and Kacperczyk et al. (2018b) provide evidence that passive ownership increases non-fundamental volatility and can lead informed investors to trade less aggressively on their private information. I use the term price informativeness to broadly capture the notion of how well a stock's price reflects information about the fundamental value of the firm. One reason for the literature's lack of consensus is that researchers do not agree on how to measure price informativeness.

In this paper, I focus on days where we know information is released to the public: earnings announcements. Motivated by Ball and Brown (1968), Fama et al. (1969), Manela (2014) and Weller (2018), I ask how returns and volatility would behave around these events if pre-earnings announcement prices were less informative. This method has three advantages: (1) it doesn't rely on a particular model's definition of price informativeness (2) there is a sharp prediction for when the information contained in the earnings announcement should be incorporated into prices and (3) all these quantities are observable, so I avoid needing to take a stand on the true – but ultimately unobservable – fundamental value of the firm.

To this end, I define two measures of price informativeness. The first is the pre earnings-drift magnitude (DM), which measures the fraction of earnings information incorporated into prices before the announcement itself, with lower values of DM implying less informative prices. The second is the quadratic variation share (QVS), which measures the share of volatility occurring on non-earnings-announcement days. The intuition is that if investors are not gathering private signals, public information releases should be the dominant source of stock volatility. Lower values of QVS, therefore, imply less informative prices.

Leveraging these metrics, I present a new set of stylized facts, showing that over the past

30 years, pre-earnings-announcement price informativeness has been declining on average. In cross-sectional regressions, passive ownership is correlated with lower pre-earnings drift and a lower quadratic variation share. These effects are economically large, with an increase in passive ownership of 15% – which is roughly the value-weighed average increase over my sample – able to explain 33%-40% of the aggregate trends.

Passive ownership, however, is not randomly assigned in the cross-section of stocks. To establish causality, I need to identify increases in passive ownership that are plausibly uncorrelated with firm fundamentals. To this end, I leverage two natural experiments based on S&P 500 index additions and Russell 1000/2000 index rebalancing. Reassuringly, the relationship between passive ownership and price informativeness remains qualitatively and quantitatively unchanged in these better-identified settings.

Finally, I aim to understand why passive ownership decreases price informativeness. The standard story in the financial press is that, "*[Passive investing] does naturally focus more attention in a more macro direction, above the single security level... The more stock prices are set by trading in ETFs, the less important are what investors call 'idiosyncratic factors'...*"[1] Fundamentally, this is a learning/attention-based mechanism, and if it's true, we would expect to see less information gathering for stocks with more passive ownership. Further, it implies this effect should be especially strong for firm-specific information.

Consistent with this, passive ownership is correlated with fewer non-robot downloads of SEC filings, evidence of less fundamental research (see e.g., Loughran and McDonald (2017)) and fewer Bloomberg terminal searches, evidence of less attention by institutional investors (see e.g., Ben-Rephael et al. (2017)). I argue this decreased demand for information has equilibrium effects on the supply of information, which I measure using data on sell-side analysts (see e.g., Martineau and Zoican (2021)). I find that passive ownership is correlated with decreased analyst coverage, increased dispersion of analysts' estimates, decreased analyst accuracy and fewer updates of analysts' forecasts. These direct effects on the supply and demand of information, however, are not the only implications passive ownership decreasing investors' attention to stock-specific news.

First, we might expect that if investors are paying less attention to firms with more passive ownership, it will be associated with more ex-ante earnings uncertainty. To quantify

[1]Source: Inigo Fraser-Jenkins, Co-Head of Institutional Solutions at Alliance Bernstein, quoted in the Wall Street Journal.

this, following Kelly et al. (2016), I construct a measure of how expensive options which span earnings announcements are, relative to those which expire just before/after the announcement. Consistent with the average decrease in price informativeness, options which span earnings announcements became relatively more expensive over my sample. The effects of passive ownership are large here, with a 15% increase in passive ownership explaining about half of the time-series trend toward increased ex-ante earnings uncertainty.

Somewhere else that decreased information gathering may manifest is in trading volume around information-release events (see e.g., Manela (2014)). One reason is that investor heterogeneity is typically cited as a source of trading volume (see e.g., Wang (1994)), and if no one gathers private information, everyone will have the same beliefs. Relatedly, as investors' signals become less precise, they may trade less aggressively, which could also lead to decreased trading volume. Therefore, if fewer investors are gathering private information, we might expect abnormally low pre-earnings trading volume. Moreover, price discovery doesn't occur in a vacuum; there must be trade for private information to be incorporated into prices, so less trading might be evidence in and of itself of decreased price informativeness. Consistent this, I show that there has been a large decline in average pre-earnings volume over the past 30 years, and that passive ownership can explain this trend.

Another implication of fewer people gathering information is that stock prices may respond more to fundamental news of a given 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. By running earnings response regressions in the spirit of Kothari and Sloan (1992), I show there has been a trend toward increased reactions to fundamental news over the past 30 years and a 15% increase in passive ownership roughly doubles a firm's responsiveness to earnings surprises. Consistent with a re-allocation of attention from firm-specific to systematic information, passive ownership's effect is concentrated in an increased reaction to negative idiosyncratic news.

Finally, I explore why passive ownership has an asymmetric effect on positive versus negative news. One natural explanation is that passive ownership has made the underlying stocks easier to short, because passive funds lend out their shares as an additional source of income. Consistent with this, I find that high passive stocks have more short interest on average – although this doesn't necessarily speak to the effect of passive ownership on price informativeness per se. I find, however, that following the release of negative earnings news, stocks with more passive ownership have a larger increase in short interest. This is consistent with previously inattentive investors adding significantly to their short positions after the bad news was publicly revealed.

Literature Review. My paper contributes to two broad sets of research questions. First, as passive ownership has grown, so has the literature on how it affects price informativeness. Many channels have been proposed, from passive ownerships influence on firms cost of capital to its effects on liquidity. This literature has found mixed results, with some papers arguing that passive ownership increases informativeness (see e.g., Buss and Sundaresan (2020), Ernst (2020), Malikov (2020), Lee (2020), Kacperczyk et al. (2018a)), some papers arguing it decreases informativeness (see e.g., Qin and Singal (2015), DeLisle et al. (2017), Garleanu and Pedersen (2018), Breugem and Buss (2019), Brogaard et al. (2019), Bennett et al. (2020), Kacperczyk et al. (2018b)), others arguing it increases the incorporation of systematic information at the expensive of idiosyncratic information (see e.g.,. Bhattacharya and O'Hara (2018), Cong et al. (2020), Antoniou et al. (2020), Glosten et al. (2021)) or vice versa (see e.g., Bond and Garcia (2018)) and yet others showing it has no effect (see e.g., Coles et al. (2020)).

I find that passive ownership decreases price informativeness. My results differ from past work because (1) I am measuring price informativeness a different way. Rather than mapping a model-based measure to the data, I leverage intuitive metrics designed to capture the fraction of information which was incorporated into prices before it was formally announced (2) I am looking at prices informativeness at a particular point in time, specifically right before earnings announcements. Focusing on the pre-earnings announcement window – and the announcements themselves – yields sharp empirical predictions for when the information should be incorporated into prices and (3) I am measuring informativeness about a different type of information. Because my results examine how prices, volatility and trading volume change around earnings announcements, I am not focused on the earnings numbers alone, but capture informativeness about any news released at the same time (e.g., in recent earnings announcements, Disney has also disclosed the number of subscribers to their online streaming platform). This is a strength of my empirical method, as I dont need to take a stand on which accounting numbers – if any – constitute the true fundamental value of the firm.

My paper also contributes to a growing literature which studies trends in price informativeness (see e.g., Bai et al. (2016), Dávila and Parlatore (2018)). This literature has consistently found that average price informativeness has been increasing over time, although this effect is not uniform across individual stocks e.g., Bai et al. (2016) and Farboodi et al. (2020) show this effect was strongest for S&P 500 firms and large growth firms, respectively. Explanations for this trend have ranged from the growth of financial data to improved liquidity. I find the opposite, showing that average price informativeness has been decreasing over the past 30 years. As mentioned above, my results differ from this literature because (1) I am measuring price informativeness differently (2) I am focusing on price informativeness specifically before earnings announcements and (3) I am examining informativeness about a different type of information.

2 Measuring price informativeness

In this section, I motivate my measures of price informativeness and show that they have been declining on average over the past 30 years.

2.1 Measure 1: Pre-earnings drift

A natural measure of price informativeness before an earnings announcement is the fraction of information that ended up being released which was incorporated into prices ahead of time. Therefore, if pre-earnings announcement prices became less informative, we would expect them to drift up relatively less ahead of good news and drift down relatively less ahead of bad news.

To visualize this, we need a definition of good and bad news, so following Novy-Marx (2015), define standardized unexpected earnings (SUE) as:

$$SUE_{i,t} = \frac{E_{i,t} - E_{i,t-4}}{\sigma_{(t-1,t-8)}(E_{i,t} - E_{i,t-4})} \tag{1}$$

where $E_{i,t}$ denotes earnings per share for firm i in quarter t in the IBES Unadjusted Detail File. In words, Equation 1 is measuring the year-over-year (YOY) change in earnings, divided by the standard deviation of YOY changes in earnings over the past 8 quarters.

Each quarter, I sort firms into deciles of SUE and calculate the cumulative market-adjusted returns of a \$1 investment over the 22 trading days before and after the earnings announcement.[2] Market-adjusted returns, $r_{i,t}$, are defined as in Campbell et al. (2001): the difference between firm i's excess return and the excess return on the market factor from Ken French's data library.

Figure 1 shows the average cumulative market-adjusted returns by SUE decile for two different time periods: 2001-2007 and 2010-2018. The brown dashed line represents the average for firms with the most positive earnings surprises, while the blue dashed line represents the average for firms with the most negative earnings surprises. Between 2010 and 2018, firms in each decile move less before earnings days than between 2001 and 2007. The decline in pre-earnings drift is even stronger when comparing to the pre-2001 period, but that may be due to Regulation Fair Disclosure (Reg FD), implemented in August 2000, which limited firms' ability to selectively disclose earnings information before it was publicly announced.

To quantify the patterns in Figure 1, I define the pre-earnings drift magnitude for firm i with an earnings announcement at time t as:

$$DM_{i,t} = \begin{cases} \frac{1+r_{i,(t-22,t-1)}}{1+r_{i,(t-22,t)}} & \text{if } r_{i,t} > 0 \\ \frac{1+r_{i,(t-22,t)}}{1+r_{i,(t-22,t-1)}} & \text{if } r_{i,t} < 0 \end{cases} \tag{2}$$

The pre-earnings drift magnitude will be near one when the earnings day move is small relative to cumulative pre-earnings returns. $DM_{i,t}$ decreases as the earnings-day return becomes large, relative to the returns over the previous 22 days. If r_t is negative, this relationship would be reversed, which is why the measure is inverted when r_t is less than zero. $DM_{i,t}$ is my first measure of pre-earnings price informativeness, and yields an empirical prediction I will use to measure the effect of passive ownership on price informativeness.

Prediction 1: *If passive ownership decreases price informativeness, it should cause $DM_{i,t}$ to decline*

Consistent with Figure 1, average $DM_{i,t}$ decreased by about -0.02 between 1990 and 2018. This drop is large, at about 40% the size of DM's whole sample standard deviation of 0.05. A natural question raised by Figure 1 is whether this decline in pre-earnings drift was

[2] While somewhat arbitrary, the choice of 22 trading-days before the earnings announcement is in line with previous literature (see e.g., Weller (2018)). A natural concern with the choice of 22 trading days is that we care about price informativeness all the time, so we should use e.g., a pre-earnings window that starts right after the last earnings announcement. Using a window this long, however, has its own issues, as it might pick up a delayed reaction to last period's earnings news and is more likely to overlap with non-earnings news releases.

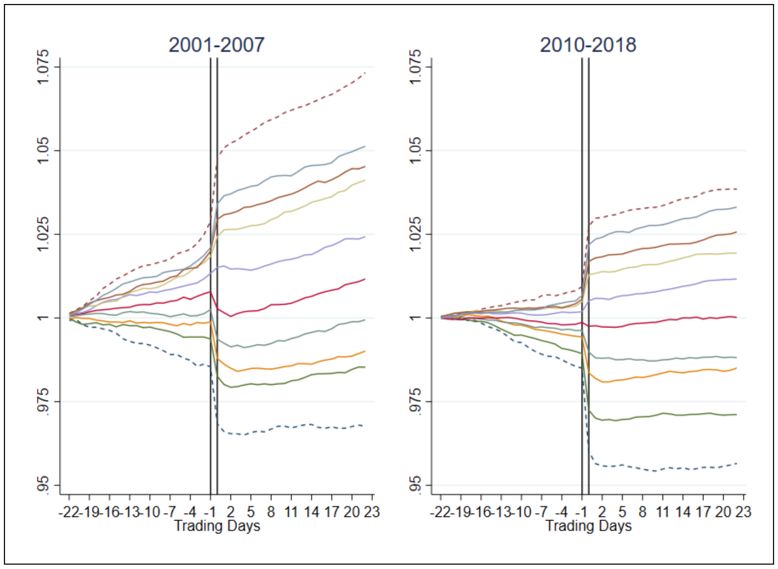

Figure 1. Decline of pre-earnings drift by SUE decile. Each quarter, firms are sorted into deciles based on standardized unexpected earnings (SUE). Each line represents the cross-sectional average market-adjusted return of \$1 invested at $t = -22$. The brown dashed line represents the average for firms with the most positive earnings surprises, while the blue dashed line represents the average for firms with the most negative earnings surprises. The solid lines represent the averages for deciles 2 to 9.

symmetric for firms which ended up releasing good and bad news. In the appendix, I show that before 2012, the decline in SUE was mostly uniform across quintiles of SUE. From 2013-2018, however, the average drift decreased relatively more for firms that ended up releasing the worst earnings news. I further explore this asymmetry, as well as its relationship to passive ownership, in Section 6.

While not identical, DM is similar to the price-jump measure of Weller (2018), which is also designed to capture the fraction of earnings information incorporated into prices before it was formally released. The difference is that DM uses gross returns, while price-jump uses net returns. If I had instead defined $\widehat{DM_{i,t}} = r_{t-22,t-1}/r_{t-22,t}$, the mean of $\widehat{DM_{i,t}}$ may not be well defined, as $r_{t-22,t}$ can be equal to zero. Weller overcomes this challenge by filtering out non-events – defined as observations with $r_{t-22,t}$ close to zero – which constitute almost 50% of earnings announcements. This filter, however, can complicate any analysis where the right-hand-side variable of interest is related to market capitalization – as is the case with passive ownership – or when using value weights, because non-events are not evenly spread across the firm size distribution.

DM is also related to the pre- and post- drug approval cumulative abnormal returns (CARs) studied in Manela (2014), as well as the absolute CARs around earnings announcements studied by e.g., Ball and Brown (1968) and Fama et al. (1969). I believe that in my setting, as discussed in Weller (2018), DM has the advantage that it captures the fraction of information incorporated into prices before it is formally announced. In the appendix, as an additional robustness check, I show that using Weller's price-jump measure or CARs does not change any of my empirical conclusions.

2.2 Measure 2: Earnings days' share of volatility

In many models, the arrival of new information is a major source of volatility in stock prices. There are two ways, however, that new information could arrive: (1) investors privately gather information, e.g., as was the case for the Lumber Liquidators scandal in 2015 or (2) firms release information to the public. As fewer investors gather private information, therefore, we would expect that public signals would be drive an increasingly large share of total stock volatility.

To capture this idea, define the quadratic variation share (QVS) for firm i around earnings announcement t as:

$$QVS_{i,t} = 1 - \left(r_{i,t}^2 / \sum_{\tau=-22}^{0} r_{i,t+\tau}^2 \right) \tag{3}$$

where r_{it} denotes a market-adjusted daily return. The numerator of the term in parenthesis is the squared earnings-day return, while the denominator is the sum of squared returns from $t-22$ to t. QVS is my second measure of price informativeness. If relatively more information is being learned and incorporated into prices on earnings announcement dates, we would expect smaller values of QVS[3], which yields a second empirical prediction for measuring the relationship between passive ownership and price informativeness.[4]

[3]This is why I define $QVS_{i,t} = 1 - \left(r_{i,t}^2 / \sum_{\tau=-22}^{0} r_{i,t+\tau}^2 \right)$ instead of $\widehat{QVS_{i,t}} = r_{i,t}^2 / \sum_{\tau=-22}^{0} r_{i,t+\tau}^2$. Lower values of QVS imply less informative prices, consistent with DM.

[4]Measures like QVS – where volatility upon information arrival is a signal about how much agents updated their beliefs – also apply in more general settings, see e.g., Ganuza and Penalva (2010) and Åstebro and Penalva (2022).

Prediction 2: *If passive ownership decreases price informativeness, it should cause* $QVS_{i,t}$ *to decline*

In this 23-day window – 22 pre-earnings days plus the earnings announcement itself – the earnings day is $1/23 \approx 4.3\%$ of observations, so values of $QVS_{i,t}$ smaller than 0.957 imply that earnings days account for a disproportionately large share of total volatility.[5] Figure 2 plots coefficients from a regression of QVS on a set of year dummy variables for all stocks in my sample. Average QVS decreased from 92.0% in 1990 to 72.6% in 2018. This 19.4% decline is about the same size as QVS's whole-sample standard deviation of 21.0%. 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.

A natural question is whether QVS captures different information than DM because e.g., they will both tend to be lower if the earnings-day return is large in absolute value. To understand where they might differ, consider the following scenario: Leading up to an earnings announcement, a stock has alternating returns of $\pm$ 5%, for a total cumulative return of 0%. Then, on the earnings day itself, the stock has a return of 5%. In this case, DM would be small at $(1+0)/(1+0.05)$, suggesting that pre-earnings prices were not informative. This, however, misses the fact that the earnings-day return itself was not large relative to the swings that the stock experienced leading up to the earnings announcement, which would be captured by QVS. In the appendix, I present more details on the relationship between DM and QVS.

2.3 Placebo tests on stylized facts

One concern is that all of the downward trends in price informativeness documented in this section could be unrelated to the information released on earnings days. To rule this out, I run the following placebo test: select the date 22 trading days before each earnings

[5] One concern is that the 22-day pre-earnings-announcement window may be too narrow to capture when information is being incorporated into prices. All results are robust to defining QVS at the annual level i.e., defining the numerator of the term in parenthesis to be the sum of squared returns on the 4 quarterly earnings days in year t, while the denominator is the sum of squared returns for all days in year t. Defining things at the annual level avoids needing to take a stand on the correct pre-earnings window length. My results are also robust to including a post-earnings announcement window in the numerator e.g., defining $\widehat{QVS_{i,t}} = 1 - \left[\left(\sum_{j=0}^{n} r_{i,t+j}^2 \right) / \left(\sum_{\tau=-22}^{0} r_{i,t+\tau}^2 \right) \right]$ for $n = 1, \ldots, 5$.

Figure 2. Increase in earnings-day volatility. This figure plots coefficients from a regression of QVS on a set of year dummy variables. The constant term for the omitted year (1989) is added to each coefficient. For firm i around earnings announcement τ the quadratic variation share (QVS) is defined as: $QVS_{i,t} = 1 - \left(r_{i,t}^2 / \sum_{\tau=-22}^{0} r_{i,t+\tau}^2 \right)$, where r denotes a market-adjusted daily return. The red bars represent 95% confidence intervals around the point-estimates. Standard errors are clustered at the firm level.

announcement – the start of each pre-earnings announcement window – to be a placebo earnings date. I then reconstruct the time-series averages of the pre-earnings drift and QVS for these placebo earnings days. In the appendix, I show that there is no downward trend in the drift or QVS for the placebo earnings dates. I repeat this exercise using randomly selected dates as placebo earnings announcements and also find no trends for either of the price informativeness measures.

As an additional placebo test, in the appendix, I examine price informativeness around scheduled Federal Open Market Committee (FOMC) meeting dates. Like earnings announcements, these are days where large quantities of information are released, but it is not specific to any particular firm. I find that there is no significant downward trend in either of the price informativeness measures around scheduled FOMC announcements. Collectively, these placebo tests confirm that the decrease in average price informativeness only applies to firm-specific information released around earnings announcements.

3 Data description

This section describes the data I use to quantify passive ownership and calculate the price informativeness measures from Section 2. It also presents summary statistics on the increase in passive ownership and decline in average price informativeness from the 1990s to the 2010s.

3.1 Defining passive ownership

Starting with the CRSP mutual fund database, passive funds are defined as all index funds – identified using the index fund flag – all ETFs and all funds with names that identify them as index funds, according to the criteria in Appel et al. (2016). Passive ownership is defined as the fraction of a stock's shares outstanding owned by passive funds. Figure 3 shows that passive ownership increased from almost zero in 1990, to now owning about 15% of the US stock market. Passive ownership is even larger as a faction of total S12 i.e., mutual fund and ETF assets, at over 40% in 2018.

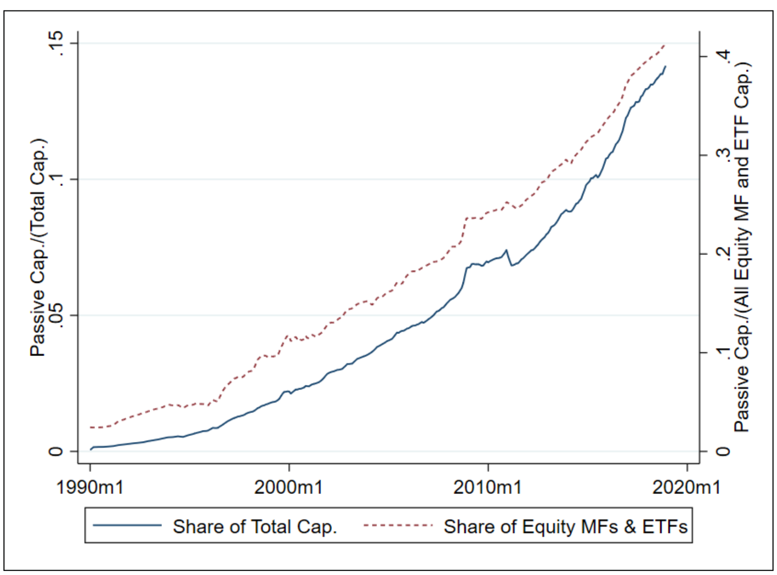

Figure 3. The rise of passive ownership: 1990-2018. Passive funds are defined as all index funds, all ETFs and all mutual funds with names that identify them as index funds. Total equity mutual fund and ETF assets is the sum of all stock holdings in the Thompson S12 data that can be matched to CRSP.

This definition likely understates the true level of passive ownership, as there are institutional investors which track broad market indexes, but are not classified as mutual funds and thus do not appear in the S12 data. Further, as discussed in Mauboussin (2019), there has been a rise of shadow indexing among self-proclaimed active managers, which is also omitted in my definition of passive management.

3.2 Data for constructing price informativeness measures

All daily stock-level data are from CRSP. I restrict to ordinary common shares (share codes 10 and 11) traded on major exchanges (exchange codes 1 to 3). I use the earnings release times in IBES to identify the first date investors could trade on earnings information during normal market hours. 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 is labeled as the effective earnings date.

3.3 Summary Statistics

Table 1 contains details on the means, standard deviations and distributions of the price informativeness measures, as well as passive ownership. Consistent with Figures 1, 2 and 3, the average of both price informativeness measures decreased between the 1990s and the 2010s, while passive ownership increased.

4 Cross-sectional relationship between passive ownership and price informativeness

In this section, I present the cross-sectional relationships between passive ownership the pre-earnings drift and QVS. Across both measures, the regressions show that higher passive ownership is correlated with decreased price informativeness. I also summarize the results of several tests designed to rule-out alternative explanations.

		25%	50%	Mean	75%	St. Dev.
DM	1990-1999	0.954	0.978	0.964	0.992	0.044
QVS		0.898	0.969	0.914	0.995	0.139
$Passive$		0.000	0.001	0.005	0.007	0.010
DM	2010-2018	0.933	0.968	0.950	0.987	0.055
QVS		0.610	0.882	0.766	0.979	0.262
$Passive$		0.000	0.014	0.052	0.093	0.068
DM	All Years	0.941	0.972	0.955	0.989	0.053
QVS		0.805	0.948	0.851	0.990	0.211
$Passive$		0.000	0.005	0.026	0.030	0.046

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

4.1 Pre-earnings drift

To test prediction 1, I run the following regression with quarterly data to measure the relationship between the pre-earnings drift and passive ownership:

$$DM_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t} \tag{4}$$

where $DM_{i,t}$ is defined as in Equation 2 . Controls in $X_{i,t}$ include time since listing (which I call firm age), one-month lagged market capitalization, returns from t-12 to t-2 (the returns typically used to form momentum portfolios), one-month lagged book-to-market ratio and total institutional ownership. $X_{i,t}$ also includes CAPM beta, CAPM R-squared, total volatility and idiosyncratic volatility, all computed over the previous 252 trading days. These controls are included because they capture many firm-specific characteristics known to be correlated with passive ownership (see e.g., Glosten et al. (2021)). Time since listing, market capitalization and past returns are computed using CRSP data. All other firm fundamental information is from Compustat. Total institutional ownership is the percent of a stock's shares outstanding held by all 13-F filing institutions. CAPM beta and R-squared are from the WRDS beta suite.

Equation 4 also includes firm fixed effects and year-quarter fixed effects. The year-quarter fixed effects, ϕ_t, ensure I am comparing firms at the same point in time with different levels of passive ownership, accounting for the time trends in price informativeness. These fixed effects also account for seasonality. The firm fixed effects, ψ_i, account for firm-specific differences in average price informativeness e.g., investors pay more attention to Apple's earnings announcements than to those of Dominion Energy. Standard errors are double-clustered at the firm and year-quarter level.

Table 2 contains the regression results. In column 1, the right-hand-side only has passive ownership and the two sets of fixed effects. I find that passive ownership is negatively correlated with the pre-earnings drift. In column 2, the right-hand-side variables are the same as column 1, but I restrict to the sample of firm-quarter observations with non-missing control variables. The coefficient is almost unchanged, so the selection effect of restricting only to observations that have non-missing controls is not driving my results. Finally, in column 3, I add in all the firm-level controls in $X_{i,t}$. The coefficient on passive ownership shrinks, but is still economically large and statistically significant. Going forward, I refer to this specification, with equal weights, all the firm-level controls and fixed effects as the baseline specification.

The coefficient on $Passive_{i,t}$ in the baseline specification (column 3) implies that a 15% increase in passive ownership would decrease the pre-earnings drift by -0.008. This effect is economically large, at about 40% the size of the average decline in the drift over my sample of 0.02. To allay concerns that small firms are driving my results, columns 4 and 5 replicate columns 2 and 3, but within each quarter, firms are weighted by their market capitalization at the end of the previous quarter. Using value weights, instead of equal weights, does not lead to a statistically significant difference in the estimated effect of passive ownership on the pre-earnings drift.

A natural question is whether passive ownership can explain the decline in the post-earnings announcement drift (PEAD), apparent in Figure 1 and documented in e.g., Mclean and Pontiff (2016) and Martineau (2018). Questions about price informativeness after earnings announcements – which includes its effects on the PEAD – are outside the scope of my empirical strategy.[6] I am using earnings announcements as a laboratory to study the incorporation of information into prices before it is released because there is a clear prediction: if the information contained in the earnings announcement was not already incorporated into

[6]Several papers (see e.g., Qin and Singal (2015) and Coles et al. (2020)) have already studied the effect of passive ownership on the PEAD and came to mixed conclusions.

	(1)	(2)	(3)	(4)	(5)
Passive Ownership	-0.0430***	-0.0468***	-0.0480***	-0.0528***	-0.0489***
	(0.006)	(0.006)	(0.006)	(0.016)	(0.012)
Observations	492,025	448,626	448,626	448,626	448,626
R-Squared	0.199	0.205	0.22	0.26	0.276
Firm + Year/Quarter FE	✓	✓	✓	✓	✓
Matched to Controls		✓	✓	✓	✓
Firm-Level Controls			✓		✓
Weight	Equal	Equal	Equal	Value	Value

Table 2 Passive ownership and pre-earnings drift. Estimates of β from:
$DM_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$
Where $DM_{i,t}$ is a measure of the pre-earnings drift. Controls in $X_{i,t}$ include firm 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.

prices, we expect it to be incorporated soon after the announcement itself. Therefore, the announcement-day return – or differences in pre vs. post announcement volatility – should reveal if this information was already in prices.

For example: suppose earnings expectations were high this quarter and then realized earnings were low. Investors now know those expectations were wrong and we have an unambiguous prediction for the direction and timing of the respective price adjustment. Predictions for the long-run implications of missing expected earnings this quarter, however, are less clear. For this reason, I restrict my empirical analysis to measuring price informativeness before information is released.

More broadly, several papers examine correlations between prices and fundamentals over longer horizons (see e.g., Bai et al. (2016), Dávila and Parlatore (2018), Dávila and Parlatore (2021), Kacperczyk et al. (2018a), Buss and Sundaresan (2020) and Farboodi et al. (2020)). While I find a different time-series trend, I do not view my results as a direct contradiction to theirs. Given the way DM is defined, I am making a narrow claim about pre-earnings announcement informativeness. Further, earnings announcements are mostly a firm-specific event, and as discussed in e.g., Glosten et al. (2021), passive ownership may increase the incorporation of systematic information at the expensive of idiosyncratic information. Therefore, it's possible that even though price informativeness about firm-specific information has trended down, there may still have been an increase in average price informativeness, especially if firm-specific information has become less important (see e.g., Bartram et al. (2019)). As an additional check on this, in Section 6, I decompose earnings information into systematic and idiosyncratic components, and show the effect of passive ownership is stronger for the firm-specific component of news.

4.2 Earnings days' share of volatility

To test prediction 2, I run the following regression to measure the relationship between earnings days' share of volatility and passive ownership:

$$QVS_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t} \tag{5}$$

where QVS is defined in Equation 3 and all the controls and fixed effects are the same as in Equation 4. The regression results are in Table 3. The baseline specification (column 3) implies that a 15% increase in passive ownership would lead to a decrease in QVS of 6.1%. This is large, at about 1/3 the magnitude of the average decline in QVS over the whole sample. This result is not significantly changed by including the firm-level controls, but it is weakened when using value weights instead of equal weights.

4.3 Ruling out alternative explanations

Over my sample, the rise of passive ownership was not the only structural change to financial markets. In this subsection, I discuss the results of placebo tests and alternative specifications designed to rule out other explanations for my main findings.

4.3.1 Placebo tests on cross-sectional regression results

One might be concerned that the regression-results above are not specific to earnings announcement days. An alternative explanation is that passive ownership increases overall volatility (see e.g., Ben-David et al. (2018)) and given the upward trend in passive ownership, this could create a spurious correlation between passive ownership and QVS. Recall

	(1)	(2)	(3)	(4)	(5)
Passive Ownership	-0.524***	-0.501***	-0.408***	-0.214*	-0.232**
	(0.027)	(0.028)	(0.031)	(0.111)	(0.094)
Observations	446,530	416,609	416,609	416,609	416,609
R-Squared	0.217	0.22	0.222	0.233	0.234
Firm + Year-Quarter FE	✓	✓	✓	✓	✓
Matched to Controls		✓	✓	✓	✓
Firm-Level Controls			✓		✓
Weight	Equal	Equal	Equal	Value	Value

Table 3 Passive ownership and earnings days' share of volatility. Table with estimates of β from:
$QVS_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$
where $QVS_{i,t}$ is a measure of earnings days' share of volatility. Controls in $X_{i,t}$ include firm 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.

$QVS_{i,t} = 1 - \left(r_{i,t}^2 / \sum_{\tau=-22}^{0} r_{i,t+\tau}^2 \right)$, so $\sum_{\tau=-22}^{0} r_{i,t+\tau}^2$ contains volatility relatively further in the past than $r_{i,t}^2$. If increases in passive ownership are persistent at the firm level, QVS might mechanically decrease over time, as volatility further in the past would tend to be lower than current volatility. To rule out this possibility and confirm that my results are specific to earnings days, in the appendix, I perform three additional placebo tests.

As before, the first set of placebo earnings dates are 22 trading days before each earnings announcement. The second are randomly selected dates each quarter. The third are all scheduled FOMC meetings. For the first two sets of placebo earnings announcements, there is no relationship between QVS and passive ownership. This suggests that my results are specific to earnings announcement dates. For the FOMC announcements, however, there is a weakly statistically significant negative relationship between passive ownership and QVS, but the magnitude is 1/20th as large as the coefficient in Table 3. This suggests that passive ownership may be correlated with decreased pre-FOMC announcement price informativeness, but the effect is quantitatively much smaller than for stock-specific news releases.

4.3.2 Regime shifts

Two additional threats to identification are (1) Regulation Fair Disclosure (Reg FD), passed in August 2000, which reduced the early release of earnings information and (2) the rise of algorithmic trading (AT), which can reduce the returns to informed trading (see e.g., Weller (2018), Farboodi and Veldkamp (2020)). The appendix shows that all the cross-sectional results are robust to only using data after Reg FD passed i.e., earnings announcements from 2001-2018. The results are also robust to controlling for the AT measures in Weller (2018).[7]

Of course, it is not possible to individually rule out every alternative explanation. To overcome this, in the next section, I replicate these main results using only quasi-exogenous increases in passive ownership arising from index additions and rebalancing.

5 Causal effect of passive ownership on price informativeness

A major concern with the results in Section 4 is that passive ownership is not randomly assigned in the cross-section of stocks. To establish causality, I need to identify increases in passive ownership which are plausibly uncorrelated with firm characteristics. To this end, I leverage S&P 500 index additions, as well as Russell 1000/2000 reconstitutions to identify quasi-exogenous increases in passive ownership. The results from Section 4 are qualitatively and quantitatively unchanged in this better-identified setting.

5.1 S&P 500 index additions

Four times a year, a committee from Standard & Poor's selects firms to be added/removed from the S&P 500 index. 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 index, it experiences a large increase in passive ownership, as many index mutual funds and ETFs buy the stock.

[7]These 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 regression specifications.

I obtain 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, I form quintiles of market capitalization. For each added 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. Because the S&P 500 is comprised of large firms, the additions are almost exclusively in the fourth and fifth market capitalization quintiles of their 3-digit SIC industry.

I also form a second control group of firms meeting the same selection criteria (3-digit SIC industry and industry market capitalization quintile), but that are already in the S&P 500 index. This results in about 500 treated firms, 600 control firms in the index and 2,000 control firms out of the index. Control firms can appear in more than once e.g., the same firm in the index can be a control firm for multiple firms added to the index at different points in time.

To identify the causal effect of passive ownership on stock price informativeness, I use index addition as an instrument for passive ownership. The first stage regression is:

$$Passive_{i,t} = \alpha + \beta_1 Post_{i,t} + \beta_2 Treated_{i,t} \times Post_{i,t} + FE + \epsilon_{i,t} \tag{6}$$

Here, $Treated_{i,t}$ is equal to one if the firm was added to the S&P 500 and zero otherwise.[8] The level of observation is firm-quarter-cohort group, where cohort group is defined by the combination of: (1) month of index addition (2) SIC-3 industry and (3) industry market capitalization quintile. There are fixed effects for each firm-cohort group to account for the fact that a firm-quarter observation can appear in multiple cohorts. $Treated_{i,t}$ is not included in Equation 6 because each firm can only be in one of the three treatment categories within each cohort group, so this term would be washed out by the fixed effects.

Figure 4 shows the level of passive ownership for the control firms and treated firms

[8] One concern with defining treatment as being added to the index and not staying in the index, is that firms may change their index status during the period of study. The results are robust to requiring treated firms to be out of the index for the whole pre-treatment period and in the index for the whole post-treatment period (and applying similar filters for both groups of control firms). This, however, is not my preferred specification, as whether or not a firm stays in/out of the index is endogenous and future index status is not known at the time of index addition.

around the month of index addition. Within each cohort group, I subtract the average level of passive ownership to make the effect comparable across cohorts 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.

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 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 4, I fix the level of passive ownership at its last reported level each month between quarter-ends. Therefore, although it appears as though passive ownership increases slowly around index addition, this is partly a function of averaging across observations with differences in time until the first set of post-index-addition S12 filings are released.

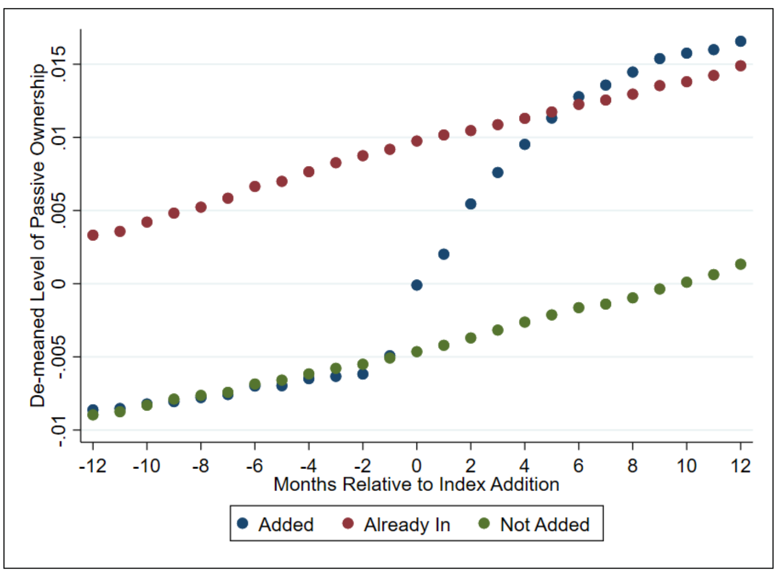

Figure 4. S&P 500 index addition and changes in passive ownership. Average level of passive ownership for control firms out of the index ("Not Added"), control firms in the index ("Already In") and treated firms ("Added"). Passive ownership is demeaned within each cohort group i.e., within each group of matched treated and control firms.

The three key pieces of my instrumental variables strategy are: (1) the instrumented change in passive ownership (2) the IV specification and (3) the reduced form specification:

$$\begin{aligned}
\widehat{Passive}_{i,t} &= \alpha + \beta_1 Post_{i,t} + \beta_2 Treated_{i,t} \times Post_{i,t} + FE + \epsilon_{i,t} \\
Outcome_{i,t} &= \alpha + \beta_3 \widehat{Passive}_{i,t} + FE + \epsilon_{i,t} \\
Outcome_{i,t} &= \alpha + \beta_4 Post_{i,t} + \beta_5 Treated_{i,t} \times Post_{i,t} + FE + \epsilon_{i,t}
\end{aligned} \tag{7}$$

Where $Outcome_{i,t}$ is the pre-earnings drift or QVS. The fixed effects are the same as in Equation 6. I restrict to data within five years before or after index addition, but exclude three months immediately before/after the event to avoid index inclusion effects (see e.g., Morck and Yang (2001)).

Because the change in passive ownership associated with being added to the S&P 500 has been increasing over time, I also run a specification that allows for heterogeneous treatment intensity:

$$\begin{aligned}
Passive_{i,t} &= \alpha + \beta_1 Post_{i,t} + \beta_2 \text{Passive Gap}_{i,t} \times Treated_{i,t} \times Post_{i,t} + FE + \epsilon_{i,t} \\
Outcome_{i,t} &= \alpha + \beta_3 \widehat{Passive}_{i,t} + FE + \epsilon_{i,t} \\
Outcome_{i,t} &= \alpha + \beta_4 Post_{i,t} + \beta_5 \text{Passive Gap}_{i,t} \times Treated_{i,t} \times Post_{i,t} + FE + \epsilon_{i,t}
\end{aligned} \tag{8}$$

Here, Passive $\text{Gap}_{i,t}$ 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, if there are not matched control firms both in and out of the index, I use the average Passive $\text{Gap}_{i,t}$ for all firms that year. Passive $\text{Gap}_{i,t}$ is designed to capture the expected increase in passive ownership from being added to the index. The average Passive Gap at the end of my sample is about 4%. The uninteracted term Passive $\text{Gap}_{i,t}$ is not included in Equation 8 because it is constant within each cohort group and would be washed out by the fixed effects. Following Coles et al. (2020), I double cluster standard errors at the firm and quarter level.

Table 4 contains the regression results. Panel A examines the effect of index addition on price informativeness using the binary treatment specification. Column 1 is the first stage regression.[9] The associated F-statistic is very large, which is not surprising given the increase

[9] Column 1 only uses observations with non-missing pre-earnings drift data. The number of observations and thus the first-stage regression is slightly different for each measure of price informativeness. The results, however, are quantitatively similar in each case, so I only report this first stage regression to avoid redundancy.

in passive ownership pictured in Figure 4.[10] The coefficient on $Post \times Treated$ implies that the average increase in passive ownership associated with index addition is about 1.3%.

Column 2 is the instrumental variables (IV) specification, where I use $Post$ and $Post \times Treated$ as instruments for passive ownership. The effect on the pre-earnings drift is negative, consistent with the cross-sectional regression results. The IV estimate of -0.131, however, is about 3 times as large as the baseline estimate of -0.048. Finally, column 3 is the reduced form (RF) regression, which provides an easy to interpret magnitude: Being added to the index is associated with an average drop in pre-earnings drift of -0.0013. Columns 4-5 repeat columns 2-3 for QVS. As with DM, the IV estimate is about triple the baseline estimate of -0.395. 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.

The reduced-form regressions for the drift and QVS are the same sign as the baseline estimates, but statistically insignificant. It is not obvious, however, that the reduced form regressions should be comparable with the baseline cross-sectional regression estimates. For the binary treatment specifications, the reduced form regression ignores the fact that the change in passive ownership associated with index addition increased from about 50bp in the early 1990s to 4% by the late 2010s. The continuous treatment specification partially addresses this issue, but given that the baseline regression estimates are about the level of passive ownership, it's not obvious why the expected change in passive ownership from index addition i.e., Passive $\text{Gap}_{i,t}$ should be informative about anything other than the sign of the treatment effect. I present a more detailed discussion of the differences between the IV and RF specifications in subsection 5.3.2 and the appendix.

Panel B repeats Panel A, but using the gap in passive ownership between the two sets of matched control firms interacted with the treatment dummy, along with $Post_{i,t}$, as instruments for passive ownership. As with the binary treatment specification, the first stage is economically large and statistically significant. The IV estimates in Panel B are quantitatively similar to the IV estimates in Panel A. With the continuous instrument, however, it more straightforward to compare the magnitude of the reduced-form estimates with the baseline cross-sectional results. For the pre-earnings drift and QVS, the coefficients are each

[10] Because I am using both $Post$ and $Post \times Treated$ as instruments for passive ownership, the time trend and the treatment effect in Figure 4 are driving the large magnitude of the F-statistic in Table 4. In a regression of passive ownership on $Post$, $Post \times Treated$ and the fixed effects, both terms are individually statistically significant, with $Post$ having a t-Statistic of 18 and $Post \times Treated$ having a t-Statistic of 15.

about half as large as the baseline estimates of -0.05 and -0.40.

		Panel A: Binary Instrument			
		Pre-Earnings Drift		QVS	
	First Stage	IV	RF	IV	RF
Post x Treated	0.0129***		-0.00132		-0.00428
	(0.001)		(0.001)		(0.005)
Passive Ownership		-0.131***		-1.902***	
		(0.038)		(0.138)	
Observations	284,094	284,094	284,094	286,053	286,053
F-statistic	299				
		Panel B: Continious Instrument			
		Pre-Earnings Drift		QVS	
	First Stage	IV	RF	IV	RF
Post x Treated	0.546***		-0.00862		-0.168
x Passive Gap	(0.048)		(0.036)		(0.162)
Passive Ownership		-0.129***		-1.895***	
		(0.037)		(0.138)	
Observations	284,094	284,094	284,094	286,053	286,053
F-statistic	428				
Cross-sectional regression estimate		-0.048	-0.048	-0.395	-0.395

Table 4 Effects of S&P 500 index addition on price informativeness. Estimates from:

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

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

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

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 regressions from the binary treatment specification, while Panel B contains regressions from the continuous treatment specification. FE are fixed effects for each firm-cohort group. Standard errors, double clustered at the firm and quarter level, are in parenthesis.

Overall, Table 4 shows that passive ownership decreases price informativeness for stocks added to the S&P 500. This is consistent with findings in several other papers including (1) Bennett et al. (2020), which shows that S&P 500 index additions tends to decrease price informativeness and (2) Bond and Garcia (2018), whose theoretical model predicts that price efficiency of assets covered by an index should be lower than assets outside the index. It is reassuring to see that even though my paper has differences both in terms of how I identify treated/control firms and how I measure price informativeness, I come to a similar conclusion.

One concern with using the S&P 500 is that because index addition is determined by a committee, the increase in passive ownership is not fully exogenous to firm fundamentals. Partially alleviating this concern is that, according to S&P (2017): "*Stocks are added to make the index representative of the U.S. economy and is not related to firm fundamentals.*" As an additional check, in the next subsection I focus on Russell 1000/2000 reconstitution, which is based on a mechanical rule, rather than discretionary selection.

5.2 Russell 1000/2000 index rebalancing

The Russell 3000 contains approximately the 3000 largest stocks in the US stock market. At the end of each May, FTSE Russell selects the 1000 largest stocks by market capitalization to be members of the Russell 1000, while it selects the next 2000 largest stocks to be members of the Russell 2000. Both of these indices are value-weighted, so moving from the 1000 to the 2000 increases the fraction of a firm's shares owned by passive funds. This is because switchers go from being the smallest stock in an index of big stocks, to the biggest stock in an index of small stocks, significantly boosting their index weight.

To reduce turnover between the two indices, in 2007 Russell switched to a bandwidth rule, rather than using a sharp cutoff. As long as a potential switcher's market capitalization is within ± 2.5% of the Russell 3000E's total market capitalization, relative to the market capitalization of the 1000th ranked stock, it will remain in the same index it was in the previous year.

The ideal experiment is to compare potential switchers to those that actually switched. This, however, is not straightforward, as the data that Russell uses to compute May market capitalizations is not made available to researchers. I follow the method in Coles et al. (2020) to compute a proxy for the Russell May market capitalizations.[11] Between 2007 and 2020, I am able to correctly predict Russell 1000/2000 index membership for 99.63% of Russell 3000 stocks overall and 98.27% of Russell 3000 stocks within 100 ranks of the upper and lower bands.

[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. For more details, see e.g., Chang et al. (2015), Wei and Young (2017), Glośner (2018), Ben-David et al. (2019) and Heath et al. (2021).

I identify treated and control firms using the method in Coles et al. (2020). Each May I identify a cohort of possible switchers: those within +/- 100 ranks around the lower threshold that were in the Russell 1000 last year.[12] The treated firms are those that were over the threshold and ended up switching, while the control firms are those that were above the threshold and stayed in the 1000. 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 stay around the lower cutoff for multiple years, but never actually switch. These filters yield about 700 treated firms and 600 control firms.

Finally, I focus on Russell index reconstitutions after the rule change in 2007 for two reasons: (1) The average increase in passive ownership is larger than earlier years, as over time, more money has started to track these indices. Specifically, for switching firms, the total average increase in passive ownership each cohort (equal weighted from 2007 to 2018) is over 4%. The same average from 1990 to 2006 is around 1%, so including these years leads to a weaker first stage. One reason for this is that the two largest Russell 1000/2000 ETFs, IWB and IWM, were not launched until May 2000 (2) Under the bandwidth regime, switching is harder to predict/manipulate, so switching is less likely to be front-run by other investors.

Figure 5 compares the level of passive ownership around the index rebalancing date between the treated and control group. Within each cohort, I subtract the mean level of passive ownership. The pre-addition changes and levels of passive ownership are similar between both groups. Unlike S&P 500 index additions, Russell reconstitutions always coincide exactly with the end of a calendar quarter. Because of this, Figure 5 only plots data points for months with S12 filings i.e., the last month of each calendar quarter.

For the Russell experiment, I use a setup similar to the S&P 500 experiment, with three key differences: (1) Passive $\text{Gap}_{i,t}$ is now 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, before index rebalancing (2) the fixed effects are identical, but cohort group is now defined only by month of index rebalancing and (3) the time period is different,

[12]Another natural set of treated/control firms are those within 100 ranks of the upper band that were in the Russell 2000 the previous year. These are possible switchers to the Russell 1000 and they experience a decrease in passive ownership if they end up moving from the Russell 2000 to the Russell 1000. In the appendix, I show that within one year of switching, the treatment effect is totally washed out by the time trend toward increased passive ownership.

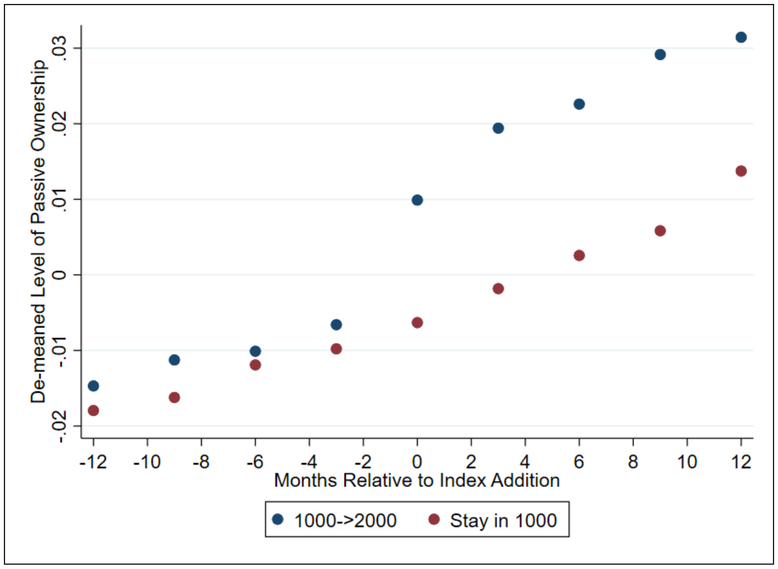

Figure 5. Russell 1000/2000 reconstitution and changes in passive ownership. Average level of passive ownership for firms that stay in the Russell 1000 (control firms) and firms that moved from the Russell 1000 to the Russell 2000 (treated firms). Passive ownership is demeaned within each cohort.

as I use Russell reconstitutions between 2007 and 2018.

Table 5 contains the regression results. The first-stage results are positive, with a large F-statistic. The estimated coefficient of about 1%, however, understates the total change in passive ownership. In Figure 5, the average total increase for treated firms is around 4%, but there is about a 3% increase for the control firms, driven by the overall trend upward in passive ownership.

All the estimated coefficients for the effect of passive ownership on price informativeness are qualitatively consistent with the cross-sectional regression estimates: switching to the Russell 2000 is associated with a drop in the pre-earnings drift and a decrease in QVS. The IV estimate of -0.09 for the drift is about twice the estimates from Table 2, while the IV estimate of -0.38 for QVS is similar in size to those in Table 3.

As with the S&P experiment, some of the reduced-form regressions are statistically insignificant. Again, this is not surprising as these reduced-form specifications do not account for (1) the time-series increase in the treatment effect even within the 2007 to 2018 sample and (2) differences in the resulting level of passive ownership across cohorts.

	Panel A: Binary Instrument				
		Pre-Earnings Drift		QVS	
	First Stage	IV	RF	IV	RF
Post x Treated	0.00976**		-0.00598**		-0.00813
	(0.005)		(0.003)		(0.012)
Passive Ownership		-0.0922***		-0.380***	
		(0.031)		(0.120)	
Observations	9,811	9,811	9,811	9,823	9,823
F-statistic	146				

	Panel B: Continious Instrument				
		Pre-Earnings Drift		QVS	
	First Stage	IV	RF	IV	RF
Post x Treated	0.759***		-0.222		-0.391
x Passive Gap	(0.217)		(0.139)		(0.583)
Passive Ownership		-0.0931***		-0.380***	
		(0.031)		(0.120)	
Observations	9,811	9,811	9,811	9,823	9,823
F-statistic	157				
Cross-sectional regression estimate		-0.048	-0.048	-0.395	-0.395

Table 5 Effects of Russell 1000/2000 index reconstitution on price informativeness. Estimates from:

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

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

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

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 regressions from the binary treatment specification, while Panel B contains regressions from the continuous treatment specification. FE are fixed effects for each firm-cohort group, as well as each month of index addition. Standard errors, double clustered at the firm and quarter level, are in parenthesis.

5.3 Robustness of quasi-experimental results

5.3.1 Passive ownership vs. Institutional ownership

One concern with the results in this section is that there may be an increase in total institutional ownership when a firm is added to the S&P 500 or when a firm switches between the Russell 1000 and the Russell 2000 (see e.g., Boone and White (2015)). Gloßner (2019), however, shows that although there is an increase in passive ownership following Russell index reconstitution events, there little change in overall institutional ownership.[13] In the appendix, I confirm that in both the case of being added to the S&P 500 and switching from the Russell 1000 to the Russell 2000, on average, there is no change in institutional ownership around the index event. While this initially seems counter intuitive, it makes sense when we think about how we expect retail (i.e., everyone who is not an institutional investor) to trade around such events. For example, when FootLocker was dropped from the S&P 500 in 8/2019, I find it unlikely that retail investors (1) read the press release on S&P's website and (2) even if they did read the press release, decided to sell their shares to institutions.

To further alleviate the concern that total institutional ownership is driving my results, in the cross-sectional regressions, I can and do explicitly control for total institutional ownership. Finally, in the appendix, I replicate all the instrumental variables and reduced form regressions in this section, including total institutional ownership on the right hand side. All the results are quantitatively unaffected by this change.

5.3.2 Significance of IV vs. reduced form

In Tables 4 and 5, the IV is always significant, while the reduced form is sometimes insignificant. The concern is that, as discussed in Chernozhukov and Hansen (2008), an insignificant reduced form is evidence of weak instruments. At a high level, this is likely not a problem in my setting, as the first stage is very strong ($F > 300$ for the S&P experiment and $F > 100$ for the Russell experiment). 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 provide a possible economic mechanism to explain this difference in statistical significance.

5.3.3 Re-use of natural experiments

A final concern with the results in this section is that many previous studies have used additions to the S&P 500 and switching between the Russell 1000 and 2000 as natural

[13] See Appel et al. (2020) for a more detailed discussion regarding the effects of Russell reconstitution on total institutional ownership and passive ownership. A related concern, also raised in Appel et al. (2020), is that for the Russell experiment, the treatment is correlated with firm size. Given that my results are robust to both S&P 500 index additions, which applies to growing firms, and switching from the Russell 1000 to the Russell 2000, which applies to shrinking firms, I find it unlikely that a pure size effect is driving my results.

experiments to study 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. A particular issue is that even if these original studies were correct, 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 5 shows that almost all of my IV t-statistics are over 3, so even if previous research has looked at the effect of Russell reconstitution on over 20 other distinct outcomes, my results are unlikely to be spurious. Further, the Russell results yield similar estimates to the S&P results, even though these experiments have different implications for other known outcomes, again allaying concerns that my results are driven by factors other than passive ownership.

6 Why passive ownership decreases price informativeness

In this section, I argue that passive ownership decreases price informativeness because passive investors gather less firm-specific information. Consistent with this, I show that passive ownership is correlated with fewer downloads of SEC filings and Bloomberg searches, as well as decreased coverage and accuracy by sell-side analysts. I present additional evidence consistent with lower information gathering, specifically that passive ownership is correlated with (1) decreased pre-earnings trading (2) larger earnings responses (3) more ex-ante earnings uncertainty and (4) larger post-earnings-announcement increases in short interest for firms that end up releasing bad news.

6.1 Information gathering

An intuitive explanation for a decrease in price informativeness is that fewer people are gathering information and/or those who gather information are expending less effort to obtain precise signals. This negative relationship between learning and price informativeness arises in many workhorse models of trade under asymmetric information (see e.g., Grossman and Stiglitz (1980) and Kyle (1985)).

Passive managers, as well as investors in passive funds, lack strong incentives to gather and consume firm-specific information. Passive funds trade on mechanical rules, such as S&P 500 index membership (SPY), or the 100 lowest volatility stocks in the S&P 500 (SPLV). Given that these trading strategies are implemented on public signals, they do not require accurate private forecasts of firm fundamentals. Further, because these funds are well diversified, even if they are traded by informed investors, they are more likely to be used for bets on systematic, rather than firm-specific, information. This yields an empirical prediction for why passive ownership decreases price informativeness.

Prediction 3: *Passive ownership should be correlated with decreased gathering of stock-specific information*

6.1.1 Bloomberg terminal searches

One way to test prediction 3 is to examine 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 stories, 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 the following regression:

$$AIAC_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t} \tag{9}$$

where $AIAC_{i,t}$ is the continuous abnormal institutional attention measure from Ben-Rephael et al. (2017).[14] All the controls and fixed effects are identical to equation 4. The sample is stock/month observations between 2010 and 2018 that can be linked between Bloomberg and CRSP on ticker.

[14] 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.

Table 6 contains the results. Consistent with passive ownership decreasing information gathering, it is correlated with fewer Bloomberg searches. In terms of magnitudes, a 15% higher level of passive ownership implies -0.22 lower institutional investor attention, which is about 20% of its whole sample mean of 0.97. If the channel 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% $\approx$ 20%, which is almost exactly what we see in Table 6.

	(1)	(2)	(3)	(4)	(5)
Passive Ownership	-1.207***	-1.354***	-1.515***	-1.885**	-2.291**
	(0.288)	(0.331)	(0.372)	(0.731)	(0.888)
Observations	62,292	58,629	58,629	58,629	58,629
R-Squared	0.529	0.529	0.542	0.483	0.496
Firm + Year/Quarter FE	✓	✓	✓	✓	✓
Matched to Controls		✓	✓	✓	✓
Firm-Level Controls			✓		✓
Weight	Equal	Equal	Equal	Value	Value

Table 6 Passive ownership and Bloomberg terminal searches. Estimates of β from:
$AIAC_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$
where $AIAC$ is the measure of continuous abnormal institutional attention from Ben-Rephael et al. (2017). Controls in $X_{i,t}$ include firm 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.

6.1.2 Downloads of SEC filings

As an alternative measure of investors' learning behavior, I examine downloads of SEC filings, with fewer downloads implying decreased gathering of fundamental information (see e.g., 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. Downloads have been trending up over time and there is typically a large increase in downloads – relative to the unconditional firm-level average – on earnings announcement dates.

To test prediction 3, I run a regression identical to equation 9, but with $Downloads_{i,t}$ on the left-hand-side. Table 7 contains the results. In column 1, which only has firm/time fixed effects, the relationship is positive and statistically significant. This could be explained, however, by the fact that high passive firms are larger on average and big firms attract more investor attention. Column 3 adds in all the firm level controls, including one-month lagged market capitalization and total institutional ownership. This flips the sign, making the coefficient negative and statistically significant, evidence that passive ownership is correlated with less investor attention conditional on observable differences between high and low passive ownership firms. This result is consistent with Israeli et al. (2017) and Coles et al. (2020), who also show that passive ownership is negatively correlated with downloads of SEC filings.

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 the whole sample standard deviation of roughly 1.3. This magnitude, however, is harder to interpret than the results in Table 6, as we don't know who is downloading these SEC filings and whether or not they themselves are investors.

6.1.3 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 people buying the individual stocks. In addition to this direct effect on information gathering, however, we could imagine there are equilibrium effects on the supply of information. Specifically, given that information is not costless to produce, a change in demand may have corresponding equilibrium effects on the supply of information. Although sell-side analysts are not necessarily investors themselves, if passive ownership leads to less demand for information, they may respond by producing less or lower quality information about stocks with higher levels of passive ownership. On the other hand, as a stock becomes more mispriced, the return to gathering fundamental information increases. Given that passive ownership

	(1)	(2)	(3)	(4)	(5)
Passive Ownership	0.555***	-0.12	-1.135***	-0.572	-1.407**
	(0.157)	(0.170)	(0.198)	(0.542)	(0.653)
Observations	640,366	533,099	533,099	533,099	533,099
R-Squared	0.784	0.807	0.81	0.889	0.892
Firm + Year/Quarter FE	✓	✓	✓	✓	✓
Matched to Controls		✓	✓	✓	✓
Firm-Level Controls			✓		✓
Weight	Equal	Equal	Equal	Value	Value

Table 7 Passive ownership and downloads of SEC filings. Estimates of β from:
$Downloads_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$
Downloads is the number of non-robot downloads from Loughran and McDonald (2017). Controls in $X_{i,t}$ include firm 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.

reduces price informativeness, analysts might decide to produce more information about the underlying stocks to e.g., attract business to their firm by increasing trading in the stocks they cover (see e.g., Martineau and Zoican (2021)).

Ex-ante, it is not obvious which of these effects will dominate in equilibrium. For example, as discussed in Coles et al. (2020), as passive ownership grows, gathering information becomes more profitable, which may lead more people to become informed. In their model, these two effects exactly offset and passive ownership has no net effect on price informativeness. To distinguish between these alternative stories, I run the following regression:

$$Outcome_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t} \tag{10}$$

where $Outcome_{i,t}$ will be measures of information production by sell-side analysts. 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 8 contains the results. Column 1 shows that higher passive ownership is correlated with lower analyst coverage. This is consistent with Israeli et al. (2017) and Coles et al. (2020), 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. This increased forecast dispersion is evidence of more uncertainty about the fundamental value of these firms (see e.g., Diether et al. (2002), Zhang (2006)).

One concern, however, is the increased standard deviation of forecasts is a mechanical function of the decrease in coverage documented in column 1. To address this, I construct a measure of analyst (in)accuracy which accounts for the increase in dispersion: 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 and/or expending less effort, we would expect their forecasts to be less accurate, even when accounting for the average increase in uncertainty. Column 3 shows that analysts' forecasts are less accurate for firms with more passive ownership.

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 estimate 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 high passive ownership.

6.1.4 Relationship between attention and index addition

The findings in Table 8 seem at odds with results in Section 5 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, in the appendix, I show that even though S&P 500 index addition leads

	Num. Est (1)	SD(Est.) (2)	Dist./SD(Est.) (3)	Updates (4)	Time (5)
Passive Ownership	-11.64***	0.719***	1.968***	-0.446***	0.354***
	(1.382)	(0.174)	(0.445)	(0.099)	(0.107)
Observations	216,805	216,805	216,805	133,176	133,176
R-squared	0.789	0.643	0.125	0.256	0.549
Mean	8.624	0.0932	2.23	2.233	3.764
St. Dev.	5.941	0.406	2.969	0.447	0.841

Table 8 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 firm 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. The last two rows of the table present the mean and standard deviation of the left-hand-side variables.

to increased analyst coverage, it also leads to (1) increased dispersion in analyst forecasts and (2) 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.

6.1.5 Why equilibrium information gathering decreases

Suppose passive investors gather no stock-specific information. Then, as passive ownership increases, we would mechanically expect total information gathering to decrease, holding the behaviour of the remaining non-passive investors fixed. 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 this, as occurs in the model of Coles et al. (2020).

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. One channel for this, as discussed in Ben-David et al. (2018), is that 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. An implication of this is that the effects I document in section 4 should be stronger for ETFs than non-ETF passive funds. In the appendix, consistent with this, I show that ETFs have a larger effect on $DM_{i,t}$ and $QVS_{i,t}$ than passive mutual funds. I caveat, however, that the effects of these two different types of passive ownership are hard to separate, as they have a correlation coefficient of almost 0.7.

Another possible channel is that passive ownership decreases pre-earnings liquidity. Consistent with this, as I show in the next subsection, stocks with more passive ownership have relatively less pre-earnings trading volume.[15] 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 that you are trading on information and push prices against you.

Risk and liquidity are just two of many possible explanations for why non-passive investors don't fully compensate for the lack of information gathering by passive investors. These general equilibrium effects, however, are hard to measure. All the evidence in tables 6, 7 and 8 only speak to net changes in information supply/demand, so it is possible that non-passive investors respond by gathering more information. Without being able to see individual investors' attention, however, it is difficult to quantify such effects.

6.2 Implications of decreased information gathering

In this subsection, I show that – consistent with decreased information gathering – stocks with more passive ownership tend to have more ex-ante earnings uncertainty, are traded less

[15] This is just one of many possible explanations for a decrease in pre-earnings trading volume. For example, in Foster and Viswanathan (1990) and Foster and Viswanathan (1993), there is less trading volume when informed investors have a relatively larger informational advantage, because this makes uninformed investors more concerned about adverse selection.

before earnings announcements and have larger responses to earnings news of a given size. I also show that passive ownership is correlated with larger increases in short interest after the release of bad earnings news.

6.2.1 Option implied volatility

Because DM and QVS are computed using earnings-day returns, they are ex-post measures of uncertainty. If fewer investors are becoming informed or investors' signals have become less precise, however, we would also expect an increase in ex-ante uncertainty. One way to measure ex-ante uncertainty about fundamentals is via the cost of options exposed to earnings announcement risk (see e.g., Dubinsky et al. (2006)). Following Kelly et al. (2016), I compute a version of their Implied Volatility Difference (IVD) measure to quantify how much more expensive options that span earnings announcements are, relative to options that expire the month before/after the announcement.[16]

Letting τ denote an earnings announcement date, I identify regular monthly expiration dates a, b and c, such that $a < \tau < b < c$. The final variable of interest, the implied volatility difference, is defined as:

$$IVD_{i,\tau} = \overline{IV}_{i,b} - \frac{1}{2}\left(\overline{IV}_{i,a} + \overline{IV}_{i,c}\right) \tag{11}$$

higher values of $IVD_{i,\tau}$ imply that options which span earnings announcements are relatively more expensive i.e., there is more ex-ante uncertainty about the earnings news.[17]

Prediction 4: *If passive ownership decreases information gathering, options which span earnings announcements should be relatively more expensive for high passive ownership stocks*

The appendix shows that average IVD is positive and has increased by about 0.05 over the past 25 years. This is evidence that there is more uncertainty about fundamentals before earnings announcements now than there was in the late 1990s. Table 9 contains the results of a regression of IVD on passive ownership, and the same controls/fixed effects as equation 4. Consistent with prediction 4, IVD is positively correlated with passive ownership. In terms of magnitudes, a 15% increase in passive ownership implies about a 0.02 higher IVD on average. This is an economically large effect, at about 40% of the increase in average

[16] See the appendix for step-by-step details on my exact variable construction procedure.

[17] 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. All the results that follow are qualitatively unchanged using $\tilde{IVD}_{i,\tau} = \overline{IV}_{i,b}/\frac{1}{2}\left(\overline{IV}_{i,a} + \overline{IV}_{i,c}\right)$.

IVD over the whole sample. In the appendix, I show that IVD increases both for stocks added to the S&P 500 index, as well as stocks that switch from the Russell 1000 to the Russell 2000.

	(1)	(2)	(3)	(4)	(5)
Passive Ownership	0.107***	0.126***	0.0958***	0.150***	0.171***
	(0.026)	(0.028)	(0.029)	(0.035)	(0.033)
Observations	118,809	111,415	111,415	111,415	111,415
R-Squared	0.273	0.281	0.286	0.416	0.423
Firm + Year/Quarter FE	✓	✓	✓	✓	✓
Matched to Controls		✓	✓	✓	✓
Firm-Level Controls			✓		✓
Weight	Equal	Equal	Equal	Value	Value

Table 9 Passive ownership and IVD for earnings announcements. Estimates of β from:
$IVD_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$
Controls in $X_{i,t}$ include firm 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.

6.2.2 Pre-earnings abnormal turnover

An additional place that decreased information gathering may manifest is in trading volume. If fewer people are gathering information, we might see relatively less trading in the pre-earnings announcement period because if fewer people have private information, there is effectively less disagreement/investor heterogeneity, which is typically cited as an explanation for trading volume (see e.g., Wang (1994)). Further, when investors have less precise information, they trade less aggressively, which may also lead to less volume (see e.g., Kyle (1985), where a decrease in signal precision decreases order size for a given signal magnitude). This pattern – of increased trading before news when investors have more information – also arises in models studying trade around other news events (see e.g., Manela (2014)). More broadly, even if investors are gathering information, price discovery doesn't occur in a vacuum and investors need to trade for information to be incorporated into prices.

So, the decrease in trading volume before earnings announcements for high passive ownership stocks may be in and of itself evidence of decreased price informativeness. Collectively, these ideas lead to a testable implication of decreased information gathering.

Prediction 5: *If passive ownership decreases information gathering, there should be a negative relationship between pre-earnings trading volume and passive ownership*

To quantify this, 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:

$$AT_{i,t+\tau} = \frac{T_{i,t+\tau}}{\overline{T_{i,t-22}}} = \frac{T_{i,t+\tau}}{\sum_{k=1}^{252} T_{i,t-22-k}/252} \tag{12}$$

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 dampening drops 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, define cumulative abnormal pre-earnings turnover as:

$$CAT_{i,t} = \sum_{\tau=-22}^{-1} AT_{i,t+\tau} \tag{13}$$

the sum of abnormal turnover from $t-22$ to $t-1$ for firm i around earnings date t. 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 the following regression with quarterly data to measure the relationship between pre-earnings turnover and passive ownership:

$$CAT_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t} \tag{14}$$

where cumulative abnormal pre-earnings turnover, $CAT_{i,t}$, is defined in Equation 13 and all the controls and fixed effects are the same as in Equation 4.[18]

The regression results are in Table 10. The baseline specification (column 3) implies that a 15% increase in passive ownership would lead to a decline in cumulative abnormal pre-earnings turnover of -1.68. This effect is economically large, especially relative to the average change in pre-earnings turnover of about -1 between the early 1990s and late 2010s. The other columns show this result is not sensitive to the inclusion of firm-level controls, or using value weights instead of equal weights. In the appendix, I show that pre-earnings abnormal trading volume also drops after being added to the S&P 500 index and when switching from the Russell 1000 to the Russell 2000.

	(1)	(2)	(3)	(4)	(5)
Passive Ownership	-12.86***	-12.19***	-11.49***	-8.658***	-10.25***
	(2.956)	(2.969)	(3.207)	(2.900)	(2.768)
Observations	428,393	407,283	407,283	407,283	407,283
R-Squared	0.06	0.061	0.082	0.144	0.145
Firm + Year/Quarter FE	✓	✓	✓	✓	✓
Matched to Controls		✓	✓	✓	✓
Firm-Level Controls			✓		✓
Weight	Equal	Equal	Equal	Value	Value

Table 10 Passive ownership and pre-earnings turnover. Estimates of β from:
$CAT_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + \phi_t + \psi_i + e_{i,t}$
where $CAT_{i,t}$ is cumulative abnormal pre-earnings turnover. Controls in $X_{i,t}$ include firm 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.

One concern is that the results in Table 10 are purely mechanical. Passive investors may trade less and therefore passive ownership just leads to less trading overall. First, I find this

[18] I am not the first researcher to use volume as an indicator of informed investor behavior. Specifically, Manela (2014) uses cumulative abnormal turnover ($CATO$) around another set of large news events – FDA drug approvals – to provide evidence on the speed with which information is incorporated into prices. In the appendix I show (1) I obtain similar results using $CATO$ instead of CAT and (2) I provide a more detailed explanation of CAT's unique advantages in my setting.

unlikely, as ETFs have higher turnover than stocks on average. Even so, suppose passive ownership did lead to less trading volume. The way CAT is defined 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 12. 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.

6.2.3 Earnings responses

An additional implication of decreased information gathering is that the response to fundamental news of a given size should increase. The intuition is that if investors have less precise beliefs before an announcement, they will update significantly afterwards, leading to a larger price change. Further, given that passive ownership is more likely to decrease attention to firm-specific information, effect should be stronger for the firm-specific component of stock news. This yields an additional testable prediction.

Prediction 6: *If passive ownership decreases information gathering, prices should respond more to earnings news of a given size. This effect should be especially strong for firm-specific news*

I use an earnings-response regression based on Kothari and Sloan (1992) to quantify the market's reaction to a standardized measure of earnings news. In the appendix, I show earnings response coefficients increased by a factor of about 3× over the past 30 years i.e., the market reaction to earnings news of a given size is about 3 times as large now as it was in the early 1990s. To test prediction 6, I run the following regression:

$$\begin{aligned} r_{i,t} = \alpha + \beta_1 SUE_{i,t} + \phi_1 Passive_{i,t} + \gamma_1 \left(SUE_{i,t} \times Passive_{i,t}\right) \\ + \gamma X_{i,t} + \phi_t + \psi_i + \epsilon_{i,t} \end{aligned} \tag{15}$$

where $r_{i,t}$ denotes the market-adjusted return on the effective quarterly earnings date i.e., the first day investors could trade on earnings information. 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. $SUE_{i,t}$ is defined identically to equation 1: 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.[19] I also run versions of Equation 15 (1) breaking SUE into positive and negative components and (2) decomposing the earnings news into a systematic and idiosyncratic component using the method in Glosten et al. (2021). This is done by regressing 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 11 contains the regression results. Columns 1-3 are a sanity check and do not include the interaction terms with passive ownership. Everything is consistent with common-sense intuition: (1) SUE is positively correlated with earnings-day returns (2) this is true individually both for the positive and negative components of SUE and (3) this is also true individually for the positive/negative components of SUE when decomposing earnings news into systematic and idiosyncratic components. Columns 4-6 add the level of passive ownership, as well as all the interaction terms with passive ownership to Columns 1-3. In Column 4, consistent with prediction 6, γ is positive and economically large. The estimates in column 5 imply that firms with a high share of passive ownership are especially responsive to negative news and column 6 shows this is driven by idiosyncratic news, also consistent with prediction 6. I explore why passive ownership has a larger effect on negative news in the next subsection.

6.2.4 Ease of shorting

Table 2 suggests that passive ownership played an important role in the average decline of the pre-earnings drift documented in Figure 1. As shown in the appendix, however, the pre-earnings drift decreased relatively more for firms which ended up releasing bad news. While this may be an artifact of differences in average returns or the nature of news between these time periods, in this subsection, I aim to understand whether passive ownership plays a role in driving this asymmetry.

One possible mechanism for this asymmetry is that an investor's hurdle rate for shorting may be higher for than long-only investments. This is because there are frictions associated

[19] 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. As a result, 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	0.528***			0.402***		
	(0.022)			(0.022)		
$SUE \times 1_{SUE>0}$		0.715***			0.655***	
		(0.028)			(0.035)	
$\lvert SUE \rvert \times 1_{SUE \leq 0}$		-0.409***			-0.253***	
		(0.023)			(0.022)	
$Sys.SUE \times 1_{Sys.SUE>0}$			0.299**			0.213
			(0.122)			(0.143)
$\lvert Sys.SUE \rvert \times 1_{Sys.SUE \leq 0}$			-0.400*			-0.417**
			(0.215)			(0.175)
$Idio.SUE \times 1_{Idio.SUE>0}$			0.707***			0.667***
			(0.031)			(0.037)
$\lvert Idio.SUE \rvert \times 1_{Idio.SUE \leq 0}$			-0.442***			-0.274***
			(0.025)			(0.024)
SUE				2.816***		
$\times Passive$				(0.261)		
$SUE \times 1_{SUE>0}$					0.889**	
$\times Passive$					(0.386)	
$\lvert SUE \rvert \times 1_{SUE \leq 0}$					-4.127***	
$\times Passive$					(0.361)	
$Sys.SUE \times 1_{Sys.SUE>0}$						1.568
$\times Passive$						(1.076)
$\lvert Sys.SUE \rvert \times 1_{Sys.SUE \leq 0}$						0.982
$\times Passive$						(2.944)
$Idio.SUE \times 1_{Idio.SUE>0}$						0.562
$\times Passive$						(0.444)
$\lvert Idio.SUE \rvert \times 1_{Idio.SUE \leq 0}$						-4.074***
$\times Passive$						(0.352)
Observations	354,799	354,799	354,799	354,799	354,799	354,799
R-squared	0.062	0.063	0.063	0.063	0.064	0.064

Table 11 Passive ownership and earnings responses. Estimates from:
$r_{i,t} = \alpha + \beta_1 SUE_{i,t} + \phi_1 Passive_{i,t} + \gamma_1 (SUE_{i,t} \times Passive_{i,t}) + \gamma X_{i,t} + \phi_t + \psi_i + \epsilon_{i,t}$
where $r_{i,t}$ is the market-adjusted return (in percentage points) on the effective earnings announcement date. Controls in $X_{i,t}$ include firm 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.

with short selling which are not present when buying a stock (e.g., the cost of borrowing the security, the possibility of a short squeeze and getting margin called). So, one explanation for the asymmetry is that when prices are too high, correcting them is harder than when prices are too low. Therefore, if passive ownership makes prices less informative, we might expect that this effect would be stronger for firms which eventually release bad news. If this is the case, however, once the bad news is released, uncertainty should go down and investors should accordingly adjust their short positions. This yields a testable prediction.

Prediction 7: *Firms with more passive ownership should have a larger increase in short interest after bad news is announced*

Prediction 7 specifically says *increase* to account for the fact that passive ownership makes it easier to short the underlying stocks because passive funds lend out their shares to generate additional income (see e.g., Beschwitz et al. (2020)).[20] Following Hanson and Sunderam (2014), I define the short interest ratio as $SR_{i,t} = SHORT_{i,t}/SHROUT_{i,t}$, where I obtain bi-weekly data on the number of shares shorted from Compustat and the number of shares outstanding from CRSP. Consistent with this, in the appendix I show that both before and after earnings announcements, passive ownership is correlated with a relatively higher level short interest.

To calculate the change in short interest around earnings announcements, I start by identifying the last date where short interest was reported which is also at least 5 days before the earnings announcement, and the first date where short interest was reported at least 5 days after the earnings announcement. I compute the change in short interest, $\Delta SR_{i,t}$ as the difference in $SR_{i,t}$ between these two dates. Prediction 7 specifies that the effect should be concentrated in firms which release bad news, so to this end, I form quintiles of SUE each quarter, where SUE is defined in equation 1.

Finally, I run a regression with $\Delta SR_{i,t}$ on the left-hand-side, dummy variables for quintiles of SUE on the right-hand-side, and the interaction between these dummy variables and passive ownership. If prediction 7 is correct, the interaction terms between the lower quintiles of SUE and passive ownership should be positive. Table 12 contains the results. In all specifications: (1) the middle quintile of SUE is the omitted group and (2) I include the same set of fixed effects as equation 4. Column 1 is a check for an unconditional correlation between passive ownership and changes in short interest around earnings announcements. It shows that on average, there is no relationship between these two quantities. Columns 2 and 3 show that, consistent with prediction 7, there is a relatively larger increase in short interest for firms that have both higher passive ownership and worse earnings news.

Columns 4-6 replicate columns 1-3, and show the results are robust to using value weights

[20]This pattern may not be specific to passive ownership, and may apply to institutional ownership more broadly, see e.g., Daniel et al. (2017).

instead of equal weights. In terms of magnitudes, a 15% higher level of passive ownership implies an additional increase in short interest of 10 basis points after releasing bad news. This effect seems small, but it is large relative to the mean change in short interest for firms in the bottom quintile of SUE, which is 5 basis points.

An alternative explanation for the asymmetric effect of passive ownership on the pre-earnings drift – which doesn't imply less informative pre-earnings-announcement prices – is that by making short selling easier, passive ownership leads negative earnings information to be incorporated into prices faster after it is released. This may also be a factor in the apparent decline in the post-earnings announcement drift visible in Figure 1.[21]

If this were the case, however, we would expect that including a sufficiently long post-earnings announcement window in the return attributed to the information release would eliminate the asymmetry between firms that end up releasing good news and bad news. One way to do this is to include $n > 0$ days in the earnings announcement return when computing $DM_{i,t}$. If this alternative story is true, increasing n should decrease the asymmetry between positive and negative announcements, as the speed with which information is incorporated after the announcement becomes irrelevant. In the appendix, and inconsistent with increased ease of shorting driving the relationship between passive ownership and the pre-earnings drift, the asymmetry remains large, even at $n = 5$ days.

7 Conclusion

In this paper, I propose two ways to measure price informativeness without relying on particular model. To do this, I focus on earnings announcements, which yields empirical predictions for how returns and volatility would look if prices were more or less informative. I show that over the past 30-years, pre-earnings announcement price informativeness has been steadily declining.

I argue passive ownership played an important role in this trend, with my cross-sectional regression results suggesting that a 15% increase in passive ownership, roughly the value-weighted average increase between 1990 and 2018, is able to explain 33%-40% of the change

[21]See e.g., Martineau (2018) for a detailed discussion of the decline in the post-earnings announcement drift over the past 40 years. Also see Dugast and Foucault (2018), who show how faster processing of new data (e.g., earnings) could make post-announcement prices more informative, while simultaneously discouraging pre-earnings announcement information production.

	(1)	(2)	(3)	(4)	(5)	(6)
	Ch. SI	Ch. SI	Ch. SI	Ch. SI	Ch. SI	Ch. SI
Passive	-0.147	-0.366*	-0.0934	-0.305*	-0.441**	0.0856
	(0.202)	(0.208)	(0.220)	(0.161)	(0.192)	(0.291)
Low SUE x Passive		0.759***	0.834***		0.737***	0.726***
		(0.183)	(0.177)		(0.209)	(0.219)
2 x Passive		0.304**	0.336**		0.135	0.11
		(0.138)	(0.140)		(0.184)	(0.185)
3 x Passive		0.0528	0.0893		0.0883	0.0791
		(0.161)	(0.161)		(0.178)	(0.188)
High SUE x Passive		-0.148	-0.147		-0.216	-0.265
		(0.186)	(0.188)		(0.196)	(0.207)
Observations	238,794	238,794	230,758	238,794	238,794	230,758
R-squared	0.043	0.044	0.046	0.033	0.033	0.035
Firm + Year/Quarter FE	✓	✓	✓	✓	✓	✓
Matched to Controls			✓			✓
Firm-Level Controls			✓			✓
Weight	Equal	Equal	Equal	Value	Value	Value

Table 12 Passive ownership and changes in short interest around earnings announcements. Estimates of β and ψs from:

$$\Delta SI_{i,t} = \alpha + \beta Passive_{i,t} + \sum_{k=1}^{5} \phi_k 1_{SUE_{i,t} \in k} + \sum_{k=1}^{5} \psi_k 1_{SUE_{i,t} \in k} \times Passive_{i,t} + \gamma X_{i,t} + \phi_t + e_{i,t}$$

where k denotes quintiles of SUE, formed within each quarter. Controls in $X_{i,t}$ include firm 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.

in average price informativeness. Passive ownership is not randomly assigned in the cross-section, but I show that my findings are robust to using only quasi-exogenous increases in passive ownership arising from S&P 500 index addition and Russell 1000/2000 index rebalancing.

My proposed mechanism is that passive investors lack strong incentives to gather firm-specific information. Consistent with this, I show that passive ownership is correlated with decreased downloads of SEC filings and fewer Bloomberg terminal searches. This may have feedback effects on the supply of information, as high passive stocks also have lower coverage and accuracy by sell-side analysts. In addition to this direct evidence on decreased information gathering, I also show indirect effects. Specifically, higher levels of passive ownership are correlated with increased ex-ante earnings uncertainty, decreased pre-earnings trading volume, increased earnings responses and bigger changes in short interest after bad news.

Relative to total institutional ownership, passive ownership is still small, owning only about 15% of the US stock market. Even at this low level, passive ownership has led to large changes in returns, trading patterns and the way stock prices respond to firm-specific news. As passive ownership continues to grow, these changes in information and trading 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).

Bartram, S. M., Brown, G. W., and Stulz, R. M. (2019). Why is there a secular decline in idiosyncratic risk in the 2000s? *Fisher College of Business Working Paper*, 1(2019-03):019.

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

Ben-David, I., Franzoni, F. A., and Moussawi, R. (2019). An Improved Method to Predict Assignment of Stocks into Russell Indexes. *SSRN Electronic Journal*.

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. (2020). Does the stock market make firms more productive? *Journal of Financial Economics*, 136(2):281–306.

Beschwitz, B. v., Honkanen, P., and Schmidt, D. (2020). Passive ownership and short selling. *Available at SSRN*.

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).

Chang, Y. C., Hong, H., and Liskovich, I. (2015). Regression discontinuity and the price effects of stock market indexing. *Review of Financial Studies*, 28(1).

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

Coles, J. L., Heath, D., and Ringgenberg, M. (2020). On Index Investing. *SSRN Electronic Journal*.

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

Daniel, K., Klos, A., and Rottke, S. (2017). Overpriced winners. *Unpublished Working Paper, Columbia University*.

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

Dávila, E. and Parlatore, C. (2021). Trading Costs and Informational Efficiency. *Journal of Finance*, 76(3).

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.

Dubinsky, A., Johannes, M., Carr, P., Chen, J., Chernov, M., Duffie, D., Glasserman, P., Heidari, M., Hull, B., Jagannatahn, R., Jones, C., Knight, J., and Linnainmaa, J. (2006). Fundamental uncertainty , earning announcements and equity options. *Wall Street Journal*, 1(August 2015).

Dugast, J. and Foucault, T. (2018). Data abundance and asset price informativeness. *Journal of Financial economics*, 130(2):367–391.

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

Fama, E. F., Fisher, L., Jensen, M. C., and Roll, R. (1969). The Adjustment of Stock Prices to New Information. *International Economic Review*, 10(1).

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).

Hanson, S. G. and Sunderam, A. (2014). The growth and limits of arbitrage: Evidence from short interest. *The Review of Financial Studies*, 27(4):1238–1286.

Heath, D., Macciocchi, D., Michaely, R., and Ringgenberg, M. C. (2021). Do Index Funds Monitor? *The Review of Financial Studies*.

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.

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. (2018). The Evolution of Market Price Efficiency Around Earnings News. *SSRN Electronic Journal*.

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*.

Mclean, R. D. and Pontiff, J. (2016). Does Academic Research Destroy Stock Return Predictability? *Journal of Finance*, 71(1).

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*.

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

S&P, D. J. I. (2017). S&p u.s. indices methodology. Technical report, S&P.

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

Wei, W. and Young, A. (2017). Selection Bias or Treatment Effect? A Re-Examination of Russell 1000/2000 Index Reconstitution. *SSRN Electronic Journal*.

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

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