Passive Ownership and the Stock Market

Marco Sammon*

September 3, 2018

Abstract

Markets have become less efficient at incorporating earnings information over the past 30 years. At the stock level, increases in passive ownership are an economically and statistically significant predictor of decreases in pre-earnings price informativeness. This result is robust to using only quasi-exogenous variation in passive ownership that arises from S&P 500 index addition, and Russell 1000/2000 index reconstitution. A potential mechanism is that high passive ownership reduces incentives to gather firm-specific information. Consistent with this mechanism, increases in passive ownership are correlated with fewer analysts covering a stock, decreased analyst accuracy, and fewer downloads of SEC filings.

1 Introduction

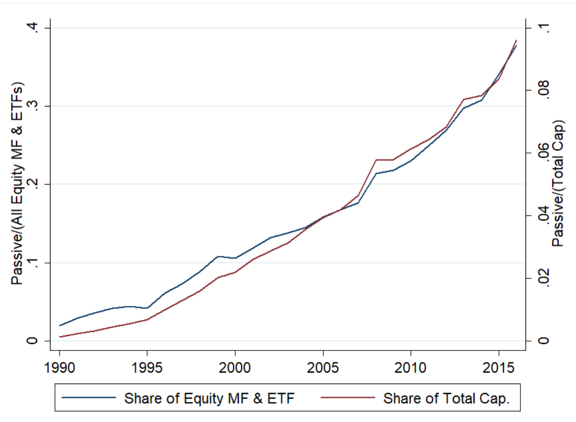
Passive ownership of stocks has grown substantially over the past 30 years. Figure 1 shows that passive ownership increased from almost zero in 1990, to nearly 40% of total mutual fund/ETF assets in 2016, owning about 10% of the total market capitalization of US firms.

Theories of asymmetric information, such as Grossman and Stiglitz (1980), predict that as the share of uninformed investors rises, prices become less infor-

*Kellogg School of Management. Email: mcsammon@gmail.com. I would like to thank Scott Baker, Ravi Jagannathan, Robert Korajczyk, Dimitris Papanikolaou and participants at the Kellogg Bag Lunch for helpful comments.

Rise of Passive Ownership

Figure 1: Passive ownership is defined as all index funds, as identified by the index fund flag in the CRSP mutual fund data, all ETFs that are not actively managed, and all mutual funds with "index" in the name. Total equity mutual fund and ETF assets is the sum of all stock holdings in the Thompson S12 data. Total market capitalization includes all CRSP firms.



mative before uncertainty is resolved. I provide empirical evidence consistent with these claims: As passive ownership in a stock increases, prices become less informative before earnings information is released.

I study three measures of pre-earnings informativeness (1) Cumulative abnormal volume before earnings announcements: If prices are less informative before earnings announcements, fear of adverse selection will deter uninformed investors from trading. (2) Pre-earnings drift: If earnings information is known before it is formally released, returns on earnings days should be small relative to the cumulative returns leading up to earnings days. (3) Earnings day share of annual volatility: The expected variance of returns on uncertainty resolution dates should be decreasing in pre-resolution price informativeness.

Over the past 30 years, pre-earnings volume, and the pre-earnings drift have been trending down, while the share of volatility on earnings days has been trending up. These trends are consistent with a decrease in the share of earnings information incorporated into prices before the announcement date.

Across all three measures, there is a negative relationship between passive ownership and pre-earnings price informativeness. This reduced-form result, however, does not rule out the possibility that unobserved or omitted factors are driving both the increases in passive ownership, and decreases in price informativeness. I exploit S&P 500 index additions, as well as Russell 1000/2000 index reconstitutions to identify increases in passive ownership which are plausibly uncorrelated with firm fundamentals. These quasi-exogenous increases in passive ownership are also negatively correlated with pre-earnings price informativeness.

One potential mechanism is that high passive ownership reduces the incentives to gather 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 rules are implemented on public signals, they reduce the need to gather accurate private forecasts of firm fundamentals. Consistent with decreased information gathering, there is a negative relationship between increases in passive ownership and analyst coverage, the accuracy of analyst forecasts, and downloads of SEC filings.

My empirical results confirm the predictions from several models of asymmetric information. Grossman and Stiglitz (1980) predicts that as the share of uninformed investors increase, a security's price will be less informative before uncertainty is resolved. I show that increases in passive ownership lead to a lower pre-earnings drift, suggesting less earnings information is incorporated into prices before the announcement date. Wang (1994) predicts that less informative prices before information releases would lead to lower trading volume through a fear of adverse selection. I find that higher passive ownership leads to less pre-earnings trading. Buffa, Vayanos, and Woolley (2014) predicts that stocks with high demand by "buy and hold" investors, whose behavior is similar to passive managers in practice, will respond more to expected cashflow shocks. I find that firms with high passive ownership respond more to a given level of earnings surprise than firms with low passive ownership.

I also contribute to the empirical literature on the effects of growing ETFs and passive ownership. Israeli, Lee, and Sridharan (2017) find that ETFs decrease the information content of the underlying securities' prices. I sharpen this result, providing causal evidence on the effect of increased passive ownership through S&P 500 index addition and Russell 1000/2000 reconstitution. Glosten, Nallareddy, and Zou (2016) find that ETFs increase the incorporation of systematic news into prices for otherwise information deficient stocks. My result does not contradict theirs, as my focus is on firm-specific information. Our results together imply a trend toward trading baskets of securities rather than individual stocks, consistent with the increasing AUM and trading volume in ETFs. Ben-David, Franzoni, and Moussawi (2014) find that ETFs increase idiosyncratic volatility by transmitting liquidity shocks from the basket to underlying stocks. I find that volatility increases on earnings days, relative to all other days. Our results together are consistent with investors trading simultaneously in the individual stocks and the basket on earnings days.

Section 2 presents evidence on the decrease in pre-earnings price informativeness, Section 3 shows the reduced-form relationships between high passive ownership and decreased information in prices. Section 4 uses index rebalancing to create a causal link between passive ownership and decreased price informativeness. Section 5 investigates potential mechanisms, while Section 6 concludes.

2 New Stylized Facts

In this section, I present three measures of pre-earnings price informativeness, and show that these quantities have decreased over the past 30 years.

2.1 Decline of Pre-Earnings Volume

Fact 1: Volume before earnings announcements has declined.

Let t denote an earnings announcement date, or the next trading date if earnings are announced when markets are closed. Define pre-earnings cumulative abnormal volume *per trading day*, for firm i, from time $t-22$ to $t-j$ as:

$$\overline{CAV_{i,j,t}} = \frac{\sum_{\tau=-22}^{-j} AV_{i,t+\tau}}{23-j} \tag{1}$$

Where abnormal volume, $AV_{i,\tau}$, is volume relative to historical average volume over the past year:

$$AV_{i,\tau} = \frac{V_{i,t+\tau}}{\overline{V_{i,t-22}}} = \frac{V_{i,t+\tau}}{\sum_{\tau=252}^{1} V_{i,t-\tau-21}/252} \tag{2}$$

In Equation 2, $V_{i,t+\tau}$ is total daily volume in CRSP. Historical average volume, $\overline{V_{i,t-22}}$, is fixed at the beginning of the 22-day window before earnings are announced to avoid mechanically amplifying drops in volume.

I run the following regression with daily data to measure abnormal volume around earnings announcements:

$$\overline{CAV_{i,j,t}} = \alpha + \sum_{\tau=-21}^{0} \beta_\tau \mathbf{1}_{\{j==-\tau\}} + \text{Fixed Effects} + e_{i,j,t} \tag{3}$$

The main right hand side variables are a set of 22 indicators for days relative to the earnings announcement. For example, $\mathbf{1}_{\{j==15\}}$ is equal to one 15 trading days before the next earnings announcement, and zero otherwise. The regression also includes firm, year and day-of-the-week fixed effects[1]. The regression includes all firms that can be matched between CRSP and I/B/E/S.

I run this regression for 4 time periods: (1) 1994-1999 (2) 2000-2004 (3) 2005-2010 (4) 2011 to 2017. Figure 2 plots the estimates of $\beta_{\{j==-\tau\}}$.

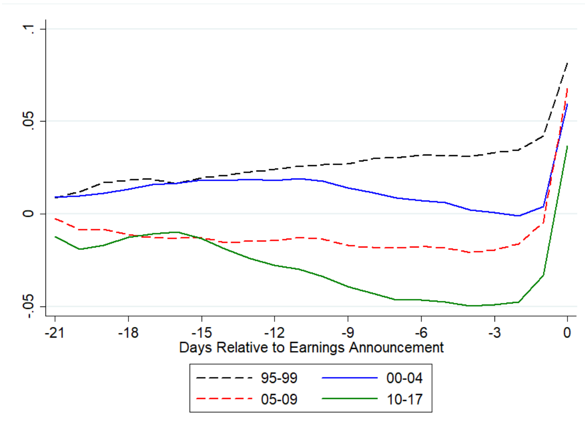
[1] Year fixed-effects are included to account for level differences in pre-earnings volume across years within each period. All results are robust to removing the year fixed effects.

Decline of Pre-Earnings Volume

Figure 2: Plot of $\beta_{\{j==-\tau\}}$ estimated from the regression:

$$\overline{CAV}_{i,j,t} = \alpha + \sum_{\tau=-21}^{0} \beta_\tau \mathbf{1}_{\{j==-\tau\}} + \text{Fixed Effects} + e_{i,j,t}$$

The units of the left-hand-side variable are average cumulative abnormal volume per trading day. For the 1995-99 period, $\beta_{\{j==1\}} = .042$, while for the 2010-2017 period, $\beta_{\{j==1\}} = -0.033$. The total difference of 0.75 is a decline of $0.75 \times 22 \approx 1.65$ days worth of pre-earnings abnormal volume.



The regression coefficients imply that between the late 90's and the present day, there was a decline of about 1.66 days worth of abnormal volume over the 22-day window before earnings announcements[2].

2.2 Pre-Earnings Drift

Fact 2: The pre-earnings drift has declined.

Let $E_{i,t}$ denote earnings per share for firm i in quarter t in the I/B/E/S Unadjusted Detail File. Define standardized unxpected earnings (SUE) as in Novy-Marx (2015): The year-over-year (YOY) change in earnings, divided by the

[2]Section D of the appendix motivates the choice of a 22-trading-day window before the announcement.

standard deviation of YOY changes in earnings over the past 8 quarters.

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

Define market-adjusted returns, $r_{i,t}$, as in Campbell, Lettau, Malkiel, and Xu (2001), as the difference between firm i's excess return and the return on the market factor from Ken French's data library.

Each quarter, I sort firms into deciles of SUE, and calculate the average cumulative market-adjusted returns over the 30 days prior to the earnings announcement. Figure 3 shows the average pre-earnings cumulative returns by SUE decile for two different time periods: 2001-2007 and 2010-2017. The black 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 2017, firms in each decile move less before earnings days, relative to the returns on earnings days themselves, than between 2001 and 2007. This 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 became public.

The apparent decline in the pre-earnings drift in Figure 3 could be driven by differences in overall return volatility or average returns between the two time periods. To quantify the decline of the pre-earnings drift, I create a drift magnitude variable designed to capture the share of earnings information already incorporated into prices before the announcement date. Let t denote an earnings announcement date. Define the pre-earnings drift for firm i as the cumulative market-adjusted from $t-30$ to $t-1$, divided by the cumulative returns from $t-30$ to t:

$$DM_{i,t} = \frac{r_{i,(t-30,t-1)}}{r_{i,(t-30,t)}} \tag{5}$$

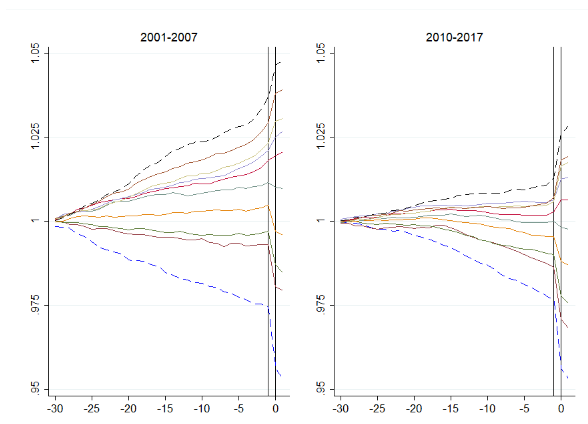
Pre-earnings drift will be near one when the earnings day move is small relative to the cumulative pre-earnings returns. $DM_{i,t}$ will be near zero when the earnings-day return is large, and in the same direction as the cumulative returns before the earnings day. $DM_{i,t}$ will be negative when the earnings day return is

Decline of Pre-Earnings Drift by SUE Decile

Figure 3: Each quarter, I sort firms into deciles on:

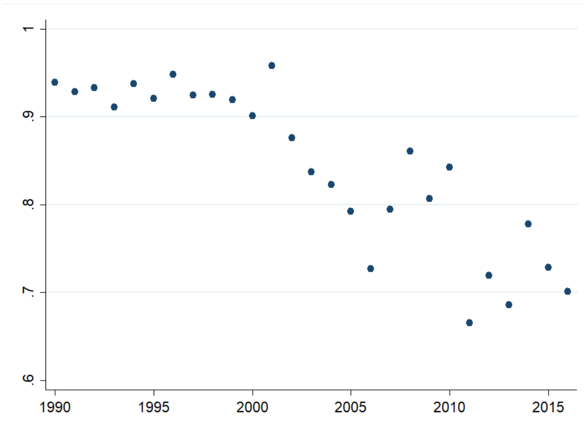
$$SUE_{i,t} = \frac{E_{i,t} - E_{i,t-4}}{\sigma_{(t-1,t-8)}(E_{i,t} - E_{i,t-4})}$$

Each line represents the cross-sectional average cumulative market-adjusted return by SUE decile. The black 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.



Decline of Average Pre-Earnings Drift

Figure 4: This figure plots the cross-sectional average of the pre-earnings drift magnitude measure, $DM_{i,t} = \frac{r_{i,(t-30,t-1)}}{r_{i,(t-30,t)}}$, by year. A value near 1 implies most earnings information is incorporated in prices before the announcement date, while lower values denote less informative pre-earnings prices.



a reversal relative to the pre-earnings return. One concern with the definition of pre-earnings drift is that $r_{i,(t-30,t)}$ may take values near zero, leading to huge values of $DM_{i,t}$. Because $r_{i,(t-30,t)}$ is a cumulative return over 30 days, fewer than 1% of all observations are smaller than 10 basis points. To further alleviate this concern, I Winsorize $DM_{i,t}$ at the 1% and 99% level by year[3]. Figure 4 shows the cross-sectional average value of $DM_{i,t}$ by year. The pre-earnings drift decreased by about 22% between 1990 and 2017.

2.3 Earnings Day Volatility

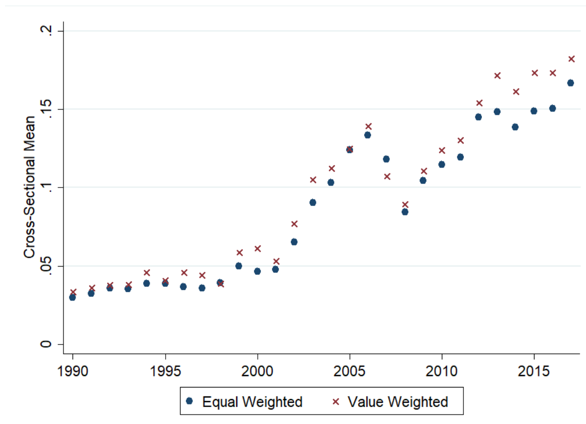
Fact 3: The share of total volatility occurring on earnings days has increased. Define the quadratic variation share (QVS) for firm i in year t as:

$$QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2 \tag{6}$$

[3]Section E of the Appendix presents alternative definitions of the pre-earnings drift, and further motivates my specification for $DM_{i,t}$.

Increase in Earnings Day Volatility

Figure 5: This figure plots the share of market-adjusted quadratic variation occurring on earnings days. For firm i in year t the quadratic variation share (QVS) is defined as: $QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, where r denotes a market adjusted daily return. The numerator sums over the 4 quarterly earnings days in year t, while the denominator includes all days in year t. The equal-weighted specification is the cross-sectional average for QVS, while the value-weighted specification uses weights proportional to lagged market capitalization.



where r denotes a market adjusted daily return. The numerator is 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. Earnings days make up roughly 1.6% of trading days, so values of $QVS_{i,t}$ larger than 0.016 imply that earnings days account for a disproportionately large share of total volatility. Figure 5 shows the cross-sectional average of $QVS_{i,t}$ by year for all CRSP firms that can be matched to 4 non-missing earnings days in a given year in I/B/E/S. Average QVS increased from 3.0% in 1990 to 16.7% in 2017.

2.4 Discussion

These downward trends in market efficiency could be unrelated to the information released on earnings days. To test this, I reconstruct the time-series averages of the pre-earnings volume, drift and share of volatility on earnings days, except I randomly assign one day each quarter for each firm to be an earnings date. In Section C of the Appendix, Figure 9 shows that there is no drop in volume before the placebo earnings dates. Figure 10 shows that there is no downward trend in the pre-earnings drift for the placebo earnings dates. Figure 11 shows there is no upward trend in the share of volatility on the placebo earnings dates. These results confirm that the changes to price informativeness are specific to earnings days.

As an additional check, Section J of the Appendix examines volume, drift and volatility around Federal Open Market Committee (FOMC) meeting dates instead of placebo earnings dates. I find no trend toward decreased efficiency in the incorporation FOMC meeting information. This further corroborates that the reduction in efficiency is only applies to firm-specific information.

3 Reduced-Form Evidence

In this section, I show the reduced-form relationship between increases in passive ownership and declines in pre-earnings volume, declines in pre-earnings drift and increases in the share of volatility on earnings days. These results suggest that passive ownership is negatively correlated with pre-earnings price informativeness.

3.1 Hypothesis Development

To motivate the empirical work, I examine the effect of increased passive management in a Grossman and Stiglitz (1980)-style model of asymmetric information[4]. The full details of the model are in Section K of the Appendix.

Consider a two period model with a single risky asset that pays a liquidating dividend of d in period 1, with total supply S. Assume d is normally distributed with mean $\overline{d}$ and variance σ_d^2. There are three types of agents: N_I of them are

[4]This section borrows heavily from Dimitris Papanikolaou's teaching notes on models with asymmetric information

informed, and get a private signal $s = d + \epsilon$ in period 0, where ϵ is normally distributed with mean zero and variance σ_ϵ^2. N_U are uninformed, and only learn about the liquidating dividend through the price. Both the informed and uninformed have exponential utility with coefficient of absolute risk aversion α.

The third type of agents are noise traders who buy and sell u shares with no regard for the price, where u is normally distributed with mean zero and variance σ_u^2. This implies that the effective supply of the asset is $S + u$. The presence of noise traders prevents the uninformed from perfectly learning the informed agents' signal from the price in a rational expectations equilibrium.

Prices are less informative if the conditional variance of the dividend given the price, $\sigma_{d|p}^2$, is high, so define price informativeness as:

$$informativeness = \frac{1}{\sigma_{d|p}^2} = \frac{\sigma_d^2 + \sigma_\epsilon^2 + \left(\frac{\alpha\sigma_\epsilon^2}{N_I}\right)^2 \sigma_u^2}{\sigma_d^2 \left(\sigma_\epsilon^2 + \left(\frac{\alpha\sigma_\epsilon^2}{N_I}\right)^2 \sigma_u^2\right)} \tag{7}$$

Equation 7 implies prices are more informative if there are more informed agents, and prices are less informative if σ_u^2 is high. To translate these comparative statics to the rise in passive ownership, we need to understand the role passive managers play in this equilibrium.

Mauboussin, Callahan, and Majd (2017) show that the large inflows into passive funds have come at the same time as large outflows from active funds. If we view active managers as informed, and passive managers as uninformed, then the model predicts that increased passive management would lead to decreased price informativeness.

Another view is that passive managers' behavior is more consistent with way the noise traders are modeled: passive demand is usually based on mechanical rules, rather than fundamentals. If increases in passive management imply an increase in σ_u^2, average informativeness would also go down, but only holding all else equal. If the marginal active manger who has exited with the rise of passive management was uninformed, then the effect on price informativeness is ambiguous.

Even in this simple setup, the effect of increasing passive ownership is not clear, so empirical work is needed to understand passive managers' roles and determine which effect dominates in practice.

3.2 Data and Definitions

Passive ownership is defined as all index funds, as identified by the index fund flag in the CRSP mutual fund data, all ETFs that are not actively managed, and all mutual funds with "index" in the name. All quarterly fund holdings are from the Thompson S12 data. I use the WRDS MF LINKS database to connect the funds identified as passive in CRSP with the holdings in the Thompson S12 data. Given that holdings are only updated quarterly, I linearly interpolate shares held between S12 filing dates[5]. If a security never appears in the S12 data, I assume the passive ownership share is zero.

I believe this is a conservative definition of passive ownership, as there are institutional investors which track indices, but are not classified as mutual funds, and do not show up in the S12 data. Further, as discussed in Mauboussin, Callahan, and Majd (2017), there has been a rise of closet indexing among self-proclaimed active managers, which will also not show up in my definition of passive management.

All return and daily volume data are from CRSP. I merge CRSP to I/B/E/S (IBES) on CUSIP, or historical CUSIP (ncusip) if available. I use the earnings release times in IBES to identify the first time market participants could trade on earnings information during normal market hours. If earnings are released before 4:00 PM EST between Monday and Friday, that day will be labeled as the effective earnings date. If earnings are released on or after 4:00 PM EST between Monday and Friday, the next trading day will be labeled as the effective earnings date. If earnings are released over the weekend, or on a trading holiday, the next trading date will be labeled as the effective earnings date.

[5] All results are robust to fixing shares held to their end of quarter value, rather than interpolating between quarters.

I define quarterly earnings per share as the "value" variable from the IBES unadjusted detail file[6]. All firm fundamental information is from Compustat.

Total institutional ownership is the sum of shares held by all 13-F filing institutions. Institutional ownership is merged to CRSP on CUSIP, or historical CUSIP if available. If a CUSIP never appears in the 13-F data, institutional ownership is assumed to be zero.

3.3 Pre-Earnings Volume

Define pre-earnings cumulative abnormal volume *per day* in the 22-day window before earnings is announced as:

$$CAV_{i,t} = \frac{\sum_{\tau=-22}^{-1} AV_{i,t+\tau}}{22} \tag{8}$$

where abnormal volume, AV, is defined in Equation 2. I run the following regression with quarterly data to measure the relationship between declines in pre-earnings volume and increases in passive ownership:

$$\Delta_{(t,t-n)}CAV_{i,t} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + \gamma X_{i,t-n} + FE + e_{i,t} \tag{9}$$

$\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. I only look at year-over-year changes to avoid differences in volume before annual earnings announcements and quarterly announcements or seasonal effects. Controls in $X_{i,t-n}$ include lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership, calculated as the sum of holdings in the 13-F filings. I also condition on the growth in market capitalization from $t-n$ to t[7]. I condition on market capitalization and growth of market capitalization because most of the increase in passive ownership has been in large stocks, and I want to minimize the influence of this trend on my results.

[6]All results are similar when using Diluted Earnings Per Share Excluding Extraordinary Items (EPSFXQ) in Compustat.

[7]The results are unchanged if all first differences are replaced with log growth rates.

Fixed effects (FE) include 2-digit SIC industry, year and security (permno). All standard errors are clustered at the firm/year level[8].

Given that passive ownership is slow moving, I examine changes in passive ownership and pre-earnings volume over 1, 3 and 5-year horizons. Table 1 contains the regression results. To interpret the magnitude of the coefficient on $\Delta_{(t,t-n)} Passive_{i,t}$: (1) The 25th percentile of passive ownership share in my sample is 0, while the 75th percentile is around 0.1 (10%) (2) The units of the left hand side variable are abnormal volume per day over the 22 days leading up to the earnings announcement (3) The decline in pre-earnings volume between the late 90's and present day was 1.65 trading days.

The coefficients on $\Delta_{(t,t-n)} Passive_{i,t}$ in the 3-year specification with firm fixed effects implies that moving from the 25th to the 75th percentile of passive ownership would explain (-0.57 $\times$ 22 $\times$0.1 = -1.254, from a total decline of -1.65) about 76% of the decline in pre-earnings abnormal volume.

3.4 Pre-Earnings Drift

Define the pre-earnings drift magnitude as in Equation 5: $DM_{i,t} = \frac{r_{i,(t-30,t-1)}}{r_{i,(t-30,t)}}$. I run the following regression with quarterly data to measure the relationship between the pre-earnings drift and passive ownership.

$$DM_{i,t} = \alpha + \beta Passive_{i,t} + \sum_{j=2}^{10} \phi_j \mathbf{1}_{SUE declie=j} + \gamma X_{i,t-1} + FE + e_{i,t} \qquad (10)$$

Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. The regression also includes a set of indicator variables for SUE decile within a given quarter. These are included because as shown in Figure 3, the size of the drift depends on the eventual earnings surprise. Fixed effects (FE) include 2-digit SIC industry, year and firm. Standard errors are clustered at the firm/year level.

[8] All results are similar when computing standard errors with panel Newey-West, setting the lags equal to 1.5x the number of overlapping observations, rounding up to the nearest integer.

Passive Ownership and Pre-Earnings Volume

Table 1: Estimates of β from:

$$\Delta_{(t,t-n)} CAV_{i,t} = \alpha + \beta \Delta_{(t,t-n)} Passive_{i,t} + X_{i,t-n} + FE + e_{i,t}$$

$CAV_{i,t}$ is average pre-earnings cumulative abnormal volume *per day* over the 22 days leading up to the earnings announcement. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. Change in passive ownership is expressed as a decimal, so $0.01 =$ 1% increase. Controls, $X_{i,t-n}$, include lagged passive ownership, market capitalization, idiosyncratic volatility and total institutional ownership. I also condition on the growth in market capitalization from $t-n$ to t. Fixed effects include 2-digit SIC industry, year and firm. Specification (1) includes all firm-level controls, plus industry and year fixed effects. Specification (2) adds includes firm fixed effects. All standard errors are clustered at the firm/year level.

		1-year		3-year		5-year	
		(1)	(2)	(1)	(2)	(1)	(2)
Increase in Passive Ownership	1-year	-1.263*	-1.090***				
		(0.651)	(0.215)				
	3-year			-0.936***	-0.573***		
				(0.309)	(0.160)		
	5-year					-1.041**	-0.545***
						(0.421)	(0.174)
Observations		272,609	272,609	191,654	191,654	139,537	139,537
R-Squared		0.009	0.010	0.007	0.008	0.025	0.025
Firm Controls		Yes	Yes	Yes	Yes	Yes	Yes
Industry/Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Firm FE		No	Yes	No	Yes	No	Yes

The regression results are in Table 2. Figure 4 shows that DM declined by about 0.2 between 1990 and 2017. The estimated coefficient on $Passive$ implies that moving from the 25th (0) to the 75th percentile (0.1) of passive ownership would explain about ($0.4 \times 0.1 = 0.04$, from a total decline of 0.2) 20% of the decline in pre-earnings drift[9].

$DM_{i,t}$ is Winsorized at the 1% and 99% level by year to minimize the effect of values of $r_{i,(t-30,t)}$ near zero. The coefficents on $Passive_{i,t}$ is more statistically significant Winzorizing at the 5% and 95% level by year, as this further reduces the standard errors. Year-over-year changes in drift are volatile, so running the regression in first differences yields similar point estimates, but larger standard errors.

3.5 Share of Volatility on Earnings Days

Define the earnings day share of annual volatility as in Equation 6: $QVS_{i,t} = \sum\limits_{\tau=1}^{4} r_{i,\tau}^2 / \sum\limits_{j=1}^{252} r_{i,j}^2$. I run the following regression with annual data to measure the relationship between changes in earnings day share of annual volatility, and changes in passive ownership:

$$\Delta_{(t,t-n)} QVS_{i,t} = \alpha + \beta \Delta_{(t,t-n)} Passive_{i,t} + \gamma X_{i,t-n} + FE + e_{i,t} \tag{11}$$

Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. I also condition on change in market capitalization from $t-n$ to t. Fixed effects (FE) include 2-digit SIC industry, year and firm. Standard errors are clustered at the firm/year

[9]One potential problem with this definition of pre-earnings drift is that the average post-earnings drift has declined (see e.g. McLean and Pontiff (2016)). If the returns that historically were realized after earnings announcements moved to the earnings days themselves, there could be a mechanical decrease in DM, as the average magnitude of $r_{i,t}$ increased. In unreported results, I find that regression estimates are similar, when defining drift as: $\widetilde{DM}_{i,t} = \frac{r_{i,(t-30,t-1)}}{r_{i,(t-30,t+k)}}$ where $k \in \{1,2,3\}$, which confirms this pattern is specific to the pre-earnings drift. Section E of the Appendix shows that the post-earnings drift is larger for firms with higher passive ownership, suggesting that high passive ownership slows the absorption of earnings information after it is released.

Passive Ownership and Pre-Earnings Drift

Table 2: Table with estimates of β from:

$$DM_{i,t} = \alpha + \beta Passive_{i,t} + \sum_{j=2}^{10} \phi_j \mathbf{1}_{SUEdeclie=j} + \gamma X_{i,t-1} + FE + e_{i,t}$$

$DM_{i,t} = \frac{r_{i,(t-30,t-1)}}{r_{i,(t-30,t)}}$, which is the ratio of the cumulative returns in the 30 days leading up to the earnings day, relative cumulative return in the 30 days up to and including the earnings day. Passive ownership is expressed as a decimal, so 0.01 = 1% of shares outstanding held by passive funds. SUE deciles are formed each quarter. $DM_{i,t}$ is Winzorized at the 1% and 99% levels. Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. Fixed effects include 2-digit SIC industry, year and firm. Specification (1) includes all firm-level controls, plus industry and year fixed effects. Specification (2) adds includes firm fixed effects. All standard errors are clustered at the firm/year level.

	Pre-Earnings Drift	
	(1)	(2)
Passive Ownership	-0.329**	-0.436**
	(0.157)	(0.177)
Observations	413,328	413,328
R-Squared	0.003	0.002
Firm Controls	Yes	Yes
SUE	Yes	Yes
Industry/Year FE	Yes	Yes
Firm FE	No	Yes

Passive Ownership and Earnings Day Share of Volatility

Table 3: Table with estimates of β from:

$$\Delta_{(t,t-n)}QVS_{i,t} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + \gamma X_{i,t-n} + FE + e_{i,t}$$

$QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, which is the ratio of the squared returns on the 4 quarterly earnings announcement days, relative to the squared returns on all days in year t. QVS takes values in [0,1]. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. Fixed effects include 2-digit SIC industry, year and firm. Specification (1) includes all firm-level controls, plus industry and year fixed effects. Specification (2) adds includes firm fixed effects. All standard errors are clustered at the firm/year level.

		1-year		3-year		5-year	
		(1)	(2)	(1)	(2)	(1)	(2)
Increase in Passive Ownership	1-year	0.0983**	0.0551*				
		(0.039)	(0.030)				
	3-year			0.103***	0.0502*		
				(0.034)	(0.026)		
	5-year					0.171***	0.0883***
						(0.042)	(0.027)
Observations		79,253	79,253	60,578	60,578	48,484	48,484
R-Squared		0.015	0.015	0.052	0.047	0.077	0.065
Firm Controls		Yes	Yes	Yes	Yes	Yes	Yes
Industry/Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Firm FE		No	Yes	No	Yes	No	Yes

level.

The regression results are in Table 3. Figure 5 shows that QVS increased by about 0.12 between 1990 and 2017. The estimated coefficient on $Passive$ implies that moving from the 25th (0) to the 75th percentile (0.1) of passive ownership would explain about ($0.1 \times 0.1 = 0.01$, from a total increase of 0.137) 7-13% of the average increase in share of annual QV on earnings days.

3.6 Placebo Tests

To confirm that my results are specific to earnings days, I re-run the three reduced-form regressions, except I select dates between the actual earnings days to represent placebo earnings dates. For example, if a firm released earnings on 12/31/2017, I would select the trading day closest to 11/15/2017 as the placebo earnings date. Appendix Tables 13, 14 and 15 compare the original regression results to the placebo results, in the specifications without firm-fixed effects, where the baseline results were strongest. All of the placebo results are insignificant, confirming that the changes in volume, drift and volatility are all specific to earnings days. In unreported results, I randomly assign one day for each firm in each quarter to be a placebo earnings day. This alternative placebo test also yields insignificant coefficients on $\Delta Passive$ and $Passive$ in all reduced-form regressions.

3.7 Addressing Competing Hypotheses

This section discusses two alternative explanations for my findings on decreased market efficiency, and its correlation with passive ownership: (1) Regulation Fair Disclosure (Reg FD), which reduced 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 (2017), Farboodi and Veldkamp (2017)). It is not possible to discuss every alternative hypothesis, so outside of explicitly testing these two alternatives, I will exploit the quasi-exogenous variation in passive ownership from index addition/rebalancing in the next section to overcome remaining identification concerns.

In addition to identification concerns, the reduced-form regressions could suffer from omitted variable bias. For example, most passive ownership is determined by mechanical rules derived from observable signals like market capitalization and past returns. This implies that it may be possible to select a large set of stock/firm characteristics to explain all of the variation in passive ownership. My results would be biased if these underlying characteristics were driving the changes in pre-earnings price informativeness. I find this unlikely as much of the differences in passive ownership across firms is determined by index membership, which is sticky for some indices, and hard to predict for others. Firms that have been in the S&P 500 index for many years would not necessarily be added to the index today, even if they meet all the criteria for index addition. For other indices like the Russell 1000, there is a sharp size cutoff in the index addition rule[10], which makes it difficult to predict index membership around the cutoff. The difficulty of predicting index membership, and as a result predicting passive ownership, reduces the likelihood that my results are driven by an omitted variables problem.

3.7.1 Reg FD

Before Reg FD was passed in August, 2000, firms would disclose earnings information to selected analysts before it became public. This information leakage could increase the share of earnings information incorporated into prices before it was formally announced. After Reg FD, firms were no longer allowed selectively disclose material information, and instead must release it to all investors at the same time.

Reg FD could be driving the trends in decreased price informativeness, as there was a large negative shock to information released by firms after it was passed. In Figures 2, 4 and 5, however, all of the information measures continue to trend in the same direction after Reg FD was implemented. Reg FD could still explain these results if the value of the inside information received by analysts before Reg FD decayed slowly. While this is possible, my prior is that information obtained in 2000 would not be relevant for more than a few years. If Reg FD totally explained the decreased pre-earnings informativeness, we would expect the trends in decreased informativeness to level out in the early 2000's. In the data, however, this leveling out does not happen for any of the three measures.

[10] There was a sharp size cutoff before the rule change in 2006, see e.g. Wei and Young (2017).

For Reg FD to be driving the reduced-form relationship between passive ownership and pre-earnings price informativeness, it would have to disproportionately affect firms with high passive ownership. This is because all the regressions have year fixed effects, which should wash out any level shifts in price informativeness before/after Reg FD was passed. To rule out this channel, the appendix contains versions of all the reduced-form regressions using only post-2000 data in Tables 19, 20 and 21. All of the results are similar using only post-2000 data, suggesting that differences between the pre and post Reg FD eras are not driving my results.

3.7.2 Rise of Algorithmic Trading Activity

Weller (2017) shows that Algorithmic Trading (AT) activity is negatively correlated with pre-earnings price informativeness. His proposed mechanism is ATs front-run informed traders, reducing the returns to gathering firm-specific fundamental information. AT activity increased significantly over my sample period, and could be responsible for some of the observed decrease in pre-earnings price informativeness.

In the context of the asymetric information model, ATs could be viewed as uninformed agents, as they learn from the price and order flow. The model also predicts that as the share of ATs increase, the price informativeness decreases.

It is difficult to measure the role of ATs in the trends toward decreased pre-earnings price informativeness as I cannot directly observe AT activity, and only have reasonable AT activity measures between 2012-2017. I can, however, measure the effect of AT activity on the reduced-form results. For AT activity to influence the regression estimates, it would have to be correlated with passive ownership, which I find plausible because: (1) Passive ownership is higher in large liquid stocks, where most AT activity occurs (2) High ETF ownership will attract ATs implementing ETF arbitrage. The effect of time trends in AT activity should be absorbed by the year fixed effects.

To rule out this channel, I construct the 4 measures of AT activity used in

Weller (2017) from the SEC MIDAS data. MIDAS has daily data for all stocks traded on 13 national exchanges from 2012 to 2017. The AT measures are (1) odd lot ratio, (2) trade-to-order ratio, (3) cancel-to-trade ratio and (4) average trade size. Measures 1 and 3 are positively correlated with AT activity, while the opposite is true for measures 2 and 4. Consistent with Weller (2017), I Winzorize each of these variables at the 1% and 99% level by year to minimize the effect of reporting errors, calculate a moving average for each of these measures in the 21 days leading up to each earnings announcement, and take logs to reduce heavy right-skewness. Only 1% of MIDAS data cannot be matched to CRSP, so the 87% drop in sample size relative to previous regressions is coming almost entirely from the year restrictions.

I run the following modified versions of the reduced-form regressions:

$$\begin{aligned}\Delta_{(t,t-n)}CAV_{i,t} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + \gamma X_{i,t-n} + \\ \phi\Delta_{(t,t-n)}ATActivity_{i,t} + FE + e_{i,t}\end{aligned} \tag{12}$$

$$\begin{aligned}DM_{i,t} = \alpha + \beta Passive_{i,t} + \sum_{j=2}^{10}\phi_j \mathbf{1}_{SUEdeclie=j} + \gamma X_{i,t-1} + \\ \psi ATActivity_{i,t} + FE + e_{i,t}\end{aligned} \tag{13}$$

$$\begin{aligned}\Delta_{(t,t-n)}QVS_{i,t} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + \gamma X_{i,t-n} + \\ \phi\Delta_{(t,t-n)}ATActivity_{i,t} + FE + e_{i,t}\end{aligned} \tag{14}$$

In regression 12 $\Delta_{(t,t-n)}ATActivity_{i,t}$ is a vector of year-over-year changes in the 4 AT activity measures. In regression 13 $ATActivity_{i,t}$ is a vector containing the levels of the 4 AT activity measures. Regression 14 is run with annual data, so I first calculate average $ATActivity_{i,t}$ across the 4 earnings announcements each year, and then calculate year-over-year changes. As a baseline, I re-run the previous regressions on the sub-sample matched to the MIDAS data – these regressions are labeled "Baseline" in the corresponding tables. The regressions with all the AT measures included are labeled "+ AT Controls". Given the limited time-series available for the AT measures, I only run the one-year and three-year difference specifications.

Tables 16, 17 and 18 contain the results. Only the results for the pre-earnings drift and the earnings day share of volatility are significant in the matched subsample. For the specifications that are significant in the subset of my original sample that I can match to MIDAS data, adding the AT activity controls does reduce the coefficient on passive ownership/change in passive ownership, but the sign and statistical significance is unchanged. This implies that increased AT activity may partially explain the observed decrease in market efficiency, but increasing passive ownership is still an important factor in decreased pre-earnings price informativeness.

4 Evidence from Index Addition/Rebalancing

In this section, I exploit plausibly exogenous variation in passive ownership to create a causal link between passive ownership and pre-earnings price informativeness.

4.1 S&P 500 Index Addition

Each 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, a specific industry classification 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 funds and ETFs buy the stock.

I obtain daily S&P 500 index constituents from Compustat. Motivated by the size, industry and financial health criteria, I select a group of control firms that reasonably could have been added to the index at the same time as the treated firms. One year before index addition, I sort firms into two-digit SIC industries. Then, within each industry, I sort firms into quintiles based on market capitalization (size) and growth rate of market capitalization over the past year. The control group is all firms in the same industry/size/growth-rate bucket that were not added to the index over the next two years[11]. For example, Sun Microsystems was added to the Index in August, 1992. while one of the corresponding control firms, Seagate Technologies, was not added to the Index until August, 1996.

For the reduced-form regressions in differences (Pre-Earnings Volume and Earnings Day Share of Volatility) I run the following regression:

$$\Delta_{(t-1,t+1)} Outcome_{i,t} = \alpha + \gamma Added_{i,t} + FE + e_{i,t} \tag{15}$$

where $\Delta_{(t-1,t+1)}$ is the change from the year before index addition to the year after index addition[12].

For the Pre-Earnings Volume regression, there are 4 observations per firm – each is the year over year change from the same fiscal quarter in the year before addition to the year after addition. For the Earnings Day Volatility regression, there is one observation per firm. $Added_{i,t}$ is an indicator variable equal to one if the firm was added to the S&P 500 index. The coefficient of interest is γ, the treatment effect of being added to the index.

For the Pre-Earnings drift regression, the reduced-form was in levels, so I run the following Difference in Differences specification:

$$Outcome_{i,t} = \alpha + \beta Added_{i,t} + \tau After_{i,t} + \gamma Added_{i,t} \times After_{i,t} + FE + e_{i,t} \tag{16}$$

In this regression, there are 8 observations for each firm: 4 in the year before

[11] It is also possible to sort on three-digit SIC industries, but this leaves many firms without a control firm as not every SIC-3 industry has at least 2 firms in each of the 25 size/growth rate buckets.

[12] The year of index addition is omitted for two reasons: (1) The increase in passive ownership associated with index addition does not all occur immediately, but rather over the three quarters after index addition (2) There are other effects associated with S&P 500 index addition, including an increase in relative valuation (see e.g. Morck and Yang (2001)) and an increase in media coverage (Engelberg and Gao (2011)). Skipping a year gives time for these index addition effects to die out. In unreported results, I find that not skipping a year yields similar point estimates, but larger standard errors.

index addition, and 4 in the year after index addition. $Added_{i,t}$ is an indicator variable equal to one if the firm was added to the S&P 500 index. $After_{i,t}$ is an indicator equal to one for the observations after the firm was added. The coefficient of interest is the interaction term, γ.

For all three regressions, the fixed effects include 2-digit SIC industry, size quintile, growth rate quintile and year/month of index addition. With these controls, the regressions are only exploiting variation between the treatment and control firms at the same point in time.

One concern is that because selection is by committee, the increase in passive ownership is not totally 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 totally mechanical rule.

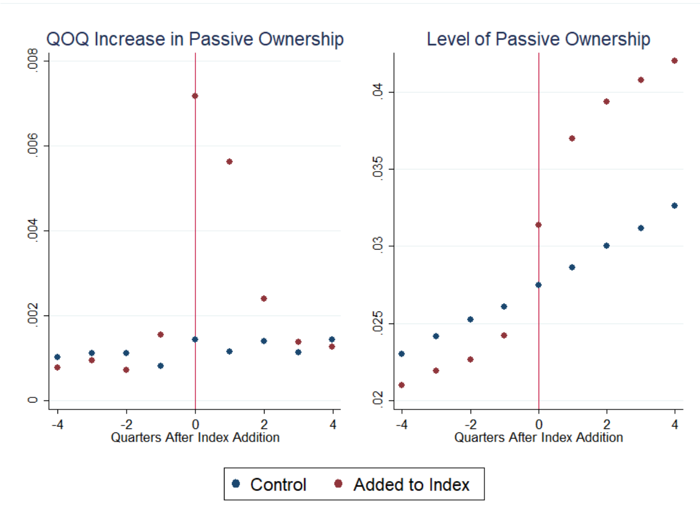
To test the similarity of the treated and control groups, Figure 6 shows the levels and changes in passive ownership for the control firms and treatment firms around the time of index addition. Both groups of firms have similar levels and pre-addition changes in passive ownership.

Table 4 contains the regression results. For comparison, I included a row with the reduced form estimates, which correspond to the 1-year changes specification with firm fixed effects estimated in Section 3. The average year-over-year increase in passive ownership for a firm added to the S&P 500 index is 2.2%, so the implied elasticity is the coefficient of interest, γ, divided by 0.022. For all three specifications, the results have the same sign and statistical significance as the reduced-form regressions. The implied elasticities are substantially larger than the reduced-form results, but my prior is that this is because 2.2% understates the true increase in passive ownership associated with index addition: There are many institutional investors which do not show up in the Thompson S12 data which track the S&P 500 index and buy these stocks after they are added.

A natural extension is to examine firms which were dropped from the S&P

S&P 500 Index Addition and Changes in Passive Ownership

Figure 6: Average level and increase in passive ownership for control firms and firms added to the S&P 500. Control firms are all firms in the same 2-digit SIC industry, in the same size and growth rate quintiles that were not added to the S&P 500 index over the next two years.



Effects of S&P 500 Index Addition

Table 4: For Pre-Earnings Volume and Earnings Day Share of QV:

$$\Delta_{(t-1,t+1)}Outcome_{i,t} = \alpha + \gamma Added_{i,t} + FE + e_{i,t}$$

For Pre-Earnings Drift:

$$Outcome_{i,t} = \alpha + \beta Added_{i,t} + \tau After_{i,t} + \gamma Added_{i,t} \times After_{i,t} + FE + e_{i,t}$$

Added firms are those which were added to the S&P 500 index. Control firms are in the same industry, same size and same growth rate quintile as the control firms. Fixed effects include industry, size quintile, growth rate quintile, and year/month of index addition.

	Pre-Earnings Volume	Pre-Earnings Drift	Earnings Day Share of QV
	Differences	Levels	Differences
Added	-0.0537**		0.0116**
	(0.026)		(0.006)
Interaction		-0.0904*	
		(0.051)	
R-squared	0.083	0.073	0.129
Reduced Form	-1.090***	-0.436**	0.0551*
First Stage	2.20%	2.20%	2.20%
Implied Elasticity	-2.44	-4.11	0.53
Year/Month FE	YES	YES	YES
Industry FE	YES	YES	YES
Size Quintile FE	YES	YES	YES
Growth Quintile FE	YES	YES	YES
Treated Firms	614	614	614
Control Firms	1,479	1,479	1,479

500 index, which experience a decrease in passive ownership. This is a less ideal experiment than index addition, as firms are usually dropped from the index for poor performance or lack of liquidity, which is related to firm fundamentals. Section H of the Appendix has more details on the effect of index deletion.

4.2 Russell 1000/2000 Index Reconstitution

The Russell 3000 contains approximately the 3000 largest stocks in the United States stock market. Each May, FTSE Russell selects the 1000 largest stocks by float to be members of the Russell 1000, while it selects the next 2000 largest stocks by float to be members of the Russell 2000[13]. Both of these indices are value-weighted, so moving from the 1000 to the 2000 significantly increases the share of passive ownership in a stock. The firm goes from being the smallest firm in an index of large firms, to the biggest firm in an index of small firms, increasing its relative weight by a factor of 10 (see e.g. Appel, Gormley, and Keim (2016)).

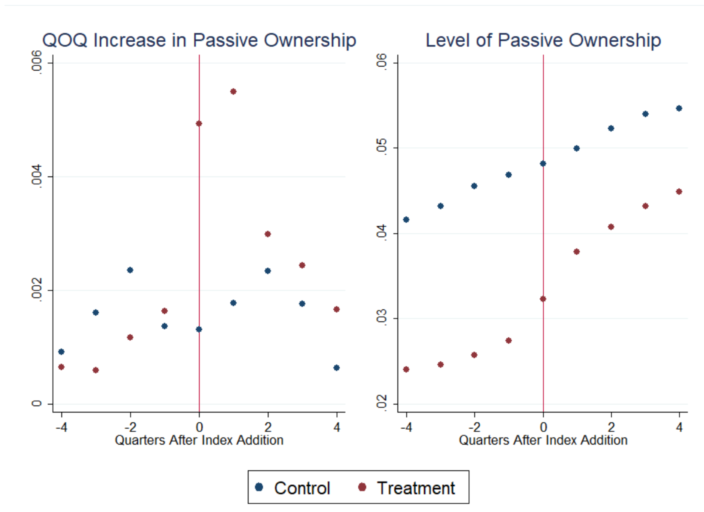
The increase in passive ownership corresponding to S&P 500 index addition is not a perfect natural experiment because being added is not random, added firms receive increased attention, and added firms may start marketing their stock differently to institutional investors. The increase in passive ownership associated with the Russell reconstitution sidesteps many of these issues, as moving from the 1000 to the 2000 is based on a mechanical rule, rather than committee selection. Further, because the firm's market capitalization shrunk, it is less likely to change the way the firm is marketing itself to institutions.

I obtain Russell 1000/2000 membership between 1996 and 2012 from the Wei and Young (2017) replication files. The treated firms are those that switched from the Russell 1000 to the Russell 2000. The control firms are the other firms

[13] This rule changed in 2006 – to reduce turnover between the two indices, Russell now has a bandwith rule: As long as the firm's market capitalization is withing 5% of the 1000th ranked stock, it will remain in the same index it was in the previous year. Given that this is still a mechanical rule, however, the increases in passive ownership are still plausibly exogenous to firm fundamentals.

Russell 1000/2000 Reconstitution and Changes in Passive Ownership

Figure 7: Average level and increase in passive ownership for control firms and firms moved from the Russell 1000 to the Russell 2000. Control firms are all firms in the Russell 3000 ranked 900 to 1100 that did not move from the 1000 to the 2000 or from the 2000 to the 1000.



with June ranks between 900 and 1100 that did not switch between the 1000 and the 2000 or between the 2000 and the 1000. Results are similar when restricting the control firms to those that stayed in the Russell 1000. This classification involves a look-ahead bias, as I am using the ex-post changes in membership to identify changes in passive ownership.

Figure 7 compares the level, and the increase in passive ownership around the index rebalancing date. While the pre-addition changes are similar, the levels are different – this is driven by the firms ranked 1001-1100 having a higher average level of passive ownership because they are the largest firms in a value-weighted index of small firms.

I re-run regressions 15 (pre-earnings volume and earnings day volatility) and 16 (pre-earnings drift) with the Russell 1000/2000 index reconstitutions, and a slightly different set of controls and time periods. I remove the industry, size quintile and growth rate quintile fixed effects, as I am not using these to select control firms. I am comparing the year before index reconstitution, ending in April, and the year following reconstitution, starting in August. This is because the rankings are determined in May, so investors may trade in advance of the actual rebalancing in June. Further, the rankings are usually released at the end of June, but sometimes they are released in early July. July is excluded to avoid any of the trading associated with index rebalancing influencing the regression results[14].

Table 5 contains the regression results. For comparison, I included a row with the reduced form estimates, which correspond to the 1-year changes specification with firm fixed effects estimated in Section 3. The average increase in passive ownership from May to August for a firm moving from the Russell 1000 to the 2000 is 1.7%, so the implied elasticity is the coefficient of interest, γ, divided by 0.017.

For pre-earnings volume and drift, the results have the same sign and statistical significance as the reduced-form regressions. The results for earnings day volatility have the opposite sign and are insignificant. Part of this could be due to the volatility of QVS, and the dependence of QVS on earnings surprises. Given the relatively short sample period (1996-2012), and the smaller number of treated and control firms, this estimate is likely noisy.

As with the S&P 500 results, the implied elasticities are substantially larger than the reduced-form estimates, but I believe 1.7% understates the true increase in passive ownership associated with index addition: There are many institutional investors which track the Russell indices which do not show up in the Thompson S12 data.

A natural extension is to look at the firms which experience a decrease in

[14]Unlike the S&P 500, where firms remain in the index for long stretches of time, the Russell indices are rebalanced annually, so one year after moving from the 1000 to the 2000, the firm may switch back. To avoid picking up the effects of firms switching back and forth, I do not skip a year after index additions.

Effects of Russell 1000/2000 Index Reconstitution

Table 5: For Pre-Earnings Volume and Earnings Day Share of QV:

$$\Delta_{(t-1,t+1)}Outcome_{i,t} = \alpha + \gamma Moved_{i,t} + FE + e_{i,t}$$

For Pre-Earnings Drift:

$$Outcome_{i,t} = \alpha + \beta Moved_{i,t} + \tau After_{i,t} + \gamma Moved_{i,t} \times After_{i,t} + FE + e_{i,t}$$

Added firms are those which were moved from the Russell 1000 to the Russell 2000. Control firms are those that were ranked 900-1100 by Russell, but did not switch from the 1000 to the 2000 or from the 2000 to the 1000. Fixed effects include year/month of index rebalancing.

	Pre-Earnings Volume	Pre-Earnings Drift	ED Share of QV
	Differences	Levels	Differences
Treated	-0.0559***		-0.00237
	(0.018)		(0.006)
Interaction		-0.100**	
		(0.043)	
R-squared	0.005	0.006	0.030
Reduced Form	-1.090***	-0.436**	0.0551*
First Stage	1.70%	1.70%	1.70%
Implied Elasticity	-3.288	-5.882	-0.139
Year/Month FE	YES	YES	YES
Treated Firms	479	479	479
Control Firms	1,284	1,284	1,284

passive ownership when they move from the Russell 2000 to the Russell 1000. In Section H I show that this treatment effect is washed out by the time trend toward increased passive ownership[15].

5 Additional Evidence

In this section, I present additional evidence on the effects of passive ownership, including changes in analyst coverage, response to cashflow news and investor attention.

5.1 Analyst Attention

One potential explanation for passive ownership decreasing price informativeness is that high passive ownership reduces the incentives to gather 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 rules are implemented on public signals, they reduce the need to gather accurate private forecasts of firm fundamentals. As a stock becomes more mispriced, however, the return to gathering fundamental information increases, so it is not obvious which effect will dominate in equilibrium. To test this hypothesis, I regress levels, and changes in analyst coverage/accuracy on passive ownership:

$$Outcome_{i,t} = \alpha + \beta Passive_{i,t} + \gamma X_{i,t} + FE + e_{i,t} \tag{17}$$

Controls in $X_{i,t}$ include market capitalization and institutional ownership. Fixed effects include industry, year and firm.

$$\Delta_{(t,t-5)} Outcome_{i,t} = \alpha + \beta \Delta_{(t,t-5)} Passive_{i,t} + \gamma X_{i,t} + FE + e_{i,t} \tag{18}$$

[15] Another plausibly exogenous change in passive ownership arises when firms move from outside the Russell 3000 to inside the Russell 3000, which results in an increase in passive ownership. While this is potentially interesting, there are sample selection issues, as these micro caps often fail to appear in IBES or Compustat.

$\Delta_{(t,t-5)}$ is the change from year t to year $t-5$. Controls in $X_{i,t}$ include institutional ownership, lagged inst. ownership, market cap, lagged market cap. Fixed effects include industry, year and firm.

In regressions 17 and 18, the outcomes of interest are (1) the number of analysts covering a stock, (2) the absolute distance between the consensus forecast and the realized earnings, divided by the absolute value of the consensus forecast, which I will call *accuracy* and (3) the probability that the consensus forecast has the same sign as the realized earnings. For the accuracy regressions, I exclude firms with a consensus forecast of 1 cent or less to minimize the effect of outliers. Accuracy is then Winzorized at the 1% and 99% level each year.

The sample is all annual earnings announcements. To determine the consensus forecast, I take the equal-weighted average of all analyst forecasts on the last statistical period in IBES before earnings are released.

Table 6 contains the regression results. Consistent with decreased information gathering, high levels and increases in passive ownership are negatively correlated with the number of analysts covering a stock, lower accuracy, and a lower likelihood of getting the sign of earnings correct.

A more mechanical explanation for the negative relationship between passive ownership and pre-earnings price informativeness is that high passive ownership decreases the number of shares available for trading (float), which leads to higher price impact and transaction costs. Section B of the Appendix shows that increases in passive ownership are correlated with increased average liquidity, but decreased liquidity around earnings announcements. This result, however is not robust using only quasi-exogenous increases in passive ownership arising from S&P 500 index addition and Russell 1000/2000 reconstitution.

5.2 Response to Earnings News

Buffa, Vayanos, and Woolley (2014) propose a model where stocks with a higher share of "buy and hold" investors are more responsive to cash flow news. In the model, buy and hold investors distort prices, so informed investors underweight

Analyst Coverage/Accuracy and Passive Ownership

Table 6: This table contains the results for two sets of regressions:

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

Controls in $X_{i,t}$ include market capitalization and institutional ownership. Fixed effects include industry, year and firm.

$$\Delta_{(t,t-5)} Outcome_{i,t} = \alpha + \beta \Delta_{(t,t-5)} Passive_{i,t} + \gamma X_{i,t} + FE + e_{i,t}$$

Controls in $X_{i,t}$ include institutional ownership, lagged institutional ownership, market capitalization, lagged market capitalization. Fixed effects include industry, year and firm.

	# Analyst Estimates		$\frac{\lvert realized-consensus \rvert}{\lvert consensus \rvert}$		P(consensus correct sign)	
	(1)	(2)	(1)	(2)	(1)	(2)
Passive Ownership	-1.307***		0.298*		-0.171***	
	(0.437)		(0.166)		(0.028)	
Increase in Passive Ownership		-5.693***		0.941***		-0.198***
		(0.704)		(0.252)		(0.043)
Unconditional Avg.	6.28	6.28	0.39	0.39	0.95	0.95
Observations	155,983	82,413	126,685	65,338	127,991	66,308
R-Squared	0.151	0.146	0.012	0.014	0.006	0.007
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Firm-Level Controls	YES	YES	YES	YES	YES	YES

these stocks. When the good cashflow news arrives, the informed investors were previously underweight these stocks, so the diversification motive is weak, and they buy. In the context of this model, I view buy and hold investors as passive owners and the cashflow news as earnings announcements.

To test the model's predictions, I run the following regression:

$$\begin{aligned} r_{i,t} = \alpha + \beta_1 SUE_{i,t} + \beta_2 1_{SUE<0} + \gamma_1 \left(SUE_{i,t} \times Passive_{i,t}\right) + \\ \gamma_2 \left(1_{SUE<0} \times Passive_{i,t}\right) + \phi X_{i,t} + FE + e_{i,t} \end{aligned} \tag{19}$$

Here, $r_{i,t}$ denotes the market-adjusted return on the effective quarterly earnings date. SUE is defined as in Novy-Marx (2015): $SUE_{i,t} = \frac{E_{i,t} - E_{i,t-4}}{\sigma_{(t-1,t-8)}(E_{i,t} - E_{i,t-4})}$[16]. Controls in $X_{i,t}$ include lagged firm size, lagged idiosyncratic volatility over the past year, and institutional ownership. Fixed effects (FE) include year, 2-digit SIC industry and firm. The coefficients of interest are γ_1 and γ_2, which capture the effect of passive ownership on responses to earnings surprises, and the asymmetric effect of responses to negative earnings surprises.

Table 7 contains the regression results. Consistent with the model, firms with a high share of passive ownership are more responsive to earnings news.

5.3 Downloads of SEC Filings

One measure of attention is the number of downloads of SEC filings (see e.g. Loughran and McDonald (2017)). If passive ownership reduces incentives to gather fundamental information, the number of downloads of SEC filings might decrease for firms with high share of passive ownership. To test this, I run the following regression:

$$\Delta_{(t-1,t)} DL_{i,t} = \alpha + \beta \Delta_{(t-1,t)} Passive_{i,t} + X_{i,t} + FE + e_{i,t} \tag{20}$$

$\Delta_{(t-1,t)}$ is the year-over-year change. $DL_{i,t}$ is the number of non-robot downloads of 10-K's, 10-Q's and 8-K's in the 30 days before earnings announcements. Robot downloads include web crawlers, index page requests and individual IPs

[16] Results are similar when calculating SUE relative to IBES estimates using the method in Anson, Chambers, Black, Kazemi, Association, et al. (2012).

Passive Ownership and Response to Earnings News

Table 7:

$$r_{i,t} = \alpha + \beta_1 SUE_{i,t} + \beta_2 1_{SUE<0} + \gamma_1 \left(SUE_{i,t} \times Passive_{i,t}\right) + \gamma_2 \left(1_{SUE<0} \times Passive_{i,t}\right) + \phi X_{i,t} + FE + e_{i,t}$$

$r_{i,t}$ is the market adjusted return on the quarterly effective earnings date. Controls in $X_{i,t}$ include lagged firm size, lagged idiosyncratic volatility over the past year, and institutional ownership. Fixed effects include year, 2-digit SIC industry and firm.

	Market-Adjusted ED Return	
	(2)	(3)
SUE	0.181***	0.210***
	(0.011)	(0.012)
$1_{SUE<0}$	-0.863***	-0.726***
	(0.033)	(0.039)
SUE x Passive Share	3.270***	2.577***
	(0.219)	(0.227)
$1_{SUE<0} \times PassiveShare$	-0.836	-3.299***
	(0.524)	(0.608)
Observations	422,569	422,569
Firm Controls	Yes	Yes
Year FE	Yes	Yes
Industry FE	Yes	Yes
Firm FE	No	Yes

Downloads and Passive Ownership

Table 8:

$$\Delta DL_{i,t} = \alpha + \beta \Delta Passive_{i,t} + X_{i,t} + FE + e_{i,t}$$

.

	Change in # of non-Robot Downloads	
	30 Days Before ED	Earnings Day
Change in Passive	-0.533***	0.378**
	(0.072)	(0.167)
Observations	119,712	119,807
R-Squared	0.076	0.049
Firm-Level Controls	YES	YES
Year/Industry FE	YES	YES
Firm FE	YES	YES

with large number of downloads in a single day. This definition is based on data made available by Bill McDonald, originally derived from the Edgar Server Log between 2003 and 2015. I exclude robot downloads as they may automatically download all filings at release, or update a database periodically for reasons other than information gathering. Controls in $X_{i,t}$ include size, idiosyncratic volatility, institutional ownership and passive ownership. Fixed effects include year, day of the week and firm. Over time, the average number of downloads has been increasing.

Table 8 contains the regression results. Consistent with decreased information gathering, firms with increases in passive ownership experience decreases in pre-earnings downloads of SEC filings. The increase in downloads on the earnings days themselves has many explanations, including the increased earnings day volatility for firms with high passive ownership, as shown in Section 3.

5.4 Investment Q Relationship

Q-theory proposes a positive relationship between marginal Q and investment. If passive ownership has made prices less informative, the market value of the firm has become a less accurate measure of the true value of the firm. This implies that investment should become less sensitive to Q for firms with high passive ownership.

I test this in reduced form, and using plausibly exogenous increase in passive ownership that arises from being added to the S&P 500 index. Because marginal Q is hard to measure, I will work with average Q, the market-to-book ratio. The reduced-form specification is based on Eberly, Rebelo, Vincent, et al. (2008):

$$\begin{aligned}\left(\frac{I}{K}\right)_{i,t} &= \alpha + \beta_1 Log(Q)_{i,t} + \beta_2 Passive_{i,t} + \\ \beta_3 Passive_{i,t} &\times Log(Q)_{i,t} + \gamma X_{i,t} + FE + e_{i,t}\end{aligned} \tag{21}$$

where K is replacement value of capital, calculated as in Salinger and Summers (1983). Q= (Market Value of Equity + Book Value of Debt - Book Value of Inventories)/Capital. The S&P addition specification is identical to the difference-in-differences specification in Section 4, except $Log(Q)$ is also included on the right-hand-side.

Table 9 contains the regression results. Consistent with decreased price informativeness, firms with a higher share of passive ownership's investment is less sensitive to average Q.

6 Conclusion

Increases in passive ownership have lead to decreased pre-earnings price informativeness. When passive ownership in a stock increases, there is less pre-earnings trading, a smaller pre-earnings drift and a larger share of volatility on earnings days. These results are robust to only exploiting quasi exogenous variation in passive ownership that arises from index addition and rebalancing.

One potential mechanism for these trends is that passive ownership reduces

Passive Ownership and the Investment-Q Relation

Table 9:

$$\left(\frac{I}{K}\right)_{i,t} = \alpha + \beta_1 Log(Q)_{i,t} + \beta_2 Passive_{i,t} + \beta_3 Passive_{i,t} \times Log(Q)_{i,t} + \gamma X_{i,t} + FE + e_{i,t}$$

K is replacement value of capital, calculated as in Salinger and Summers (1983). Q= (Market Value of Equity + Book Value of Debt - Book Value of Inventories)/Capital.

	Investment/Capital		
log(Q)	0.0371***	0.0394***	0.0418***
	(0.001)	(0.001)	(0.005)
Passive Share		0.1733***	
		(0.025)	
Passive X Log(Q)		-0.0861***	
		(0.009)	
Log(Q) x Post Add to S&P 500			-0.0140*
			(0.008)
Observations	76,227	76,227	2,574
R-squared	0.063	0.065	0.49
Year FE	YES	YES	See S&P 500
Industry FE	YES	YES	Addition Setup
Firm FE	YES	YES	

the incentives to gather firm-specific information. Consistent with this channel, firms with increases in passive ownership experience decreases in the number of analysts covering the stock, and the accuracy of the remaining analysts.

Relative to total institutional ownership, the stock market currently has a relatively small share of passive owners, at around 10% of total US market capitalization. As passive ownership continues to grow, the decrease in efficiency could happen faster than it has in the past. Models like Grossman and Stiglitz (1980) predict a convex relationship between increases in uninformed investors and volatility on information release dates. It will be interesting to re-measure these efficiency trends in the future, and see if these non-linear effects kick in.

References

Abdi, F., and A. Ranaldo (2017): "A simple estimation of bid-ask spreads from daily close, high, and low prices," *The Review of Financial Studies*, 30(12), 4437–4480.

Anson, M. J., D. R. Chambers, K. H. Black, H. Kazemi, C. Association, et al. (2012): *CAIA level I: an introduction to core topics in alternative investments*, vol. 709. John Wiley & Sons.

Appel, I. R., T. A. Gormley, and D. B. Keim (2016): "Passive investors, not passive owners," *Journal of Financial Economics*, 121(1), 111–141.

Ben-David, I., F. Franzoni, and R. Moussawi (2014): "Do ETFs increase volatility?," Discussion paper, National Bureau of Economic Research.

Buffa, A. M., D. Vayanos, and P. Woolley (2014): "Asset management contracts and equilibrium prices," Discussion paper, National Bureau of Economic Research.

Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu (2001): "Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk," *The Journal of Finance*, 56(1), 1–43.

Eberly, J. C., S. Rebelo, N. Vincent, et al. (2008): "Investment and value: A neoclassical benchmark," Discussion paper, National Bureau of Economic Research.

Engelberg, J., and P. Gao (2011): "In search of attention," *The Journal of Finance*, 66(5), 1461–1499.

Farboodi, M., and L. Veldkamp (2017): "Long run growth of financial technology," Discussion paper, National Bureau of Economic Research.

Glosten, L. R., S. Nallareddy, and Y. Zou (2016): "ETF trading and informational efficiency of underlying securities," .

Grossman, S. J., and J. E. Stiglitz (1980): "On the impossibility of informationally efficient markets," *The American economic review*, 70(3), 393–408.

Gurkaynak, R. S., B. P. Sack, and E. T. Swanson (2004): "Do actions speak louder than words? The response of asset prices to monetary policy actions and statements," .

Israeli, D., C. M. Lee, and S. A. Sridharan (2017): "Is there a dark side to exchange traded funds? An information perspective," *Review of Accounting Studies*, 22(3), 1048–1083.

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

Mauboussin, M., D. Callahan, and D. Majd (2017): "Looking for Easy Games," Discussion paper, Credit Suisse.

McLean, R. D., and J. Pontiff (2016): "Does academic research destroy stock return predictability?," *The Journal of Finance*, 71(1), 5–32.

Morck, R., and F. Yang (2001): "The mysterious growing value of S&P 500 membership," Discussion paper, National Bureau of Economic Research.

Novy-Marx, R. (2015): "Fundamentally, momentum is fundamental momentum," Discussion paper, National Bureau of Economic Research.

Salinger, M., and L. H. Summers (1983): "Tax reform and corporate investment: A microeconometric simulation study," in *Behavioral simulation methods in tax policy analysis*, pp. 247–288. University of Chicago Press.

S&P, D. J. I. (2017): "S&P U.S. Indices Methodology," Discussion paper.

Wang, J. (1994): "A model of competitive stock trading volume," *Journal of political Economy*, 102(1), 127–168.

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

Weller, B. M. (2017): "Does Algorithmic Trading Reduce Information Acquisition?," *The Review of Financial Studies*, p. hhx137.

A Data Details

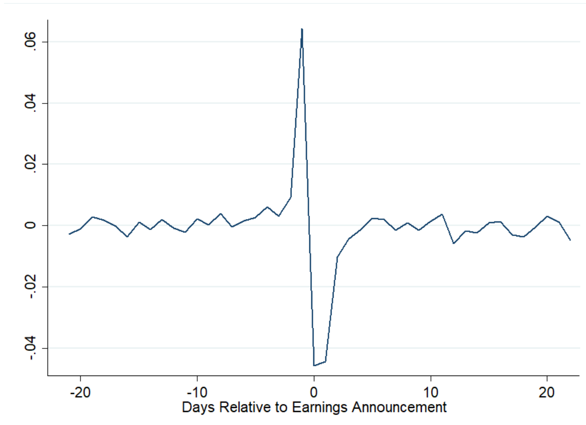
Daily Volume: Number of shares traded across all US exchanges. This quantity is not adjusted for splits during the month and it does not contain overallotments. Beginning in November 2008, volume also includes trades on the BATS Exchange, which now accounts for over 10% of all US equity trading.

I/B/E/S: Before 1998, nearly 90% of observations in IBES have an announcement time of "00:00:00", which implies the release time is missing. In 1998 this share drops to 23%, further drops to 2% in 1999, and continues to trend down to 0% by 2015. This implies that before 1998, if the earnings release date was a trading day, I will always classify that as the effective earnings date, even if earnings were released after markets closed, and it was not possible to trade on that information until the next trading day.

This time-variation in missing observations is not driving my results for two reasons: (1) I re-run every regression using only post-2000 data and the results are similar (2) For the pre-earnings drift, and pre-earnings volume, I am measuring returns/volume leading up to an earnings announcement. These missing earnings times could only move the effective earnings date earlier in time, which would bias both of my measures toward finding nothing. If volume dropped significantly on the last trading day before the earnings announcement, this would not be included in my pre-earnings volume measure for observations with a missing announcement time. For the pre-earnings drift, and the earnings day share of volatility, it would lead to selecting days where no news was released, which likely have smaller, rather than larger moves on average, pushing DM toward 1, and QVS toward 1.6%.

Change in the Bid Ask Spread Around Earnings Announcements

Figure 8: Average change in bid-ask spread. Computed as (ask-bid)/ask, using daily data in CRSP.



B Liquidity

A mechanical explanation for the negative relationship between passive ownership and large returns on earnings days is that high passive ownership decreases the number of shares available for trading (float), which leads to higher price impact and transaction costs.

I examine the effect of increases in passive ownership on liquidity, as measured by the bid-ask spread. I calculate the daily bid-ask spread for firm i at time t as: $BA_{i,t} = (ask_{i,t} - bid_{i,t})/ask_{i,t}$ using the closing *bid* and *ask* in CRSP data. All results in this section are similar using the definition of bid-ask spread from Abdi and Ranaldo (2017). Figure 8 shows that the spread expands the day before an earnings announcement, and quickly reverts to pre-announcement levels.

To understand the effect of passive ownership on average liquidity, I run the following regression:

$$\Delta_{(t,t-n)}\overline{BA}_{i,t} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + X_{i,t-n} + FE + e_{i,t} \tag{22}$$

$\overline{BA}_{i,t}$ is the average bid ask spread across all days in year t. Controls, $X_{i,t-n}$,

Passive Ownership and Bid Ask Spread (Annual Averages)

Table 10: Estimates of β from:

$$\Delta_{(t,t-n)}\overline{BA}_{i,t} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + X_{i,t-n} + FE + e_{i,t}$$

$\overline{BA}_{i,t}$ is the average bid ask spread across all days in year t. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. Change in passive ownership is expressed as a decimal, so 0.01 = 1% increase. Controls, $X_{i,t-n}$, include lagged passive ownership, market capitalization, idiosyncratic volatility and total institutional ownership. I also condition on the growth in market capitalization from $t-n$ to t. Fixed effects include 2-digit SIC industry, year and firm. Specification (1) includes all firm-level controls, plus industry and year fixed effects. Specification (2) adds includes firm fixed effects. All standard errors are clustered at the firm/year level.

		1-year		3-year		5-year	
		(1)	(2)	(1)	(2)	(1)	(2)
Increase in Passive Ownership	1-year	-0.0473***	-0.0415***				
		(0.011)	(0.001)				
	3-year			-0.0680***	-0.0460***		
				(0.016)	(0.002)		
	5-year					-0.0885***	-0.0492***
						(0.017)	(0.002)
Observations		127,270	127,270	99,109	99,109	77,920	77,920
R-Squared		0.277	0.296	0.357	0.407	0.384	0.454
Firm Controls		Yes	Yes	Yes	Yes	Yes	Yes
Industry/Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Firm FE		No	Yes	No	Yes	No	Yes

include lagged passive ownership, market capitalization, idiosyncratic volatility and total institutional ownership. I also condition on the growth in market capitalization from $t-n$ to t. Fixed effects include 2-digit SIC industry, year and firm.

The results are in Table 10. Across all specifications, increases in passive ownership are negatively correlated with the average bid-ask spread. This suggests that stocks become more liquid as passive ownership increases.

As shown in Figure 8, liquidity dries up before earnings announcements. To understand the relationship between passive ownership and pre-earnings liquidity, I run the following regression:

$$\Delta_{(t,t-n)}BA_{i,(\tau-2,\tau-1)} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + X_{i,t-n} + FE + e_{i,t} \quad (23)$$

$BA_{i,(\tau-2,\tau-1)}$ is the change in the bid-ask spread between $\tau-2$ and $\tau-1$, where τ is an earnings announcement date. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. I only look at year-over-year changes to avoid differences in bid-ask spreads before annual earnings announcements and quarterly announcements or seasonal effects. Controls in $X_{i,t-n}$ include lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership, calculated as the sum of holdings in the 13-F filings. I also condition on the growth in market capitalization from $t-n$ to t.

The results are in Table 11. Although not all specifications are significant, there is a positive relationship between increases in passive ownership, and increases in the bid-ask spread before earnings announcement. This is consistent with the results in the main body of the paper, where increased passive ownership leads to less informative prices, and more adverse selection.

When using the S&P 500 addition setup, the results have the same sign as the reduced form results, but are insignificant. When using the Russell reconstitution, the results are either insignificant or go the wrong way. These together suggest that these changes in liquidity are not entirely driven by the increase in passive ownership, but rather unobserved (or omitted) firm characteristics correlated with increases in passive ownership.

Passive Ownership and Bid Ask Spreads (Earnings Days)

Table 11: Estimates of β from:

$$\Delta_{(t,t-n)}BA_{i,(\tau-2,\tau-1)} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + X_{i,t-n} + FE + e_{i,t}$$

$BA_{i,(\tau-2,\tau-1)}$ as the change in the bid-ask spread between $\tau - 2$ and $\tau - 1$, where τ is an earnings announcement date. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. Change in passive ownership is expressed as a decimal, so $0.01 = 1\%$ increase. Controls, $X_{i,t-n}$, include lagged passive ownership, market capitalization, idiosyncratic volatility and total institutional ownership. I also condition on the growth in market capitalization from $t-n$ to t. Fixed effects include 2-digit SIC industry, year and firm. Specification (1) includes all firm-level controls, plus industry and year fixed effects. Specification (2) adds includes firm fixed effects. All standard errors are clustered at the firm/year level.

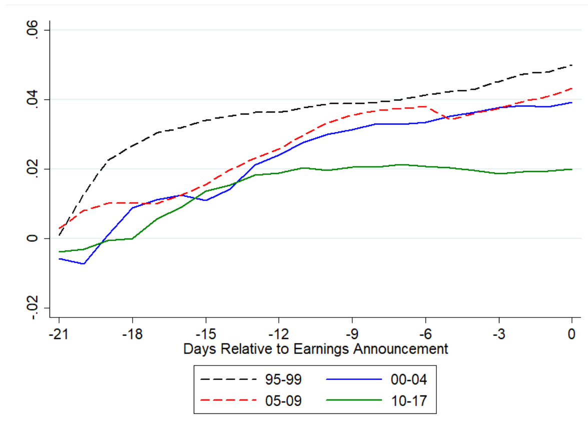
		1-year		3-year		5-year	
		(1)	(2)	(1)	(2)	(1)	(2)
Increase in Passive Ownership	1-year	0.0149***	0.011				
		(0.006)	(0.009)				
	3-year			0.00134	0.00211		
				(0.006)	(0.007)		
	5-year					0.0156***	0.0149**
						(0.005)	(0.007)
Observations		92,144	92,144	77,246	77,246	65,183	65,183
R-Squared		0.020	0.010	0.022	0.011	0.025	0.012
Firm Controls		Yes	Yes	Yes	Yes	Yes	Yes
Industry/Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Firm FE		No	Yes	No	Yes	No	Yes

Placebo Test: Pre-Earnings Volume

Figure 9: Plot of $\beta_{\{j==-\tau\}}$ estimated from the regression:

$$\overline{CAV_{i,j,t}} = \alpha + \sum_{\tau=-21}^{0} \beta_\tau \mathbf{1}_{\{j==-\tau\}} + \text{Fixed Effects} + e_{i,j,t}$$

Where t denotes a placebo earnings date. Placebo earnings dates are randomly assigned within each quarter for each firm.

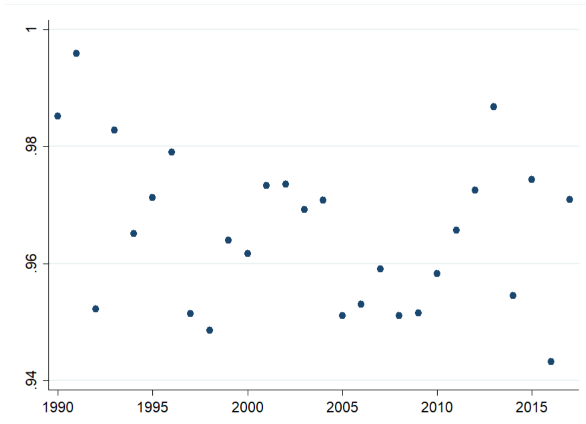


C Trend Placebo Tests

This section replicates Figures 2 (decrease in pre-earnings volume), 3 (decrease in pre-earnings drift) and 5 (increase in earnings day volatility), except replaces the true earnings dates with a randomly selected date for each firm each quarter. In all three cases, there is no trend toward decreased informativeness on the placebo earnings dates.

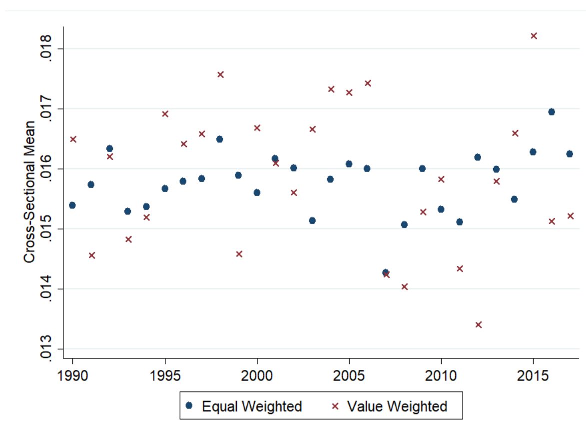
Placebo Test: Pre-Earnings Drift

Figure 10: This figure plots the cross-sectional average of the drift magnitude measure, $DM_{i,t} = \frac{r_{i,(t-30,t-1)}}{r_{i,(t-30,t)}}$, by year where t denotes a placebo earnings date. Placebo earnings dates are randomly assigned within each quarter for each firm.



Placebo Test: Earnings Day Volatility

Figure 11: This figure plots the share of market-adjusted quadratic variation occurring on placebo earnings days. For firm i in year t the quadratic variation share (QVS) is defined as: $QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, where r denotes a market adjusted daily return. The numerator is the sum of squared returns on the 4 placebo earnings dates, while the denominator is the sum of squared returns on all trading days in year t. Placebo earnings dates are randomly assigned within each quarter for each firm.

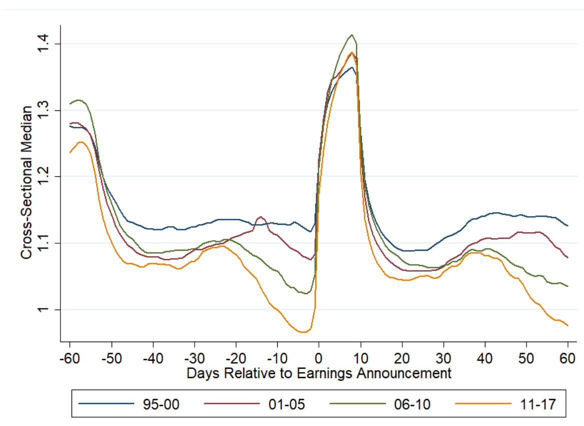


D Additional Pre-Earnings Volume Results

Rather than look at the 22 days before an earnings announcement, I expand the analysis to 60 days before the earnings announcement. This roughly corresponds to the time of the last earnings announcement. A concern with the regression specification in Equation 3 (regression of cumulative abnormal volume on day-before-earnings indicator variables) is that average earnings day volume has increased, so the relative volume on the days leading up to the earnings days would appear to *mechanically* decrease. Figure12 shows the cross-sectional median pre-earnings volume, which exhibits the same decline in pre-earnings volume as Figure 2.

Decline of Pre-Earnings Volume, Expanded Window

Figure 12: Plot of 10-day moving average of abnormal volume. Abnormal volume is volume relative to the historical average over the past year. Average historical volume is fixed at the beginning of each 10-day moving-average window to avoid mechanically amplifying drops in volume.



The figure also motivates my choise of a 22 trading-day window for the drop in pre-earnings volume: This is where there are differences across years. This is a case of looking where the effect is, but nothing in this figure suggests that this trend is driven by changes in passive ownership.

Another explanation for decreased pre-earnings volume is that informed trading before earnings announcements has moved to dark pools. This could occur because on lit exchanges, informed traders are getting front-run by algorithm traders. To test this, I obtained data on dark pool volume from FINRA. There does not appear to be an increase in dark pool volume in the weeks before earnings announcements.

Another possibility is that the increase in algorithmic trader (AT) activity over time is driving the decrease in pre-earnings volume, as AT's may trade less before earnings announcements. In years where I can construct the AT activity measures of Weller (2017), and add them to the right-hand-side of Regression 2 (regression used to compute β's for Figure 2), there is still a decline in volume before earnings announcements.

E Pre-Earnings Drift

Table 12 contains several alternative definitions of the pre-earnings drift. Only $DM_{i,t}$ is consistent with my intuition for all permutations of pre-earnings and earnings-day returns, and the pre-earnings information content of prices.

Examples of Pre-Earnings Drift Measures

Table 12: This table presents 6 hypothetical scenarios. Columns 2 and 3 contain the pre-earnings and earnings-day returns. Columns 4 and 5 calculate the drift magnitude under alternative definitions in net and gross returns, while column 6 has my definition, $DM_{i,t}$.

Case	$r_{i,(t-30,t-1)}$	$r_{i,t}$	$\frac{r_{i,t}}{r_{i,(t-30,t-1)}}$	$\frac{1+r_{i,t}}{1+r_{i,(t-30,t-1)}}$	$DM_{i,t}$	Intuition
1	-1.00%	-2.00%	2.00	1.01	0.33	Some Info.
2	-1.00%	2.00%	(2.00)	0.97	(1.00)	Low Info.
3	-1.00%	-0.50%	0.50	0.99	0.67	High Info.
4	1.00%	-2.00%	(2.00)	1.03	(1.00)	Low Info.
5	1.00%	2.00%	2.00	0.99	0.33	Some Info.
6	1.00%	0.50%	0.50	1.00	0.67	High Info.

F Reduced-Form Placebo Regressions

This section contains placebo tests for the reduced-form regressions. I select dates between the actual earnings days to represent placebo earnings dates. For example, if a firm released earnings on 12/31/2017, I would select the trading day closest to 11/15/2017 as the placebo earnings date. In all cases, the results for placebo earnings days are insignificant.

G Relationship to Competing Hypotheses

G.1 Rise of AT Activity

Tables 16, 17 and 18 contains alternative versions of the reduced-form regressions, which include controls for algorithmic trading (AT) activity.

Only the results for the pre-earnings drift and the earnings day share of volatility are significant in the matched subsample. For the specifications that are significant in the subset of my original sample that I can match to MIDAS data, adding the AT activity controls does reduce the coefficient on passive ownership/change in passive ownership, but the sign and statistical significance is unchanged. This implies that increased AT activity may partially explain the obseved decrease in market efficiency, but passive ownership is still an important factor.

G.2 Regulation Fair Disclosure

Tables 19, 20, 21 contain alternative versions of the reduced-form regressions, restricting the sample to data after 2000. The results are qualitatively similar, which alleviates concerns of the results being driven by time trends resulting from Reg FD, which was passed in August 2000.

Placebo Test: Pre-Earnings Volume

Table 13: Estimates of β from:

$$\Delta_{(t,t-n)}CAV_{i,t} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + X_{i,t-n} + FE + e_{i,t}$$

$CAV_{i,t}$ is average pre-earnings cumulative abnormal volume *per day* over the 22 days leading up to the earnings announcement. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. Change in passive ownership is expressed as a decimal, so $0.01 =$ 1% increase. Controls, $X_{i,t-n}$, include lagged passive ownership, market capitalization, idiosyncratic volatility and total institutional ownership. I also condition on the growth in market capitalization from $t-n$ to t. Fixed effects include 2-digit SIC industry and year.

The "Baseline" results are estimates from Table 1, while the "Placebo" results are the coefficient estimates when selecting dates between the actual earnings days as the placebo earnings dates.

		1-year		3-year		5-year	
		Placebo	Baseline	Placebo	Baseline	Placebo	Baseline
Increase in Passive Ownership	1-year	-0.628	-1.263*				
		(0.691)	(0.651)				
	3-year			-0.67	-0.936***		
				(0.434)	(0.309)		
	5-year					-0.502	-1.041**
						(0.382)	(0.421)
Observations		239,402	272,609	157,530	191,654	109,632	139,537
R-Squared		0.016	0.009	0.024	0.007	0.047	0.025
Firm Controls		Yes	Yes	Yes	Yes	Yes	Yes
Industry/Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Firm FE		No	No	No	No	No	No

Placebo Test: Pre-Earnings Drift

Table 14: Table with estimates of β from:

$$DM_{i,t} = \alpha + \beta Passive_{i,t} + \sum_{j=2}^{10} \phi_j \mathbf{1}_{SUEdeclie=j} + \gamma X_{i,t-1} + FE + e_{i,t}$$

$DM_{i,t} = \frac{r_{i,(t-30,t-1)}}{r_{i,(t-30,t)}}$, which is the ratio of the cumulative returns in the 30 days leading up to the earnings day, relative cumulative return in the 30 days up to and including the earnings day. Passive ownership is expressed as a decimal, so 0.01 = 1% of shares outstanding held by passive funds. SUE deciles are formed each quarter. $DM_{i,t}$ is Winzorized at the 1% and 99% levels. Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. Fixed effects include 2-digit SIC industry and year. All standard errors are clustered at the firm/year level.

The "Baseline" results are estimates from Table 2, while the "Placebo" results are the coefficient estimates when selecting dates between the actual earnings days as the placebo earnings dates.

	Pre-Earnings Drift	
	Placebo	Baseline
Passive Ownership	-0.0133	-0.329**
	(0.071)	(0.157)
Observations	394,397	413,328
R-Squared	0.000	0.003
Firm Controls	Yes	Yes
SUE	Yes	Yes
Industry/Year FE	Yes	Yes
Firm FE	No	No

Placebo Test: Earnings Day Share of Volatility

Table 15: Table with estimates of β from:

$$\Delta_{(t,t-n)} QVS_{i,t} = \alpha + \beta \Delta_{(t,t-n)} Passive_{i,t} + \gamma X_{i,t-n} + FE + e_{i,t}$$

$QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, which is the ratio of the squared returns on the 4 quarterly earnings announcement days, relative to the squared returns on all days in year t. QVS takes values in [0,1]. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. Fixed effects include 2-digit SIC industry and year. All standard errors are clustered at the firm/year level.

The "Baseline" results are estimates from Table 3, while the "Placebo" results are the coefficient estimates when selecting dates between the actual earnings days as the placebo earnings dates.

		1-year		3-year		5-year	
		Placebo	Baseline	Placebo	Baseline	Placebo	Baseline
Increase in Passive Ownership	1-year	0.00069	0.0983**				
		(0.010)	(0.039)				
	3-year			0.00329	0.103***		
				(0.005)	(0.034)		
	5-year					0.00818	0.171***
						(0.005)	(0.042)
Observations		65,951	79,253	50,151	60,578	40,215	48,484
R-Squared		0.002	0.015	0.002	0.052	0.003	0.077
Firm Controls		Yes	Yes	Yes	Yes	Yes	Yes
Industry/Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Firm FE		No	No	No	No	No	No

AT Activity: Pre-Earnings Volume

Table 16: Estimates of β from:

$$\Delta_{(t,t-n)}CAV_{i,t} = \alpha+\beta\Delta_{(t,t-n)}Passive_{i,t}+\gamma X_{i,t-n}+\phi\Delta_{(t,t-n)}ATActivity_{i,t}+FE+e_{i,t}$$

$CAV_{i,t}$ is average pre-earnings cumulative abnormal volume *per day* over the 22 days leading up to the earnings announcement. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. Change in passive ownership is expressed as a decimal, so $0.01 = 1\%$ increase. $\Delta_{(t,t-n)}ATActivity_{i,t}$ is a vector of year-over-year changes in the 4 AT activity measures from Weller (2017). Controls, $X_{i,t-n}$, include lagged passive ownership, market capitalization, idiosyncratic volatility and total institutional ownership. I also condition on the growth in market capitalization from $t-n$ to t. Fixed effects include 2-digit SIC industry and year. Only includes data from 2012-2016.

		1-year		3-year	
		Baseline	+AT Controls	Baseline	+AT Controls
Increase in Passive Ownership	1-year	-0.488	0.438		
		(0.961)	(0.913)		
	3-year			-0.496***	-0.253***
				(0.078)	(0.057)
Observations		38,081	38,048	15,681	15,665
R-Squared		0.006	0.054	0.044	0.125
Firm Controls		Yes	Yes	Yes	Yes
Industry/Year FE		Yes	Yes	Yes	Yes
Firm FE		No	No	No	No

AT Activity: Pre-Earnings Drift

Table 17: Table with estimates of β from:

$$DM_{i,t} = \alpha + \beta Passive_{i,t} + \sum_{j=2}^{10} \phi_j \mathbf{1}_{SUEdeclie=j} + \gamma X_{i,t-1} + \psi ATActivity_{i,t} + FE + e_{i,t}$$

$DM_{i,t} = \frac{r_{i,(t-30,t-1)}}{r_{i,(t-30,t)}}$, which is the ratio of the cumulative returns in the 30 days leading up to the earnings day, relative cumulative return in the 30 days up to and including the earnings day. Passive ownership is expressed as a decimal, so 0.01 = 1% of shares outstanding held by passive funds. SUE deciles are formed each quarter. $DM_{i,t}$ is Winzorized at the 1% and 99% levels. $ATActivity_{i,t}$ is a vector containing the 4 AT activity measures from Weller (2017). Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. Fixed effects include 2-digit SIC industry and year. All standard errors are clustered at the firm/year level. Only includes data from 2012-2016.

	Pre-Earnings Drift	
	Baseline	+AT Controls
Passive Ownership	-1.281***	-1.074***
	(0.379)	(0.358)
Observations	54,119	54,079
R-Squared	0.004	0.004
Firm Controls	Yes	Yes
SUE	Yes	Yes
Industry/Year FE	Yes	Yes
Firm FE	No	No

AT Activity: Earnings Day Share of Volatility

Table 18: Table with estimates of β from:

$$\Delta_{(t,t-n)}QVS_{i,t} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + \gamma X_{i,t-n} + \phi\Delta_{(t,t-n)}ATActivity_{i,t} + FE + e_{i,t}$$

$QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, which is the ratio of the squared returns on the 4 quarterly earnings announcement days, relative to the squared returns on all days in year t. QVS takes values in [0,1]. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. $\Delta_{(t,t-n)}ATActivity_{i,t}$ is a vector of year-over-year changes in the 4 AT activity measures from Weller (2017). Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. Fixed effects include 2-digit SIC industry and year. All standard errors are clustered at the firm/year level. Only includes data from 2012-2016.

		1-year		3-year	
		Baseline	+AT Controls	Baseline	+AT Controls
Increase in Passive Ownership	1-year	0.188**	0.167*		
		(0.089)	(0.092)		
	3-year			0.103***	0.0502*
				(0.034)	(0.026)
Observations		9,910	9,904	4,023	4,020
R-Squared		0.009	0.010	0.035	0.041
Firm Controls		Yes	Yes	Yes	Yes
Industry/Year FE		Yes	Yes	Yes	Yes
Firm FE		No	No	No	No

Post-2000: Pre-Earnings Volume

Table 19: Estimates of β from:

$$\Delta_{(t,t-n)}CAV_{i,t} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + X_{i,t-n} + FE + e_{i,t}$$

$CAV_{i,t}$ is average pre-earnings cumulative abnormal volume *per day* over the 22 days leading up to the earnings announcement. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. Change in passive ownership is expressed as a decimal, so 0.01 = 1% increase. Controls, $X_{i,t-n}$, include lagged passive ownership, market capitalization, idiosyncratic volatility and total institutional ownership. I also condition on the growth in market capitalization from $t-n$ to t. Fixed effects include 2-digit SIC industry and year. Only includes data from 2001-2016.

		1-year		3-year		5-year	
		(1)	(2)	(1)	(2)	(1)	(2)
Increase in Passive Ownership	1-year	-1.102*	-1.033***				
		(0.647)	(0.231)				
	3-year			-0.862***	-0.537***		
				(0.314)	(0.174)		
	5-year					-1.007**	-0.596***
						(0.434)	(0.188)
Observations		187,778	187,778	145,900	145,900	115,600	115,600
R-Squared		0.004	0.005	0.007	0.008	0.027	0.028
Firm Controls		Yes	Yes	Yes	Yes	Yes	Yes
Industry/Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Firm FE		No	Yes	No	Yes	No	Yes

Post-2000: Pre-Earnings Drift

Table 20: Table with estimates of β from:

$$DM_{i,t} = \alpha + \beta Passive_{i,t} + \sum_{j=2}^{10} \phi_j \mathbf{1}_{SUEdeclie=j} + \gamma X_{i,t-1} + FE + e_{i,t}$$

$DM_{i,t} = \frac{r_{i,(t-30,t-1)}}{r_{i,(t-30,t)}}$, which is the ratio of the cumulative returns in the 30 days leading up to the earnings day, relative cumulative return in the 30 days up to and including the earnings day. Passive ownership is expressed as a decimal, so 0.01 = 1% of shares outstanding held by passive funds. SUE deciles are formed each quarter. $DM_{i,t}$ is Winzorized at the 1% and 99% levels. Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. Fixed effects include 2-digit SIC industry and year. All standard errors are clustered at the firm/year level. Only includes data from 2001-2016.

	Pre-Earnings Drift	
	(1)	(2)
Passive Ownership	-0.260**	-0.321***
	(0.102)	(0.096)
Observations	232,301	232,301
R-Squared	0.011	0.004
Firm Controls	Yes	Yes
SUE	Yes	Yes
Industry/Year FE	Yes	Yes
Firm FE	No	Yes

Post-2000: Earnings Day Share of Volatility

Table 21: Table with estimates of β from:

$$\Delta_{(t,t-n)}QVS_{i,t} = \alpha + \beta\Delta_{(t,t-n)}Passive_{i,t} + \gamma X_{i,t-n} + FE + e_{i,t}$$

$QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, which is the ratio of the squared returns on the 4 quarterly earnings announcement days, relative to the squared returns on all days in year t. QVS takes values in [0,1]. $\Delta_{(t,t-n)}$ is the change from calendar year $t-n$ to calendar year t. Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. Fixed effects include 2-digit SIC industry and year. All standard errors are clustered at the firm/year level. Only includes data from 2001-2016.

		1-year		3-year		5-year	
		(1)	(2)	(1)	(2)	(1)	(2)
Increase in Passive Ownership	1-year	0.107***	0.0583				
		(0.041)	(0.038)				
	3-year			0.113***	0.0414		
				(0.037)	(0.032)		
	5-year					0.188***	0.0815**
						(0.044)	(0.033)
Observations		49,113	49,113	40,556	40,556	34,639	34,639
R-Squared		0.016	0.016	0.051	0.049	0.070	0.064
Firm Controls		Yes	Yes	Yes	Yes	Yes	Yes
Industry/Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Firm FE		No	Yes	No	Yes	No	Yes

H Alternative Identified Evidence

H.1 S&P 500 Index Deletions

In section 4, I use S&P 500 index additions to identify plausibly exogenous increases in passive ownership. A natural extension is to run a similar difference-in-differences regression, but use the decrease in passive ownership associated with index deletion as the treatment. In this DID setup, the parallel trends assumption is unlikely valid, because index deletion is always about fundamentals.

The next challange is identifying the control group, which should consist of firms with a similar likelihood of being dropped from the index as the treated firms. Three major reasons for S&P 500 index deletion are small market capitalization, poor performance and lack of liquidity. To facilitate a direct comparison with the index addition results, I will use the same sorting mechanism of industry, size and growth rate to identify control firms, even though it would be reasonable to remove the industry filter and replace it with a measure of liquidity.

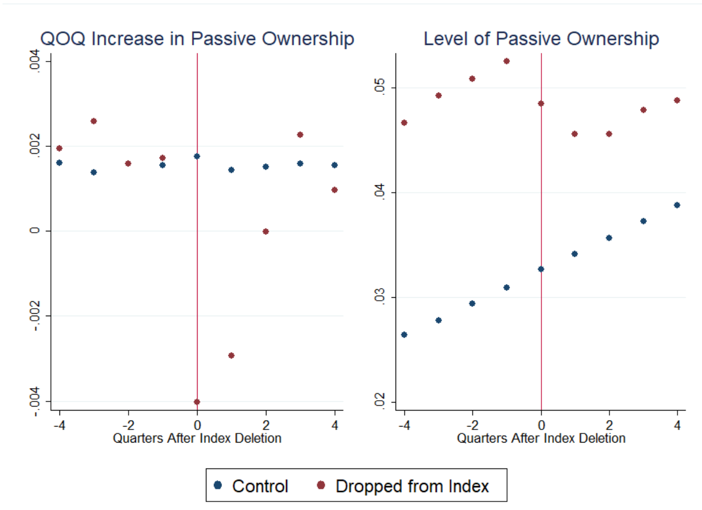
In the index deletion setup, the treatment group is all firms dropped from the S&P 500 index. The control group is all firms in the same 2-digit SIC industry, in the same size and growth rate quintiles that were initially in the S&P 500 index, and remained there over the next two years[17].

Figure 13 shows the changes in passive ownership around the index deletion date. There is a drop in passive ownerhsip in the quarter of deletion, and the quarter after deletion. Unlike 6, however, the drop is only temporary, as can be seen in the levels plot. One explanation for this is that stocks on the margin are still relatively large, and were affected by the ETF/passive management boom which increased passive ownership for all stocks. The weak and temporary

[17]All the results index deletion results are similar if the control group only includes firms were initially not in the S&P 500 index, and remained out of the index over the next two years. Results are also similar when choosing the treatement period to be the year immidetly after index deletion, instead of skipping a year.

S&P 500 Index Deletions: Testing Parallel Trends

Figure 13: Average level and increase in passive ownership for control firms and firms dropped from the S&P 500. Control firms are all firms in the same 2-digit SIC industry, in the same size and growth rate quintiles that were initially in the S&P 500 index, and remained there over the next two years.



treatment effect implies the DID results will be insignificant.

The regressions testing the effect of index deletions on pre-earnings volume, drift and volatility are in Table 22. I omit the implied elasticity because the first stage effect is near zero. The pre-earnings drift and earnings day share of volatility are now insignificant, while the volume result is significant and negative.

Because these firms are experiencing a *decrease* in passive ownership, this volume result is contrary to the results in the rest of the paper. Some explanations for this contrary result in volume, but not pre-earnings drift or earnings-day volatility is that S&P 500 funds have (1) a different propensity to lend out shares (2) a higher share of volume coming from algorithmic traders (because of ETF arbitrage, for example), but these are only conjectures.

Effects of S&P 500 Index Deletion

Table 22: For Pre-Earnings Volume and Earnings Day Share of QV:

$$\Delta_{(t-1,t+1)}Outcome_{i,t} = \alpha + \gamma Dropped_{i,t} + FE + e_{i,t}$$

For Pre-Earnings Drift:

$$Outcome_{i,t} = \alpha + \beta Dropped_{i,t} + \tau After_{i,t} + \gamma Added_{i,t} \times After_{i,t} + FE + e_{i,t}$$

Dropped firms are those which were removed to the S&P 500 index. Control firms are in the same industry, same size and same growth rate quintile as the control firms. Fixed effects include industry, size quintile, growth rate quintile, and year/month of index deletion.

	Pre-Earnings Volume	Pre-Earnings Drift	Earnings Day Share of QV
	Differences	Levels	Differences
Treated	-0.0699*		-0.0197
	(0.026)		(0.017)
Interaction		-0.0742	
		(0.100)	
R-squared	0.076	0.028	0.23
Year/Month FE	YES	YES	YES
Industry FE	YES	YES	YES
Size Quintile FE	YES	YES	YES
Growth Quintile FE	YES	YES	YES
Treated Firms	245	245	245
Control Firms	1,239	1,239	1,239

H.2 Moving from the Russell 2000 to the Russell 1000

Similar to the S&P 500 index deletion, firms experience a decrease in passive ownership after they are moved from the Russell 2000 to the Russell 1000. This is because they go from being the largest firm in a value-weighted index of small firms, to the smallest firm in a value-weighted index of large firms. Unlike the S&P 500 deletions, however, this DID setup still satisfies the exclusion restriction, as moving from firm 1001 to 999 may have nothing to do with firm fundamentals.

Similar to the setup in Section 4, I choose the control firms to be all Russell 3000 firms, with June ranks between 900 and 1100 that did not switch from the 1000 to the 2000, or from the 2000 to the 1000. Figure 7 shows the problem with this setup: the treatment is small and temporary. The common trend between moving from the Russell 2000 to the 1000, and the S&P 500 index deletion suggests that the general upward trend in passive ownership for almost all stocks drowns out the temporary change in passive ownership associated with various index rebalancing.

In Table 23, I replicate the Russell experiment results of Table 5. As with the S&P 500 deletions, the pre-earnings drift and earnings day share of volatility are now insignificant, while the volume result is significant. Again, this volume result is contrary to the results in the rest of the paper – suggesting there is not something special about deletion from the S&P 500 index and the associated drop in pre-earnings volume.

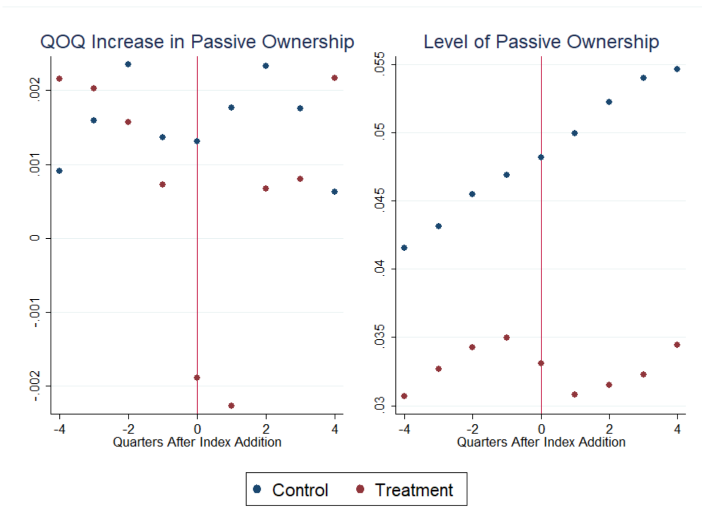
I Post-Earnings Drift

As documented in McLean and Pontiff (2016), the post-earnings drift has declined. To understand the relationship between increases in passive ownership and the post-earnings drift, I run the following regression:

$$PD_{i,t} = \alpha + \beta Passive_{i,t} + \sum_{j=2}^{10} \phi_j \mathbf{1}_{SUEdeclie=j} + \gamma X_{i,t-1} + FE + e_{i,t} \quad (24)$$

Russell 2000/1000 Reconstitution: Checking Parallel Trends

Figure 14: Average level and change in passive ownership for control firms and firms moved from the Russell 2000 to the Russell 1000. Control firms are all firms in the Russell 3000 ranked 900 to 1100 that did not move from the 1000 to the 2000 or from the 2000 to the 1000.



Effects of Russell 1000/2000 Index Reconstitution

Table 23: For Pre-Earnings Volume and Earnings Day Share of QV:

$$\Delta_{(t-1,t+1)}Outcome_{i,t} = \alpha + \gamma Moved_{i,t} + FE + e_{i,t}$$

For Pre-Earnings Drift:

$$Outcome_{i,t} = \alpha + \beta Moved_{i,t} + \tau After_{i,t} + \gamma Moved_{i,t} \times After_{i,t} + FE + e_{i,t}$$

Added firms are those which were moved from the Russell 2000 to the Russell 1000. Control firms are those that were ranked 900-1100 by Russell, but did not switch from the 1000 to the 2000 or from the 2000 to the 1000. Fixed effects include year/month of index rebalancing.

	Pre-Earnings Volume Differences	Pre-Earnings Drift Levels	ED Share of QV Differences
Treated	-0.123***		0.00498
	(0.021)		(0.006)
Interaction		-0.0531	
		(0.040)	
R-squared	0.016	0.006	0.040
Year/Month FE	YES	YES	YES
Treated Firms	974	974	974
Control Firms	1,284	1,284	1,284

The post-earnings drift $PD_{i,t}$, is defined as the ratio of the return on the earning day itself, to the cumulative return on the earnings day itself and the following 30 trading days. $PD_{i,t} = \frac{r_{i,t}}{r_{i,(t,t+30)}}$. $PD_{i,t}$ is Winzorized at the 1% and 99% levels.

Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. The regression also includes a set of indicator variables for SUE decile within a given quarter. Fixed effects (FE) include 2-digit SIC industry, year and firm. Standard errors are clustered at the firm/year level.

Table 24 shows the regression results. Firms with high passive ownership have higher post-earnings drift, suggesting that passive ownership slows the incorporation of information even after it is released to the public.

In unreported results, I also re-run the S&P 500 index addition experiment and Russell 1000/2000 reconstitution experiment with the post-earnings drift as the outcome of interest. The post-earnings drift is significantly higher after a firm is added to the S&P 500 index, but there is no effect after a firm is moved from the Russell 2000 to the Russell 1000.

J Systematic Information Announcement Days

FOMC announcement dates are from Gurkaynak, Sack, and Swanson (2004). To account for the anticipated nature of earnings announcement, I restrict the sample to scheduled FOMC meetings.

Figure 15 shows the trends in volume before FOMC announcement dates. The periodic oscillation is due to day of the week effects. Figure 16 shows the drift before FOMC announcement dates, which if anything has been trending up. Figure 17 shows a slight trend toward increased volatility on FOMC announcement dates, but may be due to increased importance of FOMC meeting dates during the financial crisis.

Passive Ownership and Post-Earnings Drift

Table 24: Table with estimates of β from:

$$PD_{i,t} = \alpha + \beta Passive_{i,t} + \sum_{j=2}^{10} \phi_j \mathbf{1}_{SUEdeclie=j} + \gamma X_{i,t-1} + FE + e_{i,t}$$

$PD_{i,t} = \frac{r_{i,t}}{r_{i,(t,t+30)}}$, which is the ratio of the return on the earnings day itself, to the cumulative returns on the earnings day itself, and the 30 following trading days. Passive ownership is expressed as a decimal, so 0.01 = 1% of shares outstanding held by passive funds. SUE deciles are formed each quarter. $PD_{i,t}$ is Winzorized at the 1% and 99% levels. Controls in $X_{i,t-1}$ include lagged institutional ownership, lagged market capitalization and lagged market-adjusted volatility over the past year. Fixed effects include 2-digit SIC industry, year and firm. Specification (1) includes all firm-level controls, plus industry and year fixed effects. Specification (2) adds includes firm fixed effects. All standard errors are clustered at the firm/year level.

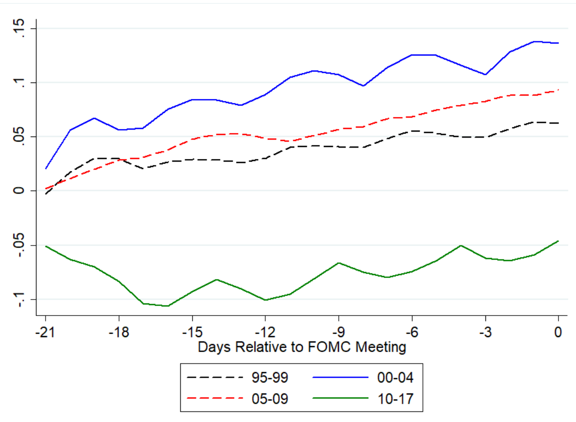
	Post-Earnings Drift	
	(1)	(2)
Passive Ownership	0.356***	0.366**
	(0.079)	(0.167)
Observations	411,019	411,019
R-Squared	0.003	0.002
Firm Controls	Yes	Yes
SUE	Yes	Yes
Industry/Year FE	Yes	Yes
Firm FE	No	Yes

FOMC Meeting Dates: Pre-Earnings Volume

Figure 15: Plot of $\beta_{\{j==-\tau\}}$ estimated from the regression:

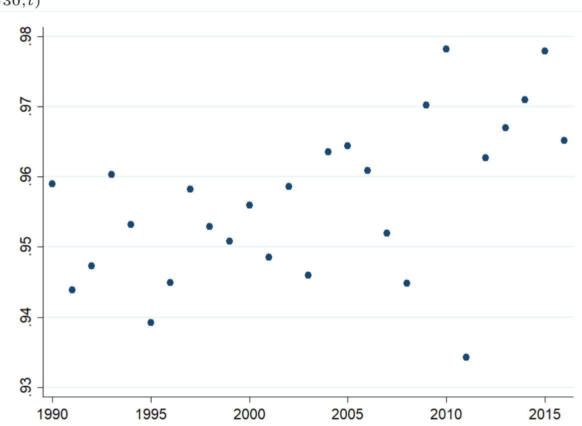
$$\overline{CAV_{i,j,t}} = \alpha + \sum_{\tau=-21}^{0} \beta_\tau \mathbf{1}_{\{j==-\tau\}} + \text{Fixed Effects} + e_{i,j,t}$$

Where t denotes a scheduled FOMC announcement date.



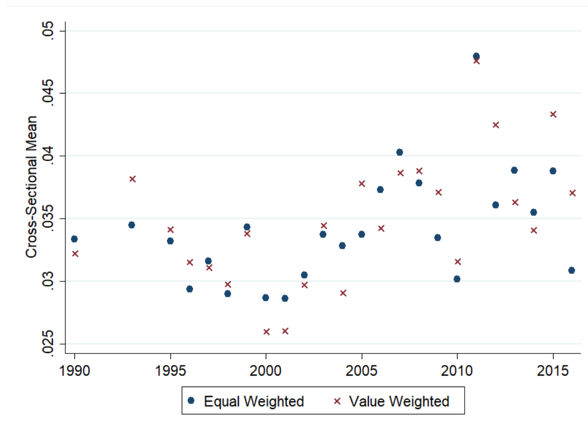
FOMC Meeting Dates: Pre-Earnings Drift

Figure 16: This figure plots the cross-sectional average of the drift magnitude measure, $DM_{i,t} = \frac{r_{i,(t-30,t-1)}}{r_{i,(t-30,t)}}$, by year where t denotes a scheduled FOMC meeting date.



FOMC Meeting Dates: Earnings Day Volatility

Figure 17: This figure plots the share of market-adjusted quadratic variation occurring on placebo earnings days. For firm i in year t the quadratic variation share (QVS) is defined as: $QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, where r denotes a market adjusted daily return. The numerator is the sum of squared returns on the 8 scheduled FOMC dates , while the denominator is the sum of squared returns on all trading days in year t.



K Model Details

The rational expectations equilibrium is a function from the informed signal s and the noisy demand u to the period 0 price that satisfies optimality and market clearing. This function takes the form:

$$price = A + B\left[(s - \overline{d}) - Cu\right] \tag{25}$$

where A, B and C are constants that depend on model parameters:

$$A = \left(\frac{\overline{d}}{1+r} - \frac{S\alpha\sigma^2_{d|s}}{(1+r)\left(N_I + N_U \frac{\sigma^2_{d|s}}{\sigma^2_{d|p}}\right)} \right)$$

$$B = \frac{N_I \beta_s}{(1+r)\left(N_I + N_U \frac{\sigma^2_{d|s}}{\sigma^2_{d|p}}\right) - N_U \frac{\sigma^2_{d|s}}{\sigma^2_{d|p}} \beta_p} \tag{26}$$

$$C = \frac{\alpha\sigma^2_\epsilon}{N_I}$$

Given the price function, we can compute the conditional mean of d given the price p:

$$E[d|p] = \overline{d} + \frac{\sigma^2}{B(\sigma^2 + \sigma^2_\epsilon + C^2\sigma^2_u)}(p - A) \tag{27}$$

Then, we can compute the conditional variance of d using the law of total variance: $V(d) = E[V(d|p)] + V(E[d|p])$. Rearranging yields:

$$\sigma^2_{d|p} = V(d) - V(E[d|p]) = \sigma^2 - \frac{\sigma^4}{\sigma^2 + \sigma^2_\epsilon + C^2\sigma^2_u} \tag{28}$$

Informativeness is the inverse of the conditional variance of d:

$$informativeness = \frac{1}{\sigma^2_{d|p}} = \frac{\sigma^2_d + \sigma^2_\epsilon + \left(\frac{\alpha\sigma^2_\epsilon}{N_I}\right)^2 \sigma^2_u}{\sigma^2_d \left(\sigma^2_\epsilon + \left(\frac{\alpha\sigma^2_\epsilon}{N_I}\right)^2 \sigma^2_u\right)} \tag{29}$$

The derivative of informativeness with respect to N_I is:

$$\frac{2\alpha^2\sigma^4_\epsilon\sigma^2_u N_I}{\left(\sigma^2_\epsilon N_I^2 + \alpha^2\sigma^4_\epsilon\sigma^2_u\right)^2} \tag{30}$$

Informativeness always increasing in N_I.

The derivative of informativeness with respect to σ^2_u is:

$$-\frac{\alpha^2\sigma^4_\epsilon N_I^2}{\left(\sigma^2_\epsilon N_I^2 + \alpha^2\sigma^4_\epsilon\sigma^2_u\right)^2} \tag{31}$$

Informativeness always decreasing in σ^2_u.