Earnings Announcements and the Rise of Passive Ownership

MARCO SAMMON*

March 3, 2020

ABSTRACT

This paper proposes a way to measure the effect of rising passive ownership on stock price informativeness that does not rely on any particular model. I examine patterns in trading volume, returns and volatility around days we know information is released: earnings announcements. Between 1990 and 2018, pre-earnings abnormal trading volume and the pre-earnings drift declined, while the share of annual volatility on earnings days increased. At the firm-level, there is a negative relationship between passive ownership and pre-earnings price informativeness. This result is robust to using only quasi-exogenous increases in passive ownership associated with S&P 500 additions and Russell 1000/2000 rebalancing. I link increases in passive ownership to decreases in information gathering: increases in passive ownership are correlated with fewer analysts covering a stock, fewer downloads of SEC filings, and a larger response to earnings news.

JEL classification: G12, G14.

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

I. Introduction

The rise of passive ownership is one of the biggest changes in asset markets over the past 30 years. Passive funds grew from owning less than 1% of the US stock market in the early 1990's to owing nearly 15% in 2018. As passive ownership continues to grow, academics and practitioners want to understand how passive investing affects stock price informativeness.

The general consensus is that: "*Passive investing is not about price discovery... it has the possibility to harm markets. More passive ownership could make markets work less efficiently.*[1]" Intuitively, the notion that passive investing decreases market efficiency makes sense. Passive investors often hold large, diversified baskets of securities, so they do not need to gather information about all the underlying firms. As fewer investors gather firm-specific information, prices become less informative. Although this seems obvious, there are two challenges in measuring this relationship empirically.

The first challenge is measuring the information content of prices. A natural starting point is a Grossman and Stiglitz (1980) style model, where price informativeness is the correlation between prices and fundamentals. The empirical analogue is a regression of future fundamentals on current prices:

$$fundamentals_{t+1} = \hat{\alpha} + \hat{\beta} \times price_t + controls + \hat{\epsilon}_t \tag{1}$$

Larger values of $\hat{\beta}$ suggest that prices are more informative: Fundamentals and prices covary more strongly with one another.

There are several difficulties, however, with estimating this regression. First, the correct measure of future fundamentals is not obvious. In a static model like Grossman-Stiglitz, there are no cashflows after $t + 1$, but in reality, firms are long-lived. Maybe the left-hand-side variable should be *all* fundamentals from $t+1$ forward, which are hard to measure. It is also not clear if earnings are the right measure of future fundamentals (see e.g. Bai, Philippon, and Savov (2016)), as management has some control over earnings growth (Schipper (1989)).

In addition, Grossman-Stiglitz defines price informativeness as a *conditional* covariance, which requires identifying the 'right' set of conditioning variables. Academic economists, however, still disagree on this. Bai et al. (2016) measure price informativeness as the vari-

[1]Malcolm Baker, director of research at Acadian Asset Management, interviewed in McElhaney (2018).

ance of fundamentals, conditional on prices. Dávila and Parlatore (2019) measure price informativeness as the variance of prices, conditional on fundamentals; effectively switching the left-hand-side and right-hand-side variables in Equation 1. This is a classic form of confusion for economists, similar to a regressions of prices-on-quantities or quantities-on-prices to estimate demand and supply curves. Swapping left-hand-side and right-hand-side variables does not solve this problem, as demand and supply cannot be the same line.

The second challenge in linking passive ownership and price informativeness is understanding the null hypothesis i.e., the predicted effect of increasing passive ownership in information-gathering models. Again, it seems natural to start from Grossman-Stiglitz: As the share of informed investors decreases, so does price informativeness. The growth of passive investing, however, has changed more than the share of informed and uninformed agents. Two often overlooked effects of growing passive ownership are (1) the ease of making factor bets and (2) increases in the total number of investors.

A common misconception is that all Exchange Traded Funds (ETFs) are like the SPDR S&P 500 Trust ETF (SPY) i.e., broad market indices. Further, it is often assumed that investors who buy ETFs plan to hold them forever in their retirement accounts. The opposite is true. Only 30% of, "*ETF investors look at these as passive funds, [and] are just there long term.*[2]" Many ETF holders are sophisticated institutional investors looking for targeted factor exposure. In reference to Global X's ETF offerings, former CEO Bruno del Ama said, "*Hedge funds tend to use our ETFs as a tactical play to get in and out of segments that are difficult for them to access directly. Greece is a good example. GREK has seen a lot [of] hedge fund trading.*[3]"

ETFs are also bringing new investors into the market. In 2017, The New York Times covered a former logistics manager at Target who made over $10 million investing in XIV, a now defunct ETF which was short the CBOE Volatility Index (VIX). Many ETF tickers are themselves used as marketing tools. For example, Defiance ETFs gave their "Next Gen Connectivity ETF" the ticker "FIVG". This is after the Securities and Exchange Commission (SEC) asked Defiance to remove "5G" from the fund's name because the agency thought it could be misleading to prospective investors (Hajric and Pickert (2019)).

Given the broader changes in financial markets brought about by the rise of ETFs, the

[2] Daniel Gamba, Global Head of Active Equity Product Strategy, BlackRock, quoted in Balchunas (2016).
[3] Bruno del Ama, former CEO of Global X, quoted in Balchunas (2016).

predicted effect of passive ownership on price informativeness is not obvious. The growth of passive funds increased the number of investors, and made it easier for informed investors to make factor bets. These changes could have led to increased incorporation of systematic information, and more informative prices.

The goal of this paper is to quantify the effect of rising passive ownership on price informativeness. Given that the model predictions are ambiguous, this question needs to be answered empirically. Given the difficulties inherent in estimating Grossman-Stiglitz style regressions, a new econometric approach is needed as well. This paper proposes a way to measure the effect of passive ownership on price informativeness. Rather than doing something totally new, however, I leverage a technique that originated in the late 1960's: Examine trading volume and returns around dates where we know firm-specific information is released – quarterly and annual earnings announcements.

Using trading volume as a measure of pre-earnings price informativeness is motivated by the literature on adverse selection (see e.g. Admati and Pfleiderer (1988) and Wang (1994)). If prices are less informative before earnings announcements, uninformed agents should be more concerned about getting a bad deal, and may delay trading until information is revealed. In a companion paper, Sammon (2020), I discuss an alternative mechanism that links trading volume to price informativeness: As agents devote less attention to learning about a specific stock, there are fewer differences of opinion, which leads to decreased trading.

Use of the pre-earnings drift is motivated by the literature on market efficiency (see e.g. Fama, Fisher, Jensen, and Roll (1969)). If less information is being gathered before an earnings announcement date, the earnings news may be more of a surprise, relative to expectations. This would be reflected in a smaller pre-earnings drift, and a relatively larger move on the earnings announcement day itself (see e.g. Foster, Olsen, and Shevlin (1984)). This arises endogenously in Sammon (2020): when agents devote less attention to learning about a specific stock, there is a decrease in pre-earnings drift and increase in earnings-day volatility.

Figure 1 shows the dynamics of cumulative abnormal returns (left panel) and abnormal trading volume (right panel) around earnings announcements. The sample is firms in the top decile of standardized unexpected earnings (SUE) i.e., firms that had the best earnings news. In the early 1990's (solid blue line), prices trend up significantly before the good news is released, and there is no slow-down in trading. The return on the earnings day itself is small, relative to the run-up over the previous 30 days. The originators of these techniques would likely argue that these patterns are evidence of informed traders getting fundamental information into prices before it is formally announced.

Compare these patterns to what we see after 2010 (dashed red line): The pre-earnings drift is smaller, and the move on earnings days is larger, relative to the pre-earnings drift. Further, investors are trading less in the weeks before the announcement and trading heavily after the information is made public. From this comparison, it appears that prices were more informative before earnings announcements in the early 1990's, when passive ownership was negligible, than they are now, when passive ownership is large.

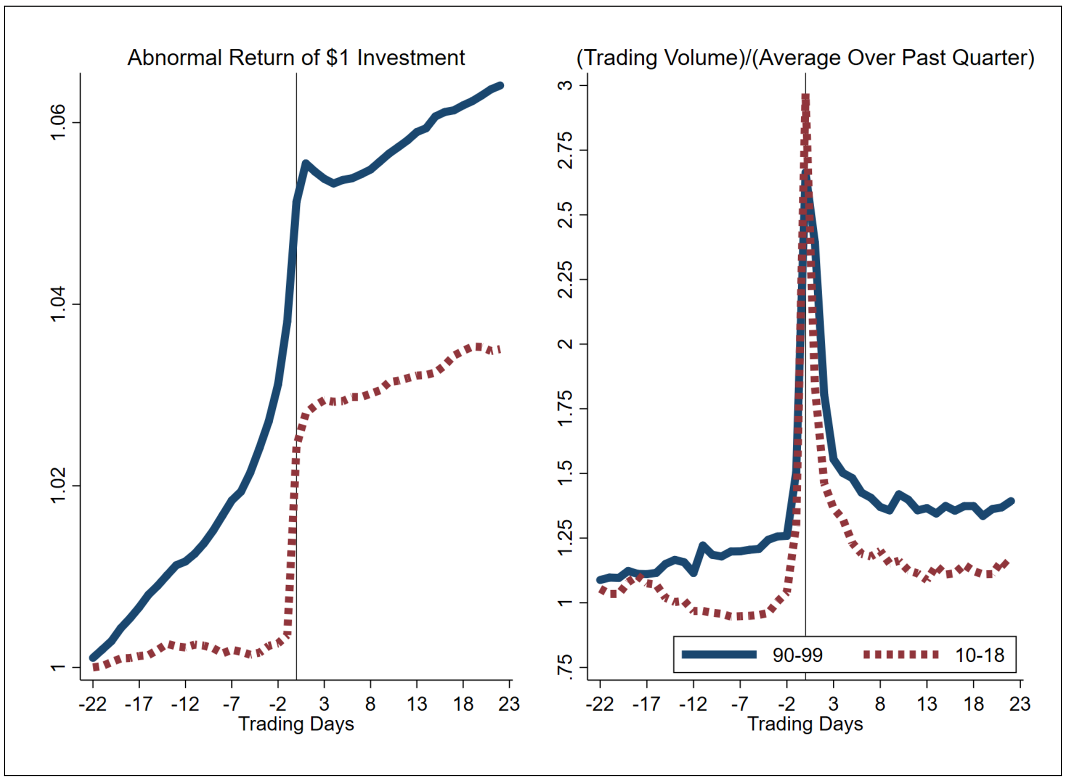

Figure 1. Returns and Trading Volume around Earnings Announcements. Each plot represents the cross-sectional average for firms in the top decile of Standardized Unexpected Earnings (SUE). SUE is defined as: $SUE_{i,t} = \frac{Earnings_{i,t} - Earnings_{i,t-4}}{\sigma_{(t-1,t-8)}(Earnings_{i,t} - Earnings_{i,t-4})}$. Deciles of SUE are calculated each quarter. Abnormal returns are returns minus the returns on the CRSP value-weighted index. Abnormal volume is volume divided by the firm-level average over the past quarter.

In cross-sectional regressions, I find that increases in passive ownership are associated with a drop in trading volume before earnings announcement dates. Passive ownership is also correlated with decreased pre-earnings drift, and increased in volatility on earnings days, further evidence of less informative prices. All of these specifications include year/quarter fixed-effects, so they are effectively comparing firms with different increases in passive ownership at the same point in time. This alleviates concerns that the results are driven by concurrent trends, or regime shifts in financial markets, such as changes in pre-earnings information leakages and the enforcement of insider trading laws.

Another concern is that passive ownership increased the most in stocks whose prices became the least informative for other reasons. To rule out reverse causality, I replicate my baseline regressions using only quasi-exogenous increases in passive ownership that arise from S&P 500 index additions and Russell 1000/2000 index rebalancing. All of the baseline results are qualitatively unchanged in these better-identified settings. These regressions include month-of-index-addition fixed effects, further ruling out the possibility that my results are driven by simultaneous trends or regime shifts.

Looking at price informativeness through the lens of earnings announcements does not rely on the precise definitions of any model objects. This allows me to provide clear evidence that the rise of passive ownership has made prices less informative about firm-specific fundamentals, consistent with the concerns raised at the beginning of this paper.

Paper Outline. I begin Section II by defining measures of price informativeness using returns and trading volume around earnings announcements. I show that between 1990 and 2018, average pre-earnings trading volume declined, as did the average pre-earnings drift. In Section III, I link these trends to the increase in passive ownership through cross-sectional regressions. In Section IV, I use S&P 500 index additions and Russell 1000/2000 index rebalancing to identify increases in passive ownership which are plausibly uncorrelated with firm fundamentals. My measures of price informativeness also decrease for firms with these quasi-exogenous increases in passive ownership. Finally, in Section V, I provide direct evidence on decreased information gathering: stocks with high passive ownership have less accurate sell-side analysts, decreased downloads of SEC filings, and larger responses to the same amount of fundamental news.

Related Literature. This paper builds on two strands of related literature. First, there is a large theoretical literature on the effects of passive ownership on price informativeness: Many papers predict that rising passive ownership should decrease price informativeness e.g. Bhattacharya and O'Hara (2018), Malikov (2018), Garleanu and Pedersen (2018), Kacperczyk, Nosal, and Sundaresan (2018). Other models, such as Cong and Xu (2016), predict that introducing ETFs reduces market frictions, and thus can increase price informativeness. My paper provides an alternative way to measure the effect of passive ownership on price informativeness that does not rely on any precise model definitions.

I also contribute to the empirical literature on the effects of passive ownership on market efficiency: Ben-David, Franzoni, and Moussawi (2018) and Chinco and Fos (2019) show that ETFs can increase non-fundamental volatility, and Israeli, Lee, and Sridharan (2017) argue that ETFs make prices less informative. On the other hand, Glosten, Nallareddy, and Zou (2016) show that ETFs can increase the systematic information content of prices. I contribute to this literature by providing an identified way, via earnings announcements, of measuring the effect of passive ownership on price informativeness.

II. Measures of Pre-Earnings Price Informativeness

In this section, I define three model-free measures of price informativeness based on trading volume, returns and volatility around earnings announcements. The cross-sectional average of all three measures has declined over the past 30 years.

A. Pre-Earnings Volume

In the presence of asymmetric information, uninformed agents are less willing to trade (Akerlof (1978)). They are concerned that the only people willing to trade with them are better informed, so any trades they make are guaranteed to be bad deals. In the stock market, an uninformed agent may prefer to delay trading until uncertainty is resolved (see e.g., Admati and Pfleiderer (1988), Wang (1994)). I define an empirical measure that captures this intuition: If prices are uninformative before earnings announcements, we would expect a decline in trading volume until after the earnings announcement date.

Let t denote an earnings announcement date, identified in I/B/E/S (IBES) data, or the next trading date if earnings are announced when markets are closed. Define abnormal volume for firm i, from time $t-22$ to $t+22$ as:

$$AV_{i,t+\tau} = \frac{V_{i,t+\tau}}{\overline{V_{i,t-22}}} = \frac{V_{i,t+\tau}}{\sum\limits_{k=1}^{63} V_{i,t-22-k}/63} \tag{2}$$

Where abnormal volume, $AV_{i,t+\tau}$, is volume divided by the historical average volume for that firm over the past quarter[4] 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:

$$AV_{i,t+\tau} = \alpha + \sum_{k=-21}^{22} \beta_k \mathbf{1}_{\{\tau=k\}} + e_{i,t+\tau} \tag{3}$$

The right-hand side variables of interest are a set of indicators for days relative to the earnings announcement. For example, $\mathbf{1}_{\{\tau=-15\}}$ is equal to one 15 trading days before the nearest earnings announcement, and zero otherwise. The regression includes all firms that can be matched between CRSP and IBES, and a ± 22 day window around each earnings announcement[5].

I run this regression for two sample periods: (1) 1990-1999 (2) 2010-2018 . Figure 2 plots the estimates of β_k for $k=-21$ to $k=-2$. For each day, the average abnormal volume is statistically significantly lower in the latter period than the former. Not pictured here, but as shown in Figure 1, abnormal volume is higher in the 2010-2018 period on the earnings day itself.

[4] All results are robust to instead using the average volume for that firm over the past year.

[5] Unless otherwise noted, all regressions and figures are equal-weighted. All results are robust to using value-weights. Value-weighted version of tables/figures can be found in the Appendix.

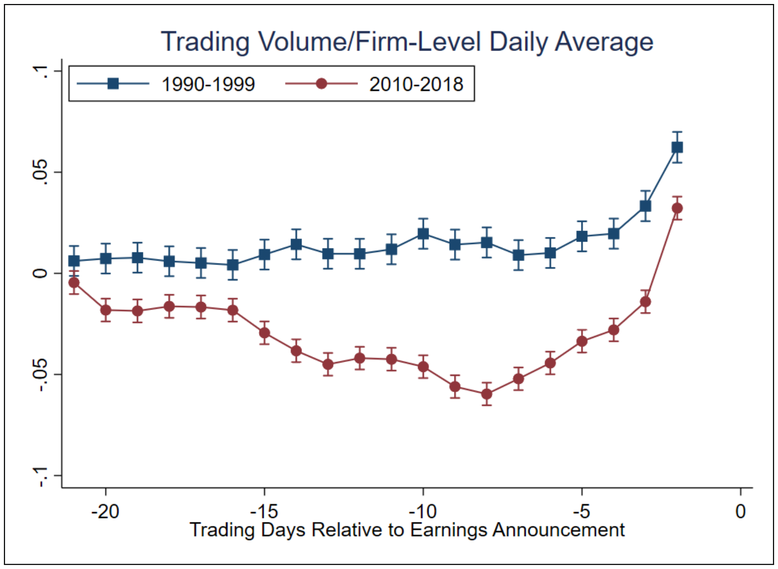

Figure 2. Decline of Pre-Earnings Volume. Plot of β_k estimated from the regression:

$$AV_{i,t+\tau} = \alpha + \sum_{k=-21}^{22} \beta_k \mathbf{1}_{\{\tau=k\}} + e_{i,t+\tau}$$

Bars represent a 95% confidence interval around the point estimate. Standard errors are clustered at the firm level.

Define cumulative abnormal pre-earnings volume as:

$$CAV_{i,t} = \sum_{\tau=-22}^{-1} AV_{i,t+\tau} \tag{4}$$

the sum of abnormal trading volume from $t-22$ to $t-1$ for firm i around earnings date t. $CAV_{i,t}$ is one of my main measures of price informativeness. Lower values of $CAV_{i,t}$ translate to less pre-earnings trading: evidence of adverse selection and less informative prices. Between the 1990's and 2010's, average $CAV_{i,t}$ declined by 2.17. This can be interpreted as as a loss of about 2 trading-days worth of volume over the 22-day window before earnings announcements.

B. Pre-Earnings Drift

The pre-earnings drift i.e., the fact that firms with strong (weak) earnings tend to have positive (negative) pre-earnings returns has been studied extensively (see e.g., Ball and Brown (1968), Foster et al. (1984)). If agents are trading on signals of good news before earnings are released, we expect prices to increase before the earnings announcement date. Prices may not move all at once, however, as informed agents want to avoid moving the market against them, as in Kyle (1985). I define an empirical measure that captures this intuition: If fewer agents are gathering information, and getting this information into prices, we would expect the pre-earnings drift to get smaller. For firms with good news, the run-up before earnings dates should get smaller, and the return on the earnings date itself should get larger.

Let $E_{i,t}$ denote earnings per share for firm i in quarter t in the IBES Unadjusted Detail File[6]. Define standardized unexpected earnings (SUE) as the year-over-year (YOY) change in earnings, divided by the standard deviation of YOY changes in earnings over the past 8 quarters.

$$SUE_{i,t} = \frac{E_{i,t} - E_{i,t-4}}{\sigma_{(t-1,t-8)}(E_{i,t} - E_{i,t-4})} \tag{5}$$

Define market-adjusted returns, $r_{i,t}$, as in Campbell, Lettau, Malkiel, and Xu (2001): 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 cumulative market-adjusted returns over the 22 trading 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 2010-2018. 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 2018, firms in each decile move less before earnings days than between 2001 and 2007. The decline in pre-earnings drift is even stronger when comparing to the pre-2001 period, but that may be due to Regulation Fair Disclosure (Reg FD), implemented in August, 2000, which limited firms' ability to selectively disclose earnings information before it was publicly announced.

[6]All results are robust to using earnings per share in Compustat. I work with IBES earnings to be consistent with analyst estimates, which I discuss in Section V

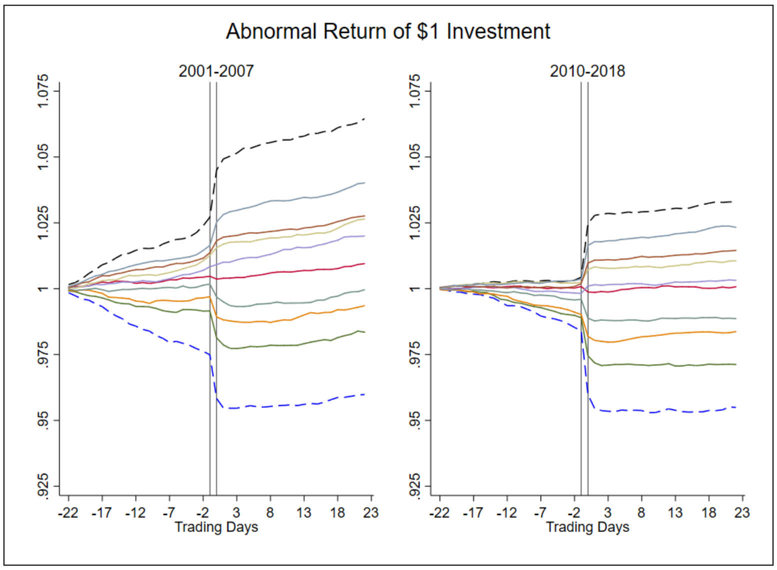

Figure 3. Decline of Pre-Earnings Drift by SUE Decile. Each quarter, I sort firms into deciles on:

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

Each line represents the cross-sectional average market-adjusted return of $1 invested at t=-22. 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. The solid lines represent the averages for deciles 2 to 9.

Figure 3's apparent decline in the pre-earnings drift could be driven by differences in overall return volatility or average returns between the two time periods. I create a drift magnitude variable designed to capture the share of earnings information 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 gross return from $t-22$ to $t-1$, divided by the cumulative returns from $t-22$ to t:

$$DM_{it} = \begin{cases} \frac{1+r_{(t-22,t-1)}}{1+r_{(t-22,t)}} & \text{if } r_t > 0 \\ \frac{1+r_{(t-22,t)}}{1+r_{(t-22,t-1)}} & \text{if } r_t < 0 \end{cases} \tag{6}$$

The pre-earnings drift will be near one when the earnings day move is small relative to cumulative pre-earnings returns. $DM_{i,t}$ will be less than one when the earning-day return is large, relative to the returns over the previous 22 days. If r_t is negative, this relationship would be reversed, which is why the measure is inverted when r_t is less than zero.[7] I work with gross returns, rather than net returns, to avoid dividing by numbers near zero[8]

$DM_{i,t}$ is going to be my second main measure of pre-earnings price informativeness. Lower values of $DM_{i,t}$ imply less information is getting into prices before earnings announcement dates. Figure 4 shows the cross-sectional average value of $DM_{i,t}$ by year. The pre-earnings drift decreased by about -0.02 between 1990 and 2018.

C. *Earnings Day Volatility*

The last two subsections showed there is less trading before earnings announcements, and the pre-earnings drift declined. If the total amount of information is not changing over time, we would expect there to be (1) relatively larger returns on earnings days and (2) relatively more trading volume on earnings days. I propose an empirical measure for testing this hypothesis: the share of total annual volatility occurring on earnings dates.

[7]Section A.A of the Appendix presents an alternative definition of the pre-earnings drift using squared returns and further motivates my specification for $DM_{i,t}$.

[8]It has been well documented (see e.g. McLean and Pontiff (2016).) that the post-earnings drift has declined. To ensure my results are not driven by this trend, I calculate alternative measures of the pre-earnings drift replacing $1+r_{(t-22,t)}$ with $1+r_{(t-22,t+n)}$ for n between 1 and 5. All my results are qualitatively and quantitative unchanged using these alternative pre-earnings drift measures.

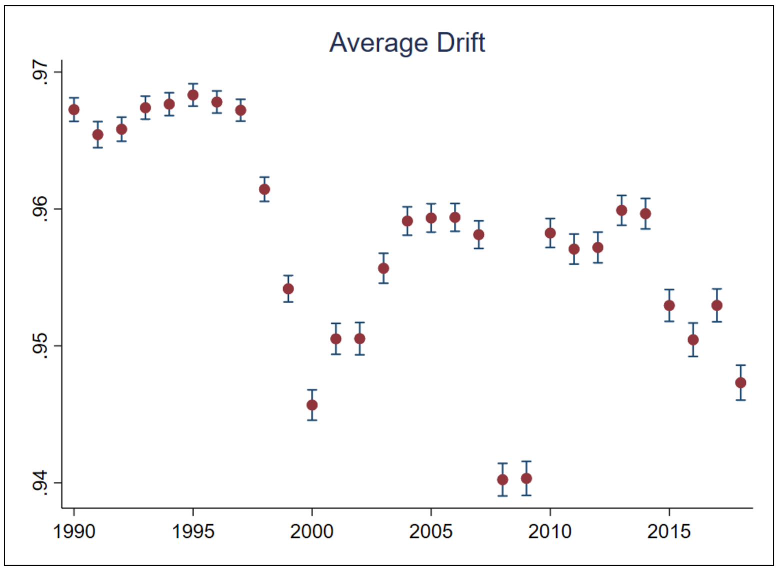

Figure 4. Decline of Average Pre-Earnings Drift. This figure plots coefficients from a regression of pre-earnings drift on a set of year dummy variables where $DM_{it} = \begin{cases} \frac{1+r_{(t-22,t-1)}}{1+r_{(t-22,t)}} & \text{if } r_t > 0 \\ \frac{1+r_{(t-22,t)}}{1+r_{(t-22,t-1)}} & \text{if } r_t < 0 \end{cases}$. A value near 1 implies most earnings information is incorporated in prices before the announcement date, while lower values denote less informative pre-earnings announcement prices. Standard errors represent 95% confidence intervals around the point-estimates. Standard errors are clustered at the firm level.

Specifically, define the quadratic variation share (QVS) for firm i in year t as:

$$QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2 \tag{7}$$

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. QVS is going to be my third main measure of price informativeness. If relatively more information is being learned and incorporated into prices on earnings dates, we would expect larger values of QVS.

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 IBES. Average QVS increased from 3.0% in 1990 to almost 16% in 2018.

An analogue of QVS to trading volume would be the share of total abnormal trading volume that occurs on earnings days. Figure 1 suggests that abnormal volume on earnings days has increased significantly between the 1990's and 2010's. While it is not the focus of this paper, it is consistent with the intuition at the start of this subsection.

D. Discussion

These downward trends in market efficiency could be unrelated to the information released on earnings days. To rule this out, I propose the following placebo test: choose the date 22 trading days before each earnings announcement to be a placebo earnings date. I then reconstruct the time-series averages of the pre-earnings volume, drift and share of volatility on these placebo earnings days. In Section A.B of the Appendix, Figure 9 shows that there is no drop in volume before the placebo earnings dates in the last third of the sample. 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 figures are similar if you use randomly selected dates as placebo earnings announcements. These results confirm that the changes in price informativeness are specific to earnings days.

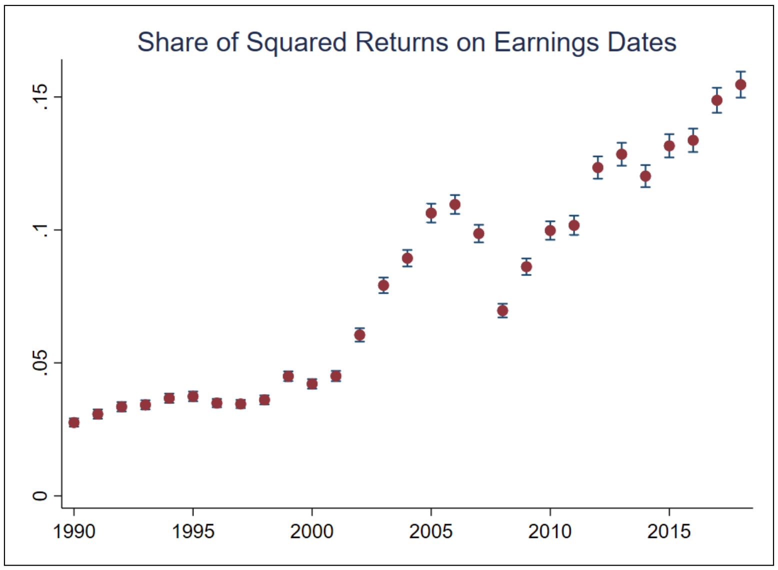

Figure 5. Increase in Earnings Day Volatility. This figure plots coefficients from a regression of QVS on a set of year dummy variables. For firm i in year t the quadratic variation share (QVS) is defined as: $QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, where r denotes a market-adjusted daily return. The numerator sums over the 4 quarterly earnings days in year t, while the denominator includes all days in calendar year t. Standard errors are clustered at the firm level.

As an additional check, Section A.G of the Appendix examines volume, drift and volatility around Federal Open Market Committee (FOMC) meeting dates instead of placebo earnings dates. The growth of index funds and ETFs has made it easier to trade on systematic information. It is possible that investors now focus on gathering information about systematic risks, so stock prices are more informative about systematic news, at the expense of firm-specific news. I find no trend toward decreased, or increased, efficiency in the incorporation FOMC meeting information at the stock-level. This also confirms that the reduction in efficiency only applies to firm-specific information.

III. Reduced-Form Evidence

In this section, I show the reduced-form relationships 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. I also address competing hypotheses for decreased pre-earnings price informativeness, including the rise of algorithmic trading and Regulation Fair Disclosure.

A. Data and Definitions

Passive ownership is defined as all index funds, all ETFs, and all mutual funds with "index" in the name[9]. Index funds are identified using the index fund flag in the CRSP mutual fund data. 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. If a security never appears in the S12 data, I assume the passive ownership share is zero. Figure 6 shows that passive ownership increased from almost zero in 1990, to over 40% of total mutual fund and ETF assets in 2018. Further, passive funds now own more than 14% of the total market capitalization of US ordinary common shares.

I believe this is a conservative definition of passive ownership, as there are institutional investors which track broad market indices, but are not classified as mutual funds, and do not appear in the S12 data. Further, as discussed in Mauboussin, Callahan, and Majd (2017),

[9] All results are robust to including only ETFs in passive ownership.

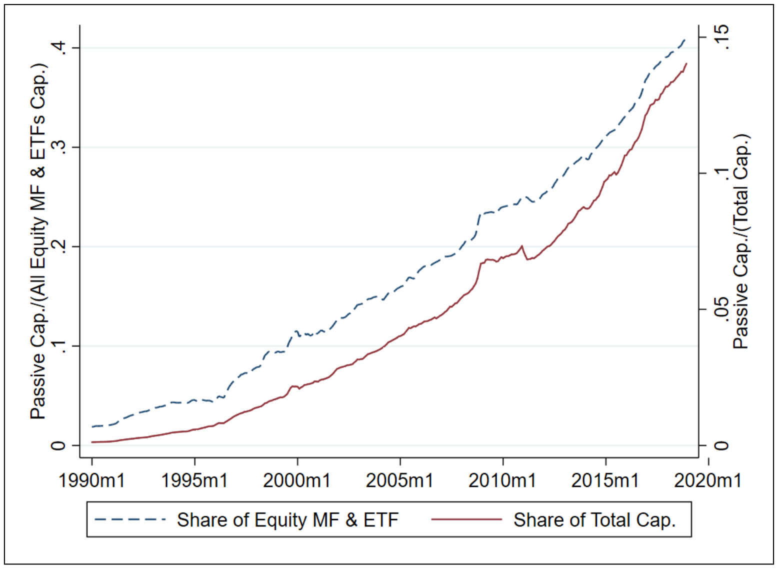

Figure 6. The Rise of Passive Ownership: 1990-2018. Passive ownership is defined as all index funds, all ETFs, and all mutual funds with "index" in the name. Index funds are identified using the index fund flag in the CRSP mutual fund data. Total equity mutual fund and ETF assets is the sum of all stock holdings in the Thompson S12 data that can be matched to CRSP. Total market capitalization includes all CRSP firms.

there has been a rise of closet indexing among self-proclaimed active managers, which is also omitted in my definition of passive management.

All return and daily volume data are from CRSP. I merge CRSP to I/B/E/S (IBES) using the WRDS linking suite. 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.

I define quarterly earnings per share as the "value" variable from the IBES unadjusted detail file.[10] All other firm fundamental information is from Compustat.

Total institutional ownership is the sum of shares held by all 13-F filing institutions.

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

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.

B. Pre-Earnings Volume

I run the following regression with quarterly data to measure the relationship between declines in pre-earnings volume and increases in passive ownership:

$$\Delta CAV_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t} \tag{8}$$

where cumulative abnormal pre-earnings volume from $t-1$ to $t-22$, $CAV_{i,t}$, is defined in Equation 4. Δ is a year-over-year change, matching on fiscal quarter. 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 $X_{i,t-1}$ include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. I also condition on the growth in market capitalization from $t-1$ to t. 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 prevent a firm-size effect driving my results. I also include firm and year/quarter fixed-effects. The time fixed effects ensure I am comparing firms at the same point in time with different increases in passive ownership. This allays concerns that my results are driven by simultaneous trends in passive ownership and price informativeness. Standard errors are computed using panel Newey-West with 8 lags[11].

Table I contains the regression results. To interpret the magnitude of the reduced-form coefficients, I consider the effect of a 10% increase in passive ownership for the average stock. The coefficients on $\Delta Passive_{i,t}$ in column 2 implies that a 10% increase in passive ownership would lead to a decline in cumulative abnormal pre-earnings volume of -1.6. This is large relative to the average decline from the 1990's to 2010's of -2.2 discussed in Section II.

[11] All results are robust to double-clustering at the firm/year level.

	(1)	(2)	(3)
Inc. Passive	-12.81***	-16.09***	-23.96***
	(1.986)	(2.487)	(5.416)
Observations	239,859	239,859	239,859
R-squared	0.022	0.04	0.112
Controls	No	Yes	Yes
Firm FE	No	Yes	Yes
Weight	Eq.	Eq.	Val.

Table I Passive Ownership and Pre-Earnings Volume. Estimates of β from:

$$\Delta CAV_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t}$$

$CAV_{i,t}$ is cumulative abnormal pre-earnings trading volume. Δ is a year-over-year change, matching on fiscal quarter. Change in passive ownership is expressed as a decimal, so $0.01 = 1\%$ increase. Controls, $X_{i,t-1}$, include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, total institutional ownership and growth in market capitalization from $t-1$ to t. Standard errors are computed using panel Newey-West with 8 lags. Column 1 is a univariate regression, while column 2 includes the controls, as well as year/quarter and firm fixed effects. Column 3 is the same as column 2, except firms are weighted by their 1-quarter lagged market capitalization relative to other firms that quarter.

	(1)	(2)	(3)
Inc. Passive	-0.0298**	-0.0322**	-0.0965***
	(0.012)	(0.014)	(0.029)
Observations	239,689	239,689	239,689
R-squared	0.02	0.045	0.063
Controls	No	Yes	Yes
Firm FE	No	Yes	Yes
Weight	Eq.	Eq.	Val.

Table II Passive Ownership and Pre-Earnings Drift. Table with estimates of β from:

$$\Delta DM_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t}$$

Where $DM_{i,t}$ is a measure of the pre-earnings drift. Passive ownership is expressed as a decimal, so 0.01 = 1% of shares outstanding held by passive funds. Controls, $X_{i,t-1}$, include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, total institutional ownership and growth in market capitalization from $t-1$ to t. Standard errors are computed using panel Newey-West with 8 lags. Column 1 is a univariate regression, while column 2 includes the controls, as well as year/quarter and firm fixed effects. Column 3 is the same as column 2, except firms are weighted by their 1-quarter lagged market capitalization relative to other firms that quarter.

C. Pre-Earnings Drift

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

$$\Delta DM_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t} \tag{9}$$

where $DM_{i,t}$ is defined as in Equation 6. Controls in $X_{i,t-1}$ include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. I also condition on the growth in market capitalization from $t-1$ to t. Fixed effects are year/quarter and firm. Standard errors are computing using panel Newey-West with 8 lags.

The regression results are in Table II. The coefficient on $\Delta Passive_{i,t}$ in column 2 implies that a 10% increase in passive ownership would decrease the pre-earnings drift by -0.003. This is about 15% of the average decline of -0.02 between the 1990's to 2010's discussed in Section II.

D. *Share of Volatility on Earnings Days*

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:

$$\Delta QVS_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t} \tag{10}$$

where QVS is defined in Equation 7. Controls in $X_{i,t-1}$ include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. I also condition on the growth in market capitalization from $t-1$ to t. Fixed effects are year/quarter and firm. Standard errors are computed using panel Newey-West with 2 lags[12]

The regression results are in Table III. Figure 5 shows that average QVS increased by about 0.12 between 1990 and 2018. The coefficients on $\Delta Passive_{i,t}$ in column 2 implies that a 10% increase in passive ownership would lead to an increase in QVS of about 0.01.

E. *Placebo Tests*

To confirm that my results are specific to earnings days, I perform two placebo tests. The first set of placebo earnings dates are 22 trading days before each earnings announcement. 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. The second are all scheduled FOMC meetings. Appendix Table VIII compares the original regression results to the placebo results in the specifications with all controls and firm-fixed effects. I focus on the earnings-day share of volatility in these tests, as I believe it offers the cleanest comparison. For the FOMC announcement dates, looking at year-over-year changes for the n^{th} annual announcement does not make much sense, as there is (1) no analogue of fiscal year to account for seasonality and (2) they do not occur at the same time every year. The first point also applies to the days between earnings announcements. All of the placebo results are insignificant, confirming that the relationship between passive ownership and volatility are all specific to earnings days. As an additional check, I randomly assign one day for each firm in each quarter to be a

[12] I use two lags, instead of 8 as in the previous two sub-sections, because this regression is run at the annual frequency.

	(1)	(2)	(3)
Inc. Passive	0.200***	0.106***	0.381**
	(0.031)	(0.036)	(0.180)
Observations	127,951	126,319	126,319
R-squared	0.011	0.03	0.035
Controls	No	Yes	Yes
Firm FE	No	Yes	Yes
Weight	Eq.	Eq.	Val.

Table III Passive Ownership and Earnings Day Share of Volatility. Table with estimates of β from:

$$\Delta QVS_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t}$$

$QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, 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. Controls in $X_{i,t-1}$ include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. I also condition on the growth in market capitalization from $t-1$ to t. Fixed effects are year and firm. Standard errors are computed using panel Newey-West with 2 lags. Column 1 is a univariate regression, while column 2 includes the controls, as well as year and firm fixed effects. Column 3 is the same as column 2, except firms are weighted by their 1-year lagged market capitalization relative to other firms that year.

placebo earnings day. This alternative placebo test also yields insignificant coefficients on $\Delta Passive$.

F. Addressing Competing Hypotheses

This subsection discusses alternative explanations for my findings on decreased market efficiency, and its correlation with passive ownership. Two major threats to identification are (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)). Section A.I discusses possible omitted variables in the baseline regressions. It is not possible to discuss every alternative hypothesis, so outside of explicitly testing alternatives, I rely on the quasi-exogenous variation in passive ownership from index addition/rebalancing in the next section to overcome any remaining identification concerns.

F.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 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, I 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.

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 account for any level shifts in price informativeness before/after Reg FD was passed. To further rule out this channel, the appendix contains versions of all the reduced-form regressions using only post-2000 data in Tables XIII, XIV and XV. 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.

F.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 algorithmic traders 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.

It is difficult to measure the role of algorithmic traders in the trends toward decreased pre-earnings price informativeness as I cannot directly observe AT activity, and only have reasonable AT activity measures between 2012-2018. 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. This, however, should not affect my results, as I condition on firm size in all the reduced-form regressions (2) High ETF ownership will attract algorithmic traders 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 2018. 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 (1) Truncate each of the AT activity variables at the 1% and 99% level by year to minimize the effect of reporting errors, (2) calculate a moving average for each of these measures in the 21 days leading up to each earnings announcement, and (3) 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 almost entirely the result of the year restrictions.

I re-run all the reduced form regressions, but restrict to the matched sample with MIDAS, and include the 4 measures of HFT activity. As a sanity check, I first re-run the baseline 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".

Tables IX, X and XI contain the regressions with AT controls. All the results in this matched sub-sample are qualitatively unchanged from Tables I, II and III. For the pre-earnings drift, and earnings day share of volatility, adding the AT activity controls does not significantly change the coefficient on change in passive ownership. For volume, the value-weighted results are robust to including the AT controls, while the equal weighted result is the right sign, but insignificant. This could imply (1) The equal-weighted volume result has become weaker over time, as the coefficient in column 2 of Table IX is 1/3 the size of the same coefficient in column 2 of Table I. Given that I can only include the AT controls in the matched subsample, it is hard to fully disentangle this effect (2) Part of the AT activity measure is mechanically correlated with passive ownership because, for example, ETFs attract algorithmic traders implementing ETF arbitrage. I show AT activity and passive ownership are positively correlated in Table XII (3) 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.

IV. Robustness to Quasi-Exogenous Variation in Passive Ownership

In this section, I exploit S&P 500 index additions, as well as Russell 1000/2000 reconstitutions to identify increases in passive ownership which are plausibly uncorrelated with firm fundamentals. These allow me to causally link increases in passive ownership and decreases in pre-earnings price informativeness.

A. 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 and industry selection criteria, I identify a group of control firms that reasonably could have been added to the index at the same time as the treated firms. At the time of index addition, I sort firms into two-digit SIC industries. Then, within each industry, I identify the 10 firms with the closest market capitalization to the firm that was eventually added. To be included in the final sample, control and added firms have non missing data in the two years before and after index addition. Also, the control firms must not be added to the index over the next two years. After applying all these filters, there are usually 2-3 control firms for each treated firm.

I then identify a second control group among firms that are already in the S&P 500 index. Within each 2-digit SIC industry industry, I identify the 10 firms with the closest market capitalization to the firm that was eventually added. I also require that these control firms do not leave the S&P 500 index over the next two years. After applying this filter, and the non-missing data filter, there is usually at least one 1 control firm for each treated firm.

In the regressions, I use index addition as an instrument for the increase in passive ownership. The first stage is:

$$\Delta Passive_{i,t} = \alpha + \beta \times Treated_{i,t} + \gamma_t + \epsilon_{i,t} \tag{11}$$

where γ_t is a month of index addition fixed effect. Including these fixed effects ensures I am only comparing treatment and control firms at the same point in time. The second stage is:

$$\Delta Outcome_{i,t} = \alpha + \beta \times \Delta \widehat{Passive}_{i,t} + \gamma_t + \epsilon_{i,t} \tag{12}$$

Where $Outcome_{i,t}$ is the average pre-earnings volume, drift, or earnings day share of volatility in the two years before or after index addition. I exclude the quarter of index addition, and

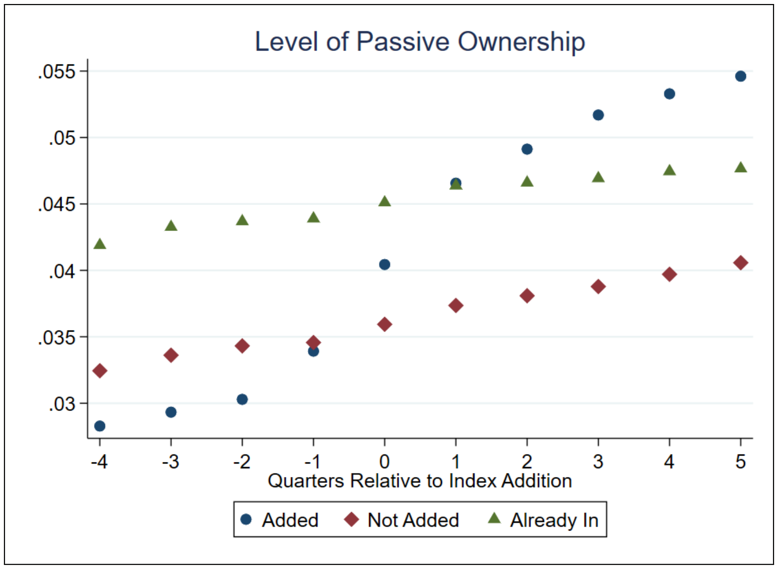

Figure 7. S&P 500 Index Addition and Changes in Passive Ownership. Average level of passive ownership for control firms out of the index ("Not Added"), control firms in the index ("Already In") and treated firms ("Added").

the quarter after index addition when computing these averages. This is to avoid index inclusion effects (see e.g. Morck and Yang (2001)), and to ensure that measures based on past averages, like abnormal volume, do not include any of the data from the pre index-addition period.

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

Figure 7 shows the level of passive ownership for the control firms and treated firms around the time of index addition. Both groups of firms have similar average pre-addition changes in passive ownership, although the firms already in the index have a higher average level of passive ownership.

Table IV contains the regression results. For comparison, I included a row with the reduced form estimates, which correspond to the equal weighted specification with all controls

	Volume	Drift	Volatility
Inc. Passive	-51.08**	-0.322**	1.924**
	(22.550)	(0.140)	(0.768)
R-squared	0.098	0.074	0.115
Reduced Form	-23.96***	-0.0965***	0.381**
Treated firms	419	419	419
Control (out)	906	906	906
Control (in)	508	508	508

Table IV Effects of S&P 500 Index Addition. Two-stage least squares estimates from first stage:

$$\Delta Passive_{i,t} = \alpha + \beta \times Treated_{i,t} + \gamma_t + \epsilon_{i,t}$$

Second stage:

$$\Delta Outcome_{i,t} = \alpha + \beta \times \Delta \widehat{Passive}_{i,t} + \gamma_t + \epsilon_{i,t}$$

Where $Treated_{i,t}$ is a dummy variable equal to one if a firm was added to the S&P 500 index, and γ_t is a month of index addition fixed effect.

and fixed effects estimated in Section III. For all three regressions, the results have the same sign as the reduced-form regressions. The estimated coefficients, however, are substantially larger than the reduced-form results. I believe the first stage 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 that are dropped from the S&P 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 A.F of the Appendix has more details on the effect of index deletion.

B. 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 firms are not added at random. Once 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 were in the Russell 1000 for at least two years, and then switched from the Russell 1000 to the Russell 2000. To be included in the regressions, treated firms must stay in the Russell 2000 for at least two years. The control firms have June ranks between 900 and 1000 at the time the treated firms are identified. The control firms must have been in the Russell 1000 for the past two years, and must stay there for the next two years, although they are allowed to have a rank lower than 900 in the pre and post periods. This classification involves a look-ahead bias, as I am using the ex-post changes in Russell index membership to identify changes in passive ownership. This method, however, avoids the issue discussed in Wei and Young (2017) with the instrumental variables approach applied in Appel et al. (2016), among others.

Figure 8 compares the level of passive ownership around the index rebalancing date between the treated and control group. The pre-addition changes and level are similar across both groups.

For the Russell experiment, I use the same two-stage-least-squares structure as I did for

[13]This rule changed in 2006 – to reduce turnover between the two indices, Russell now has a bandwidth rule: As long as the firm's market capitalization is within 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.

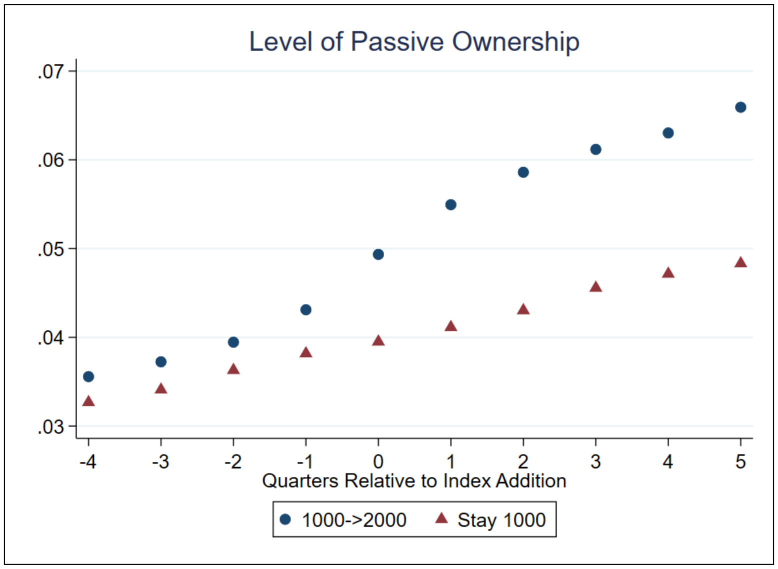

Figure 8. Russell 1000/2000 Reconstitution and Changes in Passive Ownership. Average level and increase in passive ownership for control firms and firms moved from the Russell 1000 to the Russell 2000. Control firms are all firms in the Russell 3000 ranked 900 to 1100 that did not move from the 1000 to the 2000 or from the 2000 to the 1000.

the S&P 500 experiment. The only difference, is in the timing: I am comparing the two years before index reconstitution, ending in April, and the two years following reconstitution, starting in August, and then skipping the first quarter after index reconstitution. 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 prevent the trading associated with index rebalancing from influencing the regression estimates.

Table V contains the regression results. For comparison, I included a row with the reduced form estimates, which correspond to the equal weighted specification with all controls and fixed effects estimated in Section III.

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. 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

	Volume	Drift	Volatility
Inc. Passive	-44.71**	-0.285**	0.0109
	(20.740)	(0.125)	(0.411)
R-squared	0.099	0.126	0.073
Reduced Form	-23.96***	-0.0965***	0.381**

Table V Effects of Russell 1000/2000 Index Reconstitution. Two-stage least squares estimates from first stage:

$$\Delta Passive_{i,t} = \alpha + \beta \times Treated_{i,t} + \gamma_t + \epsilon_{i,t}$$

Second stage:

$$\Delta Outcome_{i,t} = \alpha + \beta \times \Delta \widehat{Passive}_{i,t} + \gamma_t + \epsilon_{i,t}$$

Where $Treated_{i,t}$ is a dummy variable equal to one if a firm was added to the S&P 500 index, and γ_t is a month of index addition fixed effect. There are 216 treated firms and 158 control firms.

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 passive ownership when they move from the Russell 2000 to the Russell 1000. In Section A.F of the Appendix. I show that this treatment effect is washed out by the time trend toward increased passive ownership.[14]

V. Evidence on Information Gathering

Sections III and IV show the negative relationship between passive ownership and price informativeness. In this section, I present reduced-form evidence consistent with fewer investors gathering firm-specific information for stocks with high passive ownership.

[14] 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.

A. *Investor Attention*

One potential explanation for passive ownership decreasing price informativeness is that passive managers, as well as investors in passive funds, lack strong incentives to gather and consume firm-specific information. Passive funds trade on mechanical rules, such as S&P 500 index membership (SPY), or the 100 lowest volatility stocks in the S&P 500 (SPLV). Given that these trading strategies are implemented on public signals, they do not require accurate private forecasts of firm fundamentals. 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 analyst coverage/accuracy on passive ownership:

$$\Delta Outcome_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t} + \text{Fixed Effects} + e_{i,t} \tag{13}$$

Controls in $X_{i,t-1}$ include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. I also condition on the growth in market capitalization from $t-1$ to t. Fixed effects include year/quarter and firm.

In Equation 13, 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* (3) the average time (in months) between analyst updates. For the accuracy regressions, I exclude firms with a consensus forecast of 1 cent or less (in absolute value) to minimize the effect of outliers. After excluding these observations, accuracy is Winsorized at the 1% and 99% level each year. All results are robust to normalizing by the stock price instead of the consensus forecast.

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.

Another measure of investor attention is the number of downloads of SEC filings (see e.g., Loughran and McDonald (2017)). If passive managers and investors in passive funds do not gather fundamental information, the number of downloads of SEC filings might be lower for firms with high passive ownership. To test this, I run the following regression:

$$\Delta DL_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t} + \text{Fixed Effects} + e_{i,t} \tag{14}$$

where Δ is the change from year $t-1$ to year t. $DL_{i,t}$ is the number of non-robot downloads of 10-K's, 10-Q's and 8-K's in the 22 days before earnings announcements. Robot downloads include web crawlers, index page requests and individual IPs 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 robots may automatically download all filings at release, or update a database periodically. These robot downloads do not coincide with the intuition of information gathering for immediate trading. 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, so the trend would bias me against finding any results.

Table VI contains the regression results. Consistent with decreased information gathering, increases in passive ownership are correlated with the fewer analysts covering a stock, lower analyst accuracy, less frequent analyst updates. Firms with increases in passive ownership also experience decreases in pre-earnings downloads of SEC filings.

B. 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 these stocks. When the good cashflow news arrives, the informed investors were previously underweight these stocks, so their diversification motive is weak, and they buy. In relating this model to my empirical setup, I treat 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:

$$r_{i,t} = \alpha + \beta_1 SUE_{i,t} + \phi_1 Passive_{i,t} + \gamma_1 \left(SUE_{i,t} \times Passive_{i,t}\right) + \xi X_{i,t} + \text{Fixed Effects} + e_{i,t} \tag{15}$$

	# Analysts	Distance	Time Between Updates	Downloads
Inc. Passive	-8.935***	1.557***	14.93*	-3.756***
	(0.824)	(0.244)	(8.692)	(1.185)
Observations	99,004	96,365	79,131	96,380
R-squared	0.1	0.062	0.065	0.233
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Weight	Eq.	Eq.	Eq.	Eq.

Table VI Investor Attention and Passive Ownership. This table contains the estimates of β from:

$$\Delta Outcome_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t} + \text{Fixed Effects} + e_{i,t}$$

Controls in $X_{i,t-1}$ include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. I also condition on the growth in market capitalization from $t-1$ to t. Fixed effects include year/quarter and firm.

Here, $r_{i,t}$ denotes the market-adjusted return on the effective quarterly earnings date. $SUE_{i,t} = \frac{E_{i,t}-E_{i,t-4}}{\sigma_{(t-1,t-8)}(E_{i,t}-E_{i,t-4})}$.[15] Controls in $X_{i,t}$ include 1-year lagged passive ownership, market capitalization, growth of market capitalization from $t-1$ to t, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. Fixed effects include year/quarter and firm.

I also run a version of Equation 15, breaking SUE into two variables: They are equal to SUE if SUE is positive (negative), and zero otherwise. I then interact these variables with passive ownership, to account for a possibly asymmetric effect of passive ownership on positive and negative news.

Table VII contains the regression results. Consistent with the Buffa et al. (2014) model, firms with a high share of passive ownership are more responsive to earnings news, especially

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

	(1)	(2)	(3)	(4)
SUE	0.00912***		0.00314***	
	(0.000)		(0.000)	
SUE > 0		0.00745***		0.00369***
		(0.000)		(0.000)
SUE < 0		-0.00394***		0.000128
		(0.000)		(0.001)
SUE x Passive	0.0545***		0.0435***	
	(0.003)		(0.007)	
SUE > 0 x Passive		0.0217***		0.0246***
		(0.003)		(0.006)
SUE < 0 x Passive		-0.0411***		-0.0196*
		(0.004)		(0.011)
Observations	415,961	415,961	415,961	415,961
R-squared	0.068	0.069	0.039	0.041
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Weight	Eq.	Eq.	Val.	Val.

Table VII Passive Ownership and Response to Earnings News. This table contains the results of the following regression:

$$r_{i,t} = \alpha + \beta_1 SUE_{i,t} + \phi_1 Passive_{i,t} + \gamma_1 \left(SUE_{i,t} \times Passive_{i,t}\right) + \xi X_{i,t} + \text{Fixed Effects} + e_{i,t}$$

Here, $r_{i,t}$ denotes the market-adjusted return on the effective quarterly earnings date. $SUE_{i,t} = \frac{E_{i,t}-E_{i,t-4}}{\sigma_{(t-1,t-8)}(E_{i,t}-E_{i,t-4})}$. Controls in $X_{i,t}$ include 1-year lagged passive ownership, market capitalization, growth of market capitalization from $t-1$ to t, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. Fixed effects include year/quarter and firm. Columns 1 and 2 are equal weighted, while Columns 3 and 4 are value weighted.

if that news is negative[16].

VI. Conclusion

The goal of this paper is to measure the effect of rising passive ownership on price informativeness, without relying on a model. I accomplish this by examining changes in trading volume and returns around earnings announcement dates.

Increases in passive ownership have lead to decreased pre-earnings-announcement price

[16] An alternative explanation is that passive ownership, in particular ETFs, has reduced short sale constraints, as in Palia and Sokolinski (2019). In unreported results, I calculate the shorting of individual stocks through ETFs, and I find this cannot explain the larger response of stocks with high passive ownership to negative earnings news.

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 is that passive managers, as well as investors in passive funds, lack strong incentives to gather and consume firm-specific information. Consistent with this channel, firms with increases in passive ownership experience decreases in the number of analysts covering the stock, decreases in the accuracy of the remaining analysts, and fewer downloads of SEC filings.

Relative to total institutional ownership, passive ownership is still relatively small, owning around 14% of total US market capitalization. Even at this low level, passive ownership has lead to drastic changes in trading patterns, liquidity and response to firm-specific news. As passive ownership continues to grow, these changes in information and trading may be amplified, further changing the way equity markets reflect firm-specific information.

Appendix A. Appendix

Appendix A. Pre-Earnings Drift

We might be concerned that $DM_{i,t}$ has a Cauchy-like distribution i.e., very fat tails. One way to address this is to define a pre-earnings drift measure that uses squared returns, rather than signed returns. Define the drift volatility, DV, as the ratio of the squared earnings day return to the squared return on all days since last earnings announcement plus squared earnings day return. This is similar to QVS, but is defined for each individual earnings announcement, rather than a whole year.

If DV is near one, almost all volatility occurs on earnings days, while if it is near zero, almost all volatility happens between earnings days. In a Kyle (1985)-type model, DV would decrease with the precision of the insider's signal[17]. All my baseline pre-earnings drift results are robust to switching any drift magnitude measure for DV.

Appendix B. 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 placebo earnings dates 22 days before the actual announcements. In all three cases, there is no trend toward decreased informativeness on the placebo earnings dates.

Appendix C. Additional Pre-Earnings Volume Results

Rather than look at the 22 days before an earnings announcement, I expand the analysis to 60 trading days before the earnings announcement. 60 trading days roughly corresponds to the time between earnings announcements. A concern with the regression specification in Equation 3 (the regression of cumulative abnormal volume on days-before-earnings-announcement 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 in a regression with year fixed effects. Figure 12 shows the cross-sectional median pre-earnings volume, which exhibits the same decline in pre-earnings volume as Figure 2.

[17] I thank Alex Chinco for making his two-period Kyle code available on his website. I am happy to share my version of Alex's code with these simulations upon request.

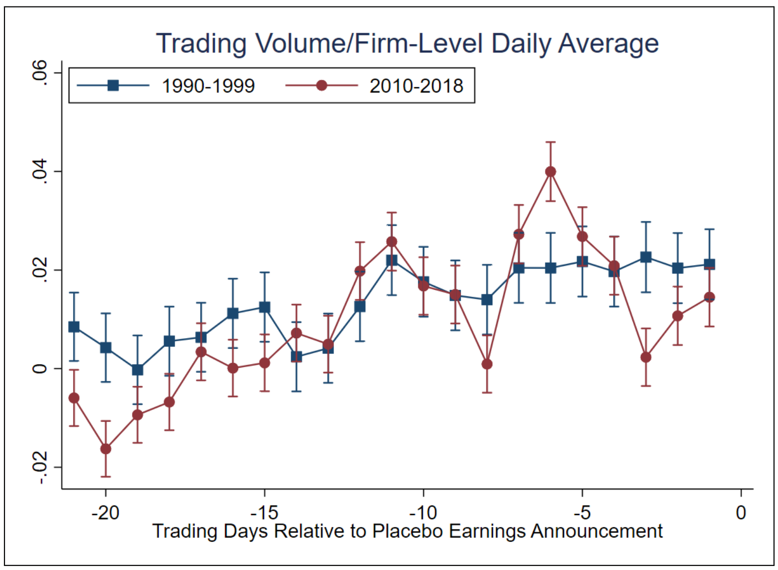

Figure 9. Placebo Test: Pre-Earnings Volume. Plot of β_k estimated from the regression:

$$AV_{i,t+\tau} = \alpha + \sum_{k=-21}^{22} \beta_k \mathbf{1}_{\{\tau=k\}} + e_{i,t+\tau}$$

where t represents a placebo earnings date, 22 days before the actual earnings announcement date. Bars represent a 95% confidence interval around the point estimate. Standard errors are clustered at the firm level.

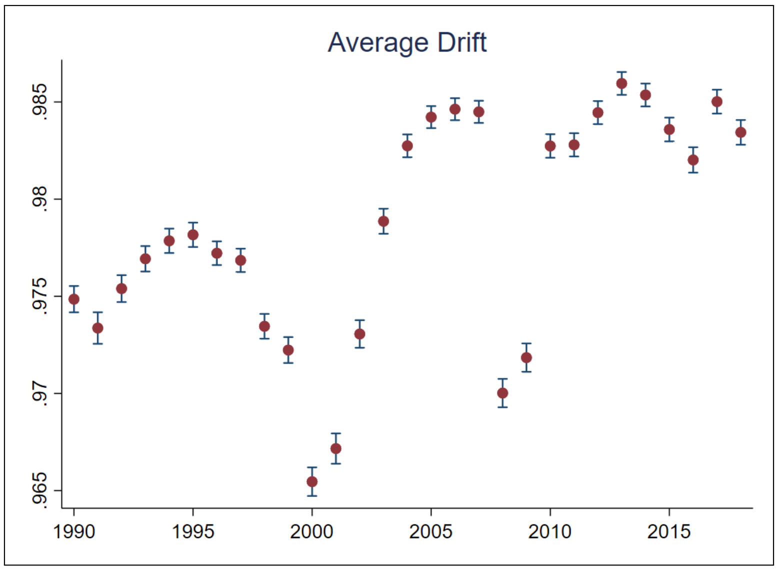

Figure 10. Placebo Test: Pre-Earnings Drift. This figure plots coefficients from a regression of pre-earnings drift on a set of year dummy variables where $DM_{it} = \begin{cases} \frac{1+r_{(t-22,t-1)}}{1+r_{(t-22,t)}} & \text{if } r_t > 0 \\ \frac{1+r_{(t-22,t)}}{1+r_{(t-22,t-1)}} & \text{if } r_t < 0 \end{cases}$. except t is a placebo earnings date 22 days before the actual earnings announcement date. Standard errors represent 95% confidence intervals around the point-estimates. Standard errors are clustered at the firm level.

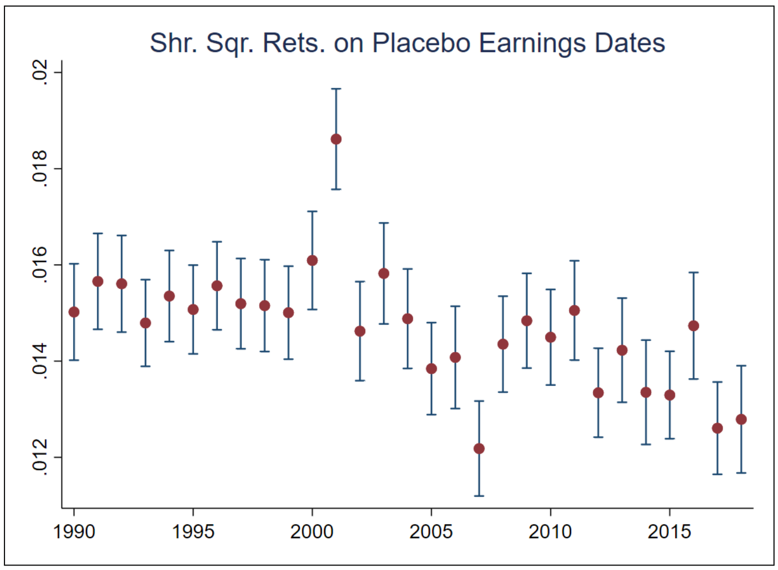

Figure 11. Placebo Test: Earnings Day Volatility. This figure plots the share of market-adjusted quadratic variation occurring on placebo earnings days. For firm i in year t the quadratic variation share (QVS) is defined as: $QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, where r denotes a market-adjusted daily return. The numerator is the sum of squared returns on the 4 placebo earnings dates, each of which are 22 days before the actual earnings announcement dates, while the denominator is the sum of squared returns on all trading days in calendar year t.

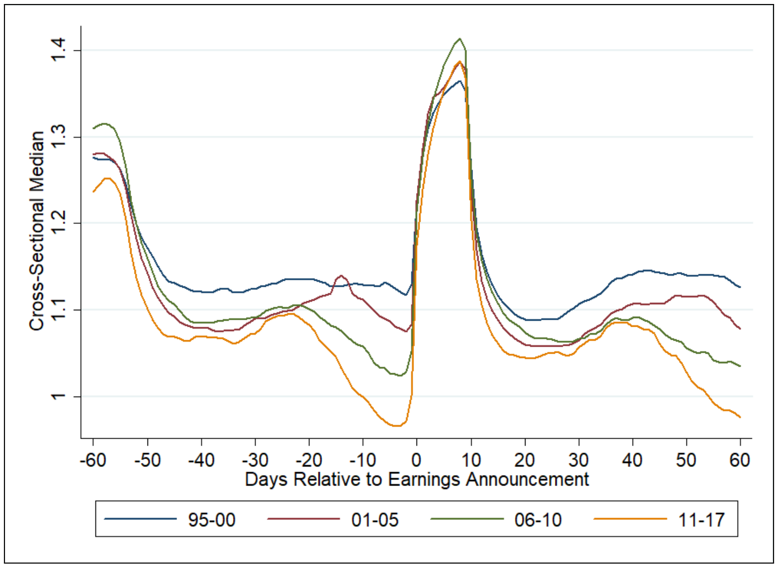

Figure 12. Decline of Pre-Earnings Volume, Expanded Window. Plot of 10-day moving average of abnormal volume. Abnormal volume is volume relative to the historical average over the past year. Average historical volume is fixed at the beginning of each 10-day moving-average window to avoid mechanically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

The figure also motivates my choice 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, either in aggregate, or for stocks with high passive ownership.

	Baseline	Placebo t=-22	Placebo FOMC	Baseline	Placebo t=-22	Placebo FOMC
Inc. Passive	0.106***	-0.00474	0.00709	0.382**	-0.0501	0.0159
	(0.036)	(0.013)	(0.009)	(0.180)	(0.042)	(0.023)
Observations	126,319	157,769	126,654	126,319	157,769	126,654
R-squared	0.03	0.031	0.03	0.035	0.035	0.034
Controls/FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Equal	Equal	Equal	Value	Value	Value

Table VIII Placebo Test: Earnings Day Share of Volatility. Table with estimates of β from:

$$\Delta QVS_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t}$$

$QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, 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. Controls in $X_{i,t-1}$ include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. I also condition on the growth in market capitalization from $t-1$ to t. Fixed effects are year and firm. Standard errors are computed using panel Newey-West with 2 lags. The "Baseline" columns use actual earnings dates, while the "Placebo" results are the coefficient estimates when selecting dates between the actual earnings days $t = -22$, or the FOMC meeting dates.

Appendix D. Reduced-Form Placebo Regressions

Appendix E. Relationship to Competing Hypotheses

Appendix E.1. Rise of AT Activity

Tables IX, X and XI 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 subsample that can be matched to the 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 passive ownership is still an important factor.

One possible reason for the drop in statistical significance when including the AT activity measures is a strong correlation between passive ownership and AT activity. To test this, I calculate an AT activity score as the first principal component of the 4 AT measures in Weller (2017). Table XII runs a regression of the AT activity score on the level of and changes in passive. Across almost all specifications, the relationship is positive and statistically significant. In unreported results, I find AT activity also increases in stocks after they are added to the S&P 500.

Appendix E.2. Regulation Fair Disclosure

Tables XIII, XIV, XV 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.

	Baseline			AT Measures	
	(1)	(2)	(3)	(4)	(5)
Ch. Passive	-7.602***	-7.261**	-19.28***	-1.651	-13.59**
	(2.313)	(3.385)	(6.339)	(3.289)	(6.338)
Ch. Oddlot				-3.428***	-2.596***
				(0.272)	(0.712)
Ch. Trade/Order				6.009***	3.371***
				(0.245)	(0.479)
Ch. Trade/Cancel				-1.941***	-4.248***
				(0.245)	(0.595)
Ch. Tradesize				1.247***	2.789***
				(0.452)	(1.015)
Obs	44,544	44,542	44,542	44,542	44,542
R-squared	0.053	0.078	0.171	0.19	0.262
Controls	No	Yes	Yes	Yes	Yes
Firm FE	No	Yes	Yes	Yes	Yes
Weight	Eq.	Eq.	Vw.	Eq.	Vw.

Table IX AT Activity: Pre-Earnings Volume. Estimates of β from:

$$\Delta CAV_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t}$$

$CAV_{i,t}$ is cumulative abnormal pre-earnings trading volume. Δ is a year-over-year change, matching on fiscal quarter. Change in passive ownership is expressed as a decimal, so $0.01 = 1\%$ increase. Controls, $X_{i,t-1}$, include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, total institutional ownership and growth in market capitalization from $t-1$ to t. Standard errors are computed using panel Newey-West with 8 lags. Only uses data that can be matched to MIDAS from 2012-2018. Column 1 is a univariate regression, while column 2 includes the controls, as well as year/quarter and firm fixed effects. Column 3 is the same as column 2, except firms are weighted by their 1-quarter lagged market capitalization relative to other firms that quarter. Columns 4 and 5 add the AT activity measures from Weller (2017).

		Baseline		AT Measures	
	(1)	(2)	(3)	(4)	(5)
Ch. Passive	-0.0531***	-0.0770***	-0.0994***	-0.0790***	-0.103***
	(0.017)	(0.024)	(0.037)	(0.024)	(0.037)
Ch. Oddlot				0.00408***	0.00158
				(0.001)	(0.003)
Ch. Trade/Order				0.0012	0.00128
				(0.001)	(0.003)
Ch. Trade/Cancel				0.00328***	0.00781***
				(0.001)	(0.003)
Ch. Tradesize				0.00202	0.00551
				(0.002)	(0.005)
Obs	44,527	44,525	44,525	44,525	44,525
R-squared	0.013	0.059	0.05	0.06	0.053
Controls	No	Yes	Yes	Yes	Yes
Firm FE	No	Yes	Yes	Yes	Yes
Weight	Eq.	Eq.	Vw.	Eq.	Vw.

Table X AT Activity: Pre-Earnings Drift. Table with estimates of β from:

$$\Delta DM_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t}$$

Where $DM_{i,t}$ is a measure of the pre-earnings drift. Passive ownership is expressed as a decimal, so 0.01 = 1% of shares outstanding held by passive funds. Controls, $X_{i,t-1}$, include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, total institutional ownership and growth in market capitalization from $t-1$ to t. Standard errors are computed using panel Newey-West with 8 lags. Column 1 is a univariate regression, while column 2 includes the controls, as well as year/quarter and firm fixed effects. Column 3 is the same as column 2, except firms are weighted by their 1-quarter lagged market capitalization relative to other firms that quarter.

		Baseline		AT Measures	
	(1)	(2)	(3)	(4)	(5)
Ch. Passive	0.312***	0.255***	0.789**	0.237***	0.805**
	(0.050)	(0.077)	(0.328)	(0.077)	(0.334)
Ch. Oddlot				-0.00439	-0.00685
				(0.005)	(0.030)
Ch. Trade/Order				-0.00875	-0.0384
				(0.006)	(0.043)
Ch. Trade/Cancel				-0.0106**	-0.0653
				(0.005)	(0.053)
Ch. Tradesize				-0.0223**	0.0211
				(0.009)	(0.067)
Obs	19,220	18,815	18,815	18,815	18,815
R-squared	0.006	0.065	0.084	0.066	0.087
Controls	No	Yes	Yes	Yes	Yes
Firm FE	No	Yes	Yes	Yes	Yes
Weight	Eq.	Eq.	Vw.	Eq.	Vw.

Table XI AT Activity: Earnings Day Share of Volatility. Table with estimates of β from:

$$\Delta QVS_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t}$$

$QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, 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. Controls in $X_{i,t-1}$ include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. I also condition on the growth in market capitalization from $t-1$ to t. Fixed effects are year and firm. Standard errors are computed using panel Newey-West with 2 lags. Only uses data that can be matched to MIDAS from 2012-2018. Column 1 is a univariate regression, while column 2 includes the controls, as well as year/quarter and firm fixed effects. Column 3 is the same as column 2, except firms are weighted by their 1-quarter lagged market capitalization relative to other firms that quarter. Columns 4 and 5 add the AT activity measures from Weller (2017).

	AT Activity Score		1-Year Ch. In Score		3-Year Ch. In Score	
Level of Passive Ownership	1.585***	0.318**				
	(0.155)	(0.158)				
1-Year Inc. in Passive			0.736***	0.369*		
			(0.199)	(0.212)		
3-Year Inc. in Passive					0.619***	0.462
					(0.211)	(0.290)
Observations	17,210	17,210	17,068	17,068	12,783	12,783
R-squared	0.3	0.086	0.093	0.098	0.125	0.108
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

Table XII Relationship Between Passive Ownership and AT Activity I calculate the AT activity score as the first principal component of the 4 AT measures in Weller (2017). This score is normalized to have mean zero and standard deviation one. The effect in levels of moving from the 25th percentile (0.0) to 75th percentile (0.1) of passive ownership is a 0.15 standard deviation increase in AT activity score. Although a 0.15 standard deviation increase may seem small, this is about half the mean for firms with high passive ownership. Controls include firm market capitalization, institutional ownership, idiosyncratic volatility and market cap. growth.

Appendix F. Alternative Identified Evidence

Appendix F.1. S&P 500 Index Deletions

In Section IV, 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 exogeneity assumption is likely violated, because index deletion is always about firm fundamentals.

The next challenge 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 sort on industry, size and growth rate to identify control firms, even though removing the industry filter and replacing it with a measure of liquidity would probably yield a more appropriate control group.

	(1)	(2)	(3)
Inc. Passive	-12.26***	-13.97***	-15.91**
	(1.923)	(2.214)	(6.207)
Observations	151,068	151,064	151,064
R-squared	0.03	0.031	0.181
Controls	No	Yes	Yes
Firm FE	No	No	Yes
Weight	Eq.	Eq.	Val.

Table XIII Post-2000: Pre-Earnings Volume. Estimates of β from:

$$\Delta CAV_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t}$$

$CAV_{i,t}$ is cumulative abnormal pre-earnings trading volume. Δ is a year-over-year change, matching on fiscal quarter. Change in passive ownership is expressed as a decimal, so $0.01 = 1\%$ increase. Controls, $X_{i,t-1}$, include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, total institutional ownership and growth in market capitalization from $t-1$ to t. Standard errors are computed using panel Newey-West with 8 lags. Column 1 is a univariate regression, while column 2 includes the controls, as well as year/quarter and firm fixed effects. Column 3 is the same as column 2, except firms are weighted by their 1-quarter lagged market capitalization relative to other firms that quarter.

	(1)	(2)	(3)
Inc. Passive	-0.0282**	-0.0273**	-0.0660**
	(0.012)	(0.014)	(0.029)
Observations	151,023	151,020	151,020
R-squared	0.022	0.023	0.072
Controls	No	Yes	Yes
Firm FE	No	No	Yes
Weight	Eq.	Eq.	Val.

Table XIV Post-2000: Pre-Earnings Drift. Table with estimates of β from:

$$\Delta DM_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t}$$

Where $DM_{i,t}$ is a measure of the pre-earnings drift. Passive ownership is expressed as a decimal, so $0.01 = 1\%$ of shares outstanding held by passive funds. Controls, $X_{i,t-1}$, include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, total institutional ownership and growth in market capitalization from $t-1$ to t. Standard errors are computed using panel Newey-West with 8 lags. Column 1 is a univariate regression, while column 2 includes the controls, as well as year/quarter and firm fixed effects. Column 3 is the same as column 2, except firms are weighted by their 1-quarter lagged market capitalization relative to other firms that quarter.

	(1)	(2)	(3)
Inc. Passive	0.216***	0.140***	0.365**
	(0.032)	(0.035)	(0.172)
Observations	68,142	68,126	67,224
R-squared	0.012	0.013	0.037
Controls	No	Yes	Yes
Firm FE	No	No	Yes
Weight	Eq.	Eq.	Val.

Table XV Post-2000: Earnings Day Share of Volatility. Table with estimates of β from:

$$\Delta QVS_{i,t} = \alpha + \beta \Delta Passive_{i,t} + \gamma X_{i,t-1} + \text{Fixed Effects} + e_{i,t}$$

$QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, 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. Controls in $X_{i,t-1}$ include 1-year lagged passive ownership, market capitalization, idiosyncratic volatility, calculated as the sum of squared market-adjusted returns over the past year, and total institutional ownership. I also condition on the growth in market capitalization from $t-1$ to t. Fixed effects are year and firm. Standard errors are computed using panel Newey-West with 2 lags. Column 1 is a univariate regression, while column 2 includes the controls, as well as year and firm fixed effects. Column 3 is the same as column 2, except firms are weighted by their 1-year lagged market capitalization relative to other firms that year.

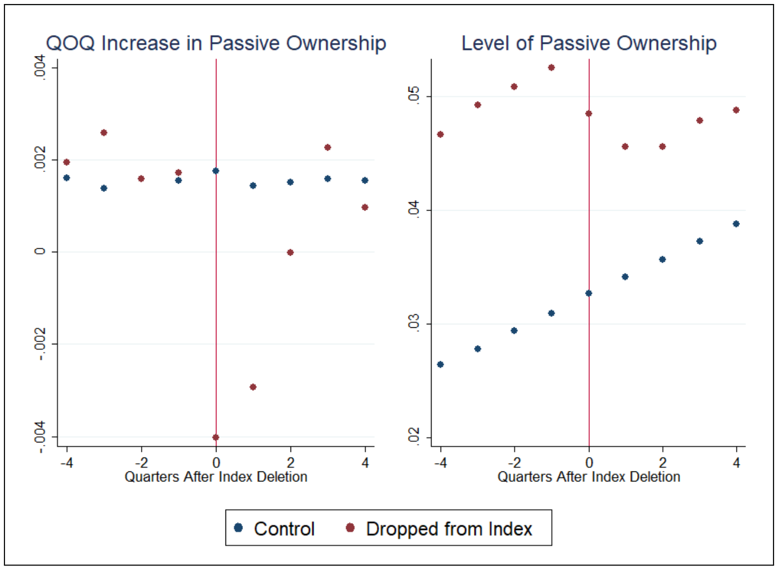

Figure 13. S&P 500 Index Deletions: Testing Parallel Trends. Average level and increase in passive ownership for control firms and firms dropped from the S&P 500. Control firms are all firms in the same 2-digit SIC industry, in the same size and growth rate quintiles that were initially in the S&P 500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

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.[18]

Figure 13 shows the changes in passive ownership around the index deletion date. There is a drop in passive ownership in the quarter of deletion, and the quarter after deletion. Unlike the increase in passive ownership after index addition, however, the decrease after index deletion is only temporary, as can be seen in the levels plot. One explanation is that stocks on the margin are still relatively large. Passive ownership increased for all stocks over my sample, especially the larger ones. The weak and temporary treatment effect suggests that index deletion would be a weak instrument for change in passive ownership.

[18] 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 treatment period to be the year immediately after index deletion, instead of skipping a year.

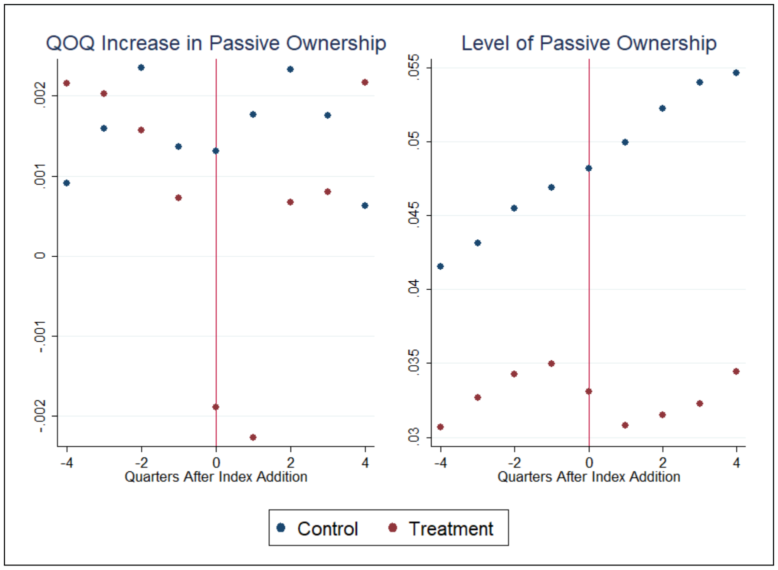

Figure 14. Russell 2000/1000 Reconstitution: Checking Parallel Trends. Average level and change in passive ownership for control firms and firms moved from the Russell 2000 to the Russell 1000. Control firms are all firms in the Russell 3000 ranked 900 to 1100 that did not move from the 1000 to the 2000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

Appendix F.2. Moving from the Russell 2000 to the Russell 1000

Similar to 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 exogenity assumption, as moving from firm 1001 to 999 may have nothing to do with firm fundamentals.

Similar to the setup in Section IV, 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 8 shows the problem with this setup: the treatment is small and temporary. The common pattern between moving from the Russell 2000 to the 1000, and 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 index rebalancing.

Appendix F.3. Blackrock's Acquisition of Barclays Global Investors

Another well-known source of quasi-exogenous variation in passive ownership is Blackrock's acquisition of Barclays' iShares ETF business in December 2009. This is not an ideal setting for testing my hypothesis because: (1) My theory has no predictions for the effects of increased concentration of ownership among passive investors (Azar, Schmalz, and Tecu (2018), Massa, Schumacher, and Wang (2018)) (2) While there may have been a *relative* increases in flows to iShares ETFs, relative to all other ETFs (Zou (2018)), I do not find a significant increase in overall ETF ownership for the stocks owned by iShares funds. Given that my right-hand-side variable of interest is the percent of shares owned by informed or uninformed investors, the model has no predictions for the effect of moving dollars from iShares ETFs to non-iShares ETFs.

Appendix G. Systematic Information Announcement Days

I obtain FOMC announcement dates from Gorodnichenko and Weber (2016). To create an apples-to-apples comparison with the anticipated nature of earnings announcements, I restrict the sample to scheduled FOMC meetings. I compute versions of pre-earnings volume, and pre-earnings drift for these FOMC dates. The only difference is that I use a $\pm$ 10 day around each announcement, instead of $\pm$ 22 days, to avoid overlap as there are 8 scheduled meetings per year. Share of volatility on FOMC meeting dates is the sum of squared returns on those dates, divided by the sum of squared returns on all dates.

Figure 15 shows the trends in volume around FOMC announcement dates. There is no drop before the announcement in the last third of the sample. Figure 16 shows the pre-FOMC announcement drift. There is no upward/downward trend throughout the sample. Figure 17 shows a slight trend toward increased volatility on FOMC announcement dates, but this may be due to the increased importance of FOMC meetings during the global financial crisis.

Appendix H. Options Results

Given the increase in volatility on earnings days, especially for stocks with high passive ownership, it is natural to believe that ex-ante measures of uncertainty, like options prices, should also reflect this change. To quantify the change in option prices around earnings

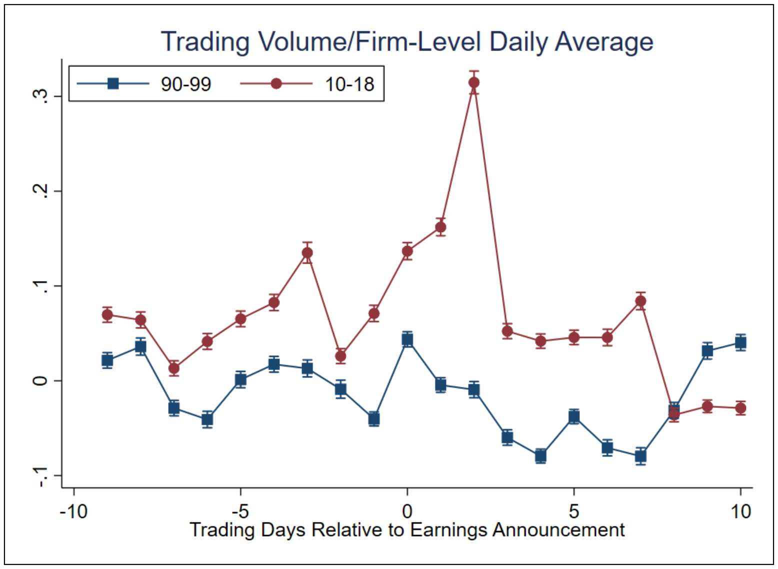

Figure 15. FOMC Meeting Dates: Pre-Earnings Volume. Plot of β_k estimated from the regression:

$$AV_{i,t+\tau} = \alpha + \sum_{k=-10}^{10} \beta_k \mathbf{1}_{\{\tau=k\}} + e_{i,t+\tau}$$

Bars represent a 95% confidence interval around the point estimate. Standard errors are clustered at the firm level. t represents a scheduled FOMC meeting date. The firm-level daily average is computed over the previous quarter. Observations are weighted by a firm's 1-year lagged market capitalization.

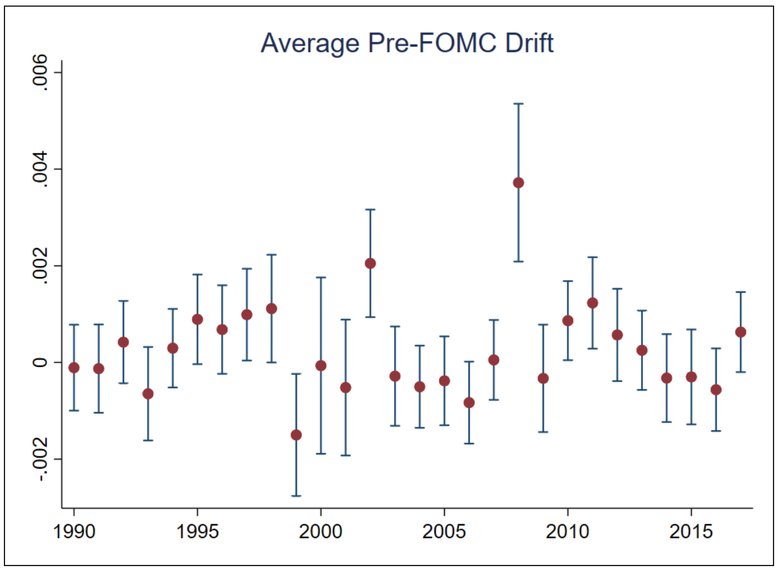

Figure 16. FOMC Meeting Dates: Pre-Earnings Drift. This figure plots coefficients from a regression of pre-earnings drift on a set of year dummy variables where $DM_{it} = \begin{cases} \frac{1+r_{(t-10,t-1)}}{1+r_{(t-10,t)}} & \text{if } r_t > 0 \\ \frac{1+r_{(t-10,t)}}{1+r_{(t-10,t-1)}} & \text{if } r_t < 0 \end{cases}$. Standard errors represent 95% confidence intervals around the point-estimates. Standard errors are clustered at the firm level. t denotes a scheduled FOMC meeting date. Observations are weighted by a firm's 1-year lagged market capitalization.

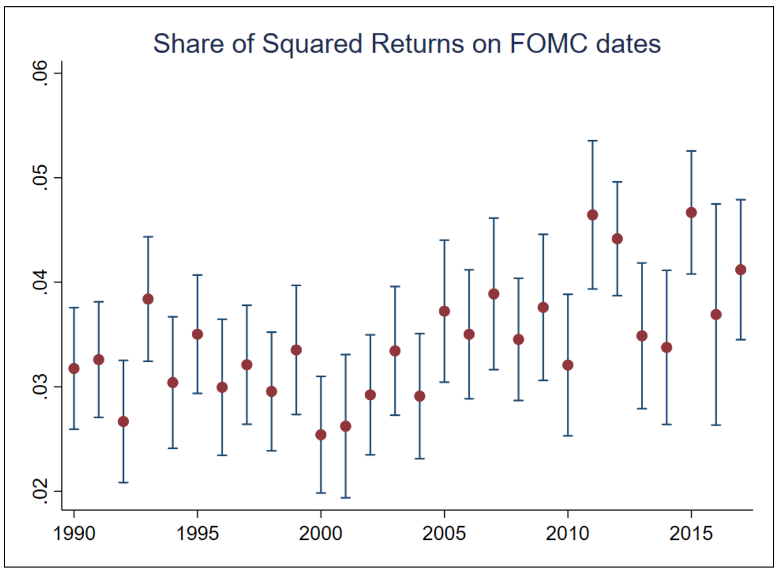

Figure 17. FOMC Meeting Dates: Earnings Day Volatility. This figure plots coefficients from a regression of QVS on a set of year dummy variables. For firm i in year t the quadratic variation share (QVS) is defined as: $QVS_{i,t} = \sum_{\tau=1}^{8} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, where r denotes a market-adjusted daily return. The numerator sums over the 8 scheduled FOMC meeting days in year t, while the denominator includes all days in calendar year t. Standard errors are clustered at the firm level. Observations are weighted by a firm's 1-year lagged market capitalization.

announcements, I follow the method in Kelly, Pástor, and Veronesi (2016). For each earnings announcement τ, I select 3 expiration dates, a, b, and c, where a the last expiration before the earnings announcement date, b is the first expiration after the announcement and c is expiration after b. For each of these dates, I average implied volatility (IV) across all options with $|\Delta| \in (0.4, 0.5)$ in a 20-day window starting 1 day before τ. I also choose corresponding 20-day windows before a and c to match time to expiration.

This measure was originally designed for S&P 500 options, so I have to modify it to be comparable across firms:

$$IVD_\tau = \frac{\overline{IV}_b}{0.5\left(\overline{IV}_a + \overline{IV}_c\right)} \tag{A1}$$

I also calculated the *Variance Risk Premium* and *Slope* measures from Kelly et al. (2016), but given that I am using individual equity options, I had the longest/most reliable sample using the *IVD* measure.

I constructed a secont measure of ex-ante uncertainty based on Dubinsky, Johannes, et al. (2006). Suppose there is one predictable announcement before an option expires:

$$\text{Before event IV} = \sigma^2 + \frac{1}{T_i}(\sigma_j^{\mathbb{Q}})^2 \tag{A2}$$

We can estimate earnings uncertainty using two options with different maturities, $T_1 < T_2$, on the day before earnings are released:

$$(\sigma_{term}^{\mathbb{Q}})^2 = \frac{\sigma_{t,T_1}^2 - \sigma_{t,T_2}^2}{T_1^{-1} - T_2^{-1}} \tag{A3}$$

For each firm/earnings announcement, I average this measure over the 3 strikes closest to the money. To make this comparable across firms, I divide by average IV: $\frac{(\sigma_{term}^{\mathbb{Q}})^2}{\overline{IV}_b}$:

To construct these option-based measures of earnings uncertainty, I used daily implied volatility data from OptionMetrics. To make sure I am working with reliable data, I restrict to S&P 500 firms, discard all options with $\frac{ask}{bid} > 5$, or zero open interest and filter for firms which have at least 15 years of non-missing options data.

Figure 18 shows that both measures of option prices around earnings announcements increased substantially between the 1990's and 2010's.

I run the following regression to see if changes in passive ownership can explain the

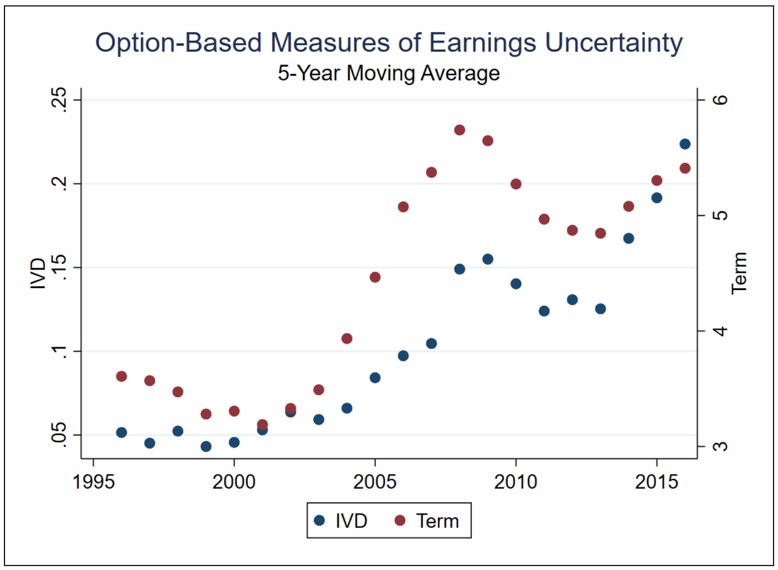

Figure 18. Option-Based Measures of Earnings Uncertainty. Each point represents the cross-sectional average of each option-based measure of earnings uncertainty. For earnings date τ, $IVD = \frac{\overline{IV}_b}{0.5\left(\overline{IV}_a+\overline{IV}_c\right)}$ and Term$=\frac{(\sigma^{\mathbb{Q}}_{term})^2}{\overline{IV}_b}$. Sample includes 306 S&P 500 firms with options that meet the filters described in Kelly et al. (2016), and have at least 16 years with 4 non-missing earnings announcements.

increases in IVD and $Term$:

$$\Delta_{(t,t-5)}Outcome_{i,t} = \alpha + \beta\Delta_{(t,t-5)}Passive_{i,t} + \gamma X_{i,t} + \text{Fixed Effects} + e_{i,t} \tag{A4}$$

Controls in $X_{i,t}$ include institutional ownership, lagged institutional ownership, market capitalization, lagged market capitalization. Fixed effects include industry, year and firm. The results are in Table XVI. Although the effects are not always statistically significant, increases in passive ownership are correlated with increases in both the IVD and $Term$ measures of earnings uncertainty.

Appendix I. Possible Omitted Variables

In addition to identification concerns, the reduced-form regressions could suffer from omitted variable bias. 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 a significant amount of the differences in passive ownership across stocks 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[19], 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.

Another possibly omitted variable is the quantity of ETF rebalancing. The changes in pre-earnings volume could be driven by mechanical changes in systematic trading rules by ETFs, as in Chinco and Fos (2019). I find this unlikely, as ETFs typically rebalance on a calendar frequency, not around particular firms' earnings announcements, which may be scattered throughout a calendar quarter. The drop in volume before earnings announcements can not be explained by the ETF rebalancing or imbalance measures of Chinco and Fos

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

		IVD					
		1-year		3-year		5-year	
Increase in Passive Ownership	1-year	0.325	0.334*				
		(0.225)	(0.191)				
	3-year			0.173	0.100		
				(0.154)	(0.158)		
	5-year					0.328**	0.300*
						(0.141)	(0.157)
Observations		4,519	4,519	3,979	3,979	3,496	3,496
		Term					
		1-year		3-year		5-year	
Increase in Passive Ownership	1-year	9.375	9.155**				
		(5.873)	(4.495)				
	3-year			8.789*	8.278**		
				(4.758)	(4.000)		
	5-year					5.489	7.653*
						(4.547)	(4.121)
Observations		4,457	4,457	3,916	3,916	3,441	3,441
Firm Controls		Yes	Yes	Yes	Yes	Yes	Yes
Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Industry FE		Yes	Yes	Yes	Yes	Yes	Yes
Firm FE		No	Yes	No	Yes	No	Yes

Table XVI Option Regressions. Estimates of β from:

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

Where the outcomes are $IVD = \frac{\overline{IV}_b}{0.5\left(\overline{IV}_a+\overline{IV}_c\right)}$ and Term= $\frac{(\sigma^{\mathbb{Q}}_{term})^2}{\overline{IV}_b}$. Sample includes 306 S&P 500 firms with options that meet the filters described in Kelly et al. (2016), and have at least 16 years with 4 non-missing earnings announcements. Controls in $X_{i,t}$ include institutional ownership, lagged institutional ownership, market capitalization, lagged market capitalization. Fixed effects include industry, year and firm.

(2019).

It is also possible that firms with high passive ownership anticipate increased volatility on earnings days, or amplified responses to earnings news (as I show in Section V) and as a result, release earnings information when the market is closed. To test this, I form three groups of earnings announcements (1) Weekday, before 4PM EST (2) Weekday, after 4PM EST (3) Weekend or trading holiday. I then run a multinomial logistic regression of these categories on passive ownership. Passive ownership does not have significant predictive power in this regression, once I control for time trends, and differences in firm characteristics correlated with passive ownership like market capitalization.

A final omitted variable is the composition and importance of systematic and idiosyncratic information in the economy. The growth of passive management could be a response to a secular increase in the importance of systematic risks, and increased demand for exposure to these factors. If this were true, it might be rational to focus on learning about factor risk, rather than firm-specific risk, which could explain my results. I find this unlikely for several reasons.

Figure 19 replicates the volatility decomposition in Campbell et al. (2001), extending the analysis to 2017. Between 1990 and 2017, there does not appear to have been a significant increase in the market or industry components of total risk. It is still possible, however, that this does not fully account for changes in systematic risk, if those changes occur on a small number of days, like FOMC announcements, as the figure is composed of slow-moving averages. Figure 20 re-constructs Figure 1, but using FOMC announcement dates instead of earnings announcement dates. Between the 1990's and 2010's there was no major change in returns or trading volume around these systematic information release dates at the stock level, also inconsistent with an increase in importance of systematic risk.

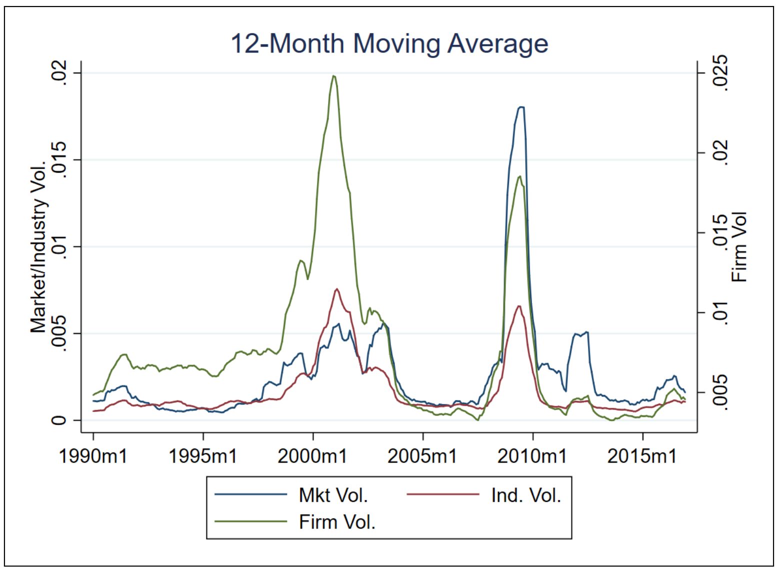

Figure 19. Replication and Expansion of Campbell et al. (2001). This figure plots a 12-month moving average for the 3 components of total volatility. The market component in period t is $\sum_{s \in t} (R_{ms} - \mu_m)^2$, where R_{ms} is the return on the market at time s and μ_m is the mean return on the market over the sample. Industry returns are assumed to follow $R_{it} = R_{mt} + \epsilon_{it}$, so the industry component is the sum of squared ϵ_{it}'s. The firm component is any residual volatility not explained by the industry and market components.

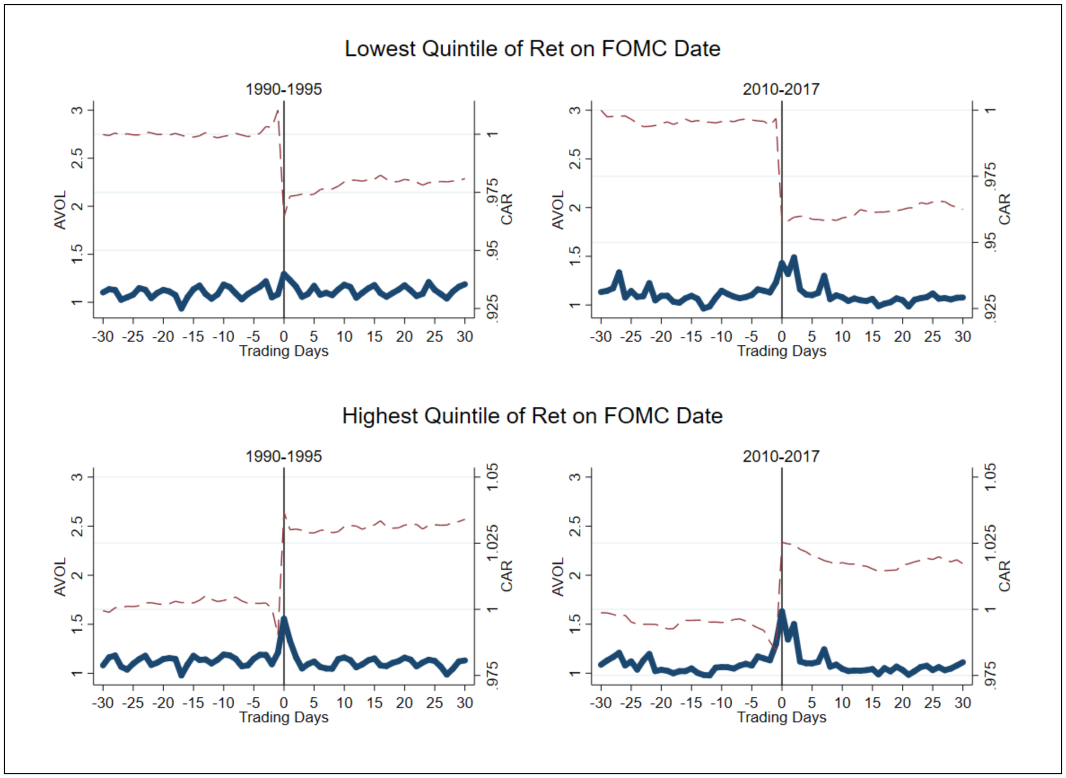

Figure 20. Returns and Trading Volume around FOMC meeting dates. Each plot represents the cross-sectional average for each quintile of Return on FOMC dates over the specified years. Abnormal returns are defined as returns relative to the returns on the CRSP value-weighted index, as in Campbell et al. (2001). Cumulative abnormal returns (CAR) are calculated by compounding net abnormal returns. Cumulative abnormal volume is volume relative to the firm-level average over the past year, see Equation 2.

Appendix J. Robustness of Stylized Facts

Appendix K. Data Details

Daily Volume (from CRSP): Number of shares traded across all US exchanges.

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.

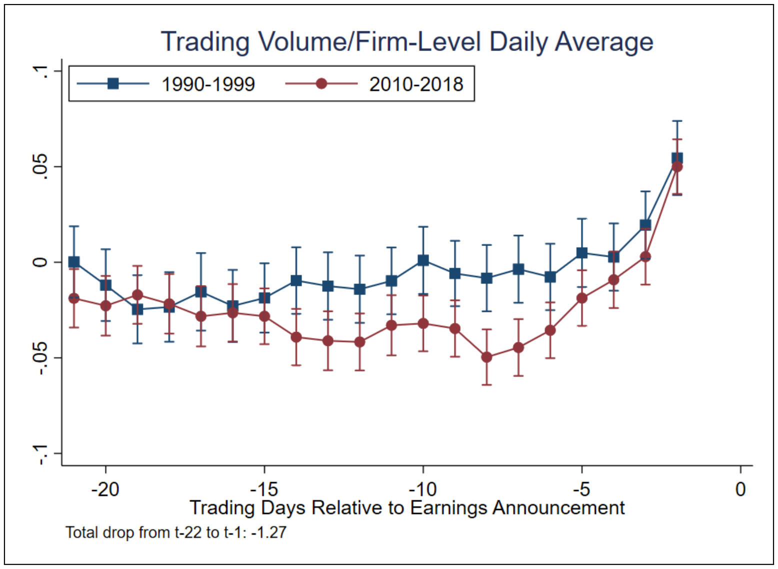

Figure 21. Decline of Pre-Earnings Volume (Value Weighted). Plot of β_k estimated from the regression:

$$AV_{i,t+\tau} = \alpha + \sum_{k=-21}^{22} \beta_k \mathbf{1}_{\{\tau=k\}} + e_{i,t+\tau}$$

Bars represent a 95% confidence interval around the point estimate. Standard errors are clustered at the firm level. Observations are weighted by a firm's 1-year lagged market capitalization.

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 when ruling out the influence of Regulation Fair Disclosure 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%.

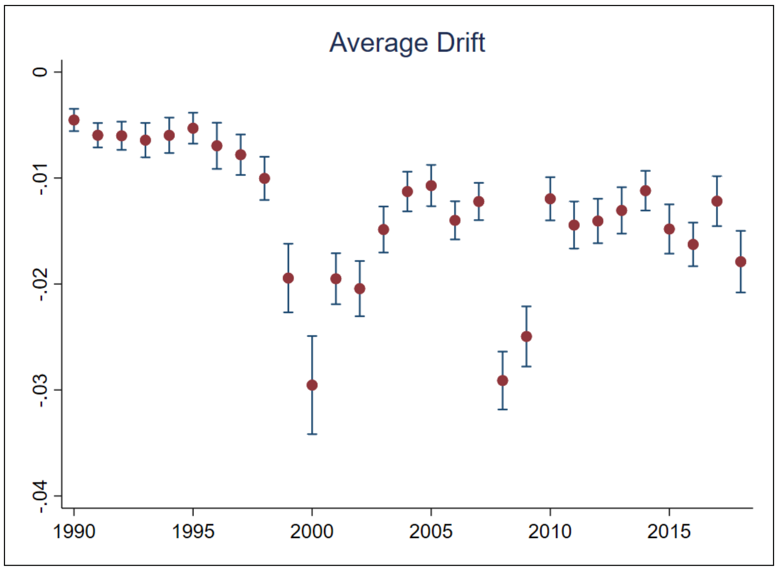

Figure 22. Decline of Average Pre-Earnings Drift (Value Weighted). This figure plots coefficients from a regression of pre-earnings drift on a set of year dummy variables where $DM_{it} = \begin{cases} \frac{1+r_{(t-22,t-1)}}{1+r_{(t-22,t)}} & \text{if } r_t > 0 \\ \frac{1+r_{(t-22,t)}}{1+r_{(t-22,t-1)}} & \text{if } r_t < 0 \end{cases}$. A value near 1 implies most earnings information is incorporated in prices before the announcement date, while lower values denote less informative pre-earnings announcement prices. Standard errors represent 95% confidence intervals around the point-estimates. Standard errors are clustered at the firm level. Observations are weighted by a firm's 1-year lagged market capitalization.

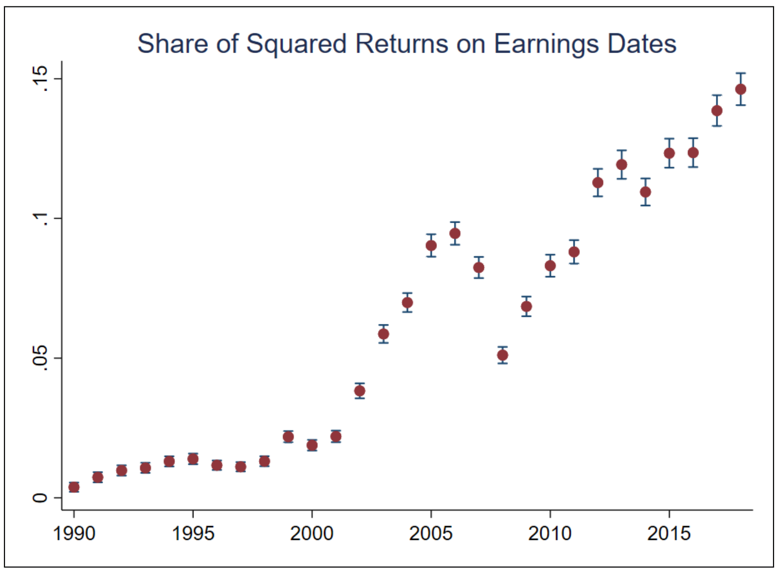

Figure 23. Increase in Earnings Day Volatility (Value Weighted). This figure plots coefficients from a regression of QVS on a set of year dummy variables. For firm i in year t the quadratic variation share (QVS) is defined as: $QVS_{i,t} = \sum_{\tau=1}^{4} r_{i,\tau}^2 / \sum_{j=1}^{252} r_{i,j}^2$, 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 calendar year t. Standard errors are clustered at the firm level. Observations are weighted by a firm's 1-year lagged market capitalization.

REFERENCES

Admati, Anat R, and Paul Pfleiderer, 1988, A theory of intraday patterns: Volume and price variability, *The Review of Financial Studies* 1, 3–40.

Akerlof, George A, 1978, The market for "lemons": Quality uncertainty and the market mechanism, in *Uncertainty in economics*, 235–251 (Elsevier).

Anson, Mark JP, Donald R Chambers, Keith H Black, Hossein Kazemi, CAIA Association, et al., 2012, *CAIA level I: an introduction to core topics in alternative investments*, volume 709 (John Wiley & Sons).

Appel, Ian R, Todd A Gormley, and Donald B Keim, 2016, Passive investors, not passive owners, *Journal of Financial Economics* 121, 111–141.

Azar, José, Martin C Schmalz, and Isabel Tecu, 2018, Anticompetitive effects of common ownership, *The Journal of Finance* 73, 1513–1565.

Bai, Jennie, Thomas Philippon, and Alexi Savov, 2016, Have financial markets become more informative?, *Journal of Financial Economics* 122, 625–654.

Balchunas, Eric, 2016, *The Institutional ETF Toolbox: How Institutions Can Understand and Utilize the Fast-Growing World of ETFs* (John Wiley & Sons).

Ball, Ray, and Philip Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of accounting research* 159–178.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2018, Do etfs increase volatility?, *The Journal of Finance* 73, 2471–2535.

Bhattacharya, Ayan, and Maureen O'Hara, 2018, Can etfs increase market fragility? effect of information linkages in etf markets, *Effect of Information Linkages in ETF Markets (April 17, 2018)* .

Buffa, Andrea M, Dimitri Vayanos, and Paul Woolley, 2014, Asset management contracts and equilibrium prices, Technical report, National Bureau of Economic Research.

Campbell, John Y, Martin Lettau, Burton G Malkiel, and Yexiao Xu, 2001, Have individual stocks become more volatile? an empirical exploration of idiosyncratic risk, *The Journal of Finance* 56, 1–43.

Chinco, Alexander, and Vyacheslav Fos, 2019, The sound of many funds rebalancing .

Cong, Lin William, and Douglas Xu, 2016, Rise of factor investing: asset prices, informational efficiency, and security design, in *29th Australasian Finance and Banking Conference*.

Dávila, Eduardo, and Cecilia Parlatore, 2019, Trading costs and informational efficiency, Technical report, National Bureau of Economic Research.

Dubinsky, Andrew, Michael Johannes, et al., 2006, Fundamental uncertainty, earning announcements and equity options, in *Working Paper*.

Fama, Eugene F, Lawrence Fisher, Michael C Jensen, and Richard Roll, 1969, The adjustment of stock prices to new information, *International economic review* 10, 1–21.

Farboodi, Maryam, and Laura Veldkamp, 2017, Long run growth of financial technology, Technical report, National Bureau of Economic Research.

Foster, George, Chris Olsen, and Terry Shevlin, 1984, Earnings releases, anomalies, and the behavior of security returns, *Accounting Review* 574–603.

Garleanu, Nicolae, and Lasse Heje Pedersen, 2018, Active and passive investing, *Available at SSRN 3253537* .

Glosten, Lawrence R, Suresh Nallareddy, and Yuan Zou, 2016, Etf trading and informational efficiency of underlying securities .

Gorodnichenko, Yuriy, and Michael Weber, 2016, Are sticky prices costly? evidence from the stock market, *American Economic Review* 106, 165–99.

Grossman, Sanford J, and Joseph E Stiglitz, 1980, On the impossibility of informationally efficient markets, *The American economic review* 70, 393–408.

Hajric, Vildana, and Reade Pickert, 2019, Sec cracks down on etf names that could be misleading investors, *Bloomberg* .

Israeli, Doron, Charles MC Lee, and Suhas A Sridharan, 2017, Is there a dark side to exchange traded funds? an information perspective, *Review of Accounting Studies* 22, 1048–1083.

Kacperczyk, Marcin T, Jaromir B Nosal, and Savitar Sundaresan, 2018, Market power and price informativeness, *Available at SSRN 3137803* .

Kelly, Bryan, L'uboš Pástor, and Pietro Veronesi, 2016, The price of political uncertainty: Theory and evidence from the option market, *The Journal of Finance* 71, 2417–2480.

Kyle, Albert S, 1985, Continuous auctions and insider trading, *Econometrica: Journal of the Econometric Society* 1315–1335.

Loughran, Tim, and Bill McDonald, 2017, The use of edgar filings by investors, *Journal of Behavioral Finance* 18, 231–248.

Malikov, George, 2018, Information, participation, and passive investing, *Participation, and Passive Investing (December 15, 2018)* .

Massa, Massimo, David Schumacher, and Yan Wang, 2018, Who is afraid of blackrock? .

Mauboussin, Michael, Dan Callahan, and Darius Majd, 2017, Looking for easy games, Technical report, Credit Suisse.

McElhaney, Alicia, 2018, Is passive investing making markets dumb?, *Institutional Investor* .

McLean, R David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *The Journal of Finance* 71, 5–32.

Morck, Randall, and Fan Yang, 2001, The mysterious growing value of s&p 500 membership, Technical report, National Bureau of Economic Research.

Palia, Darius, and Stanislav Sokolinski, 2019, Passive asset management, securities lending and stock prices, *Securities Lending and Stock Prices (February 15, 2019)* .

Sammon, Marco, 2020, Etf introduction, learning and price informativeness, Technical report, Kellogg School of Management.

Schipper, Katherine, 1989, Earnings management, *Accounting horizons* 3, 91.

S&P, Dow Jones Indices, 2017, S&p u.s. indices methodology, Technical report.

Wang, Jiang, 1994, A model of competitive stock trading volume, *Journal of political Economy* 102, 127–168.

Wei, Wei, and Alex Young, 2017, Selection bias or treatment effect? a re-examination of russell 1000/2000 index reconstitution .

Weller, Brian M, 2017, Does algorithmic trading reduce information acquisition?, *The Review of Financial Studies* hhx137.

Zou, Yuan, 2018, Lost in the rising tide: Etf flows and valuation .