Limits to Diversification: Passive Investing and Market Risk[1]

Lily Fang
INSEAD

Hao Jiang
Michigan State University

Zheng Sun
University of California, Irvine

Ximing Yin
Hunan University

Lu Zheng
University of California, Irvine

This Draft: November 15, 2023

Abstract

We hypothesize and show that the rise of passive investing contributed to higher correlations among stocks and higher market volatility. The degree to which a stock is held by passive funds (index mutual funds and ETFs) strongly predicts its beta and correlations with other stocks but not its idiosyncratic volatility. Difference-in-difference analyses around three market shocks—September 11 (2001), Lehman collapse (September 2008), and Covid (March 2020)—show that stocks with high passive holdings contributed more to market volatility. These results are neither subsumed by common institutional holdings or performance benchmarking, nor explained by correlation in earnings. The rise of passive investing thus could limit its own benefit of diversification.

[1] We thank Emma Qiu for superb research assistance. Fang thanks INSEAD R&D for funding support. All errors and omissions are our own.

Introduction

Can index investing be too much of a good thing? Due to the rapid growth and increasing dominance of index investing, market participants and researchers are beginning to question whether there are unintended implications. For instance, John Bogle, founder of Vanguard who is known as the "father of index funds," cautioned in a 2019 WSJ article published shortly before his passing that "Public policy cannot ignore this growing dominance, and consider its impact on the financial markets, corporate governance, and regulation."[2] In this paper, we examine the implications of index investing for the benefits of portfolio diversification, which is not only a key proposed advantage of index investing and the original rationale behind its creation, but also a cornerstone of modern portfolio theory.

The invention of index investing is one of the most significant financial innovations in the 20th century. It revolutionized investing by dramatically reducing the costs and simplifying the process for everyday investors to invest in a well-diversified portfolio. Investing in index funds that passively track the market performance is a game-changing idea not only because it gave millions of investors cheap access to diversification, but also that the portfolio they would hold—the "market portfolio"—has been shown by modern portfolio theory and the Efficient Market Hypothesis to be the *optimal* portfolio from a risk-return trade-off point of view. However, index investing, by definition, means securities are combined into baskets and trading of different securities are more likely to be in unison rather than purely idiosyncratic. When asset markets are not frictionless and trading affects asset prices, such correlated trading can generate common price movements beyond what is explained by fundamentals, i.e., excess co-movements of asset prices. Since in the limit market volatility is driven solely by co-movements (i.e., correlations) among security prices and not the idiosyncratic movements

[2] We will provide a detailed review of the academic literature later in the introduction.

of individual securities which are diversified away, this means that the rise of index investing can, in theory, lead to higher volatility of the market portfolio. We thus face a paradox of index investing for the purpose of diversification: Investor hold the market portfolio to achieve diversification as prescribed by modern portfolio theory; yet as more and more investors do the same and hold the same basket(s) of assets, the benefit of diversification diminishes.

This paradox is the focus of our paper. We study how index-based investing – both index funds and index ETFs, which we collectively refer to as "passive investing" in this paper – affects the correlation structure between assets and, ultimately, the overall market volatility. Along the rationale outlined above, we hypothesize that passive investing increases correlation among assets and, since correlation among assets is what determines the aggregate market volatility, it also increases overall market volatility.

Our hypothesis has profound implications for market efficiency and the cost and benefit of passive investing. A large literature exemplified by Jensen (1968) and Carhart (1997) has presented robust evidence of the benefits of passive investing: Actively managed mutual funds generally do not outperform passive funds after fees and expenses. This robust finding has contributed to the rise of passive investing in recent decades, which now account for 43% of the total assets under management in US equity funds.[3] But our hypothesis suggests an important downside to the rise of passive investing: Its rise could lead to higher *market-level* volatility, limiting the power of diversification and the virtue of passive investing itself. Moreover, the information structure in the market will change: the shift from active to passive investing might lead to a reduction in the generation and incorporation of unique, company-specific information into stock prices.[4]

[3] Investment Company Institute (ICI) Factbook: https://www.icifactbook.org/pdf/2022_factbook.pdf.
[4] The literature and popular press has discussed another cost associated with the rise of passive investing: The risk of ownership concentration. See, for example, Azar et al. (2018) and Bogel (2018). We do not address this issue in this paper.

A seminal paper by Campbell et al. (2001) documents that between 1962 and 1997, correlation among stocks dropped while individual stock volatilities increased; these two effects balanced each other out so that the overall market volatility remained stable.[5] While we are able to replicate and confirm these well-known results (see Figure 1), we show that there has been a significant reversal of these patterns in the period since the publication of the paper. Figure 2 extends the sample period to 2020 and reveals a stark contrast between the pre-1997 period and the time since. Panel A shows that individual stock volatility continued to rise until about 2001 but has since declined. Strikingly, Panel B shows an *increase* in pair-wise correlation among stocks post 1997. Indeed, the average pair-wise correlation is 13.4% in the period of 1998-2020, more than doubling the 5.7% for the period 1962-1997 (t-statistic=17.85 for the difference). Panel C shows that the net effect of these two forces is a slight increase in market-level variance in the post 1997 era; the average pre- (post-) 1997 market annualized return volatility is 12.51%, and 19.56% respectively, and the difference is statistically significant (t-statistic = 5.44).

While there can be multiple factors contributing to the increase in correlations among stocks, our paper primarily examines this phenomenon in the context of the growing popularity of passive investing - index funds and index ETFs. Figure 3 presents a first view of the time-series relation between passive investing and correlations among stocks. The left-hand-side scale of the figure plots the average pair-wise correlation among all stocks in the CRSP universe while the right-hand-side scale plots the "*Passive-to-Market*" ratio, which is the total assets under management by index funds and index ETFs, divided by the total market capitalization of all stocks. The figure shows that both series rose in tandem over this period. The correlation between the two series is 68% and statistically

[5] In the limit market return variance is the average pair-wise covariance between stocks, which in turn is the product of the average pair-wise correlation between stocks and the average stock return variance.

significant at the 1% level, suggesting at least an association between passive investing and market-level risk.

To investigate the relation between passive investing and measures of risk in greater granularity, we use panel data for the entire CRSP universe. For each stock in each quarter, we calculate the percentage of the stock's market capitalization that is held by index funds and ETFs in our sample. We term this measure a stock's "*PASSIVE_EXP*" (exposure to passive investing). We then examine the relation between stocks' *PASSIVE_EXP* and four risk measures of the stock return: 1) market beta, 2) average correlation with all other stocks, 3) average covariance with all other stocks, 4) idiosyncratic volatility (relative to the Fama-French 4-factor model). We focus on these measures because they reflect different aspects of a stock's "risk".

Our main empirical results are as follows. On the one hand, *PASSIVE_EXP* strongly and significantly predicts a stock's beta, correlation, and covariance with other stocks. On the other hand, *PASSIVE_EXP* has a less robust relation with measures of stocks' idiosyncratic movement such as its idiosyncratic volatility. In fact, in most specifications, we find a *negative* (albeit not always significant) relation between *PASSIVE_EXP* and a stock's idiosyncratic volatility. Overall, the robust conclusion is that there is a strong relation between a stock's exposure to passive investing and how this stock comoves with other stocks.

We use a difference-in-differences approach to illustrate that stocks with higher PASSIVE_EXP *contribute* more to overall market volatility. We focus on three episodes of market-wide shocks – the 9/11 event in September 2001, the collapse of Lehman Brothers and onset of the financial crisis in September 2008, and the Covid crash in March 2020. Each episode represents unanticipated, market-wide shocks that resulted in high market-level volatility. We sort stocks by their pre-event *PASSIVE_EXP* and study the changes to stocks' risk measures before and during the crisis periods. We find that stocks with high *pre-crisis PASSIVE_EXP* exhibit significantly higher *increase*s in their betas, and in their average correlation and covariance with other stocks. They also exhibit higher increases in their idiosyncratic volatility. Thus, high *PASSIVE_EXP* stocks contribute more to overall market risk during the crisis period, precisely when investors need the benefit of diversification the most.

Our hypothesis is based on the notion that passive investing often involves the simultaneous buying and selling of securities within an index, and this correlated trading can lead to increases in stocks' systematic risk measures such as beta and correlation with other stocks. To directly shed light on this mechanism, we examine the effect of trading induced by index and ETF fund flows (we call this passive-flow-induced trading). Xu (2021) shows that majority of the trading by index funds are induced by fund flows. We construct a stock-level measure that captures the passive-flow-induced net trading of the stock across all index funds and ETFs in our sample. Presumably, the buying pressure induced by the inflows of some funds could be offset by the selling pressure driven by the outflows from other funds. Our measure captures the net amount of flow-induced trading that cannot be absorbed within the index fund sector, thus reflects a net liquidity demand by index funds to other investors. We find that this passive-flow-induced-trading is significantly correlated with stocks' beta and their average correlation coefficients with other stocks, which is consistent with the hypothesis that indexing increases correlation among stocks. We also find that passive-flow-induced-trading is negatively correlated with stocks' idiosyncratic volatility, which is consistent with the hypothesis that indexing reduces the incorporation of company-specific information into stock prices. Moreover, the contributing effects of passive-flow-induced-trading on beta, correlations, and covariance, is especially strong during the crisis periods defined above.

Our analysis has focused on passive funds with strict mandates to track pre-specified market indices. In practice, large groups of investors have more flexible investment mandates, but their performance is benchmarked against specific indices. As shown by Pavlova and Sikorskaya (2023) and Kashyap et al. (2023), such performance benchmarking leads stocks in the indices to be the preferred habitat for these investors. Sharing a similar mechanism, the common demand for stocks in the indices due to performance benchmarking can lead to excess price co-movements. To investigate this hypothesis, we expand our baseline regressions by adding the Benchmark Intensity Measure (BMI) in Pavlova and Sikorskaya (2023) alongside the *PASSIVE_EXP* measure. The BMI measure indicates both active and passive funds' aggregate inelastic demand due to benchmarking. Since the BMI measure and our *PASSIVE_EXP* both reflect the degree of passive holdings in a stock, not surprisingly they are highly correlated with a correlation coefficient of 61%. Thus, in the regressions, we use the residual BMI measure after orthogonalizing it against the *PASSIVE_EXP* measure.[6] If both explicit and implicit passive holdings have similar effects on market risk, we expect these two variables to exhibit the same signs in the regressions. This is indeed what we find: Both variables are positive and significant in the regressions of beta, average correlation and covariance among stocks. Furthermore, the effects are separate and additive: A one standard deviation increase in *PASSIVE_EXP* is associated with a 0.096 increase in beta; a one standard deviation increase in residual BMI is associated with an additional 0.057 increase in beta. Thus, a one standard deviation increase in both measures would together be associated with 0.153 (15.3%) increase in beta, relative to the baseline level of 1.

The same mechanism also applies to institutional investors in general, including active funds. Therefore one concern is whether our results is driven by these funds. To check this, we also expand

[6] As an alternative measure we use the difference between the BMI measure and our *PASSIVE_EXP* measure and find qualitatively similar results.

our baseline specification by including institutional ownership and ownership by active funds. Our results are robust in these settings, indicating that our results are not driven by other institutional investors and active funds.

In further analyses, we show that the effects we document are not driven by increased correlations of company fundamentals such as earnings, and the influence of mega stocks in recent years. Consistent with the notion that the rise of passive investing could lead to changes in the information structure in the market, we find that stocks with higher *PASSIVE_EXP* have significantly less idiosyncratic movements around earnings release.

Our paper contributes to the literature on the effect of passive investing on market outcomes. Early papers in this literature focus on stock price movements (first moment) around the rebalancing of stock indices (e.g., Harris and Gurel, 1986; Shleifer, 1986). More recent papers focus on changes in co-movement patterns (second moment) around index rebalancing. An important theoretical paper is Barberis and Shleifer (2003). In their model, excess stock price co-movement arises from the correlated trading by investors who invest in the stocks in a given category, i.e., the common trading by style investors. Supporting this hypothesis, Barberis, Shleifer and Wurgler (2005) find that a stock's beta with the S&P 500 index increases following its inclusion into the index; Greenwood (2008) finds that stocks overweighted in the Nikkei 225 index tend to have higher co-movements with other stocks in the index; Boyer (2011) finds that reclassifying a stock from growth to value within the S&P500 index increases (decreases) its co-movement with the value (growth) index. Our hypothesis is based on the same theoretical foundation as Barberis and Shleifer (2003). The difference between our paper and prior work is that to the best of our knowledge ours is the first that examines the implication of passive investing on the entire universe of stocks and on the aggregate market outcomes, rather than on individual stocks being added/removed from specific indices.

Our paper also contributes to the growing literature on the impact of exchange-traded funds (ETFs). This literature has studied the effects of ETF expansion on the market efficiency, stock price correlation, and volatility of individual stocks (see, e.g., Israeli et al., 2017; Ben-David, Franzoni, and Moussawi, 2018; Da and Shive, 2018). However, this literature has not investigated the effect of ETFs on the aggregate stock market volatility, which is the goal of our paper.

In a recent follow up to their 2001 paper, Campbell et al. (2022) document the same reversal of volatility and correlation patterns since the publication of their earlier paper. The authors did not provide an explanation to this phenomenon but indicated that micro-structure changes in the market is unlikely the explanation. Our paper highlights the rise of passive investing, arguably a most prominent trend in asset management in the last two decades, as a potential explanation.

Finally, there is a growing literature that studies how inelastic investor demand influences stock prices and returns. For instance, Gabaix and Koijen (2021) develop a tractable framework to study how institutional constraints lead to price inelastic demand in stock markets, which results in stock price fluctuations. Pavlova and Sikorskaya (2023) examine how inelastic investor demand arising from institutional performance benchmarking drives up the prices of stocks in the benchmark indices. In a dynamic framework, Parker, Schoar and Sun (2023) study how the portfolio rebalancing by target date funds leads to higher flows into stock markets following stock price declines, which tends to dampen the effect of shocks to the stock market. Our study focuses on the impact of passive investing on the comovement of individual stocks with the stock market and the resulting aggregate volatility effect, which provides interesting evidence resonating with the theme of this literature.

The rest of the paper is organized as follows. Section I describes our data and variables. Section II presents our main empirical results on market volatility, cross-sectional patterns, and difference-in-differences analysis. Section III examines the mechanism - the role of fund flows induced by passive investing. Section IV extends our analysis from passive investing to benchmarking and common holdings in general. Section V examines competing mechanisms and Section VI concludes.

I. Data and Variables

Our paper focuses on risk measures of stocks and the market or the second moment of stock returns, i.e., variances and covariance properties. Our goal is to investigate whether these risk measures are related to the extent of passive investing. For individual stocks, we examine four dependent variables: a stock's beta, its average correlation with all other stocks in the dataset, its average covariance with all other stocks, and its idiosyncratic volatility. Because fund holding and ownership data have a quarterly frequency, we measure the second-moment variables over quarterly horizons for all stocks. We use CRSP's daily stock return file from 1964-2020 for these calculations. A stock's quarterly beta is estimated as the regression coefficient on market excess return in the CAPM model of the stock's daily returns. A stock's average correlation (covariance) with other stocks in a quarter is the equal-weighted average of the stock's correlations (covariances) with all other stocks in the CRSP universe, calculated using daily returns over the quarter. A stock's idiosyncratic volatility in each quarter is the standard deviation of the regression residuals from a Fama-French 4-factor model of the stock's daily returns.

To measure the extent of passive investing at the stock level, we calculate, for each stock in each quarter, the percentage of the stock's market cap that is held in all index funds and index ETFs in our sample. We call this variable PASSIVE_EXP, and it is defined as follows:

$$PASSIVE\_EXP_{i,t} = \frac{1}{MV_{i,t}} \sum_{j=1}^{N} Holdings_{i,j,t} \qquad \text{(Eq. 1)}$$

where $MV_{i,t}$, is stock i's market capitalization in quarter t, and $Holdings_{i,j,t}$ is the dollar amount of holdings of stock i by passive (index mutual fund or ETF) fund j, in quarter t.

Data on index fund and ETF holdings are from Thomson Reuters (TR) S12 Mutual Fund Holdings and CRSP Mutual Fund database. Zhu (2020) points out that 58% of the newly-founded domestic equity mutual funds are missing from the Thomson Reuters database in the recent years. Hence, to maximize our coverage, we only use TR S12 database to calculate the holdings prior to 2010 but rely on the CRSP Mutual Fund database for the post-2010 period. We identify index funds using the index fund flag from CRSP. Our sample of ETFs consists of a manually compiled list of 1,799 ETFs. We restrict our funds to domestic equity funds (Lipper Class in 'EIEI','G','LCCE', 'LCGE', 'LCVE', 'MCCE', 'MCGE', 'MCVE', 'MLCE', 'MLGE', 'MLVE', 'SCCE', 'SCGE', 'SCVE').

Table 1 presents summary statistics of our sample. The average PASSIVE_EXP is 3.8% across all stocks over the entire sample period. But as Figure 4 shows, this measure increased steadily over time, from nearly zero in 1980 and through the 1990s to roughly 10% in 2020. It is also interesting to note that the rise in passive investing did not start in earnest until around year 2000, the period after the sample period of Campbell et al. (2001) paper. The average beta is 1.04 in our sample and the average pair-wise correlation is 12%. The average daily stock volatility is 3.2%.

II. Main Findings

A. Market-level risk analysis

We begin by examining the relation between a market-level measure of the degree of passive investing and measures of market-level risk. To measure the market-level degree of passive investing, we use the total net assets of index funds and index ETFs divided by the total market capitalization of all stocks in the CRSP universe. This ratio indicates the percentage of total market cap invested in passive funds, and we denote it by *Passive-to-Market*. Within our sample period of 1980 to 2020, the average value of *Passive-to-Market* is 1.76% and the standard deviation is 2.30%. The time-series of

Passive-to-Market is plotted in Figure 3 and we observe a clear upward trend throughout years. It reached the maximum of 7.60% in 2020Q1.

We examine five second-moment (i.e., risk) measures of the market return to capture different components that contribute to the overall market risk. These measures distinguish between the part of the market risk that is due to stock-level volatility, and the part that is driven by co-movements among stocks. Specifically:

- The first measure is the market volatility. The literature has found that many factors contribute to market volatility, but at different frequencies. Since the degree of passive investing is a relatively slow-moving variable, we expect it to affect the market volatility at a low frequency. To tease out the low frequency movement of the market volatility, we follow Engle and Rangel (2008) and extract quarterly component of the realized market volatility using Spline-GARCH model (technical details of model estimation appear in Appendix A).
- The second market-level risk measure is the average pairwise correlation among stocks. We first calculate the correlations between a stock and each of the other stocks in our sample (the entire CRSP universe) using daily returns for each month and average the monthly correlations by calendar quarter and average this again across stocks.
- Our third measure is the average pairwise covariance among stocks. This measure is related to the correlation measure and is similarly calculated. But unlike correlation which is scale-free, the magnitude of covariance matters. In the limiting case, market variance is the average covariance of all securities.
- The fourth measure is the average idiosyncratic volatility. Idiosyncratic volatility is defined as the standard deviation of the residuals of individual stock return relative to the Fama-

French 4-factor model. We estimate this model each month for each stock using daily returns and aggregate it by calendar quarter and across stocks.

Finally, as an aggregate measure of the diversification benefit, in some analyses we examine the "volatility gap", which we define as the difference between the quarterly firm-level volatility and market-level volatility. This captures the "benefit of diversification" – i.e., the amount of volatility reduction that is achieved by moving from a randomly chosen single stock to the market portfolio. The larger this measure, the more gain there is to diversification.

Panel A of Table 2 reports the cross-sectional correlation matrix between the *Passive-to-Market* ratio and market-level risk measures. While *Passive-to-Market* is positively correlated with low-frequency market volatility (correlation = 0.327, p-value<1%) and average pairwise correlation (correlation = 0.737; p-value<1%), it is negatively correlated with firm-level total volatility (correlation = -0.413, p-value<1%), idiosyncratic volatility (correlation = -0.567, p-value<1%) and the volatility gap (correlation = -0.608, p-value<1%). The negative correlation with the volatility gap means that as the overall extent of indexing increases, there is a smaller difference between firm-level volatility and market-level volatility; i.e., less benefit to diversification.

Panels B-D present the time-series regressions of the various market-level risk measures on *Passive-to-Market*. In Panel B, we use the full sample period from 1980Q1 to 2020Q4. Panel C and D separates the full sample into pre- and post-1997; the goal is to benchmark with Campbell et al. (2001)'s sample, which is pre-1997. We lag our key independent variable *Passive-to-Market* by one quarter and the standard errors are corrected using Newy-West method with one lag in all specifications.

In the full sample (Panel B), we find that low-frequency market volatility is positively and significantly correlated with lagged *Passive-to-Market* (coefficient = 0.012, *t*-statistic = 2.82), suggesting that the rise of passive investment is associated with higher future market-level volatility. However, the relation between *Passive-to-Market* and the different components of the overall market volatility is very different. Lagged *Passive-to-Market* positively predicts next quarter's average correlation among stocks (coefficient = 2.459, *t*-statistic = 8.20), but it negatively predicts next quarter's average total stock volatility (coefficient = -0.175, *t*-statistic = -4.25) and average idiosyncratic stock volatility (coefficient = -0.205, *t*-statistic = -7.11). This is consistent with the correlation measures presented in Panel A and indicates that passive investing affects different components of the market level risk differently; it contributes to the part of the market risk that arises from comovements among stocks.

Interestingly, sub-period results in Panels C and D show that the above pattern is only found in the post 1997 period (Panel D) when index investing started to grow rapidly and became a major investment strategy. Pre-1997, *Passive-to-Market* is negatively related to overall market volatility (coefficient = -0.768, *t*-statistic = -8.14); it is negatively related to pair-wise correlation (coefficient = -31.485, *t*-statistic = -3.51), but positively related to total and idiosyncratic firm-level volatility. Post-1997, we find passive investing is positively correlated with market volatility and return comovement and negatively correlated with the benefit of diversification. The results are consistent with the hypothesis that at a significant scale passive investing contributing to overall market risk through a correlation/covariance channel instead of a volatility channel.

B. Cross-sectional evidence - Panel regressions

While the time-series relation between the market-level extent of passive investing and market-level risk measures is informative, it cannot rule out that the effect is driven by a missing market trend. To sharpen the identification, we take advantage of the granular firm-level panel data. If the higher market volatility is indeed due to index investing, then firms that are more heavily invested by index funds and ETFs should have stronger comovment with the other stocks in the market.

Table 3 reports results from panel regressions where we regress each of the four stock-level second-moment return measures on one-quarter lagged PASSIVE_EXP. Specifically, we estimate the following panel regressions:

$$Y_{i,t} = b_1 + b_2 * PASSIVE\_EXP_{i,t-1} + controls_{i,t} + \epsilon_{i,t} \qquad \text{(Eq. 2)}$$

where $Y_{i,t}$ is one of the four return second-moment measures: 1) a stock's beta, 2) a stock's average correlation with all other stocks; 3) a stock's average covariance with all other stocks; and 4) its idiosyncratic volatility. The key coefficient of interest is b_2, the coefficient on the lagged *PASSIVE_EXP* measure. For each dependent variable, we show two regression specifications. In the first one we include only the key variable *PASSIVE_EXP*; in the second one we include standard control variables such as a stock's size, book-to-market ratio, momentum, financial leverage, and institutional ownership. To control for firm-invariant traits and commonalities across stocks in a given period, we also include separate firm and year fixed effect for all regressions.

Results in Table 3 show that lagged *PASSIVE_EXP* has a strong and positive predictive power on three of the four dependent variables: beta, average correlation, and average covariance. The coefficient on *PASSIVE_EXP* in these regressions are always positive and highly statistically significant. In the beta regressions (columns (1) and (2)), the coefficient on *PASSIVE_EXP* does not change much between the two specifications. This means that the effect of *PASSIVE_EXP* on beta is stable and largely unaffected by the inclusion of firm-level control variables. In the regressions pertaining to correlations (columns (3) and (4)) and covariances (columns (5) and (6)) the magnitude of the coefficient on PASSIVE_EXP drops slightly when the control variables are included, but they remain highly significant throughout.

In terms of economic magnitude, the estimation implies that a one-standard-deviation increase in lagged *PASSIVE_EXP* (i.e., a 5.5% increase in holdings by passive funds) is associated with a 0.1 increase in beta.[7] Since the global average beta is our sample is 1.04, this means a 10% increase. The magnitude is not only statistically significant but also economically large. Columns (3) and (4) of Table 3 shows that *PASSIVE_EXP* also has a strong and significant predictive power for stocks' average correlation with other stocks. The coefficient on *PASSIVE_EXP* is highly significant across all specifications. Using these coefficients, we estimate that a one-standard-deviation increase in *PASSIVE_EXP* is associated with an increase of 11.7% in the average correlation with other stocks.[8] In contrast, Columns (7) and (8) show that *PASSIVE_EXP* generally has a negative relation with a stock's idiosyncratic volatility, although the result is insignificant at 10% level when we include firm-level control variables. This is consistent with our findings in the previous sub-section for the market level: passive investing seems to contribute to overall market risk through a correlation/covariance channel, rather than a volatility channel.

C. Contribution to Market Risk: Difference-in-differences analysis

To provide more direct evidence that exposure to passive investing contributes to stocks' systematic risk, we conduct a difference-in-differences analysis. We exploit three unanticipated market-level crisis within our sample: the 9/11 terror attack in 2001, the onset of the Global Financial Crisis in Sep 2008, and the Covid 19-led market panic in March 2020. These are periods of significant market turmoil driven by external shocks that no individual firm can cause or anticipate. We sort

[7] The standard deviation of PASSIVE_EXP is 0.055 (Table 1). The average coefficient across the columns (1) and (2) in Table 3 is 1.78. Therefore, a one-standard-deviation increase in PASSIVE_EXP is associated with a 0.055*1.78=0.10 increase in beta.

[8] The standard deviation of *PASSIVE_EXP* is 0.055 (Table 1). The average coefficient across columns (3) and (4) in Table 3 is 0.27. Therefore, a one-standard-deviation increase in *PASSIVE_EXP* is associated with a 0.055*0.28=0.015 (or 11.7%) increase in average correlation.

stocks according to their *PASSIVE_EXP* in the quarter *prior* to these events and examine the changes in their betas and correlation with other stocks, as well as their volatilities during the crisis period relative to before. If high *PASSIVE_EXP* stocks contribute more to the market crisis, we should see these stocks exhibiting larger increases in betas and correlation with other stocks than low *PASSIVE_EXP* stocks. On the other hand, *PASSIVE_EXP* may not affect changes in individual stock volatility.

Table 4 reports the DID results. Panels A, B, C, and D pertain to beta, average correlation, average covariance, and idiosyncratic volatility, respectively. Panel A shows that the average beta of high *PASSIVE_EXP* stocks (quintile 5) increased from 1.117 before the crisis period to 1.256 during the crisis period. In contrast, the average beta of low *PASSIVE_EXP* stocks (quintile 1) *declined* from 0.989 to 0.912. The difference-in-differences is 0.215 with a *t*-statistic of 2.09. Panel B focuses on stocks' average correlation with other stocks. Across all quintiles, the average correlation with other stocks increases during market crisis, but the increase in high *PASSIVE_EXP* stocks is significantly larger. Low *PASSIVE_EXP* stocks' average correlation with other stocks increased from 0.097 to 0.177, an increase of 0.08; whereas high *PASSIVE_EXP* stocks' average correlation with other stocks increased from 0.152 to 0.276, an increase of 0.124, almost 50% larger than the low *PASSIVE_EXP* stocks, and the difference-in-differences between the two group is significant at the 1% level (*t*-statistic = 3.63). Panel C shows that the average covariance of high *PASSIVE_EXP* stocks increased four-fold from 0.00012 to 0.00048 whereas that of low *PASSIVE_EXP* stocks increased from 0.00009 to 0.00032. Both increases are significant at the 1% level and indicates that during crisis stocks tend to comove more together; but the difference in differences is also significant at the 1% level, indicating that the increase in co-movement is higher for stocks with high *PASSIVE_EXP*.

In contrast to the results in Panels A, B and C, in Panel D we find that high *PASSIVE_EXP* stocks do *not* experience larger increases in volatility in crisis periods compared to low *PASSIVE_EXP* stocks. High *PASSIVE_EXP* stocks' average volatility increased from 0.030 to 0.040 during crisis periods, while low *PASSIVE_EXP* stocks' average volatility increased from 0.037 to 0.048 in the same time frame. The difference-in-differences of the two groups is insignificant.

Overall, results from the DiD analysis in Table 4 are consistent with the panel regressions in Table 3 and illustrate that stocks' exposure to passive investing is strongly related to properties connected to stocks' systematic risk measures such as beta, correlation, and covariance with other stocks. However, the extent of passive investing is not strongly correlated with stock-level volatility.

Table 5 examines the difference-in-differences analysis in panel regression setting. We examine beta, average correlation, and idiosyncratic volatility as the dependent variable, respectively. The main independent variables are one-quarter lagged *PASSIVE_EXP*, and its interaction term with the crisis indicator. Columns (1) to (4) show that the lagged *PASSIVE_EXP* is always highly significant in predicting beta and average correlation with other stocks. The interaction term, *PASSIVE_EXP**Crisis is also always positive and highly significant. This means that during crisis period, stocks with high exposure to passive investing exhibit significantly higher *increases* in their betas and average correlations with other stocks than stocks with low exposure, thus contributing more to the overall market volatility during crisis periods.

Columns (7) and (8) examine idiosyncratic volatility. The crisis indicator is positive (though insignificant) in both models, meaning that stocks' idiosyncratic volatilities increase during crisis periods but not salient. However, the interaction term between *PASSIVE_EXP* and the crisis indicator is negative and significant in both models, indicating that stocks with higher *PASSIVE_EXP* experience smaller increases in idiosyncratic volatility. Since the coefficient on *PASSIVE_EXP* itself is smaller than that of the interaction terms, the combined effect is that during crisis period high *PASSIVE_EXP* stocks' idiosyncratic volatility increases less than low PASSIVE_EXP stocks. These results are in contrast to those for beta and average correlations but consistent with results shown earlier in this paper: In general *PASSIVE_EXP* does not positively contribute to stock idiosyncratic volatility but positively contributes to measures of comovement with other stocks.

Overall, results in Tables 3, 4, and 5 present consistent evidence that *PASSIVE_EXP* is strongly predictive of those second-moment stock return measures that are related to systematic risk – beta, correlation, and covariance with other stocks, but it is not highly predictive of stocks' idiosyncratic volatility. Stocks with high exposure to passive investing contribute disproportionally to the overall market risk through larger increases in their correlations with other stocks, but not through higher volatilities.

III. Mechanism – Passive-Flow-induced Trading

In our hypothesis, the main mechanism for indexing to drive systematic risk is correlated trading. This suggests that fund flows into and from passive funds – which drive trades by these funds – should contribute to the relation between risk and exposure to passive investing. To test this mechanism, we estimate the following panel regression:

$$\beta_{i,t} = b_1 + b_2 * PASSIVE\_FLOW_{i,t} + controls_{i,t-1} + \epsilon_{i,t} \qquad \text{(Eq. 3)}$$

Where $PASSIVE\_FLOW_{i,t}$ is the trading of stock *i* induced by the flows to all the index funds and ETFs holding stock *i* during quarter *t.* Specifically, $PASSIVE\_FLOW_{i,t}$ is defined as:

$$PASSIVE\_FLOW_{i,t} = \frac{1}{MV_{i,t-1}} abs\left[\sum_{j=1}^{N}\left(Holdings_{i,j,t-1} \cdot Flow_{j,t}\right)\right] \qquad \text{(Eq. 4)}$$

The measure, $PASSIVE\_FLOW$, takes into account a stock's exposure to passive index funds, index fund flow, and the cross-trading among index funds. First, for each index fund *j* holding stock *i,* we multiply the net percentage flow of fund *j* by the dollar amount that fund *j* holds in stock *i.* This captures the amount of flow-induced trading by fund j in stock *i.* The measure will be positive (negative) if fund *j* has a net inflow(outflow), thus must buy (sell) stock *i.* Next, we add up the flow-induced trading for all the index funds holding stock *i.* This step will net out any cross-trading within the index fund sector due to opposite flows. Finally, since we are not concerned with the direction of trading, we take the absolute value of the net trading by index funds and scale it by the market value of the stock. Overall, the measure reflects the net amount of flow-induced trading that cannot be absorbed within the index fund sector; it is thus a net liquidity demand by index funds to other types of investors.

We regress stock beta on $PASSIVE\_FLOW$, keeping the rest of the specification the same as our baseline regressions in Table 3. The results are presented in Table 6 Panel A. We find that $PASSIVE\_FLOW$ is significantly positively related to the beta of the stock, as indicated by the high *t*-statistics. The economic magnitude is also meaningful: A one-standard-deviation increase in $PASSIVE\_FLOW$ increases the beta of the stock by 0.025, which is 2.5% increase compare to the sample average. These magnitudes are obtained after controlling separately for quarter fixed effects and firm fixed effects and are thus economically significant.

Earlier we find that stocks with higher exposure to index funds experience larger increase in stock beta during the crisis period. To examine whether the passive-flow-induced trading plays a role, we add to the above regression the interaction between $PASSIVE\_FLOW$ and the *Crisis* indicators. The results in Table 6 Panel B show that the effect of flow induced trading on stock beta is particularly strong during the crisis period.

Overall, the analyses in this section provide an important mechanism of why index funds increase the systematic risk of stocks. Correlated trading by index funds due to fund flows generated correlated price impact, increasing the return correlation among individual stocks.

IV. Extension: Indexing vs. Benchmarking and Common Holdings in general

A. Passive Investing vs. Benchmarking

Our paper has focused on index funds and ETFs which have an explicit adherence to indices and therefore are the purest setting to examine the impact of common holdings. However, the literature has long documented implicit forms of indexing known as closet indexing (Cremers and Petajesto, 2009). In particular, Pavlova and Sikorskaya (2023) and Kashyap et al. (2023) argue that benchmarking – a common practice in performance evaluation of mutual funds – creates incentives for fund managers to stay close to the benchmark index, creating inelastic demand for stocks in indices, making active funds less active than in name.

Given the presence of closet indexing, one concern is that we are over-stating the effect of passive investing on risk. Our key input measure *PASSIVE_EXP* captures only explicit passive holdings; but the outcome measures – the 2^{nd} moment return statistics – reflect impacts from active investors too, therefore we might be over-attributing or mis-attributing effects due to closet indexing or benchmarking to passive investing.

First, it is important to note that this possibility is consistent with our central hypothesis. We do not claim the effects we document in this paper to be unique to passive funds. We focus on passive funds because they have emerged as an important asset class and offer the purest setting to study common ownership and correlated trading. But the same mechanism arising from common ownership and correlated trading applies to other forms of implicit indexing, due to benchmarking or any other reason. This means that the true economic effect of both explicit and implicit indexing could far exceed the impact of passive funds alone; what we document could be a lower bound of the total effect arising from common ownership and correlated trading.

To investigate this broader hypothesis, we augment our baseline analysis by adding the Benchmark Intensity Measure (BMI) of Pavlova and Sikorskaya (2023). The BMI measure indicates the total inelastic demand for a stock due to benchmarking by both passive and active funds.[9] Since both *BMI* and *PASSIVE_EXP* indicate the extent of a stock's passive holdings, not surprisingly the two measures are highly correlated; we find a correlation coefficient of 0.61. To address multicollinearity, we thus use the residual *BMI* in the regression along with *PASSIVE_EXP.*[10] If the two measures have qualitatively similar effects on risk measures, we expect them to have the same signs in the regressions. The results are reported in Table 7. For each dependent variable, the first model includes only the residual *BMI* measure and the second model includes both.

We find that by itself, the residual *BMI* measure has positive and significant signs in all three sets of regressions pertaining to measures that reflect co-movement among stocks— beta, correlation, and covariance. It has an insignificant coefficient on total volatility and a negative coefficient on idiosyncratic volatility. These results are qualitatively the same as what we observed for the *PASSIVE_EXP* measure alone. When we add the *PASSIVE_EXP* measure into the model, we find that both the residual BMI and *PASSIVE_EXP* have the same, positive signs on the three dependent variables that reflect co-movement. The two variables' effects on total or idiosyncratic volatility are less robust.

---

[9] Specifically, $BMI_{it} = \sum_{j=1}^{J} \lambda_{jt} w_{ijt}$ where λ_{jt} is the assets under management (AUM) of mutual funds benchmarked to index *j* as a fraction of the total equity market cap in month *t* and w_{ijt} is the weight of stock *i* in index *j* in month *t*.

[10] We first regress BMI on *PASSIVE_EXP* and then use the residual from the estimation in the regression of risk variables.

These results provide strong evidence that similar mechanisms are at work for both *PASSIVE_EXP* and *BMI*, which are common holdings. Whether driven explicitly by indexing or more implicitly by benchmarking, both measures strongly and positively contribute to measures reflecting co-movements. Furthermore, since we utilize the residual *BMI* measure, the results indicate that these two effects are separable and additive. Coefficient estimates in the beta regressions indicate that a one-standard deviation increase in *PASSIVE_EXP* leads to a 0.098 (9.8%) increase in beta, similar to the baseline figures in Table 3. In addition, a one-standard deviation increase in residual BMI would lead to a 0.057 increase in beta.[11] Therefore a one-standard deviation increase in both variables would be associated with a total increase of 0.153 in beta, equivalent to over 15% of the baseline beta of 1.

B. Passive Investing vs. Institutional and Active ownership

The last sub-section shows that PASSIVE_EXP and benchmarking intensity have separate and additive effects on the correlation structures among stocks. A related concern is whether what we document is driven by similar effects from common institutional investor ownership, including ownership by active funds. Anton and Polk (2014) find that common mutual fund holdings have an effect of increasing pair-wise correlation.

It is again important to point out that if general institutional ownership or common ownership among active managers increases correlations and covariances among stocks, such a finding does not contradict our central thesis. Indeed, the same mechanism through common trading should apply to these investors as well, which again means the effects due to passive funds alone is only a lower bound of the total common-holding-induced effect. Nevertheless, to check the robustness of our results and

[11] The standard deviation of *PASSIVE_EXP* is 0.055. Thus, according to the coefficient in equation (2) of Table 7, the increase in beta is 1.787*0.055=0.098. Similarly, the standard deviation of residual BMI is 0.073. Thus, according to the regression coefficient in equation (2), the increase in beta is 0.779*0.073=0.057.

take other types of ownership into account, we use the same augmented regression framework as in the last sub-section with BMI, by including a) orthogonalized institutional ownership and b) the active fund common ownership measure in Anton and Polk (2014).

Table 8 reports these augmented regression results. Panels A and B pertain to the results when we add stocks' orthogonalized institutional ownership measure or active common ownership measure based on Anton and Polk (2014) respectively. [12] From Panel A, it is immediately clear that general institutional ownership does not dissipate the effect we document in the baseline. *PASSIVE_EXP* remains positive and significant in explaining stocks' beta, correlation and covariance with other stocks, but not idiosyncratic volatility. Even quantitatively, the effect of *PASSIVE_EXP* remains similar to the baseline. For instance, Equation (2) shows that the coefficient of *PASSIVE_EXP* is 1.846 in the beta regression, very similar to the coefficient of 1.778 in Table 3, which means its contribution to beta is similar, at about 10%.[13]

Panel B shows that the results are generally robust when controlling for residual active investor common ownership. However the data on active common ownership (the Anton and Polk (2014) measure) is sparse and massively reduces our sample and the results are more noisy. These results show that *PASSIVE_EXP* remains significant in the beta and covariance regressions. Its significance on correlation disappears. Residual active common ownership is significant in the beta, correlation, covariance, and idiosyncratic volatility regressions, indicating that common active ownership has significant effects on the second-moment measures of asset returns, which is broadly consistent with our overall thesis.

[12] Since the Anton and Polk common holdings is a pair-wise measure and our regression is at the stock level, for each stock, we calculate its "average common holding with all other stocks" by averaging all its pair-wise measure with all other stocks and orthogonalized this measure.
[13] The standard deviation of PASSIVE_EXP is 0.055. Therefore a one standard deviation increase in PASSIVE_EXP is associated with 0.055*1.846 = 0.10, or 10%.

V. Alternative Hypotheses

A) Correlation in fundamentals

One alternative explanation for the increased correlations among stocks we observed over time is increased correlation in firm fundamentals: for example, with more integrated supply chains and economic activities, firms earnings are more correlated with one another. For this explanation to drive the effect of *PASSIVE_EXP*, it would mean that stocks with high *PASSIVE_EXP* also have stronger correlation in earnings growth with other stocks.

To test this hypothesis, we estimate panel regressions of firms' quarterly earnings growth on market-wide earnings growth (to proxy for earnings growth in other firms), *PASSIVE_EXP* and the interaction between the two. If correlations in fundamentals drive our results, we would expect that the interaction term between market-wide earnings growth and *PASSIVE_EXP* to be positive and significant in explaining firm level earnings growth. Results in Table 9 reject this hypothesis: the interaction term has no explanatory power for firm-level earnings growth: Earnings of stocks with high passive exposure are not more correlated with market-wide earnings.

B) Influence of Very Large Companies

In the last ten years, the rise of very large firms has attracted research attention. For instance, FAANG stocks, including Meta (previously Facebook), Apple, Amazon, Netflix, and Alphabet (previously Google) collectively account for 19% of the S&P as of August 2021.[14] More recently, the stellar performance of the leading tech stocks, the so-called "Magnificent Seven" drives the returns on

[14] https://investorpedia.com, accessed July 21, 2022.

the S&P 500 index in 2023.[15] If a few stocks dominate the entire market and these large stocks are also broadly held by passive funds, then what we document could be driven by a few very large stocks.

We check the robustness of our results by excluding the top 10 largest stocks by average market capitalization over our sample period and re-estimating the baseline regression in Table 3.[16] The results are reported in Table 10. For brevity, we only report the key coefficients. Results in Table 10 show that all our baseline results in Table 3 hold qualitatively and quantitatively after excluding large stocks from the sample. *PASSIVE_EXP* continues to be a highly significant predictor of firms' beta and average correlation with other stocks, but not of idiosyncratic volatility. The coefficients' magnitudes are also similar to those of Table 3.

C) Firm-specific vs. market information discovery

If exposure to passive investing makes stocks more correlated to other stocks and by implication the overall stock market, it follows that price movement of stocks highly exposed to passive investing can reflect more market-level movements rather than idiosyncratic firm-level information during information sensitive times such as around earnings announcements. This would indicate that price discovery is less efficient to firm-specific information.

To investigate this hypothesis, we examine total and idiosyncratic volatility of stocks around earnings announcements. Table 11 reports our findings. In this analysis, we examine idiosyncratic volatility in the period between [-5, +5] days of earnings announcement. We sort stocks by their *PASSIVE_EXP* in the quarter prior to the earnings announcements and tabulate the average volatility

[15] https://www.wsj.com/livecoverage/fed-meeting-interest-rate-decision-fomc-november-2023/card/thanks-magnificent-seven-tech-stocks-UXz0V603bOawqvMNfWn4

[16] The 10 excluded stocks are: Apple, Microsoft, Amazon, Tesla, Facebook, Alphabet, Johnson & Johnson, JP Morgan Chase & Co, as of December 2020.

by quintiles. We do this for the whole sample, as well as two subsamples: from 1980 to 1997; and from 1998 to 2020.

Table 11 shows that overall *PASSIVE_EXP* is negatively related to idiosyncratic volatility: High *PASSIVE_EXP* stocks have average idiosyncratic volatility of 0.031 around earnings announcement, compared to low *PASSIVE_EXP*'s average of 0.034. This difference is highly significant. This means that there is less stock-specific movements or information discovery for highly index stocks. This result is consistent with our findings in Tables 3, 4, and 5, which show that *PASSIVE_EXP* has either a negative, or at most insignificant relation with idiosyncratic volatility.

The sub-sample analysis reveals however that this result is entirely driven by the second half of our sample: from 1998 to 2020. In this period, the idiosyncratic volatility of the high *PASSIVE_EXP* stocks is 0.030, 20% lower than low *PASSIVE_EXP* stocks' average idiosyncratic volatility of 0.036, with the difference being highly significant. In the earlier sample of 1980-1997, the relationship between *PASSIVE_EXP* and idiosyncratic volatility is the reverse: High *PASSIVE_EXP* stocks (which tend to be larger) have a higher average idiosyncratic volatility of 0.034, compared to 0.028 for stocks with low *PASSIVE_EXP*. Looking across the different sample period, we see that high PASSIVE_EXP stocks' idiosyncratic volatility *dropped* from 0.034 to 0.030 while low *PASSIVE_EXP* stocks' idiosyncratic volatility increased from 0.028 to 0.036. This result is consistent with our hypothesis that there is less firm-specific price discovery with the rise of passive investing.

VI. Conclusion

In the past twenty years, one of the most salient trends in the asset management industry is the rise of passive investing. In this paper, we provide evidence pointing to a potential downside of the ever-rising extent of passive investing: as a popular strategy with many followers, it can undo its own benefits of diversification and lead to increased market-level volatility.

We document that since around 2000, concurrent with the rise of passive investing, there has been a strong trend of higher market-level volatility, which is driven by higher correlations among individual firms. We construct a firm-level measure of its exposure to indexing, and find that this measure is highly related to the systematic (aka, undiversifiable, or, beta) portion of a stock's risk; and it is unrelated to a stock's idiosyncratic risk. Examining three episodes of sudden and largely exogenous rise in market volatility – the post 9/11 period, the 2008 financial crisis, and the 2020 Covid pandemic – we find that it is the stocks that have high exposure to indices that contributed strongly to the increased market-wide volatilities in these periods. We provide evidence that our observed concurrence between index exposure and higher systematic risk is not explained by increased correlation in fundamentals such as firm earnings; it is not subsumed by general, including active institutional holdings; and it is not driven by the influence of very large firms. We also find that information discovery around earnings announcements also became more driven by market movements rather than idiosyncratic information, the more a stock is exposed to indexing.

Our results raise a number of additional questions such as a theoretical estimation of the "maximum" market-wide volatility and loss of diversification benefit that could result from indexing. We leave these interesting and ambitious questions to future research.

References

Anton, Miguel, and Christopher Polk, 2014, Connected stocks, *Journal of Finance* 69, 1099-1127.

Azar, Jose, Martin C. Schmalz, and Isabel Tecu, 2018, Anticompetitive effects of common ownership, *Journal of Finance* 73, 1513-1563.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing. *Journal of Financial Economics* 68, 161-199.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, Comovement. *Journal of Financial Economics* 75, 283-317.

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

Bogle, John C., Bogle sounds a warning on index funds, the *Wall Street Journal*, Nov 29, 2018.

Boyer, Brian H., 2011, Style-related comovement: fundamentals or labels?. *Journal of Finance* 66, 307-332.

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

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and YeXiao Xu, 2022, Idiosyncratic equity risk two decades later, *NBER Working Paper.*

Cremers, Martijn, and Antti Petagisto, 2009, How active is your mutual fund manager? A new measure that predicts performance, *Review of Financial Studies* 29, 3329-3365.

Da, Zhi, and Sophie Shive, 2018, Exchange traded funds and asset return correlations. *European Financial Management* 24, 136-168.

Gabaix, Xavier and Koijen, Ralph S. J., 2021, In Search of the Origins of Financial Fluctuations: The Inelastic Markets Hypothesis, *NBER Working Paper.*

Grossman, Sanford J., and Joseph Stiglitz, 1980, On the impossibility of informationally efficient market, *Journal of Finance* 70, 393-408.

Greenwood, Robin, 2008, Excess comovement of stock returns: Evidence from cross-sectional variation in Nikkei 225 weights. *Review of Financial Studies* 21, 1153-1186.

Harris, Lawrence, and Eitan Gurel, 1986, Price and volume effects associated with changes in the S&P 500 list: New evidence for the existence of price pressures. *Journal of Finance* 41, 815-829.

Israeli, Doron, Charles 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.

Kashyap, Anil K., Natalia Kovrijnykh, Jian Li, and Anna Pavlova, 2023, Is there too much benchmarking in asset management? *American Economic Review* 113, 1112-1141.

Parker, Jonathan A., Antoinette Schoar, and Yang Sun, 2023. Retail financial innovation and stock market dynamics: The case of target date funds. *Journal of Finance*, 78(5), 2673-2723.

Pavlova, Anna, and Taisiya Sikorskaya, 2023, Benchmarking intensity, Review of Financial Studies 36, 859-903.

Pirinsky, Christo, and Qinghai Wang, 2006, Does corporate headquarters location matter for stock returns? *Journal of Finance* 61, 1991-2015.

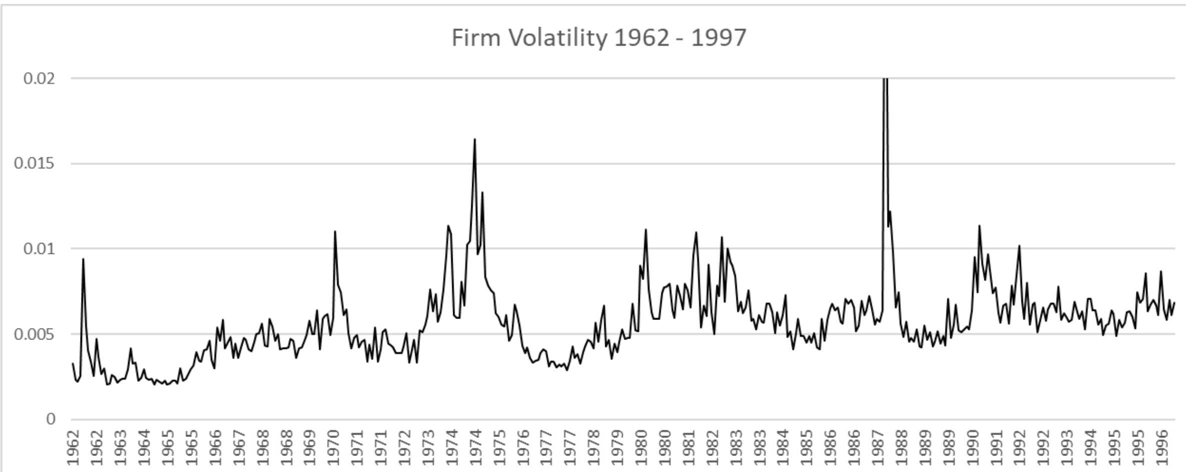
Shleifer, Andrei, 1986, Do demand curves for stocks slope down? *Journal of Finance* 41, 579-590.

Zhu, Qifei, 2020, The missing new funds. *Management Science* 66, 1193-1204.

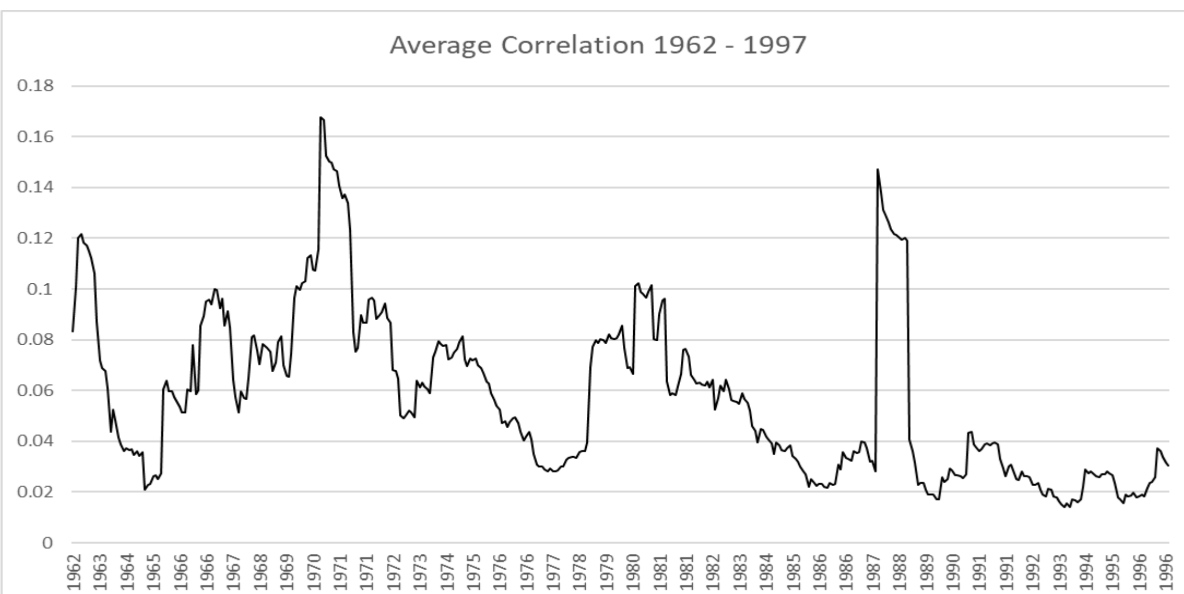
Figure 1. Replicating Campbell et al. (2001): 1962-1997

This figure replicates the results in Campbell et al. (2001). Panel A replicates Figure 4(A) of Campbell et al. (2001) and plots the average annualized firm-level return variance within each month calculated from daily stock returns. Panel B replicates Figure 5(A) of Campbell et al. (2001) and plots the equal-weighted pairwise correlation across stocks traded on NYSE, AMEX, and Nasdaq. Panel C replicates Figure 2(A) of Campbell et al. (2001) and plots the annualized market return variance within each month calculated using daily market returns.

A. Average firm-level return variance



B. Average pair-wise return correlation



C. Market return variances

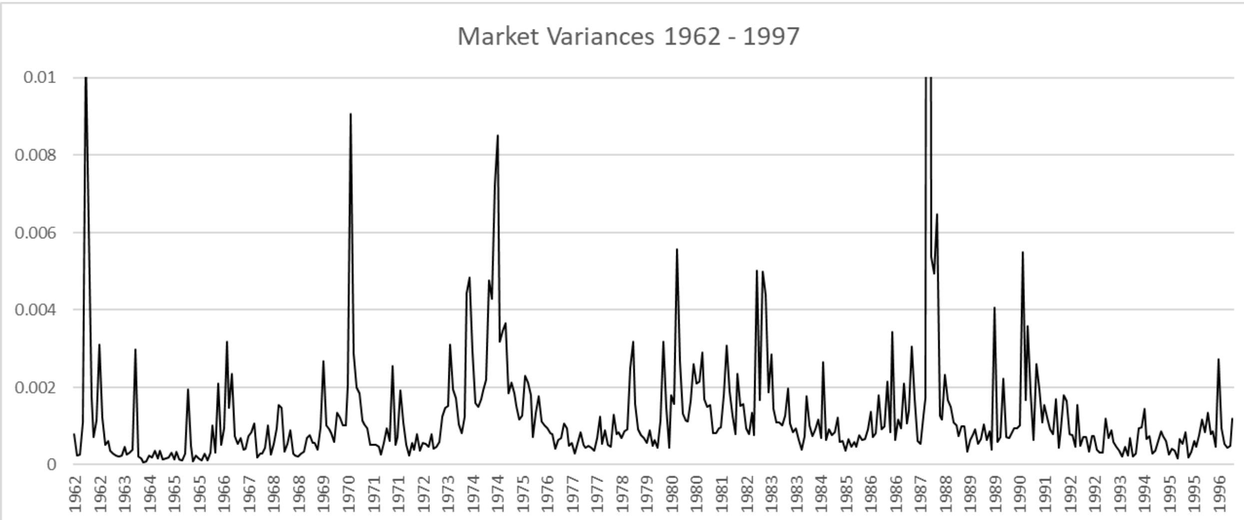
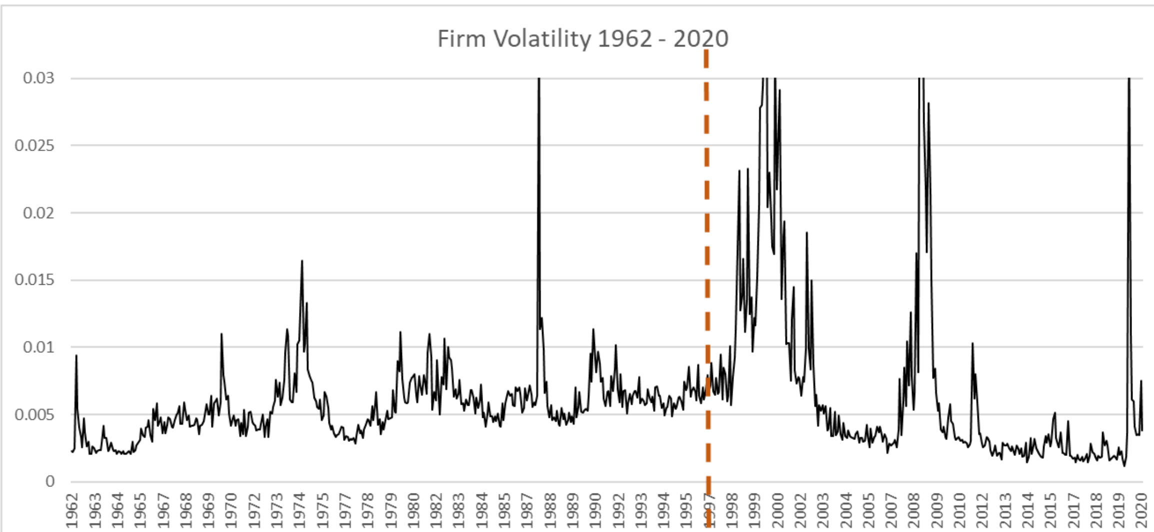


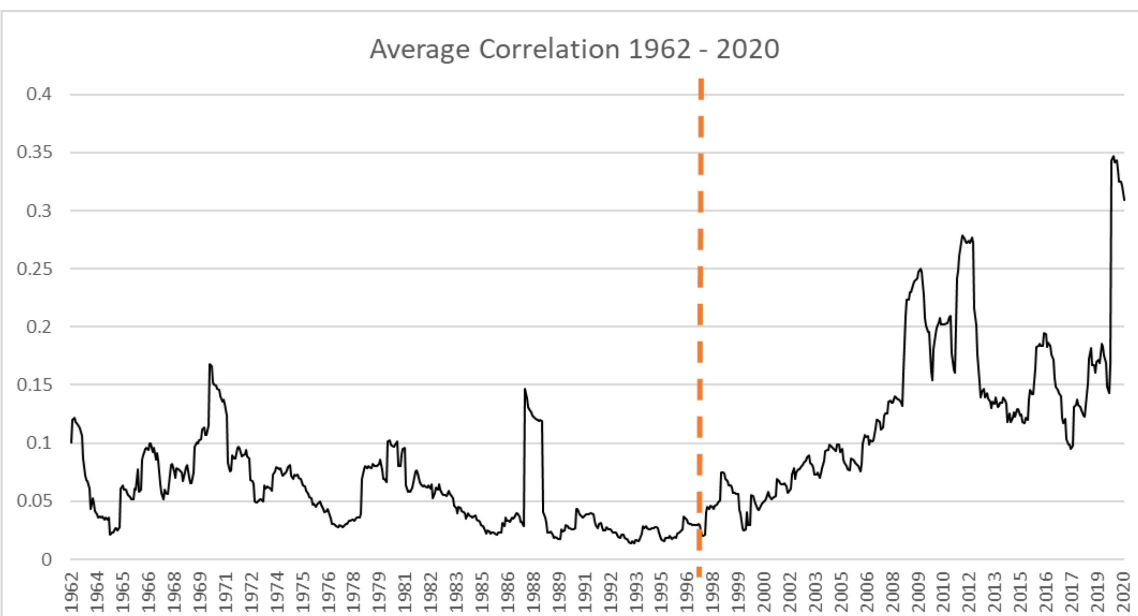
Figure 2. Extending Campbell et al. (2001): 1962 – 2020

This figure replicates the analyses in Campbell et al. (2001) with the sample period extended to 2020. Panel A replicates and extends figure 4(A) of Campbell et al. (2001) and plots the average annualized firm-level return variance within each month calculated from daily stock returns. Panel B replicates Figure 5(A) of Campbell et al. (2001) and plots the equal-weighted pairwise correlation across stocks traded on NYSE, AMEX, and Nasdaq. Panel C replicates Figure 2(A) of Campbell et al. (2001) and plots the annualized market return variance within each month calculated using daily market returns. The dashed line divides the pre- and post-Campbell (2001) sample periods.

A. Average firm-level return variance



B. Average pair-wise return correlation



C. Market Variances

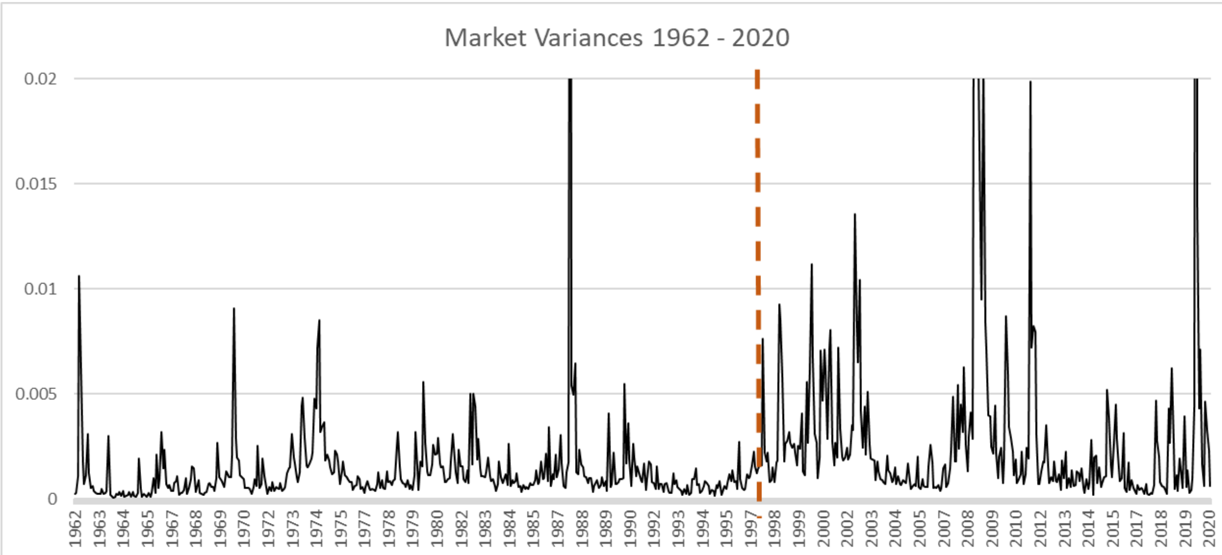


Figure 3. Passive-to-Market ratio and Average Pairwise Correlation

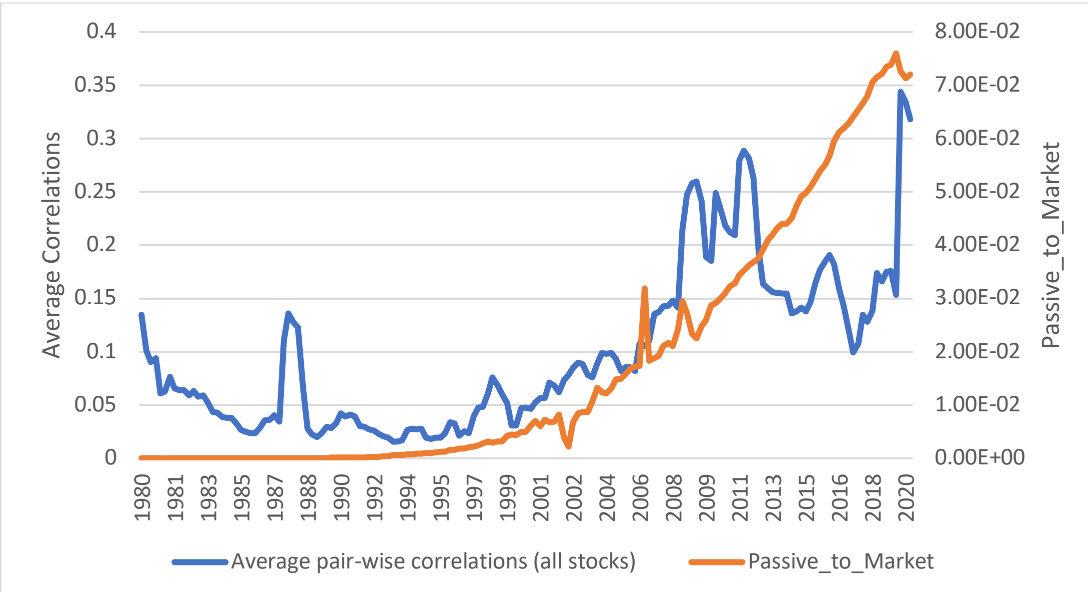


Figure 4. Average *PASSIVE_EXP* Measure over time

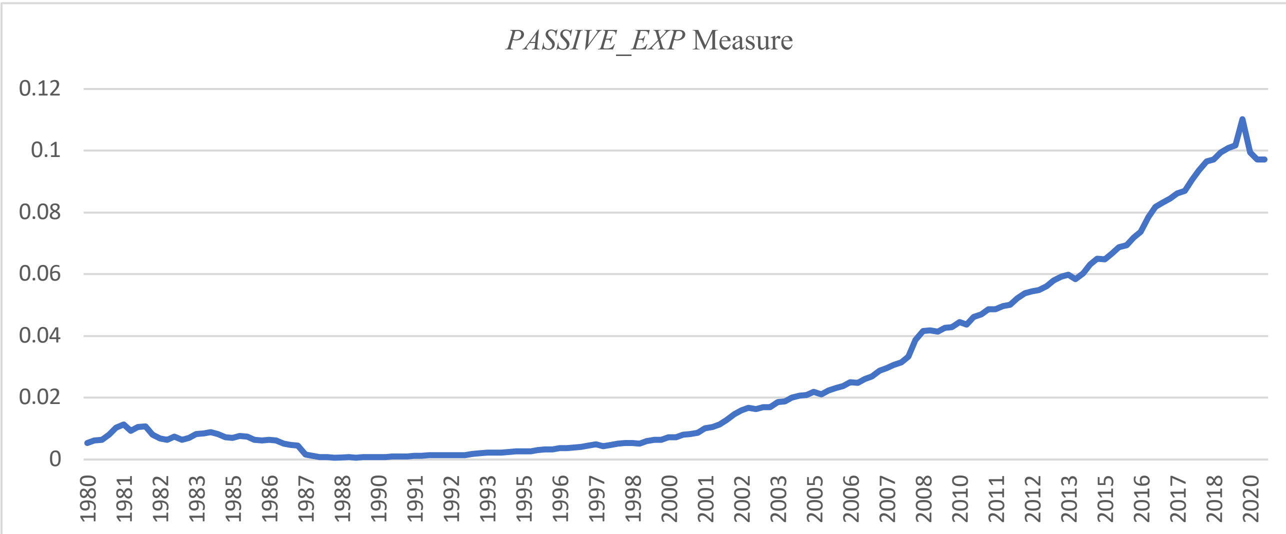


Table 1. Summary Statistics

This table presents summary statistics of our sample. *PASSIVE_EXP* is calculated according to Eq. (1); it is the percentage of a stock's total market capitalization held by all index and ETF funds. *PASSIVE_FLOW* is calculated according to Eq. (4); it is the trading of stocks induced by the flows to all the index funds and ETFs holding this stock. Beta is the quarterly beta coefficient calculated using daily stock returns. Correlation is the average quarterly correlation of each stock with all the other stocks in the CRSP universe calculated using daily stock returns. Covariance is the average quarterly covariance of each stock with all the other stocks in the CRSP universe, calculated using daily stock returns. Volatility is the quarterly stock volatility calculated using daily returns. Size is measured as the market capitalization of equity (in USD millions). BM is the ratio between book-value of equity and market value of equity. MOM is the 12-month momentum factor with one-month reversal. IO is the total institutional ownership as percentage of shares outstanding. Leverage is the long-term and short-term debt, divided by total assets. All variables are winsorized at 1% and 99% level. *, **, *** denote statistical significance at the 10%, 5%, and 1%, respectively.

Panel A: Summary Stats

	Mean	Median	Min	Max	Std	Obs
PASSIVE_EXP	0.038	0.013	0.000	0.845	0.055	474,483
PASSIVE_FLOW	0.001	0.000	0.000	0.158	0.002	426,514
Beta	1.044	0.932	-16.904	24.970	1.306	474,483
Correlation	0.128	0.102	-0.684	0.693	0.106	474,707
Covariance	0.0001	0.0001	-0.002	0.005	0.0002	474,714
Volatility	0.031	0.025	0.000	2.316	0.025	475,986
Idiosyncratic Volatility	0.024	0.019	0.000	2.001	0.021	475,773
Size	3139.138	444.736	8.149	63634.87	8867	433,330
BM	0.872	0.565	0.039	11.087	1.358	433,330
MOM	0.127	0.067	-0.822	2.497	0.528	433,330
Leverage	0.211	0.174	0.000	0.769	0.192	401,319
IO	0.502	0.506	0.001	1.000	0.305	415,045

Panel B: Correlations Matrix

	PASSIVE_EXP	Size	BM	MOM	Leverage	IO
PASSIVE_EXP	1.000					
Size	0.029***	1.000				
BM	-0.081***	-0.087***	1.000			
MOM	-0.026***	0.047***	-0.170***	1.000		
Leverage	0.065***	0.081***	0.065***	-0.057***	1.000	
IO	0.500***	0.173***	-0.230***	0.077***	0.081***	1.000

Table 2. Extent of Passive Investing and Market-level Risk

This table presents the analysis on the relation between market-wide second-moment measures and a measure of the overall level of passive investing in the market. We examine five market-wide second moment measures: overall market volatility (Market_lvol), the average pair-wise correlation (Avg_Corr), the average stock volatility (Avg_Firmvol), the average idiosyncratic volatility (Avg_idio_vol), and the volatility gap (Vol_gap) defined as the difference between the quarterly firm-level volatility and market-level volatility. Mkt_lvol is the low-frequency market-level volatility calculated using Spline-Garch model (Engle and Rangle (2008)). To measure the overall level of passive investing, we use the total amount invested in index funds and ETFs, divided by the total market cap of all stocks (Passive_to_Market). The sample period is 1980-2020 with quarterly frequency. For each of the risk measures, we take simple averages of the monthly values by calendar quarter to create a quarterly data series that matches with the holdings data. Panel A reports cross-sectional correlation between the variables. Panel B reports the predictive regressions of each of the market-wide risk measure on the extent of passive investing. Panel C and D represent the sub-sample analysis for pre- and post-1997 period, respectively. Standard errors are corrected with Newy-West method of lag one. *, **, *** denote statistical significance at the 10%, 5%, and 1%, respectively.

Panel A: Cross-sectional correlation matrix

	Mkt_lvol	Avg_Corr	Avg_Firmvol	Avg_idio_vol	Vol_gap	Passive_to_Market
Mkt_lvol	1					
Avg_Corr	0.642***	1				
Avg_Firmvol	-0.167**	-0.285***	1			
Avg_idio_vol	-0.335***	-0.480***	0.957***	1		
Vol_gap	-0.434***	-0.578***	0.832***	0.943***	1	
Passive_to_Market	0.327***	0.737***	-0.413***	-0.567***	-0.608***	1

Panel B: Market-level, time-series regressions

VARIABLES	Mkt_lvol_t	Avg_Corr_t	$Avg\_Firmvol_t$	$Avg\_idio\_vol_t$	Vol_gap_t
$Passive\_to\_Market_{t-1}$	0.012***	2.459***	-0.175***	-0.205***	-0.204***
	(2.82)	(8.20)	(-4.25)	(-7.11)	(-8.06)
Constant	0.009***	0.056***	0.035***	0.030***	0.026***
	(115.46)	(11.51)	(33.12)	(33.14)	(30.39)
Observations	163	163	163	163	163
R-squared	0.107	0.547	0.169	0.32	0.37

Panel C: Time-series regressions before 1997

VARIABLES	Mkt_lvol_t	Avg_Corr_t	$Avg\_Firmvol_t$	$Avg\_idio\_vol_t$	Vol_gap_t
$Passive\_to\_Market_{t-1}$	-0.768***	-31.485***	5.849***	6.191***	8.501***
	(-8.14)	(-3.51)	(3.46)	(3.76)	(4.08)
Constant	0.009***	0.053***	0.033***	0.028***	0.025***
	(198.92)	(8.79)	(27.48)	(26.56)	(23.06)
Observations	67	67	67	67	67
R-squared	0.641	0.146	0.105	0.153	0.247

Panel D: Time-series regressions after 1997

VARIABLES	Mkt_lvol_t	Avg_Corr_t	$Avg\_Firmvol_t$	$Avg\_idio\_vol_t$	Vol_gap_t
$Passive\_to\_Market_{t-1}$	0.011**	2.015***	-0.218***	-0.229***	-0.194***
	(2.09)	(5.65)	(-3.71)	(-5.36)	(-5.76)
Constant	0.009***	0.078***	0.037***	0.031***	0.026***
	(48.61)	(7.42)	(17.02)	(17.66)	(17.92)
Observations	96	96	96	96	96
R-squared	0.052	0.378	0.192	0.334	0.367

Table 3 – Panel Regressions of Risk Measures on PASSIVE_EXP

This table reports panel regression results of stocks' beta, average pair-wise correlation with other stocks, average pair-wise covariance with other stocks, idiosyncratic volatility and average covariance with the market on lagged PASSIVE_EXP. The sample period is 1980-2020. Beta is measured from a market model using daily stock returns for each quarter. Corr is the equal-weighted average of a stock's pair-wise daily-return correlation with all other stocks in the CRSP universe in a quarter. Cov is the equal-weighted average of a stock's pair-wise daily-return covariance with all other stocks in the CRSP universe in a quarter (scaled up by 1,000). Vol is a stock's volatility, measured as the standard deviation of daily stock returns in a quarter. In all panels, the main independent variable is PASSIVE_EXP$_{t-1}$, stocks' one-quarter lagged exposure to passive investing, calculated using Eq. (1) and measures the percentage of a stock's market capitalization that is held by all index funds and ETFs in our sample. Size is the natural log of quarter-end price times total shares outstanding. BM is the book-to-market ratio. MOM is the 12-month momentum factor with one-month reversal. IO is the total institutional ownership as percentage of shares outstanding. Leverage is the long-term and short-term debt, divided by total assets. All variables are winsorized at 1% and 99% level. All firm-level control variables are lagged by one quarter. Year and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. *t*-statistic are in parenthesis. ***, ** and * indicate significant levels at 1%, 5% and 10% respectively.

	Beta	Beta	Corr	Corr	Cov	Cov	Idio_Vol	Idio_Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PASSIVE_EXP$_{t-1}$	1.789***	1.778***	0.317***	0.229***	0.298***	0.228***	-0.019***	0.001
	(7.00)	(5.99)	(10.46)	(7.52)	(5.10)	(3.46)	(-5.96)	(0.51)
Size$_{t-1}$		0.040**		0.017***		0.007**		-0.005***
		(2.39)		(16.41)		(2.35)		(-17.93)
BM$_{t-1}$		0.097***		-0.001		0.010***		0.003***
		(6.71)		(-0.94)		(3.83)		(10.01)
MOM$_{t-1}$		0.101***		-0.008***		-0.006		0.000
		(3.45)		(-6.69)		(-1.17)		(1.47)
Leverage$_{t-1}$		0.189***		-0.005**		0.013**		0.008***
		(4.37)		(-2.04)		(2.25)		(9.54)
IO$_{t-1}$		0.024		0.028***		0.021**		-0.002***
		(0.60)		(5.20)		(2.06)		(-3.42)
Constant	0.976***	0.584***	0.118***	-0.002	0.101***	0.032**	0.025***	0.049***
	(62.32)	(5.98)	(51.78)	(-0.32)	(23.22)	(2.21)	(90.48)	(30.99)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	456,772	377,822	457,215	377,852	457,215	377,852	458,053	343,560
R-squared	0.160	0.169	0.800	0.816	0.692	0.706	0.488	0.562

Table 4 – Difference-in-Differences Analysis around Crisis Period: Univariate Sorting

This table reports difference-in-differences analyses of changes in stocks risk measures from before-crisis to crisis periods. We sort stocks based on their before-crisis *PASSIVE_EXP* measures and compare their before-crisis and during-crisis risk measures. The three risk measures we examine are beta (Panel A), average correlation with other stocks (Panel B), average covariance with other stocks (Panel C) and idiosyncratic volatility (Panel D). The three crisis periods are 2001Q3-2001Q4, 2008Q4-2009Q3, and 2020Q2-2020Q4. The three before-crisis periods are 2000Q3-2001Q2, 2007Q4-2008Q3, and 2019Q1-2020Q1.

Panel A: Beta

	Before Crisis	During Crisis	During - Before	*p*-value
Quintile 1 (low)	0.989	0.912	-0.076	0.40
Quintile 2	1.118	1.043	-0.075	0.40
Quintile 3	1.141	1.146	0.005	0.95
Quintile 4	1.080	1.208	0.128	0.12
Quintile 5 (high)	1.117	1.256	0.139	0.16
High - Low	0.128	0.343	0.215	0.05

Panel B: Correlation

	Before Crisis	During Crisis	During - Before	*p*-value
Quintile 1 (low)	0.097	0.177	0.080	0.01
Quintile 2	0.087	0.173	0.086	0.00
Quintile 3	0.131	0.242	0.111	0.00
Quintile 4	0.146	0.264	0.118	0.01
Quintile 5 (high)	0.152	0.276	0.124	0.01
High - Low	0.055	0.099	0.044	0.00

Panel C: Covariance

	Before Crisis	During Crisis	During - Before	*p*-value
Quintile 1 (low)	0.00009	0.00032	0.00023	0.00
Quintile 2	0.00011	0.00037	0.00025	0.00
Quintile 3	0.00013	0.00046	0.00033	0.00
Quintile 4	0.00013	0.00049	0.00036	0.00
Quintile 5 (high)	0.00012	0.00048	0.00036	0.00
High - Low	0.00002	0.00016	0.00014	0.00

Panel D: Idiosyncratic Volatility

	Before Crisis	During Crisis	During - Before	*p*-value
Quintile 1 (low)	0.034	0.043	0.008	0.05
Quintile 2	0.037	0.044	0.007	0.06
Quintile 3	0.028	0.033	0.004	0.26
Quintile 4	0.025	0.029	0.004	0.23
Quintile 5 (high)	0.023	0.027	0.003	0.28
High - Low	-0.011	-0.016	-0.005	0.01

Table 5. Difference-in-Difference Regression around Crisis Periods

This table reports the results of difference-in-difference regressions of stocks' risk measures around market crises. We analyze beta, average correlation with other stocks, average covariance and volatility as the risk measure, respectively. The main independent variable is the interaction term between lagged *PASSIVE_EXP* and the indicator variable for the crisis period. *PASSIVE_EXP* is calculated as Eq. (1) and measures the fraction of shares outstanding held by all index funds and ETFs in our sample. The three separate crisis periods are: 2001Q3-2001Q4, 2008Q4-2009Q3, and 2020Q2-2020Q4. The three corresponding before-crisis periods are 2000Q3- 2001Q2, 2007Q4 to 2008Q3, and 2019Q1 to 2020Q1. Crisis is an indicator variable that equals one for the crisis periods and zero for the before-crisis periods. Size is the quarte-end price times total shares outstanding and we take log transformations. BM is the book-to-market ratio. MOM is the 12-month momentum factor with one-month reversal. IO is the total institutional ownership as percentage of shares outstanding. Leverage is the long-term and short-term debt, divided by total assets. All variables are winsorized at 1% and 99% level. Year and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. We report *t*-statistic in parenthesis. ***, ** and * indicate significant levels at 1%, 5% and 10%.

	Beta	Beta	Corr	Corr	Cov	Cov	Idio_Vol	Idio_Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PASSIVE\_EXP_{t-1}$*$Crisis_t$	1.528***	1.686**	0.507***	0.620***	1.401***	1.606***	-0.045**	-0.058***
	(3.08)	(2.21)	(3.79)	(3.84)	(4.27)	(3.52)	(-2.09)	(-2.93)
$PASSIVE\_EXP_{t-1}$	1.183**	1.190**	0.005	0.008	-0.273	-0.086	0.012	0.032**
	(2.81)	(2.34)	(0.06)	(0.09)	(-1.71)	(-0.69)	(0.95)	(2.69)
$Crisis_t$	-0.140**	-0.158**	0.060***	0.045**	0.132**	0.086*	0.004	0.004
	(-2.22)	(-2.34)	(3.11)	(2.45)	(2.64)	(1.78)	(0.82)	(0.85)
$Size_{t-1}$		0.027		0.020***		0.014*		-0.006***
		(0.52)		(7.83)		(2.04)		(-9.80)
BM_{t-1}		0.129***		0.000		0.020***		0.003***
		(5.07)		(0.37)		(5.08)		(7.33)
MOM_{t-1}		-0.070		-0.012***		-0.020		0.001
		(-0.78)		(-3.88)		(-1.34)		(1.30)
$Leverage_{t-1}$		0.151		-0.013**		0.034		0.014***
		(1.14)		(-2.53)		(1.39)		(5.89)
IO_{t-1}		-0.078		0.029*		-0.005		-0.002
		(-0.67)		(1.82)		(-0.20)		(-1.20)
Constant	1.055***	0.779**	0.136***	-0.010	0.180***	0.061*	0.030***	0.061***
	(26.13)	(2.71)	(13.08)	(-0.67)	(7.12)	(1.74)	(13.98)	(18.24)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	87,895	67,555	87,944	67,563	87,944	67,563	88,082	71,514
R-squared	0.278	0.305	0.850	0.856	0.763	0.785	0.504	0.536

Table 6. Mechanisms – Passive Fund Flows and Risk

The sample spans from 1991 to 2020, with quarterly frequency. We analyze beta, average correlation with other stocks, average covariance and volatility as the risk measure, respectively. PASSIVE_FLOW is defined as the absolute value of the total holdings from passive funds times the percentage of fund flows during the current quarter, divided by the total market value of the stocks. The raw fund flows are winsorized at 5% and 95% level. Panel A reports the results on panel regressions and Panel B is on difference-in-differences regressions. We include a series of control variables. Size is the quarter-end price times total shares outstanding and we take log transformations. BM is the book-to-market ratio. MOM is the 12-month momentum factor with one-month reversal. IO is the total institutional ownership as percentage of shares outstanding. Leverage is the long-term and short-term debt, divided by total assets. All variables are winsorized at 1% and 99% level. We also lag these firm-level control variables by one quarter. Year and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. We report t-statistic in parenthesis. ***, ** and * indicate significant levels at 1%, 5% and 10%.

Panel A: Panel Regressions

	Beta	Beta	Corr	Corr	Cov	Cov	Idio_Vol	Idio_Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(5)	(6)
PASSIVE_FLOW$_t$	10.876**	12.312***	1.629**	1.148	1.841	1.727	-0.240***	-0.149*
	(2.57)	(2.91)	(2.21)	(1.33)	(1.09)	(0.91)	(-3.22)	(-1.80)
Size$_{t-1}$		0.039**		0.020***		0.009***		-0.005***
		(2.09)		(20.58)		(3.02)		(-16.38)
BM$_{t-1}$		0.098***		-0.000		0.012***		0.003***
		(6.07)		(-0.36)		(4.16)		(9.54)
MOM$_{t-1}$		0.094***		-0.010***		-0.009*		0.001**
		(2.96)		(-7.23)		(-1.67)		(2.15)
Leverage$_{t-1}$		0.189***		-0.006**		0.021**		0.008***
		(4.06)		(-2.02)		(2.56)		(9.00)
IO$_{t-1}$		0.076		0.029***		0.016**		-0.001***
		(1.65)		(8.23)		(2.52)		(-3.15)
Constant	1.043***	0.616***	0.134***	-0.002	0.116***	0.041**	0.024***	0.049***
	(67.67)	(5.47)	(62.18)	(-0.42)	(26.15)	(2.45)	(77.06)	(28.31)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	424,349	313,710	424,664	313,718	424,666	313,718	425,440	313,728
R-squared	0.164	0.182	0.778	0.811	0.678	0.700	0.483	0.556

Panel B: Difference-in-Difference Regression

	Beta	Beta	Corr	Corr	Cov	Cov	Idio_Vol	Idio_Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PASSIVE\_FLOW_t$*Crisis	50.712***	52.841***	9.826***	10.979***	0.034***	0.036***	-0.889	-1.036
	(4.79)	(5.09)	(3.49)	(3.19)	(5.40)	(5.14)	(-1.66)	(-1.61)
$PASSIVE\_FLOW_t$	-22.627***	-26.870***	-3.335	-4.675	-0.011**	-0.013*	0.107	0.219
	(-2.99)	(-3.16)	(-1.40)	(-1.44)	(-2.23)	(-2.06)	(0.22)	(0.35)
Crisis	-0.138*	-0.162**	0.070***	0.065***	0.000**	0.000**	0.004	0.003
	(-1.96)	(-2.63)	(2.91)	(3.00)	(2.60)	(2.58)	(0.65)	(0.58)
$Size_{t-1}$		0.023		0.020***		0.000*		-0.006***
		(0.49)		(8.17)		(1.89)		(-9.77)
BM_{t-1}		0.122***		0.001		0.000***		0.003***
		(4.83)		(0.42)		(5.79)		(6.99)
MOM_{t-1}		-0.102		-0.017***		-0.000*		0.001
		(-1.22)		(-4.20)		(-1.84)		(1.21)
$Leverage_{t-1}$		0.195		-0.009		0.000**		0.014***
		(1.59)		(-1.69)		(2.61)		(5.98)
IO_{t-1}		0.033		0.026***		-0.000		-0.002
		(0.32)		(3.34)		(-0.11)		(-1.29)
Constant	1.153***	0.845***	0.142***	0.006	0.000***	0.000**	0.031***	0.061***
	(27.00)	(3.16)	(11.02)	(0.36)	(5.80)	(2.27)	(15.12)	(17.90)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	87,863	70,839	87,929	70,842	87,929	70,842	88,076	70,842
R-squared	0.277	0.293	0.835	0.848	0.749	0.767	0.494	0.527

Table 7. Passive Investing vs. Benchmarking

This table reports augmented regressions results by adding to the baseline regression (Table 3) orthogonalized measures of BMI (Benchmark Intensity measure of Pavlova and Sikorskaya (2023)). The data is from 1998 to 2020. $PASSIVE\_EXP_{t-1}$ is stocks' one-quarter lagged exposure to passive investing, calculated using Eq. (1) and measures the percentage of a stock's market capitalization that is held by all index funds and ETFs in our sample. BMI_Resid_{t-1} is the one-quarter lagged residual from a panel regression of stocks' BMI measure on their *PASSIVE_EXP* measure. Size is the quarter-end price times total shares outstanding and we take log transformations. BM is the book-to-market ratio. MOM is the 12-month momentum factor with one-month reversal. IO is the total institutional ownership as percentage of shares outstanding. Leverage is the long-term and short-term debt, divided by total assets. All variables are winsorized at 1% and 99% level and lagged by one quarter. Year and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. We report *t*-statistics in parenthesis. ***, ** and * indicate significant levels at 1%, 5% and 10%.

	Beta	Beta	Corr	Corr	Cov	Cov	Idio_Vol	Idio_Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI_Resid_{t-1}	0.641***	0.779***	0.185***	0.198***	0.239***	0.256***	-0.006***	-0.006***
	(4.70)	(5.69)	(12.90)	(13.41)	(5.42)	(5.59)	(-4.32)	(-4.35)
$PASSIVE\_EXP_{t-1}$		1.787***		0.166***		0.205***		0.002
		(4.86)		(7.36)		(4.29)		(0.62)
$Size_{t-1}$	0.014	0.007	0.016***	0.015***	0.003	0.002	-0.003***	-0.003***
	(0.55)	(0.27)	(12.86)	(12.66)	(0.77)	(0.58)	(-9.15)	(-9.10)
BM_{t-1}	0.164***	0.161***	-0.004***	-0.004***	0.019***	0.019***	0.004***	0.004***
	(6.93)	(6.87)	(-3.05)	(-3.26)	(3.68)	(3.65)	(8.73)	(8.73)
MOM_{t-1}	0.051	0.059	-0.007***	-0.006***	-0.008	-0.007	0.001***	0.001***
	(1.41)	(1.64)	(-4.34)	(-3.85)	(-1.39)	(-1.24)	(3.43)	(3.47)
$Leverage_{t-1}$	0.323***	0.315***	-0.014***	-0.015***	0.013	0.012	0.006***	0.006***
	(5.11)	(5.02)	(-3.96)	(-4.16)	(1.42)	(1.33)	(6.37)	(6.38)
IO_{t-1}	-0.001	-0.074	0.014***	0.006*	-0.012*	-0.021***	-0.002***	-0.002***
	(-0.03)	(-1.62)	(3.67)	(1.67)	(-1.69)	(-2.67)	(-4.76)	(-4.66)
Constant	0.853***	0.844***	0.058***	0.057***	0.113***	0.112***	0.039***	0.039***
	(5.41)	(5.34)	(6.74)	(6.78)	(4.82)	(4.77)	(16.60)	(16.62)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Observations	205,981	205,981	178,917	178,917	178,917	178,917	178,918	178,918
R-squared	0.200	0.201	0.808	0.810	0.757	0.758	0.616	0.616

Table 8. Passive Investing vs. Institutional Ownership and Active Holdings

This table reports augmented regressions results by adding to the baseline regression (Table 3) orthogonalized measures of a) stocks' institutional ownership (Panel A), and b) the Anton and Polk (2014) common ownership by active mutual funds (Panel B). Since the Anton and Polk (2014) measure is pairwise but the regression is at the stock level, we calculate, for each stock, its average co-ownership with other stocks. IO_Resid is the one-quarter lagged residual from a panel regression of stocks institutional ownership on their *PASSIVE_EXP* measure and AO_Reside is the one-quarter lagged residual from a panel regression of stocks' average pairwise active ownership measure on their *PASSIVE_EXP* measure. All variables and specifications are otherwise the same as Table 7.

Panel A: Institutional ownership								
	Beta	Beta	Corr	Corr	Cov	Cov	Idio_Vol	Idio_Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IO_Resid_{t-1}	-0.630***	0.024	-0.081***	0.028***	-0.081***	0.021***	-0.001	-0.002***
	(-14.30)	(1.10)	(-20.95)	(18.45)	(-17.07)	(8.72)	(-0.64)	(-4.70)
$PASSIVE\_EXP_{t-1}$		1.846***		0.310***		0.287***		-0.003
		(13.98)		(26.88)		(20.04)		(-1.28)
$Size_{t-1}$	0.040***	0.040***	0.017***	0.017***	0.007***	0.007***	-0.005***	-0.005***
	(6.12)	(6.12)	(39.40)	(39.40)	(9.12)	(9.12)	(-36.21)	(-36.21)
BM_{t-1}	0.097***	0.097***	-0.001**	-0.001**	0.010***	0.010***	0.003***	0.003***
	(14.09)	(14.09)	(-2.08)	(-2.08)	(9.11)	(9.11)	(17.01)	(17.01)
MOM_{t-1}	0.101***	0.101***	-0.008***	-0.008***	-0.006***	-0.006***	0.000***	0.000***
	(16.91)	(16.91)	(-28.51)	(-28.51)	(-10.38)	(-10.38)	(4.59)	(4.59)
$Leverage_{t-1}$	0.189***	0.189***	-0.005**	-0.005**	0.013***	0.013***	0.008***	0.008***
	(6.00)	(6.00)	(-2.42)	(-2.42)	(3.70)	(3.70)	(14.33)	(14.33)
IO_{t-1}	0.654***		0.110***		0.102***		-0.001	
	(13.98)		(26.88)		(20.04)		(-1.28)	
Constant	0.334***	0.593***	-0.034***	0.010***	-0.000	0.040***	0.048***	0.048***
	(7.13)	(13.42)	(-9.62)	(3.26)	(-0.06)	(7.54)	(53.03)	(55.51)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	377,822	377,822	377,852	377,852	377,852	377,852	343,560	343,560
R-squared	0.169	0.169	0.816	0.816	0.706	0.706	0.563	0.563

Panel B: Active Ownership								
	Beta	Beta	Corr	Corr	Cov	Cov	Idio_Vol	Idio_Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AO_Resid$_{t-1}$	-0.134	0.418*	0.040**	0.050***	0.021	0.143***	-0.003*	0.004**
	(-0.66)	(1.89)	(2.53)	(3.14)	(0.81)	(5.11)	(-1.71)	(2.25)
PASSIVE_EXP$_{t-1}$		3.258***		0.060		0.721***		0.047***
		(4.05)		(0.79)		(6.29)		(5.35)
Size$_{t-1}$	0.005	0.015	0.005***	0.006***	0.009**	0.011***	-0.001***	-0.001***
	(0.22)	(0.63)	(3.73)	(3.75)	(2.36)	(2.88)	(-4.19)	(-3.65)
BM$_{t-1}$	0.229***	0.232***	-0.005**	-0.005**	0.057***	0.058***	0.004***	0.004***
	(5.73)	(5.72)	(-2.49)	(-2.44)	(5.18)	(5.17)	(6.46)	(6.46)
MOM$_{t-1}$	-0.006	-0.012	-0.002*	-0.002*	-0.011***	-0.013***	-0.000	-0.000
	(-0.28)	(-0.54)	(-1.68)	(-1.80)	(-4.41)	(-4.91)	(-0.75)	(-1.17)
Leverage$_{t-1}$	-0.023	-0.032	-0.015**	-0.016***	0.012	0.010	0.002**	0.002*
	(-0.24)	(-0.32)	(-2.56)	(-2.59)	(0.87)	(0.73)	(1.97)	(1.82)
IO$_{t-1}$	-0.006	-0.094	-0.001	-0.003	-0.007	-0.027***	0.000	-0.001
	(-0.09)	(-1.25)	(-0.18)	(-0.49)	(-0.82)	(-2.80)	(0.29)	(-1.25)
Constant	0.848***	0.739***	0.122***	0.120***	-0.007	-0.031	0.021***	0.020***
	(3.48)	(3.03)	(8.46)	(8.01)	(-0.16)	(-0.73)	(7.69)	(7.06)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39,442	39,442	39,446	39,446	39,446	39,446	37,446	37,446
R-squared	0.224	0.225	0.886	0.886	0.783	0.784	0.528	0.530

Table 9. Correlation in Firm Fundamentals

This table reports the results on pooled panel regressions on earnings growth, from 1980 to 2020, with quarterly frequency. Dependent variable is the earnings (Net Income) changes from the previous quarter scaled by lagged market value of equity. The main independent variable is the index exposure, calculated as the fraction of shares outstanding held by index funds and ETFs; the market value-weighted market-level earnings change from the last quarter; and the interaction between the two variables. We include a series of control variables. Size is the quarte-end price times total shares outstanding and we take log transformations. BM is the book-to-market ratio. MOM is the 12-month momentum factor with one-month reversal. IO is the total institutional ownership as percentage of shares outstanding. Leverage is the long-term and short-term debt, divided by total assets. We lag all these firm-level control variables for one quarter. Year and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. We report t-statistic in parenthesis. ***, ** and * indicate significant levels at 1%, 5% and 10%.

	Earnings Growth	Earnings Growth	Earnings Growth	Earnings Growth
	(1)	(2)	(3)	(4)
$PASSIVE\_EXP_{t-1}$*Mkt Earnings Growth	0.333	0.564	0.340	0.597
	(0.96)	(1.02)	(0.97)	(1.04)
$PASSIVE\_EXP_{t-1}$	-0.032	-0.075	-0.223	-0.171
	(-0.47)	(-0.66)	(-1.61)	(-1.02)
Mkt Earnings Growth	-0.001	-0.005	-0.001	-0.006
	(-0.23)	(-0.87)	(-0.28)	(-0.89)
$Size_{t-1}$		0.001		-0.017***
		(0.58)		(-4.56)
BM_{t-1}		-0.013		-0.072**
		(-1.12)		(-2.34)
MOM_{t-1}		-0.006		-0.015*
		(-1.02)		(-1.97)
$Leverage_{t-1}$		0.012		0.052**
		(1.24)		(2.00)
IO_{t-1}		-0.009		-0.022
		(-0.65)		(-0.47)
Constant	-0.005	0.005	0.003	0.170***
	(-1.02)	(0.41)	(0.59)	(3.27)
Year FE	No	No	Yes	Yes
Firm FE	No	No	Yes	Yes
Observations	417,146	339,882	416,643	339,275
R-squared	0.001	0.005	0.064	0.036

Table 10. The Influence of Very Large Companies

This table reports the results on pooled regressions, from 1980 to 2020. We exclude the 10 largest firms each quarter in the analysis. The main independent variable is the passive exposure measure, calculated as the fraction of shares outstanding held by index funds and ETFs. We lag the exposure measure for one quarter. We include a series of control variables. Size is the quarte-end price times total shares outstanding and we take log transformations. BM is the book-to-market ratio. MOM is the 12-month momentum factor with one-month reversal. IO is the total institutional ownership as percentage of shares outstanding. Leverage is the long-term and short-term debt, divided by total assets. We also lag these firm-level control variables for one quarter. Year and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. We report t-statistic in parenthesis. ***, ** and * indicate significant levels at 1%, 5% and 10%.

	Beta	Beta	Corr	Corr	Cov	Cov	Idio_Vol	Idio_Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PASSIVE\_EXP_{t-1}$	1.778***	1.766***	0.318***	0.231***	0.297***	0.227***	-0.019***	0.001
	(6.98)	(5.97)	(10.46)	(7.54)	(5.08)	(3.44)	(-5.99)	(0.41)
Size$_{t-1}$		0.040**		0.017***		0.007**		-0.005***
		(2.37)		(16.44)		(2.34)		(-17.87)
BM$_{t-1}$		0.097***		-0.001		0.010***		0.003***
		(6.71)		(-0.90)		(3.82)		(9.98)
MOM$_{t-1}$		0.101***		-0.008***		-0.006		0.000
		(3.45)		(-6.72)		(-1.18)		(1.47)
Leverage$_{t-1}$		0.192***		-0.005**		0.014**		0.008***
		(4.44)		(-2.08)		(2.27)		(9.55)
IO$_{t-1}$		0.024		0.028***		0.021**		-0.002***
		(0.58)		(5.21)		(2.06)		(-3.46)
Constant	0.976***	0.587***	0.118***	-0.002	0.101***	0.032**	0.025***	0.049***
	(62.16)	(6.00)	(51.63)	(-0.39)	(23.20)	(2.23)	(90.39)	(30.91)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	455,246	376,406	455,671	376,426	455,671	376,426	456,509	342,149
R-squared	0.160	0.169	0.800	0.816	0.692	0.706	0.488	0.559

Table 11. Information discovery around earnings announcements

This table reports the result on idiosyncratic volatility around quarterly earnings announcement day. The sample period is from 1980 to 2020, with quarterly frequency. Idiosyncratic volatility is calculated as the standard deviation of residual returns in [-5, +5] window around firm's quarterly earnings announcement day. Residual returns is calculated by first estimating coefficients using Fama-French 3-factor models from [-35,-5] window; then obtain the difference between actual return and model predicted returns. The t-statistic are reported in parenthesis and it's Newey-West corrected with 4 lags.

PASSIVE_EXP Quintile	Full Sample	1980 to 1997	1998 to 2020
Low	0.034	0.028	0.036
2	0.032	0.024	0.034
3	0.031	0.028	0.032
4	0.030	0.029	0.031
High	0.031	0.034	0.030
High-Low	-0.004***	0.006***	-0.006***
t-statistic	(-20.55)	(19.04)	(-31.98)