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

March 12, 2023

Lily Fang
INSEAD

Hao Jiang
Michigan State University

Zheng Sun
University of California, Irvine

Ximing Yin
Hunan University

Lu Zheng
University of California, Irvine

Abstract

Campbell et al. (2001) documents that between 1962 and 1997 correlation among stocks dropped. This pattern reversed post 1997: correlation among stocks roughly doubled in 1998-2020 from previous levels. We hypothesize and provide evidence that the rise of passive investing contributed to higher correlations among stocks and in turn higher market volatility. We find that the degree to which a stock is held by passive (index and ETF) funds strongly predicts its beta and correlation with other stocks. Difference-in-difference analyses around three market shocks – 9/11 in 2001, Lehman collapse in Sep 2008, and Covid shock in March 2020 – show that stocks with high passive holdings contributed more to market volatility. Our results are not subsumed under common holdings by institutions in general and are not explained by increases in earnings correlations. We conclude that the rise of passive investing could lead to higher correlation among stocks and higher market volatility, limiting 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? Researchers, practitioners, and regulators are beginning to question the unintended implications of index investing due to its rapid growth and significant market dominance. Prior to his passing in 2019, John Bogle, known as the "father of index funds," cautioned in a WSJ article about what might happen if index investing became too successful for its own good. He wrote: “Public policy cannot ignore this growing dominance, and consider its impact on the financial markets, corporate governance, and regulation." Indeed, the dominance of index investing may have many significant implications. In this paper, we focus on its implications for diversification, a cornerstone of modern portfolio theory and one of the most important benefits of indexing.

The invention of mutual funds changed people’s way of investing from individual assets to a basket of assets. The invention of indexing reduced the number of baskets that investors hold. When investors trade a basket of assets, they trade the underlying assets in the same direction at the same time. If investors hold and trade a few similar baskets (indices), the trading activities become more correlated across individual assets. In the extreme case, when all investors hold and trade only one basket of assets, the trading effects will be in the same direction for all assets all the time. The resulting effect is perfect correlations among all asset prices. The rapid growth of indexing funds/ETFs and the dominance of a few key indices push markets toward the direction of the limit case of trading a single basket. As a result, we may face the paradox of indexing for diversification: we hold the market portfolio to achieve diversification as prescribed by theory; yet as more and more investors do the same and hold the same basket of assets, we lose the benefit of diversification.

In this paper, we study how indexing affects correlation between assets and eventually aggregate market volatility. We hypothesize that more indexing increases correlation across assets and aggregate

market volatility. The influential 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 and the overall market volatility remained stable.[2] Figure 1 is our exact replication of Campbell et al. (2001) and confirms these well-known results.

We show that there has been a powerful reversal of these patterns in the period since the publication of Campbell 2001. Figure 2 extends the original Campbell et al. analysis 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 a dramatic *increase* in pair-wise correlation among stocks since the original Campbell et al. (2001). The average pair-wise correlation is 13.4% in the period of 1998-2020, more than double the 5.7% for the period 1962-1997 (*t*-stats=17.85 for the difference). Panel C shows that the net effect between 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 (t-stats = 5.44 for the difference).

While multiple explanations are possible for higher correlations among stocks, in this paper, we focus on the rise of index funds and ETFs, which we collectively refer to as passive investing. Our hypothesis stems from the notion that passive investing frequently entails buying and selling all securities in an index together, which can lead to higher return correlation between stocks held in the index beyond the correlation driven by fundamentals.

If true, our hypothesis has profound implications for market efficiency and the cost and benefit of passive investing. One of the central tenets of modern portfolio theory and the Efficient Market Hypothesis is the optimality of the market portfolio. A large literature (e.g., Jensen, 1968; Carhart,

[2] 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.

1997) has established the advantages of passive investing: active stock picking on average does not out-perform passive indexing while costing more in fees and expenses. This robust result has been reflected in the real world by the rise of passive investing: from the emergence of Vanguard as the first index fund in 1976, index funds and ETFs now collectively account for 43% of the total assets under management in US equity funds.[3] But our hypothesis suggests an important downside to the dominance of passive investing: this dominance could lead to increases in correlations among stocks which can result in higher *market-level* volatility, limiting the power of diversification and the virtue of passive investing. Moreover, information structure in the market will change: the rise of passive investing could mean less idiosyncratic information being produced and reflected in stock prices.[4]

Figure 3 presents a first view of the relationship between the extent of passive investing and the correlation structure 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 amount invested in the index funds and ETF, 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%, significantly different from 0 at the 1% level. This visual evidence indicates that passive investing is at least correlated with the extent of indexing.

To systematically investigate the relation between passive investing and measures of risk, 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 call this measure a stock's "PASSIVE_EXP" (exposure to passive investing). We then examine the relation between stocks' PASSIVE_EXP and four second-moment measures of the stock return: 1) market

[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.

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* relation between PASSIVE_EXP and a stock's idiosyncratic volatility, although this result is not always statistically significant. Overall, the robust conclusion is that there is a strong relation between a stock's exposure to passive investing and how it comoves with other stocks.

We use a difference-in-differences approach to illustrate that the above finding means that stocks with higher PASSIVE_EXP *contribute* more to overall market volatility. We focus on three episodes of market-wide shocks – Sep 2001 9/11 event, Sep 2008 the collapse of Lehman and onset of the financial crisis, and March 2020 Covid shock. 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.

Our hypothesis is based on the notion that correlated trading by index funds and ETFs increase stocks' systematic risk measures. To directly shed light on this mechanism, we examine the effect of trading which is 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 and ETF funds in our sample. It reflects the net amount of flow-induced trading that cannot be absorbed within the index fund sector, thus 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. It is negatively correlated with stocks' idiosyncratic volatility. The contributing effect of passive-flow-induced-trading on beta, correlations, and covariance, is especially strong during the crisis periods defined above.

We show that the effects we document is not subsumed under general institutional investing and common ownership among institutional investors, and is 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. For these stocks, price movements around earnings are primarily driven by market movements.

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 the risk of passive investing on the entire universe of stocks and on the aggregate market outcomes, rather than on individual stocks being added/removed from the index.

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 supplies the rise of passive investing as a potential explanation, which is perhaps the single most robust trend in asset management in the last two decades.

The rest of the paper is organized as follows. Section I describes out 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 contains additional analysis such as market information structure and Section V concludes.

I. Data and Variables

Our paper is focused on the second moment of stock returns, i.e., variances and covariance Properties. Our goal is to investigate whether these second-moment 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 a given 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 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 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 reaches maximum of 7.60% in 2020Q1.

We examine five second-moment (i.e., risk) measures of market return to capture the different components that contribute to overall market risk. These measures distinguish between the part of the

market risk that is due to stock-level volatility, and the other component that is driven by correlation among stocks. Specifically:

- The first measure is simply the market volatility. Treating this as a macroeconomic variable and to reduce noise in its estimation, we follow Engle and Rangel (2008) and extract the low-frequency (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 firm-level total volatility, which is the equal-weighted average of each stock's monthly volatility calculated from daily stock returns, and then averaged by calendar quarter and across stocks.
- 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. This is estimated each month using daily returns for each stock and aggregated into a market-level quarterly series in the same way as with our third measure.
- As our fifth measure, 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 moving from a 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 *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) and the volatility gap (correlation = -0.608). 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., there is 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,* for one quarter and 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*-stats = 2.82), suggesting that the rise of passive investment is associated with an increase the market-level volatility over time. 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*-stats = 8.20), but it negatively predicts next quarter's average total stock volatility (coefficient = -0.175, *t*-stats = -4.25) and average idiosyncratic stock volatility (coefficient = -0.205, *t*-stats = -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 comes from correlations 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). Pre-1997, *Passive-to-Market* is negatively related to overall market volatility (coefficient = -0.768, *t*-stats = -8.14); it is negatively related to pair-wise correlation (coefficient = -31.485, *t*-stats = -3.51), but positively related to total and idiosyncratic firm-level volatility. The evidence in Panel C is consistent with Campbell et al. (2001): pre-1997, since correlations among stocks were dropping but passive investing was increasing, the latter did not contribute to the market level risk through the correlation channel. Post-1997, we find strong evidence that it did, and the channel of its contribution is through correlations among stocks.

B. Cross-sectional evidence - Panel regressions

Having examined the time-series relation between the market-level degree of passive investing and market level risk measures, we now turn to cross-sectional analysis.

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 a number of

control variables which include 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 year, we include separate firm and year-quarter 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 is associated with a 0.1 increase in beta.[5] Since the global average beta is 1.04, this means a 10% increase. The magnitude is not only statistically significant but 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.[6] In contrast, Columns (7) and (8) show that PASSIVE_EXP generally has a negative

[5] 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.

[6] 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.

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 crisis periods 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. There are periods of significant market volatility driven by external shocks that no individual firm can cause or anticipate. We sort stocks according to the 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 high PASSIVE_EXP 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*-stats 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*-stats = 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. For example, 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 indistinguishable.

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 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 always positive though insignificant, meaning that stocks' idiosyncratic volatilities increase during crisis periods but not salient. The interaction term between PASSIVE_EXP and the crisis indicator is also always negative and significant with firm-level control variables, indicating that stocks with higher PASSIVE_EXP experience larger decreases in idiosyncratic volatility. Considering that the PASSIVE_EXP variable itself has a positive sign but the magnitude is smaller than that of the interaction terms, it indicates that exposure to passive investing is overall negatively related to stock-level idiosyncratic volatility, once other stock-level characteristics are controlled for. These results are in contrast to those in beta and average correlations but consistent with results shown earlier in this paper: In general PASSIVE_EXP does not positively predict stock idiosyncratic volatility but is highly positively predictive of correlations with other stocks.

Overall, results in Tables 3, 4, and 5 present consistent evidence that PASSIVE_EXP is strongly predictive of stocks' second-moment measures 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 of their correlations with other stocks, but not through higher volatilities.

IV. Mechanism – Passive Flow-induced Trading s

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 drives trades – should contribute to the relation between risk and exposure to passive investing. To test this mechanism, we estimate the following panel regression:

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

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:

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

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 fund *j* hold in stock *i.* This captures the amount 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.* Then 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. Lastly, since we are not concerned with the direction of trading, we take the absolute value of the net trading by index funds and scaled 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, thus a net liquidity demand by index funds to other investors.

We regress stock beta on $PASSIVE\_FLOW$, with the rest of the specification the same as 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 over 2.5% increase compare to the sample average. These magnitudes are obtained after controlling separately for year-quarter fixed effects and firm fixed effects, and are thus economically significant.

Earlier we find that stocks with a higher exposure to index funds experience larger increase in stock beta during the crisis period. To examine whether the 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 especially strong during the crisis period.

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

V. Additional Analysis – Alternative Explanations and Mechanism

A) Active Institutional ownership

One concern with our results is that we may not have sufficiently controlled for other institutional investing, including active funds. If stocks' PASSIVE_EXP is highly correlated with their overall institutional ownership, then the single sorting results can reflect an overall institutional holding effect, rather than a passive investing effect.

We should first point out that our baseline regressions in Table 3 already include general institutional ownership as a control variable. Second, our hypothesis is fully consistent with a more general notion that common holdings among institutions – active or passive can lead to increased correlations among stocks and hence contribute to overall market volatility. Anton and Polk (2014) find that common mutual fund holdings have the effect of increasing pair-wise correlation. The

presence of "closet indexers" – i.e., active funds that implicitly hold similar stocks as indexes – and the widespread practice of benchmarking imply that the true extent of "passive" investing is higher than indicated by the AUMs of passive funds. If the mechanism we propose in this paper is true, then the concern regarding market risk can be even more significant.

To check the robustness of our main results to further controlling for institutional ownership, we repeat the difference-in-differences analysis in Table 4 after double-sorting the stocks. We first sort stocks into terciles based on their overall institutional ownership (or the Anton and Polk (2014) pairwise correlation measure). Then, within each tercile, we further sort the stocks into quintiles based on the *PASSIVE_EXP* measure.

Table 7 reports the results of the difference-in-differences analysis of stocks' average correlation with all other stocks after double sorting. For brevity, we only report the results on average correlations. Un-reported results on beta and volatility are qualitatively similar. Panel A pertains to the double-sorting based on institutional ownership and *PASSIVE_EXP*. We find that the general conclusion from Table 4 holds within each tercile: high *PASSIVE_EXP* stocks exhibit larger *increases* in correlations during crisis period than low *PASSIVE_EXP* stocks. The effect is larger for the low- and median-IO terciles (terciles 1 and 2) and smaller for the high-IO tercile (tercile 3). Within the low-IO tercile, high *PASSIVE_EXP* stocks experience an 0.1071 (i.e., 10.71%) increase in average correlation with other stocks, compared with an increase of 0.0825 (i.e., 8.25%) for low *PASSIVE_EXP* stocks. The difference is highly significant. Within the high-IO tercile, high *PASSIVE_EXP* stocks experience an increase of 0.1262 (i.e., 12.62%) in average correlation with other stocks, which low *PASSIVE_EXP* stocks experience an increase of 0.1165 (i.e., 11.65%). The difference is smaller but remains significant at the 10% level. The larger increase across the board among the high-IO (tercile 3) stocks indicate that there is an effect related to general institutional

ownership, but the robust difference-in-differences results within each tercile point to the separate effect due to *PASSIVE_EXP*.

Results in Table 7 Panel B pertaining to double-sorting based on the Anton and Polk (2014) measure and *PASSIVE_EXP* are qualitatively similar. Here, the magnitude of the *PASSIVE_EXP* effect is remarkably consistent across Polk-measure terciles. In Tercile 1 (low Polk measure), the difference in the increases in average correlations between high- and low-*PASSIVE_EXP* stocks is 0.0221 with a p-value of 0.01, while the same difference-in-difference is 0.024 (0.0293) with a *p*-value of 0.03 (0.04) in the middle- and high-terciles, respectively. Therefore, we conclude that there is a distinctive effect associated with the exposure to passive investing, above and beyond a general institutional (or active) ownership effect.

B) Correlation in fundamentals

A second concern with our analysis is the possibility that the increased correlations among stocks can be due to an increase in the correlation of firm fundamentals. For this concern to be drive our results, it would be argued that stocks with high *PASSIVE_EXP* also have stronger correlation in earnings growth with other stocks.

To check this hypothesis, we estimate panel regressions of firms' quarterly earnings growth on *PASSIVE_EXP*, market-wide earnings growth (to proxy for earnings growth in other firms), and the interaction between the two. If correlations in fundamentals drive our results, we should find that the interaction term between *PASSIVE_EXP* and the market-wide earnings growth to be significant in explaining firm level earnings growth. Results in Table 8 clearly reject this hypothesis: the interaction term has no explanatory power for firm-level earnings growth.

C) Influence of Very Large Companies

In the last ten years, the rise of very large firms has attracted research attention (citations). FAANG stocks, including Meta (previously Facebook), Apple, Amazon, Netflix, and Alphabet (previously Google) collectively account for 19% of the S&P as of Aug 2021.[7] 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.[8] The results are reported in Table 9. For brevity, we only report the key coefficients. Results in Table 9 shows 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 that of Table 3.

D) 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, more systematic, but less idiosyncratic news is impounded into prices.

[7] https://investorpedia.com, accessed July 21, 2022.
[8] The 10 excluded stocks are: Apple, Microsoft, Amazon, Tesla, Facebook, Alphabet, Johnson & Johnson, JP Morgan Chase & Co, as of December 2020.

To investigate this hypothesis, we examine total and idiosyncratic volatility of stocks around earnings announcements. Table 10 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 by quintiles. We do this for the whole sample, as well as two subsamples: from 1980 to 1999; and from 2001 to 2020.

Table 10 shows that overall *PASSIVE_EXP* is negatively related to idiosyncratic volatility: High *PASSIVE_EXP* stocks have average volatility of 0.030 around earnings announcement, compared to low *PASSIVE_EXP*'s average of 0.034. This difference is highly significant with a t-stats of over 40. 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. Table 10 shows that this is particularly true around earnings event, which is very informationally rich for individual stocks.

The sub-sample analysis reveal however that this result is entirely driven by the second half of our sample: from 2000 to 2020. In this period, the idiosyncratic volatility of the high *PASSIVE_EXP* stocks *dropped* to 0.028, over 20% lower than low *PASSIVE_EXP* stocks' average idiosyncratic volatility of 0.034, and the difference is highly significant with a t-stats of over 40. In fact, in the earlier sample, the relationship between *PASSIVE_EXP* and idiosyncratic volatility is the reverse: High *PASSIVE_EXP* stocks (which tend to be larger stocks) have a higher average idiosyncratic volatility of 0.035. Looking across the different sample period, we see that low *PASSIVE_EXP* stocks' average idiosyncratic volatility stayed around 0.034 throughout the 40-year period; high *PASSIVE_EXP* stocks on the other hand experienced a significant drop in idiosyncratic volatility from the first half of the

sample (before 2000), to the second half of the sample (after 2000). This result is consistent with our hypothesis that there is less firm-specific price discovery with the rise of passive investing.

Conclusion

In the past twenty years, one of the most salient trends in the asset management industry is the rise of passive investing. Index investing and ETFs are significant financial innovations in the 20th century that brought affordable, broad diversification benefits to the general investing public. In this paper, we provide evidence that cautions against a potential dark side of the ever-rising extent of passive investing: 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 index 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 that could result from indexing. We leave these interesting and ambitious questions to future research.

References

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.*

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

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.

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

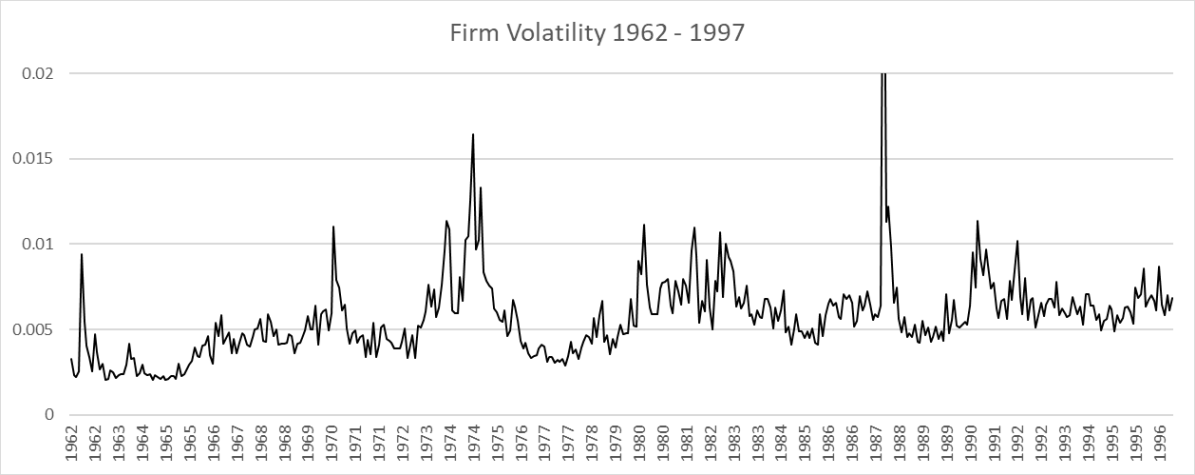
Shleifer, Andrei, 1986, Do demand curves for stocks slope down? *Journal of Finance* 41, 579-590.

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

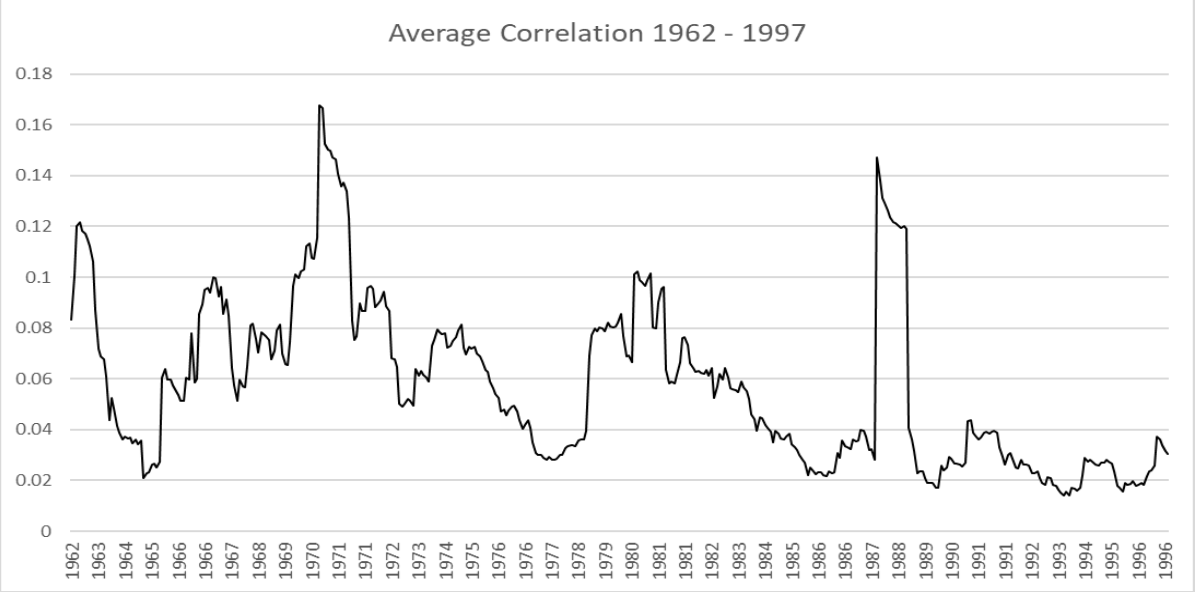
Figure 1. Replicating Campbell et al. (2001): 1962-1997

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 annualize stock 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

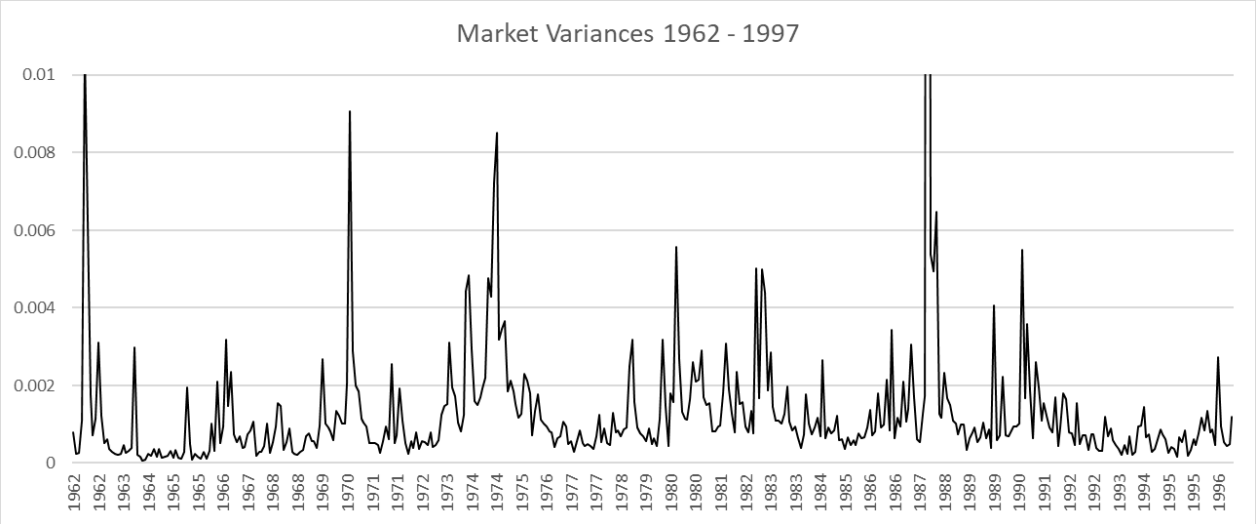
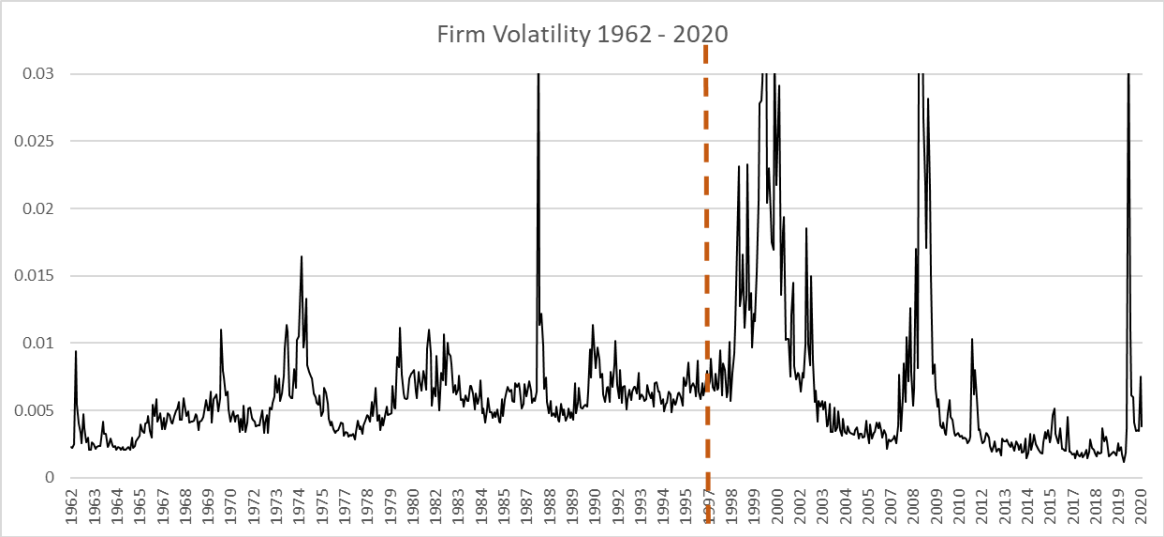


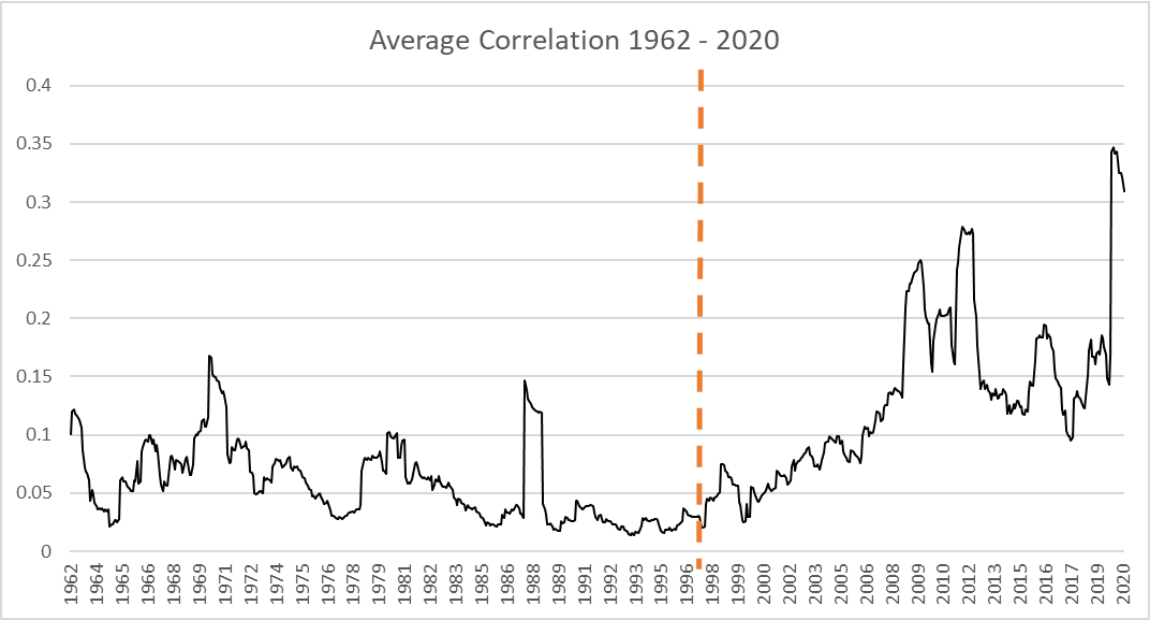
Figure 2. Extending Campbell et al. (2001): 1962 – 2020

This figure extends the results in Campbell et al. (2001) to 2020. Panel A replicates and extends figure 4(A) of Campbell et al. (2001) and plots the average annualize stock 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 marks the pre- and post-Campbell (2001) sample period.

A. Average firm-level return variance



B. Average pair-wise return correlation



C. Market Variances

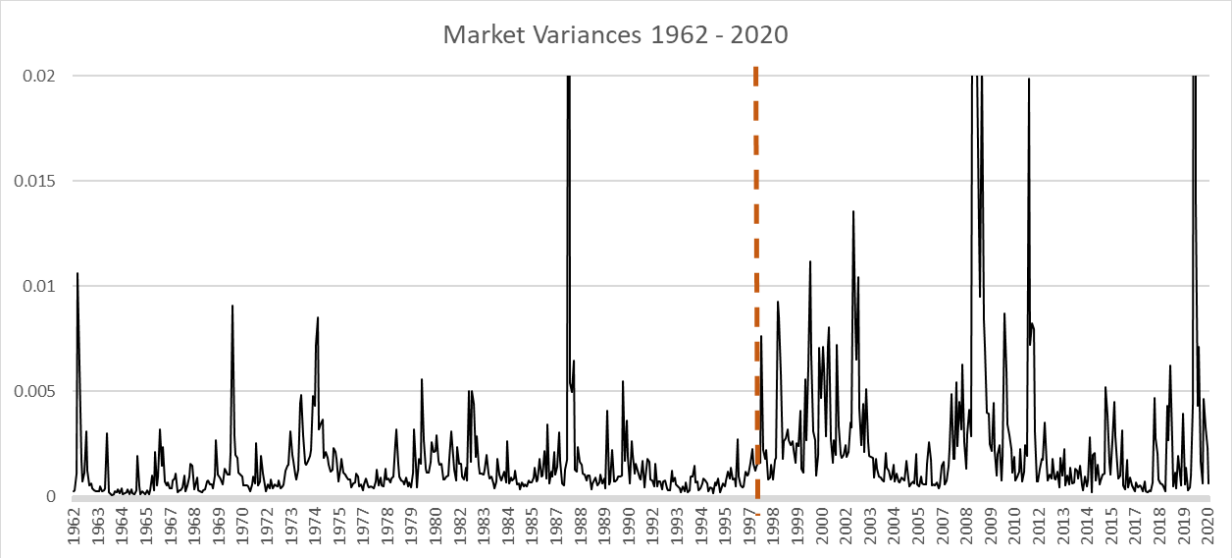


Figure 3. Passive-to-Market ratio and Average Pairwise Correlation

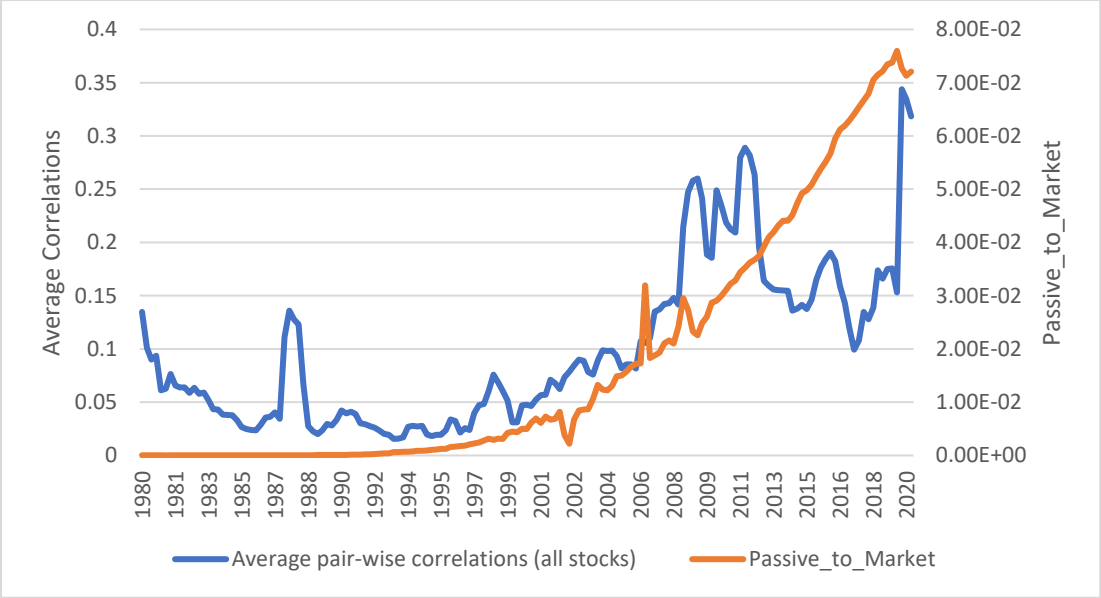


Figure 4. Average PASSIVE_EXP Measure over time

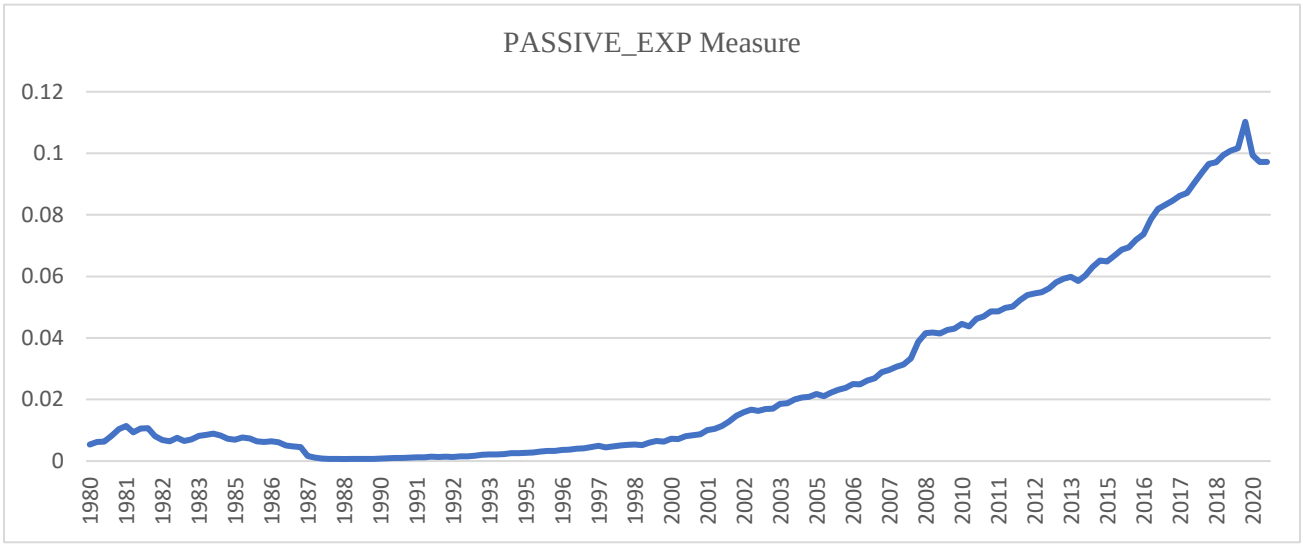


Table 1. Summary Statistics

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. 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 Second Moments

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 three market-wide second moment measures: overall market volatility (Market_lvol), the average pair-wise correlation (Avg_Corr), the average stock volatility (Avg_Firmvol). 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 report the predictive regressions of each of the market-wide risk measure on the extent of passive investing. We also regress volatility gap (Vol_gap) on the Passive_to_Market ratio, where volatility gap is defined as average firm volatility minus average market volatility each quarter. 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	Avg_Corr	Avg_Firmvol	Avg_idio_vol	Vol_gap
Passive_to_Market	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	Avg_Corr	Avg_Firmvol	Avg_idio_vol	Vol_gap
Passive_to_Market	-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	Avg_Corr	Avg_Firmvol	Avg_idio_vol	Vol_gap
Passive_to_Market	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. 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 (multiplied by 1,000). 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-quarter and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. *t*-stats 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-QTR 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), and idiosyncratic volatility (Panel C). 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. DiD 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, 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-quarter and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. We report *t*-stats 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	1.528***	1.686**	0.244***	0.302**	0.921***	1.233**	-0.006	-0.015**
	(3.08)	(2.21)	(4.47)	(2.43)	(4.60)	(2.42)	(-0.77)	(-2.20)
$PASSIVE\_EXP_{t-1}$	1.183**	1.190**	0.109**	0.072	-0.097	-0.017	-0.005	0.014*
	(2.81)	(2.34)	(2.64)	(1.51)	(-0.97)	(-0.25)	(-0.56)	(1.83)
Crisis	-0.140**	-0.158**	0.032**	0.032**	0.078	0.068	0.009	0.008
	(-2.22)	(-2.34)	(2.63)	(2.52)	(1.49)	(1.43)	(1.53)	(1.55)
$Size_{t-1}$		0.027		0.023***		0.017**		-0.006***
		(0.52)		(12.60)		(2.68)		(-9.24)
BM_{t-1}		0.129***		-0.000		0.020***		0.003***
		(5.07)		(-0.37)		(4.77)		(8.79)
MOM_{t-1}		-0.070		-0.013***		-0.021		0.001
		(-0.78)		(-4.21)		(-1.40)		(1.72)
$Leverage_{t-1}$		0.151		-0.014**		0.031		0.013***
		(1.14)		(-2.37)		(1.29)		(5.89)
IO_{t-1}		-0.078		0.018*		-0.018		-0.004***
		(-0.67)		(1.81)		(-1.09)		(-3.82)
Constant	1.055***	0.779**	0.149***	-0.010	0.211***	0.069*	0.028***	0.059***
	(26.13)	(2.71)	(31.15)	(-0.80)	(10.09)	(1.96)	(12.16)	(17.15)
Year-QTR FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	87,895	67,555	83,541	63,659	83,541	63,659	83,675	67,847
R-squared	0.278	0.305	0.872	0.875	0.776	0.788	0.527	0.565

Table 6. Mechanisms – Fund flows

The sample spans from 1991 to 2020, with quarterly frequency. We analyze beta, average correlation with other stocks, and volatility as the risk measure, respectively. PASSIVE_FLOWs 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. 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. All variables are winsorized at 1% and 99% level. We also lag these firm-level control variables for one quarter. Year-quarter and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. We report t-stats 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)	(7)	(8)
$PASSIVE\_FLOWS_t$	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-QTR 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)
$PASSIVE\_FLOWS_t$*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)
$PASSIVE\_FLOWS_t$	-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-QTR 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. Controlling for Institutional and Active Mutual Fund Ownership: Double-sorting

This table reports the difference-in-differences results in stocks' average pairwise correlations with all other stocks in the CRSP universe, after double-sorting stocks with institutional ownership measure and our PASSIVE_EXP measure (Panel A), and after double-sorting stocks with the Anton and Polk (2014) measure of mutual fund common ownership and the PASSIVE_EXP measure (Panel B). We first sort stocks into terciles according to their institutional ownership or the Anton and Polk (2014) measure each quarter. Then within each tercile, we sort stocks according to the average pairwise index exposure measure. we report the average pairwise correlations. Before Crisis periods are 2000Q3 to 2001Q2, 2007Q4 to 2008Q3 and 2019Q1 to 2020Q1; During Crisis periods are 2001Q3 to 2001Q4, 2008Q4 to 2009Q3 and 2020Q2 to 2020Q4.

Panel A: Double-sorting using Institutional Ownership (IO) and PASSIVE_EXP

Panel A1: IO Tercile 1 (Low)

	Before Crisis	During Crisis	During - Before	*p*-value
Quintile 1 (low)	0.0917	0.1742	0.0825	0.00
Quintile 2	0.0791	0.1555	0.0765	0.00
Quintile 3	0.0685	0.1321	0.0636	0.00
Quintile 4	0.0863	0.1616	0.0753	0.00
Quintile 5 (high)	0.1207	0.2278	0.1071	0.00
High - Low	0.0290	0.0536	0.0246	0.03

Panel A2: IO Tercile 2 (Medium)

	Before Crisis	During Crisis	During - Before	*p*-value
Quintile 1 (low)	0.1084	0.1959	0.0875	0.00
Quintile 2	0.1190	0.2350	0.1161	0.00
Quintile 3	0.1309	0.2504	0.1194	0.00
Quintile 4	0.1397	0.2676	0.1278	0.00
Quintile 5 (high)	0.1526	0.2802	0.1276	0.00
High - Low	0.0441	0.0843	0.0401	0.00

Panel A3: IO Tercile 3 (High)

	Before Crisis	During Crisis	During - Before	*p*-value
Quintile 1 (low)	0.1405	0.2540	0.1135	0.00
Quintile 2	0.1491	0.2742	0.1251	0.00
Quintile 3	0.1526	0.2782	0.1256	0.00
Quintile 4	0.1547	0.2847	0.1300	0.00
Quintile 5 (high)	0.1525	0.2787	0.1262	0.00
High - Low	0.0121	0.0247	0.0126	0.09

Panel B: Double-sorting using Anton and Polk (2014) measure of Active Mutual Fund Common Ownership and PASSIVE_EXP

Panel B1: Anton & Polk Tercile 1 (Low)

	Before Crisis	During Crisis	During - Before	*p*-value
Quintile 1 (low)	0.2850	0.4092	0.1243	0.13
Quintile 2	0.2628	0.4075	0.1447	0.10
Quintile 3	0.2496	0.4035	0.1540	0.08
Quintile 4	0.2492	0.4059	0.1567	0.07
Quintile 5 (high)	0.2617	0.4080	0.1464	0.08
High - Low	-0.0233	-0.0012	0.0221	0.01

Panel B2: Anton & Polk Tercile 2 (Medium)

	Before Crisis	During Crisis	During - Before	*p*-value
Quintile 1 (low)	0.2722	0.4126	0.1404	0.07
Quintile 2	0.2503	0.4067	0.1564	0.06
Quintile 3	0.2396	0.4030	0.1634	0.06
Quintile 4	0.2374	0.4030	0.1656	0.05
Quintile 5 (high)	0.2477	0.4121	0.1644	0.04
High - Low	-0.0245	-0.0004	0.0240	0.03

Panel B3: Anton & Polk Tercile 3 (High)

	Before Crisis	During Crisis	During - Before	*p*-value
Quintile 1 (low)	0.2468	0.3959	0.1492	0.06
Quintile 2	0.2300	0.3940	0.1640	0.05
Quintile 3	0.2175	0.3950	0.1775	0.04
Quintile 4	0.2209	0.4008	0.1799	0.03
Quintile 5 (high)	0.2332	0.4117	0.1785	0.03
High - Low	-0.0135	0.0158	0.0293	0.04

Table 8. Correlation in Fundamentals

This table reports the results on pooled panel regressions, from 1980 to 2020, with quarterly frequency. Dependent variable is 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-quarter and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. We report t-stats in parenthesis. ***, ** and * indicate significant levels at 1%, 5% and 10%.

	Earnings Growth (1)	Earnings Growth (2)	Earnings Growth (3)	Earnings Growth (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-QTR 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 9. 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 index exposure, 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-quarter and firm-fixed effects are included. Standard errors are clustered at firm and year-quarter level. We report t-stats 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-QTR 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 10. 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.

Passive_Exp Quintile	Full Sample	1980 to 1999	2001 to 2020
Low	0.034	0.032	0.034
2	0.033	0.027	0.035
3	0.030	0.030	0.030
4	0.030	0.031	0.029
High	0.030	0.035	0.028
High-Low	-0.004***	0.003***	-0.006***
t-stats	(-41.70)	(13.03)	(-54.50)