Passive Investing and the Rise of Mega-Firms

HAO JIANG
Michigan State University

DIMITRI VAYANOS
London School of Economics, CEPR and NBER

LU ZHENG
University of California–Irvine

June 1, 2022*

Abstract

We show that flows into passive funds raise disproportionately the prices of the largest stocks in the index, while also making them more volatile. If, in addition, stocks are mispriced because of noise traders, then passive flows raise disproportionately the prices of the overvalued stocks among the index’s largest. Passive flows generally drive the aggregate market up even when they are entirely due to a switch from active to passive. Underlying these effects is that passive flows make prices more sensitive to news about idiosyncratic future cashflows. We provide empirical evidence in support of our model’s main predictions.

JEL: G10, G11, G12, G23

Keywords: Market efficiency; mutual funds; indexing; limits of arbitrage

*jiangh@broad.msu.edu, d.vayanos@lse.ac.uk, luzheng@uci.edu. We thank seminar participants at Panagora Asset Management, Q-Group, Shanghai Advanced Institute of Finance, and conference participants at the American Finance Association for helpful comments. We are grateful to the LSE Paul Woolley Centre for financial support, and to Farbod Ekbatani and Pete Trairatanobhas for excellent research assistance. A previous version of this paper was circulated with the title “Tracking Biased Weights: Asset Pricing Implications of Value-Weighted Indexing.”
1 Introduction

One of the most important capital-market developments of the past thirty years has been the growth of passive investing. Passive funds, such as index mutual funds and index exchange-traded funds (ETFs), track market indices and charge lower fees than active funds. In 1993, passive funds invested in US stocks managed $23 billion of assets. That was 3.7% of the combined assets managed by active and passive funds invested in US stocks, and 0.44% of the US stock market as a whole. By 2021, passive assets had risen to $8.4 trillion. Moreover, passive funds had overtaken active funds: passive assets had risen to 53% of combined active and passive, and to 16% of the stock market. The S&P500 index attracts the bulk of passive investing: as of 2021, 42% of index mutual funds invested in US stocks were tracking that index.\(^1\)

The growth of passive investing has stimulated academic and policy interest in how passive investing affects asset prices and the real economy. In this paper we show that flows into passive funds tracking value-weighed indices raise disproportionately the prices of the largest stocks within the indices. If, for example, passive funds tracking the S&P500 index receive inflows, then the largest S&P500 stocks will experience higher returns than smaller S&P500 stocks. The same largest stocks will experience an increase in return volatility and sensitivity to cashflow shocks. If, in addition, stocks are mispriced because of noise traders, then flows into passive funds will raise disproportionately the prices of the overvalued stocks among the index’s largest. Moreover, flows have an asymmetric effect in the cross-section, driving the aggregate market up: even when flows are entirely due to a reallocation from active to passive within the stock market, the gains in market capitalization from the stocks that rise exceed the losses from the stocks that drop. Our theory implies that passive investing is not neutral, but reduces primarily the cost of capital of the largest firms and makes the size distribution of firms more skewed.

To explain the intuition for our results, we first describe how passive flows affect prices in a simple CAPM world, in which the index tracked by passive funds is the market portfolio. If passive

\(^1\)The data come from the 2022 Investment Company Institute Factbook (Figure 2.9 and Tables 11 and 42), and from https://data.worldbank.org/indicator/CM.MKT.LCAP.CD?locations=US. The growth of passive investing is even more dramatic when accounting for the decline in active share and rise of closet indexing (Petajisto (2013)), and for the passive investments of pension funds, sovereign-wealth funds and other institutions that are made outside mutual funds and ETFs.
flows are due to increased stock-market participation, then their effect is to drive down the market risk premium. Therefore, stock prices rise, and the effect is more pronounced for high CAPM-beta stocks. Since small stocks have higher CAPM beta than large stocks (Fama and French (1992)), small stocks experience higher returns than large stocks. If instead passive flows are due to a switch from active to passive, then they have no effect on stock prices because active and passive funds hold the same portfolio.

While the simple CAPM argument describes correctly how passive flows affect price levels, it fails to account for the flows’ effect on price volatility, which feeds back into levels. To best illustrate these effects, we introduce noise traders and short-sales, but drop these features later in the Introduction. Suppose that a stock is in high demand by noise traders. Suppose also that the additional demand generated by passive flows is such that smart-money investors must short-sell the stock in equilibrium. The high demand not only raises the stock’s price, but also makes it more sensitive to cashflow shocks. Indeed, following a positive shock to idiosyncratic cashflows, the stock accounts for a larger fraction of market movements. Therefore, smart-money investors find themselves with a riskier short position and become more willing to unwind it, amplifying the effect of the cashflow shock. Higher price sensitivity implies more risk, making smart-money investors even more willing to unwind their position, and causing the price to rise further and become even more sensitive to cashflow shocks. This mechanism is quantitatively significant for large stocks, as their idiosyncratic risk is non-negligible, and explains why passive flows raise their prices the most while also raising their volatility.

Section 2 presents the model. Agents can invest in a riskless asset and in multiple stocks over an infinite horizon. The riskless rate is constant. Each stock’s dividend flow is the sum of a constant component, a systematic component and an idiosyncratic component. The systematic and idiosyncratic components follow independent square-root processes. Agents can be experts or non-experts. Experts observe dividend flows, and can invest in all assets without constraints. They can be interpreted as active-fund investors. Non-experts do not observe dividend flows, and can invest in the riskless asset and in a stock portfolio that tracks an index. They can be interpreted as passive-fund investors. Experts and non-experts maximize a mean-variance objective over instantaneous changes in wealth. Noise traders can also be present, and hold a number of shares of each stock that
is exogenous and constant over time. The presence of noise traders strengthens our main results but is not required for them to hold.

[[[TO BE COMPLETED]]]

Our model builds on Buffa, Vayanos, and Woolley (2022, BVW), who examine how constraints on asset managers’ deviations from market indices affect equilibrium prices. We focus on the special case of BVW where the constraints are infinitely tight, i.e., managers must track indices perfectly. We extend BVW by introducing correlation across assets through a systematic and an idiosyncratic component of asset payoffs. That extension is important for our analysis and results.

2 Model

Time $t$ is continuous and goes from zero to infinity. The riskless rate is exogenous and equal to $r > 0$. There are $N$ risky assets, which we interpret as stocks. Stock $n = 1, \ldots, N$ pays a dividend flow $D_{nt}$ per share and is in supply of $\eta_n > 0$ shares. The dividend flow of stock $n$ is

$$D_{nt} = \bar{D}_n + b_n D_s^t + D_{nti}, \quad (2.1)$$

the sum of a constant component $\bar{D}_n \geq 0$, a systematic component $b_n D_s^t$ and an idiosyncratic component $D_{nti}$. The systematic component is the product of a systematic factor $D_s^t$ times a factor loading $b_n \geq 0$. The systematic factor follows the square-root processes

$$dD_s^t = \kappa^s (\bar{D}_s^t - D_s^t) \, dt + \sigma^s \sqrt{D_s^t} \, dB_s^t, \quad (2.2)$$

where $(\kappa^s, \bar{D}_s^t, \sigma^s)$ are positive constants and $B_s^t$ is a Brownian motion. The idiosyncratic component follows the square-root process

$$dD_{nti} = \kappa_n^i (\bar{D}_n^i - D_{nti}^i) \, dt + \sigma_n^i \sqrt{D_{nti}^i} \, dB_{nti}^i, \quad (2.3)$$

where $\{\kappa_n^i, \bar{D}_n^i, \sigma_n^i\}_{n=1, \ldots, N}$ are positive constants and $\{B_{nti}^i\}_{n=1, \ldots, N}$ are Brownian motions that are mutually independent and independent of $B_s^t$. By possibly redefining factor loadings, we set the
long-run mean $\bar{D}^s$ of the systematic factor to one. By possibly redefining the supply $\eta_n$, we set the long-run mean $\bar{D}_n = b_n + \bar{D}^i_n$ of the dividend flow of stock $n$ to one for all $n$. With these normalizations, we can write the dividend flow of stock $n$ as

$$D_{nt} = 1 + b_n(D^s_t - 1) + (D^i_{nt} - \bar{D}^i_n).$$

The square-root specification (2.2) and (2.3) ensures that dividends and prices are always positive. It also implies that the volatility of dividends per share increases with the dividend level, a property which is important for our results. This property is realistic: if a firm becomes larger and keeps the number of its shares constant, then its dividends per share become more uncertain. A geometric Brownian motion specification for dividends, which is commonly used in the literature, would also imply the above properties. We adopt the square-root specification because it yields closed-form solutions.

Denoting by $S_{nt}$ the price of stock $n$, the stock’s return per share in excess of the riskless rate is

$$dR_{nt}^{sh} \equiv D_{nt}dt + dS_{nt} - rS_{nt}dt,$$

and the stock’s return per dollar in excess of the riskless rate is

$$dR_{nt} \equiv \frac{dR_{nt}^{sh}}{S_{nt}} = \frac{D_{nt}dt + dS_{nt}}{S_{nt}} - rdt.$$

We refer to $dR_{nt}^{sh}$ as share return, omitting that it is in excess of the riskless rate. We refer to $dR_{nt}$ as return, omitting that it is per dollar and in excess of the riskless rate.

Agents are competitive and form overlapping generations living over infinitesimal time intervals. Each generation includes experts and non-experts. Experts observe dividend flows, and can invest in the riskless asset and in the stocks without constraints. These agents can be interpreted as investors who invest in active funds. Non-experts do not observe dividend flows, and can invest in the riskless asset and in a stock portfolio that tracks an index. These agents can be interpreted as investors who invest in passive funds.
In addition to experts and non-experts, noise traders can be present. These agents generate an
exogenous demand for each stock, which is constant over time when expressed in number of shares,
and is smaller than the supply coming from the issuing firm. When noise traders are absent, or when
their demand is proportional to issuer supply in the cross-section, experts and non-experts hold
the same portfolio of stocks in equilibrium. When instead noise-trader demand is non-proportional
to issuer supply, experts hold a superior portfolio. Our main result on how the price impact of
passive flows depends on stock size does not require noise traders. The presence of noise traders
strengthens that result and yields additional implications.

The index does not include a possibly empty subset of stocks. It is value-weighted over the
remaining stocks, i.e., weights them proportionately to their market capitalization. We refer to the
included and the non-included stocks as index and non-index stocks, respectively. Denoting by $\mathcal{I}$
the subset of index stocks and by $\eta'_n$ the number of shares of stock $n$ included in the index, $\eta'_n = 0$
for $n \notin \mathcal{I}$. Since, in addition, the index is value-weighted over the stocks that it includes, $\eta'_n$ for
$n \in \mathcal{I}$ is constant over time and proportional to the number of shares $\eta_n$ coming from the issuer.
By possibly rescaling the index, we set $\eta'_n = \eta_n$ for $n \in \mathcal{I}$.

We denote by $W_{1t}$ and $W_{2t}$ the wealth of an expert and a non-expert, respectively, by $z_{1nt}$ and
$z_{2nt}$ the number of shares of stock $n$ that these agents hold, and by $\mu_1$ and $\mu_2$ these agents’ measure.
A non-expert thus holds $z_{2nt} = \lambda\eta'_n$ shares of stock $n$, where $\lambda$ is a proportionality coefficient that
the agent chooses optimally. We denote by $u_n < \eta_n$ the number of shares of stock $n$ held by noise
traders. The special case where noise traders are absent corresponds to $u_n = 0$ for all $n$.

Experts and non-experts born at time $t$ are endowed with wealth $W$. Their budget constraint
is

$$dW_{it} = \left( W - \sum_{n=1}^{N} z_{int} S_t \right) r dt + \sum_{n=1}^{N} z_{int} (D_t dt + dS_t) = W r dt + \sum_{n=1}^{N} z_{int} dR_{nt}^{sh}, \quad (2.7)$$

where $dW_{it}$ is the infinitesimal change in wealth over their life, $i = 1$ for experts, and $i = 2$ for
non-experts. They have mean-variance preferences over $dW_{it}$. These preferences can be derived
from any VNM utility $u$, as can be seen from the second-order Taylor expansion

$$u(W + dW_{it}) = u(W) + u'(W)dW_{it} + \frac{1}{2} u''(W)dW_{it}^2 + o(dW_{it}^2). \tag{2.8}$$

Experts, who observe $\{D_{nt}\}_{n=1,...,N}$, maximize the conditional expectation of (2.8). This is equivalent to maximizing

$$E_t(dW_{1t}) - \frac{\rho}{2}\text{Var}_t(dW_{1t}) \tag{2.9}$$

with $\rho = -\frac{u''(W)}{u'(W)}$, because infinitesimal $dW_{1t}$ implies that $E_t(dW_{1t}^2)$ is equal to $\text{Var}_t(dW_{1t})$ plus smaller-order terms. Non-experts, who do not observe $\{D_{nt}\}_{n=1,...,N}$, maximize the unconditional expectation of (2.8). This is equivalent to maximizing

$$E(dW_{2t}) - \frac{\rho}{2}\text{Var}(dW_{2t}), \tag{2.10}$$

because infinitesimal $dW_{1t}$ implies that $E(dW_{1t}^2)$ is equal to $\text{Var}(dW_{1t})$ plus smaller-order terms.

### 3 Equilibrium

We look for an equilibrium where the price $S_{nt}$ of stock $n$ is

$$S_{nt} = \bar{S}_n + b_nS^s(D^s_{nt}) + S^i_n(D^i_{nt}), \tag{3.11}$$

the sum of a constant term $\bar{S}_n$, a term $b_nS^s(D^s_{nt})$ corresponding to the systematic component of dividends and a term $S^i_n(D^i_{nt})$ corresponding to the idiosyncratic component. Assuming that the functions $(S^s(D^s_{t}), S^i_n(D^i_{nt}))$ are twice continuously differentiable, we can write the share return $dR^s_{nt}$ as

$$dR^s_{nt} = (\bar{D}_n + b_nD^s_{nt} + D^i_{nt})dt + (b_n dS^s(D^s_{nt}) + dS^i_n(D^i_{nt})) - r (\bar{S}_n + b_nS^s(D^s_{nt}) + S^i_n(D^i_{nt})) dt$$

$$= \mu_{nt}dt + b_n\sigma^s\sqrt{D^s_{nt}(S^s)'(D^s_{nt})}dB^s_{nt} + \sigma^i_n\sqrt{D^i_{nt}(S^i_n)'(D^i_{nt})}dB^i_{nt}, \tag{3.12}$$
where

\[
\mu_{nt} \equiv \frac{E_t(dR_{nt}^{sh})}{dt} = D_n - rS_n
\]
\[+ \nu_n \left[ D_t^s + \kappa^s (1 - D_t^s)(S^s)'(D_t^s) + \frac{1}{2}(\sigma^s)^2 D_t^s(S^s)''(D_t^s) - rS^s(D_t^s) \right]
\[+ \left[ D_{nt}^i + \kappa_n^i (D_n^i - D_{nt}^i)(S_n^i)'(D_{nt}^i) + \frac{1}{2}(\sigma_n^i)^2 D_{nt}^i(S_n^i)''(D_{nt}^i) - rS_n^i(D_{nt}^i) \right],
\] (3.13)

and the second step in (3.12) follows from (2.2), (2.3) and Ito’s lemma.

Using (2.7) and (3.12), we can write the objective (2.9) of experts as

\[
\sum_{n=1}^{N} z_{1nt} \mu_{nt} - \frac{\rho}{2} \left[ \left( \sum_{n=1}^{N} z_{1nt} \nu_n \right)^2 (\sigma^s)^2 D_t^s([S^s]'(D_t^s)]^2 + \sum_{n=1}^{N} z_{1nt}^2 (\sigma_n^i)^2 D_{nt}^i([S_n^i]'(D_{nt}^i)]^2 \right] dt,
\] (3.14)

Experts maximize (3.14) over positions \( \{z_{1nt}\}_{n=1}^{N} \). Their first-order condition is

\[
\mu_{nt} = \rho \left[ \nu_n \left( \sum_{n=1}^{N} z_{1nt} \nu_n \right) (\sigma^s)^2 D_t^s([S^s]'(D_t^s)]^2 + \sum_{n=1}^{N} z_{1nt}^2 (\sigma_n^i)^2 D_{nt}^i([S_n^i]'(D_{nt}^i)]^2 \right].
\] (3.15)

Using (2.7), (3.12) and \( z_{2nt} = \lambda \eta_n \), we can write the objective (2.10) of non-experts as

\[
\sum_{n=1}^{N} \lambda \eta_n \mu_n - \frac{\rho}{2} \lambda^2 \left[ \left( \sum_{n=1}^{N} \eta_n \nu_n \right)^2 (\sigma^s)^2 E[D_t^s([S^s]'(D_t^s)]^2 + \sum_{n=1}^{N} \eta_n^2 (\sigma_n^i)^2 E[D_{nt}^i([S_n^i]'(D_{nt}^i)]^2 \right) \right],
\] (3.16)

where \( \mu_n = \frac{E_t(dR_{nt}^{sh})}{dt} = E(\mu_{nt}) \). Non-experts maximize (3.16) over \( \lambda \). The first-order condition is

\[
\sum_{n=1}^{N} \eta_n \mu_n = \rho \lambda \left[ \left( \sum_{n=1}^{N} \eta_n \nu_n \right)^2 (\sigma^s)^2 E[D_t^s([S^s]'(D_t^s)]^2 + \sum_{n=1}^{N} \eta_n^2 (\sigma_n^i)^2 E[D_{nt}^i([S_n^i]'(D_{nt}^i)]^2 \right) \right].
\] (3.17)

Market clearing requires that the demand of experts, non-experts and noise traders equals the
supply coming from issuers:

$$\mu_1 z_{1nt} + \mu_2 \lambda \eta'_n + u_n = \eta_n. \tag{3.18}$$

Solving for $z_{1nt} = \frac{\eta_n - \mu_2 \lambda \eta'_n - u_n}{\mu_1}$, and substituting into the first-order condition (3.15) of experts, we find

$$\mu_{nt} = \rho \left[ b_n \left( \sum_{n=1}^{N} \frac{\eta_n - \mu_2 \lambda \eta'_n - u_n}{\mu_1} b_n \right) (\sigma^s)^2 D^*_t [(S^s)'(D^*_t)]^2 + \frac{\eta_n - \mu_2 \lambda \eta'_n - u_n}{\mu_1} (\sigma^i_n)^2 D^i_{nt} [(S^i_n)'(D^i_{nt})]^2 \right]. \tag{3.19}$$

We look for functions $(S^s(D^*_t), S^i_n(D^i_{nt}))$ that are affine in their arguments,

$$\begin{align*}
S^s(D^*_t) &= a^s_0 + a^s_1 D^*_t, \tag{3.20} \\
S^i_n(D^i_{nt}) &= a^i_{n0} + a^i_{n1} D^i_{nt}, \tag{3.21}
\end{align*}$$

for positive constants $(a^s_0, a^s_1, \{a^i_{n0}, a^i_{n1}\}_{n=1,...,N})$. Substituting (3.13), (3.20) and (3.21) into (3.19), we can write (3.19) as

$$\begin{align*}
\bar{D}_n - r \bar{S}_n + b_n \left[ (D^*_t + \kappa^s a^s_1 (1 - D^*_t)) - r (a^s_0 + a^s_1 D^*_t) \right] \\
+ (D^i_{nt} + \kappa^i_n a^i_{n1} (\bar{D}^i_n - D^i_{nt})) - r (a^i_{n0} + a^i_{n1} D^i_{nt}) \\
= \rho \left[ b_n \left( \sum_{n=1}^{N} \frac{\eta_n - \mu_2 \lambda \eta'_n - u_n}{\mu_1} b_n \right) (\sigma^s a^s_1)^2 D^*_t + \frac{\eta_n - \mu_2 \lambda \eta'_n - u_n}{\mu_1} (\sigma^i_n a^i_{n1})^2 D^i_{nt} \right]. \tag{3.22}
\end{align*}$$

Identifying terms in $D^*_t$, yields a quadratic equation that determines $a^s_1$. Identifying terms in $D^i_{nt}$, yields a quadratic equation that determines $a^i_{n1}$. Identifying the remaining terms, yields $\bar{S}_n + b_n a^s_0 + a^i_{n0}$. Substituting $(a^s_1, \{a^i_{n1}\}_{n=1,...,N})$ into the first-order condition (3.17) of non-experts, yields an equation for $\lambda$, whose solution completes our characterization of the equilibrium.

**Proposition 3.1.** In equilibrium, the price of stock $n$ is

$$S_{nt} = \frac{\bar{D}_n + b_n \kappa^s a^s_1 + \kappa^i_n a^i_{n1} \bar{D}^i_n}{r} + b_n a^s_1 D^*_t + a^i_{n1} D^i_{nt}, \tag{3.23}$$
where

\[ a_1^s = \frac{2}{r + \kappa^s + \sqrt{(r + \kappa^s)^2 + 4\rho \left( \sum_{m=1}^{N} \frac{\eta_m - \mu_2^s \lambda \eta_m' - u_m}{\mu_1} b_m \right) (\sigma^s)^2}}, \] (3.24)

\[ a_n^i = \frac{2}{r + \kappa_n^i + \sqrt{(r + \kappa_n^i)^2 + 4\rho \frac{\eta_n - \mu_2^i \lambda \eta_n' - u_n}{\mu_1} (\sigma_n^i)^2}}, \] (3.25)

and \( \lambda > 0 \) solves

\[
\left( \sum_{m=1}^{N} \eta_m' b_m \right) \left( \sum_{m=1}^{N} (\eta_m - u_m) b_m \right) (\sigma^s a_1^s)^2 + \sum_{m=1}^{N} \eta_m'(\eta_m - u_m)(\sigma_m a_m^1)^2 \bar{D}_m^i = \lambda(\mu_1 + \mu_2) \left[ \left( \sum_{m=1}^{N} \eta_m' b_m \right)^2 (\sigma^s a_1^s)^2 + \sum_{m=1}^{N} (\eta_m')^2 (\sigma_m a_m^i)^2 \bar{D}_m^i \right].
\] (3.26)

The price depends on \((\mu_1, \mu_2, \sigma^s, \{b_m, \sigma_m^i, \eta_m, \eta_m', u_m\}_{m=1,...,M})\) only through \(\left( \sum_{m=1}^{N} \frac{\eta_m - \mu_2^s \lambda \eta_m' - u_m}{\mu_1} b_m \right) (\sigma^s)^2\) and \(\frac{\eta_n - \mu_2^i \lambda \eta_n' - u_n}{\mu_1} (\sigma_n^i)^2\), and is decreasing and convex in the latter two variables.

The price of stock \(n\) depends on two measures of supply: systematic supply and idiosyncratic supply. Systematic supply is \(\left( \sum_{m=1}^{N} \frac{\eta_m - \mu_2^s \lambda \eta_m' - u_m}{\mu_1} b_m \right) (\sigma^s)^2\), the aggregate risk-adjusted supply of all stocks that each expert holds in equilibrium. The supply of stock \(m\) held by all experts combined is equal to the supply \(\eta_m\) coming from the issuer, minus the demand \(\mu_2^s \lambda \eta_m'\) and \(u_m\) coming from non-experts and noise traders, respectively. It is expressed in per-expert terms by dividing by the measure \(\mu_1\) of experts, is risk-adjusted by multiplying by the factor loading \(b_m\) of stock \(m\) and by the diffusion parameter \(\sigma^s\) of the systematic factor, and is aggregated across all stocks. Idiosyncratic supply is \(\frac{\eta_n - \mu_2^i \lambda \eta_n' - u_n}{\mu_1} (\sigma_n^i)^2\), the risk-adjusted supply of stock \(n\) that each expert holds in equilibrium. Risk adjustment is made by multiplying by the diffusion parameter \(\sigma_n^i\) of the idiosyncratic component of stock \(n\)’s dividends.

An increase in systematic or idiosyncratic supply causes the price of stock \(n\) to drop. This is the usual risk-premium channel. A supply increase has the additional effect that the price of stock \(n\) becomes less sensitive, in absolute terms, to dividend shocks. The intuition is as follows. A positive shock to dividends not only raises expected future dividends but also makes them riskier. (The
square-root specification implies that the diffusion coefficient of dividends per share increases with the level of dividends.) If the supply held by experts is positive, i.e., experts hold a long position, then the increase in risk makes them more willing to unwind their position by selling stock \( n \). This results in a smaller price increase compared to the case where supply is zero. If supply is negative, i.e., experts hold a short position, then the increase in risk makes them more willing to unwind their position by buying stock \( n \). This results in a larger price increase compared to the case where supply is zero.

4 Calibrated Example

We next characterize how flows into passive funds affect stock prices, and show our main results. In this section we show the results within a calibrated example. In Section 5 we consider general parameter values.

4.1 Parameter Values

The model parameters are the riskless rate \( r \), the number \( N \) of stocks, the parameters \((\kappa^s, \bar{D}^s, \sigma^s)\) and \((b_n, \kappa^i_n, \bar{D}^i_n, \sigma^i_n)_{n=1,...,N}\) of the dividend processes, the supply parameters \((\eta_n, \eta'_n, u_n)_{n=1,...,N}\), the measures \((\mu_1, \mu_2)\) of experts and non-experts, and the risk-aversion coefficient \( \rho \).

We set the sum \( \mu_1 + \mu_2 \) to one in the baseline case. This is a normalization because we can redefine \( \rho \). We set \( \rho \) to one. This is also a normalization because we can redefine the numeraire in the units of which wealth is expressed. Since the dividend flow is normalized by \( \bar{D}_n + b_n + \bar{D}^i_n = 1 \), redefining the numeraire amounts to rescaling the numbers of shares \((\eta_n, \eta'_n, u_n)_{n=1,...,N}\). We set the riskless rate \( r \) to 3%.

We assume that in the baseline case \( \mu_1 = 0.9 \) and \( \mu_2 = 0.1 \), i.e., non-experts hold 10% of total wealth. We examine how stock prices change when \( \mu_2 \) is raised to 0.6, i.e., non-experts’ wealth rises six-fold. We consider two polar cases for experts’ wealth, as well as all cases in-between. The first polar case is when flows into passive funds are due to increased participation by households in asset markets, and not to a switch from active to passive. In that case, experts’ wealth does not change and \( \mu_1 \) remains equal to 0.9. Non-experts’ wealth becomes two-thirds \((\frac{0.6}{0.9})\) of experts’ wealth, and
total investable wealth rises by 50% \((0.9 + 0.6)\). The second polar case is when flows into passive funds are due to a switch by households from active to passive. In that case, total investable wealth does not change and \(\mu_1 + \mu_2\) remains equal to one. Non-experts’ wealth becomes 50% larger than experts’ wealth \((0.6)\). We also consider all cases in-between, by setting \(\mu_1 = 0.9 - \zeta \times 0.5\), where \(\zeta \in [0, 1]\) is equal to zero in the first polar case and to one in the second polar case.

We calibrate the number \(N\) of stocks and the supply \(\eta_n\) coming from issuers based on the distribution of firms’ market capitalization in the US stock market. Averaging market capitalization across the ten largest US firms yields approximately one trillion dollars (per firm). That average scales down by a factor of approximately five when computed across the next 50 firms, then by five again when computed across the next 250 firms, then by five again when computed across the next 1250 firms, and then by five again when computed across the next 1250 firms. Consequently, we assume that there are ten stocks in supply of \(3125 \times \eta\) shares each, 50 stocks in supply of \(625 \times \eta\) shares each, 250 stocks in supply of \(25 \times \eta\) shares each, 1250 stocks in supply of \(5 \times \eta\) shares each, and 1250 stocks in supply of \(\eta\) shares each. We refer to the smallest stocks as size group 1 and to the largest stocks as size group 5.

We consider two cases about noise trader demand. The baseline case is that noise-trader demand is equal to zero for all stocks. The second case is that noise-trader demand \(u_n\) is equal to zero for one-half of the stocks in each size group, and is equal to a constant fraction \(\Delta u > 0\) of the supply coming from issuers for the remaining half. The former stocks are the low-demand ones and the latter stocks are the high-demand ones. We set \(\Delta u = 0.4\), i.e., noise traders demand 40% of the available supply of the high-demand stocks \((u_n = 0.4 \times \eta_n)\).

We consider two cases about index composition. The baseline case is that the index includes all stocks and is thus the true market portfolio, i.e., \(\eta_n' = \eta_n\) for all \(n\). The second case is that the index includes only the stocks in the top three size groups, i.e., \(\eta_n' = \eta_n\) for the 310 stocks in size groups 3, 4 and 5, and \(\eta_n' = 0\) for the 2500 stocks in size groups 1 and 2. Under the second assumption, the index is a large-stock index such as the S&P500.

---

\(^2\)As of 13 March 2022, the average market capitalization across the ten largest US firms was \$1.01 trillion; across the next 50 firms was \$207 billion; across the next 250 firms was \$48.1 billion; across the next 1250 firms was \$6.71 billion; and across the next 1250 firms was \$815 million. See https://companiesmarketcap.com/usa/largest-companies-in-the-usa-by-market-cap/.
We set the mean-reversion parameters $\kappa^i$ and $\{\kappa^i_n\}_{n=1}^{N}$ to a common value $\kappa$. We set the long-run means $\bar{D}_i^i$ and diffusion parameters $\{\sigma_i^i\}_{n=1}^{N}$ of the idiosyncratic components to common values $\bar{D}^i$ and $\sigma^i$, respectively. The stationary distribution of $D_{nt}^i$ is gamma with support $(0, \infty)$ and density

$$f(D_{nt}^i) = \frac{(\beta_i)^{\alpha^i}}{\Gamma(\alpha^i)} (D_{nt}^i)^{\alpha^i-1} e^{-\beta^i D_{nt}^i}, \quad (4.27)$$

where

$$\alpha^i \equiv \frac{2\kappa \bar{D}^i}{(\sigma^i)^2},$$
$$\beta^i \equiv \frac{2\kappa}{(\sigma^i)^2},$$

and $\Gamma$ is the Gamma function. The distribution of $D_{nt}^s$ is also gamma, with density given by (4.27) in which $D_{nt}^i$ is replaced by $D_{nt}^s$, $\alpha^i$ by $\alpha^s \equiv \frac{2\kappa \bar{D}^s}{(\sigma^s)^2} = \frac{2\kappa}{(\sigma^s)^2}$, and $\beta^i$ by $\beta^s \equiv \frac{2\kappa}{(\sigma^s)^2}$. We set $\frac{\sigma_i^i}{\sqrt{\bar{D}^i}} = \frac{\sigma_s^i}{\sqrt{\bar{D}^s}} = \sigma^s$. This ensures that the distributions of $D_{nt}^s$ and $D_{nt}^i$ are the same when scaled by their long-run means: $\frac{D_{nt}^i}{\bar{D}^i}$ has the same distribution as $\frac{D_{nt}^s}{\bar{D}^s} = D_{nt}^s$.

We allow for correlation between size and systematic risk. We assume that the value of the loading $b_n$ on the systematic factor for stocks in size group $m = 1, \ldots, 5$ is $b_n = \bar{b} - (m - 3)\Delta b \geq 0$. The relationship between size and systematic risk is negative when $\Delta b$ is positive, and vice-versa.

The parameters left to calibrate are $(\kappa, \bar{D}^i, \bar{b}, \Delta b, \sigma^s, \eta)$. We calibrate them based on stocks’ expected return, return variance, CAPM beta, and CAPM $R$-squared (fraction of return variance explained by index movements). We compute unconditional versions of these moments. We use the values in the baseline case as calibration targets. The formulas are in Appendix ?? and the values in the baseline case are in Table 1.

The effects of changing $\kappa$ on return moments and other numerical results are similar to those of changing the remaining parameters. We set $\kappa = 4\%$.

The values of $(\bar{D}^i, \bar{b}, \Delta b)$ must satisfy $\bar{b} + (m - 3)\Delta b + \bar{D}^i \leq 1$ for all $m = 1, \ldots, 5$ because of $\bar{D}_n \geq 0$ and the normalization $\bar{D}_n + b_n + \bar{D}^i = 1$. Inequality $\bar{b} + (m - 3)\Delta b + \bar{D}^i \leq 1$ for all $m = 1, \ldots, 5$ is equivalent to $\bar{b} + 2\Delta b + \bar{D}^i \leq 1$. We assume that the latter inequality holds as an
equality for the stocks with largest $b_n$. This minimizes the constant component $\bar{D}_n \geq 0$ (which becomes zero for the largest $b_n$ stocks). Minimizing $\bar{D}_n$ maximizes return variances by maximizing leverage, and brings them closer to their empirical counterparts.

We choose $\Delta b$ to be positive, consistent with the empirical negative relationship between size and CAPM beta. We set $\Delta b = 0.025$, to generate a spread in CAPM betas between small and large stocks of 0.40: CAPM beta averages 1.35 for the stocks in size group 1, and 0.95 for the stocks in size group 5. This is in line with the spread of 0.45 in Fama and French (1992): CAPM beta averages 1.42 for the stocks in size deciles 1 and 2, and 0.97 for the stocks in size deciles 9 and 10.

We determine the relative size of $\bar{b}$ and $\bar{D}^i$ based on CAPM $R^2$. We set $\bar{b} = 0.85$ and $\bar{D}^i = 0.10$, to generate a CAPM $R^2$ that averages to 22.69% across all stocks, and to 26.83% when stocks are weighted by number of shares. By comparison, the average adjusted $R^2$ from a CAPM regression with monthly returns and a five-year lookback window across all CRSP stocks in our sample period is 16.7% and the market-capitalization weighted average is 27.1%. Lowering the $R^2$ (by lowering $\bar{b}$ or raising $\bar{D}^i$) strengthens our results.

We determine the supply parameter $\eta$ based on stocks’ expected returns (in excess of the riskless rate). We set $\eta = 0.00003$, to generate expected returns across size groups that lie between 4-6%. Expected return ranges from 5.61% for the stocks in size group 1 to 4.09% for the stocks in size group 5.

We determine the diffusion parameter $\sigma^s$ based on stocks’ return variances. Raising $\sigma^s$ (and $\sigma^i$ through $\frac{\sigma^i}{\sqrt{D^i}} = \sigma^s$) has a non-monotone effect on variances. For given values of $D^s_t$ and $\{D^i_{nt}\}_{n=1,...,N}$, variances rise. At the same time, the stationary distributions of $D^s_t$ and $\{D^i_{nt}\}_{n=1,...,N}$ shift more weight towards very small or very large values, for which variances are low. We choose $\sigma^s$ to maximize return variances. Return volatility (square root of the variance) ranges from 21.12% for stocks in size group 1 to 11.58% for stocks in size group 5. These values are two to three times smaller than their empirical counterparts. The discrepancy is partly due to discount-rate shocks in our model being perfectly correlated with cashflow shocks and attenuating them. (Since experts hold a long position in the systematic component of dividends, they become more willing to sell stocks following a positive shock to that component and the resulting increase in risk.) Allowing for independent discount-rate shocks would increase return volatilities. Our results strengthen when
raising volatilities (i.e., when raising $\sigma^*$ towards its volatility-maximizing value).

4.2 No Noise Traders

Table 1 shows return moments in the baseline case, in which there are no noise traders, the index includes all stocks, and non-experts hold 10% of total wealth. Expected return, return volatility, and market beta decline when moving from the smallest to the largest size group. The decline in CAPM beta is built into our calibrated example because we set $\Delta b$ to a positive value. The decline in expected return reflects the decline in CAPM beta because without noise traders the conditional CAPM holds in our model. The decline in return volatility reflects partly the decline in CAPM beta. It also reflects that shocks to the idiosyncratic component of dividends have larger effects on the prices of small stocks. This is because Proposition 3.1 implies that discount-rate shocks attenuate idiosyncratic cashflow shocks to a lesser extent when idiosyncratic supply is small. Because the idiosyncratic component of dividends has larger effects on small stocks, CAPM $R^2$-squared rises when moving from small to large stocks, consistent with the empirical evidence.

Table 1: Return Moments.

<table>
<thead>
<tr>
<th>Size Group</th>
<th>Expected Return (%)</th>
<th>Return Volatility (%)</th>
<th>CAPM Beta</th>
<th>CAPM $R^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Smallest)</td>
<td>5.61</td>
<td>21.12</td>
<td>1.35</td>
<td>22.68</td>
</tr>
<tr>
<td>2</td>
<td>4.94</td>
<td>18.19</td>
<td>1.16</td>
<td>22.45</td>
</tr>
<tr>
<td>3</td>
<td>4.45</td>
<td>16.01</td>
<td>1.02</td>
<td>22.70</td>
</tr>
<tr>
<td>4</td>
<td>4.17</td>
<td>13.98</td>
<td>0.95</td>
<td>25.79</td>
</tr>
<tr>
<td>5 (Largest)</td>
<td>4.09</td>
<td>11.58</td>
<td>0.95</td>
<td>37.21</td>
</tr>
</tbody>
</table>

Table 2 shows how flows into passive funds affect stock prices. We compute the percentage change in the price $S_{nt}$ of stock $n$ assuming that the systematic component $D^s_t$ and idiosyncratic component $D^i_{nt}$ of dividends are equal to their long-run means. Since the price is linear in $D^s_t$ and $D^i_{nt}$, this amounts to computing the percentage change in the average price of stock $n$.

The second and third columns of Table 2 report the percentage price change when $\mu_2$ is raised...
to 0.6 and $\mu_1$ is held equal to 0.9. This corresponds to increased participation in the stock market through passive funds. The second column assumes that the index includes all stocks, and the third column assumes that only size groups 3, 4 and 5 are included. The fourth and fifth columns are counterparts of the second and third columns when $\mu_2$ is raised to 0.6 and $\mu_1$ is lowered to 0.4. This corresponds to a switch from active to passive.

Table 2: Percentage Price Change Following Flows into Passive Funds.

<table>
<thead>
<tr>
<th>Size Group</th>
<th>Increase in Market Participation</th>
<th>Switch from Active to Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Stocks in Index</td>
<td>Size Groups 3-5 in Index</td>
</tr>
<tr>
<td>1 (Smallest)</td>
<td>6.51</td>
<td>6.33</td>
</tr>
<tr>
<td>2</td>
<td>5.60</td>
<td>5.29</td>
</tr>
<tr>
<td>3</td>
<td>5.44</td>
<td>5.67</td>
</tr>
<tr>
<td>4</td>
<td>6.54</td>
<td>7.57</td>
</tr>
<tr>
<td>5 (Largest)</td>
<td>7.71</td>
<td>9.84</td>
</tr>
</tbody>
</table>

Table 2 shows our main results. Consider first the case where flows into passive funds are due to increased participation in the stock market, and where the index includes all stocks. As shown in the second column of Table 2, stock prices increase. Moreover, the effect is non-monotone with size. The percentage price increase becomes smaller when moving from size group 1 to size group 3. It becomes larger when moving from size group 3 to size group 5, and is the largest across all five groups for size group 5.

The non-monotone effect is surprising. Indeed, consider an one-period CAPM, in which stock $n$ pays expected dividend $\bar{D}_n$ and has CAPM beta $b_n$. The stock's expected return is $r + b_n\text{MRP}$, where $r$ and MRP are the one-period riskless rate and market risk premium, respectively. The price of stock $n$ is

$$S_n = \frac{\bar{D}_n}{1 + r + b_n\text{MRP}}.$$
Since flows into passive funds lower MRP, their effect is proportional in the cross-section to

\[- \frac{1}{S_n} \frac{\partial S_n}{\partial \text{MRP}} = \frac{b_n}{1 + r + b_n \text{MRP}},\]

and is increasing in \( b_n \). It should thus be the largest for small stocks, which have the highest \( b_n \) in our calibrated example. Intuitively, flows into passive funds have a larger effect on stocks with higher \( b_n \) because their expected returns are more sensitive to changes in MRP. This mechanism explains why the price increase in the second column of Table 2 becomes smaller when moving from small to medium-size stocks.

When moving from medium-size to large stocks, a different mechanism is at work and its effects are opposite to the standard mechanism. To understand the new mechanism, we use Proposition 3.1 and note that flows into passive funds amount to raising \( \mu^2 \lambda \) (the measure \( \mu^2 \) of non-experts times the fraction \( \lambda \) of the index that they hold). Differentiating (3.23) with respect to \( \mu^2 \lambda \), we find that the percentage change in the price of stock \( n \) in response to an increase in \( \mu^2 \lambda \) is

\[
\frac{1}{S_{nt}} \frac{\partial S_{nt}}{\partial (\mu_2 \lambda)} = \frac{1}{S_{nt}} \left[ b_n \left( \frac{\kappa^s}{r} + D_t^s \right) \frac{\partial a_1^s}{\partial (\mu_2 \lambda)} + \left( \frac{\kappa^i}{r} D_n^i + D_{nt}^i \right) \frac{\partial a_1^i}{\partial (\mu_2 \lambda)} \right] = \frac{\rho}{S_{nt} \mu^1} \left[ b_n \left( \sum_{m=1}^N b_m \eta_m \right) (\sigma^s a_1^s)^2 \left( \frac{\kappa^s}{r} + D_t^s \right) + \frac{\eta_n^i (\sigma_n^i a_1^i)^2 \left( \frac{\kappa^i}{r} D_n^i + D_{nt}^i \right)}{\sqrt{(r + \kappa^s)^2 + 4 \rho \left( \sum_{m=1}^N \eta_m - \mu_2 \lambda \eta_m - \mu_1 \right) b_m (\sigma^s)^2}} \right].
\]

The price effect in (4.28) is proportional in the cross-section to a modified CAPM beta. Relative to the standard beta, the modified beta gives greater weight to the part of the covariance between stock \( n \) and the index that arises because of the idiosyncratic component of dividends. Equations (3.12), (3.20) and (3.21) imply that the covariance between the return of stock \( n \) and the share return of the index is

\[
\text{Cov} \left( dR_{nt}, \sum_{m=1}^N \eta_m' dR_{mt}^h \right) = \frac{1}{S_{nt}} \left[ b_n \left( \sum_{m=1}^N b_m \eta_m \right) (\sigma^s a_1^s)^2 D_t^s + \eta_n^i (\sigma_n^i a_1^i)^2 D_{nt}^i \right].
\]

The first term in the square bracket in (4.29) is the systematic part of the covariance, arising
because of the systematic component of stock $n$’s dividends. The second term is the idiosyncratic part, arising because of the idiosyncratic component of dividends. Both parts are present in (4.28), as the numerators of the respective terms in the square bracket. Relative to (4.29), (4.28) gives greater weight to the idiosyncratic part because the denominator is smaller than for the systematic part.

The economic intuition is as follows. Passive flows reduce the systematic and idiosyncratic supply held by experts, causing the price of stock $n$ to rise. This is the standard mechanism. The effects of systematic and idiosyncratic supply on the price are proportional to the systematic and idiosyncratic parts, respectively, of the covariance between stock $n$ and the index. The idiosyncratic part is much smaller than the systematic part, even for the largest stocks: in Table 2, it is smaller by a factor of approximately twenty for the stocks in size group 5.

The new mechanism is that passive flows make the price of stock $n$ more sensitive to dividend shocks. The increase in price sensitivity renders a position (of a fixed number of shares) in stock $n$ riskier. This affects the experts’ willingness to hold the stock and feeds back into the stock’s price rise. While the feedback due to the reduction in systematic supply always attenuates the price rise, the feedback due to the reduction in idiosyncratic supply attenuates it less and can even amplify it. It is because of this feedback that the modified beta gives greater weight than the standard beta to the idiosyncratic part of the covariance.

Consider first the feedback due to the reduction in systematic supply. Since systematic supply is always positive, i.e., experts hold a long position in the systematic component of dividends, an increase in price sensitivity to shocks to that component makes experts more willing to unwind their position by selling stock $n$. This attenuates the price rise.

Consider next the feedback due to the reduction in idiosyncratic supply, and suppose that idiosyncratic supply is positive, i.e., experts hold a long position in stock $n$. Since idiosyncratic supply is smaller than systematic supply, an increase in price sensitivity to shocks to the idiosyncratic component of dividends has a smaller effect in experts’ willingness to sell stock $n$ than the same increase in price sensitivity pertaining to the systematic component. Therefore, the attenuation effect is weaker. Because of the weaker attenuation, the idiosyncratic part of modified beta in Table 2 is smaller than the systematic part by a factor of only two for the stocks in size group 5.
Suppose next that idiosyncratic supply is negative, i.e., experts hold a short position in stock \( n \) (while being long in the systematic component of dividends). This happens for stocks in high noisetrader demand in our calibrated example (Table 6). Attenuation then turns into amplification. The increase in price sensitivity to shocks to the idiosyncratic component makes experts more willing to unwind their position by buying stock \( n \). This amplifies the price rise.

For small stocks, modified beta is close to standard beta because the idiosyncratic part of the covariance is negligible. (The second term in brackets in (4.28) and (4.29) is close to zero.) For large stocks instead, the idiosyncratic part of the covariance causes modified beta to differ significantly from standard beta. The effect of passive flows on price for large stocks is driven by the reduction in idiosyncratic supply and the increase in risk that it generates.

Consider next the case where the index includes only stocks in size groups 3, 4 and 5 (and flows into passive funds are still due to increased participation in the stock market). As shown in the third column of Table 2, the percentage price increase remains non-monotone when moving across size groups. Relative to the case where the index includes all stocks, the effect rises more sharply with size when moving from size group 3 to size group 5. This is because non-experts establish larger positions in the more restricted set of stocks, causing the reduction in idiosyncratic supply to be larger.

Consider finally the case where flows into passive funds are due to a switch from active to passive. When the index includes all stocks, stock prices do not change. This is because experts and non-experts hold the same portfolio, which is the index. When instead the index includes only stocks in size groups 3, 4 and 5, prices do not change for size groups 1, 2 and 3, but rise for size group 4 and especially 5. This is because the higher demand by non-experts is not offset by lower demand by non-experts, causing idiosyncratic supply to drop for stocks in size groups 3, 4 and 5.

Flows into passive funds affect not only stock prices but also expected returns and return volatilities. Table 3 shows the effect of flows on expected returns. When flows into passive funds are due to increased participation in the stock market, expected returns drop across all size groups. The drop is more pronounced for size group 1 because of that group’s highest CAPM beta. (The percentage price rise in Table 2 is more pronounced for size group 5 because expected return is lower for these stocks, so a given drop in expected return yields a larger percentage price rise.)
When flows into passive funds are due to a switch from active to passive, and the index includes all stocks, expected returns do not change. When instead the index includes only stocks in size groups 3, 4 and 5, expected returns rise for size groups 1 and 2, and drop for size groups 4 and especially 5.

Table 3: Change in Expected Return Following Flows into Passive Funds.

<table>
<thead>
<tr>
<th>Size Group</th>
<th>Baseline Expected Return</th>
<th>Change in Expected Return</th>
<th>Increase in Market Participation</th>
<th>Switch from Active to Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>All Stocks in Index</td>
<td>Size Groups 3-5 in Index</td>
</tr>
<tr>
<td>1 (Smallest)</td>
<td>5.61</td>
<td>-0.77</td>
<td>-0.76</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>4.94</td>
<td>-0.63</td>
<td>-0.62</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4.45</td>
<td>-0.55</td>
<td>-0.55</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>4.17</td>
<td>-0.52</td>
<td>-0.54</td>
<td>0</td>
</tr>
<tr>
<td>5 (Largest)</td>
<td>4.09</td>
<td>-0.53</td>
<td>-0.57</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4 shows the effect of passive flows on return volatilities. Volatilities do not change when the index includes all stocks and flows into passive funds are due to a switch from active to passive. In all other cases, volatilities do not change for the small size groups but rise significantly for size groups 4 and especially 5. The effect is driven by the increased price sensitivity to shocks to the idiosyncratic component of dividends. For small stocks, the reduction in idiosyncratic supply due to passive flows has a negligible effect on the price and on the price sensitivity to the idiosyncratic component of dividends. The reduction in systematic supply raises the price and the price sensitivity to the systematic component of dividends, but the two effects cancel out in the calculation of return volatility: the price sensitivity per unit of price stays constant. For large stocks instead, the price sensitivity to the idiosyncratic component of dividends rises significantly as well, and causes the rise in volatility.
Table 4: Change in Return Volatility Following Flows into Passive Funds.

<table>
<thead>
<tr>
<th>Size Group</th>
<th>Baseline Return Volatility</th>
<th>Change in Return Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Increase in Market Participation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>All Stocks in Index</td>
</tr>
<tr>
<td>1 (Smallest)</td>
<td>21.12</td>
<td>-0.04</td>
</tr>
<tr>
<td>2</td>
<td>18.19</td>
<td>0.11</td>
</tr>
<tr>
<td>3</td>
<td>16.01</td>
<td>0.22</td>
</tr>
<tr>
<td>4</td>
<td>13.98</td>
<td>0.39</td>
</tr>
<tr>
<td>5 (Largest)</td>
<td>11.58</td>
<td>0.65</td>
</tr>
</tbody>
</table>

4.3 Noise Traders

Table 5 is the counterpart of Table 1 with noise traders. Stocks within each size group are split equally across those without noise traders and those for which noise traders hold 40% of the supply coming from issuers. This yields ten groups of stocks. The effects across size groups are similar to those in Table 1. The effects within size groups depend on size. Within size groups 1 and 2, expected return and volatility are independent of noise-trader demand. Within size groups 3, 4 and 5 instead, expected return declines and volatility rises when moving from low to high noise-trader demand. The effect of noise-trader demand becomes larger as stock size increases because noise-trader demand is proportional to supply and thus largest for the largest stocks.

Table 5 implies that the risk-return relationship is positive across size groups, is zero within small stocks and is negative within large stocks. The positive risk-return relationship across size groups is driven by fundamentals: high CAPM beta of small stocks implies high expected return and high volatility. The negative risk-return relationship within large stocks is driven by noise-trader demand: high demand implies low expected return and high volatility. Noise-trader demand raises the volatility of large stocks for the same reason why flows into passive funds raise their volatility (Table 4): the price becomes more sensitive to the idiosyncratic component of dividends.
Table 5: Return Moments with Noise Traders.

<table>
<thead>
<tr>
<th>Size Group</th>
<th>Noise-Trader Demand</th>
<th>Expected Return (%)</th>
<th>Return Volatility (%)</th>
<th>Market Beta</th>
<th>CAPM $R^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Smallest)</td>
<td>Low</td>
<td>5.17</td>
<td>21.10</td>
<td>1.34</td>
<td>24.95</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>5.17</td>
<td>21.10</td>
<td>1.34</td>
<td>24.93</td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
<td>4.58</td>
<td>18.25</td>
<td>1.16</td>
<td>24.78</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>4.58</td>
<td>18.25</td>
<td>1.16</td>
<td>24.69</td>
</tr>
<tr>
<td>3</td>
<td>Low</td>
<td>4.16</td>
<td>16.10</td>
<td>1.03</td>
<td>25.11</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>4.13</td>
<td>16.16</td>
<td>1.02</td>
<td>24.70</td>
</tr>
<tr>
<td>4</td>
<td>Low</td>
<td>3.91</td>
<td>14.10</td>
<td>0.96</td>
<td>28.40</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>3.84</td>
<td>14.31</td>
<td>0.95</td>
<td>26.88</td>
</tr>
<tr>
<td>5 (Largest)</td>
<td>Low</td>
<td>3.86</td>
<td>11.75</td>
<td>0.95</td>
<td>40.06</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>3.73</td>
<td>12.19</td>
<td>0.94</td>
<td>36.72</td>
</tr>
</tbody>
</table>

Table 6 is the counterpart of Table 2 with noise traders. When flows into passive funds are due to increased participation in the stock market, their effect varies across size groups in a manner similar to Table 2. The effect within size groups 1, 2 and 3 is independent of noise-trader demand. Within size groups 4 and 5 instead, flows have a larger effect on the prices of high-demand stocks. This is because idiosyncratic supply is smaller for these stocks, so the attenuation effect is weaker.

When flows into passive funds are due to a switch from active to passive, their effect is somewhat different than in Table 2. The difference is clearest in the case where the index includes all stocks. Without noise traders, flows have no effect on prices. With noise-traders instead, stocks in low noise-trader demand drop in price, stocks in high demand rise, and the effect is asymmetric implying an overall rise when averaged within size groups. Stocks in low noise-trader demand drop because they are in high demand by experts, so a switch from experts to non-experts lowers their net demand. Conversely, stocks in high noise-trader demand rise because they are in low demand by experts, so a switch raises their net demand. The intuition why the effect is asymmetric is easier to see in the case where the index includes only size groups 3, 4 and 5. In that case, flows into passive...
Table 6: Percentage Price Change Following Flows into Passive Funds with Noise Traders.

<table>
<thead>
<tr>
<th>Size Group</th>
<th>Noise-Trader Demand</th>
<th>Increase in Market Participation</th>
<th>Switch from Active to Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All Stocks in Index</td>
<td>Size Groups 3-5 in Index</td>
</tr>
<tr>
<td>1 (Smallest)</td>
<td>Low</td>
<td>6.97</td>
<td>6.79</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>6.97</td>
<td>6.79</td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
<td>5.98</td>
<td>5.71</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>5.97</td>
<td>5.70</td>
</tr>
<tr>
<td>3</td>
<td>Low</td>
<td>5.66</td>
<td>5.80</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>5.65</td>
<td>5.81</td>
</tr>
<tr>
<td>4</td>
<td>Low</td>
<td>6.36</td>
<td>7.07</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>6.72</td>
<td>7.72</td>
</tr>
<tr>
<td>5 (Largest)</td>
<td>Low</td>
<td>7.13</td>
<td>8.49</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>8.94</td>
<td>12.08</td>
</tr>
</tbody>
</table>

Funds result in experts holding short positions in stocks in high noise-trader demand. This implies an amplification effect for high-demand stocks, which can raise their price substantially, as Table 6 shows for size group 5.

5 General Results

[[[TO BE WRITTEN]]]

6 Empirical Evidence

In this section we show that our model’s main prediction that flows into passive funds raise disproportionately the prices of the largest stocks in the index holds in the data. We take the index to be the S&P500, and passive flows to be into index mutual funds and index ETFs tracking the S&P500. The S&P500 index attracts the bulk of passive investing in US stocks.
6.1 Data and Descriptive Statistics

Our data on stock returns and firm accounting variables come from the Center for Research in Security Prices (CRSP) and Compustat. Our data on assets and flows for index mutual funds tracking the S&P500 come from the Investment Company Institute (ICI). ICI does not report data on S&P500 ETFs. We instead collect those data from CRSP. CRSP reports data on domestically listed ETFs. We include in our analysis only plain-vanilla ETFs, excluding alternative ETFs such as leveraged ETFs, inverse ETFs and buffered ETFs. Our ETF sample consists of the SPDR S&P 500 ETF Trust, the iShares Core S&P 500 ETF, and the Vanguard S&P 500 Index Fund ETF, which collectively account for almost all of the plain-vanilla S&P500 ETF market.

Table 7 reports descriptive statistics for our main variables, for the sample of S&P500 stocks and the period July 2000 to June 2019. The descriptive statistics in Panel A of Table 7 concern firm-level variables. The average stock earns an average monthly return of 0.91%, with a standard deviation of 9.76%. It has an average market capitalization of $27 billion and an average index weight of 0.18%. The distributions of market capitalization and index weight are skewed to the right, with high skewness and kurtosis.

Table 7: Descriptive Statistics

<table>
<thead>
<tr>
<th>Panel A: Firm-Level Variables</th>
<th>Mean</th>
<th>Std Dev</th>
<th>25th Pctl</th>
<th>50th Pctl</th>
<th>75th Pctl</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly Return ((R_t \times 100))</td>
<td>0.91</td>
<td>9.76</td>
<td>-3.75</td>
<td>1.06</td>
<td>5.62</td>
<td>0.46</td>
<td>10.93</td>
</tr>
<tr>
<td>Market Cap ((\text{millions}))</td>
<td>27,393</td>
<td>51,383</td>
<td>5,9021</td>
<td>11,827</td>
<td>25,305</td>
<td>5.71</td>
<td>50.56</td>
</tr>
<tr>
<td>Index Weight ((\times 100))</td>
<td>0.18</td>
<td>0.32</td>
<td>0.04</td>
<td>0.08</td>
<td>0.16</td>
<td>5.05</td>
<td>34.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Time-Series Variables (Quarterly, (\times 100))</th>
<th>Index Fund Holdings</th>
<th>Flow_1</th>
<th>Flow_2</th>
<th>Dispersion</th>
<th>HHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>4.17</td>
<td>0.05</td>
<td>0.03</td>
<td>0.35</td>
<td>0.81</td>
</tr>
<tr>
<td>Std Dev</td>
<td>1.08</td>
<td>0.10</td>
<td>0.09</td>
<td>0.03</td>
<td>0.11</td>
</tr>
<tr>
<td>25th Pctl</td>
<td>3.40</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.32</td>
<td>0.72</td>
</tr>
<tr>
<td>50th Pctl</td>
<td>4.11</td>
<td>0.04</td>
<td>0.02</td>
<td>0.34</td>
<td>0.79</td>
</tr>
<tr>
<td>75th Pctl</td>
<td>4.88</td>
<td>0.10</td>
<td>0.07</td>
<td>0.37</td>
<td>0.88</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.46</td>
<td>0.51</td>
<td>0.27</td>
<td>0.38</td>
<td>0.52</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.57</td>
<td>3.44</td>
<td>3.73</td>
<td>-0.79</td>
<td>-0.63</td>
</tr>
</tbody>
</table>

This table shows descriptive statistics for the main variables in our sample from July 2000 to June 2019. Panel A includes the following firm-level variables: monthly stock return in percent, market capitalization in millions of dollars, and weight of a stock in the S&P500 index in percent. Panel B includes the following aggregate variables: index fund holdings (ratio of S&P500 index fund net assets to index value), changes in index fund holdings (Flow_1), dollar flows into index funds divided by index value (Flow_2), cross-sectional standard deviation of S&P500 index weights (Dispersion), and Herfindahl-Hirschman Index (HHI) of S&P500 index weights. The variables in Panel B are sourced at a quarterly frequency and are multiplied by 100.
In our sample period, the growth of passive funds was substantial. As shown in Figure 1, the assets of passive funds tracking the S&P500 more than tripled, growing from less than $500 billion in July 2000 to more than $1.5 trillion in July 2019. As a result, the funds’ ownership of S&P500 stocks more than doubled, expanding from 2.5% to more than 6%. Because index fund holdings exhibit a secular trend during our sample period, we focus on fund flows in our empirical tests.

Figure 1: Assets of Index Mutual Funds and ETFs Tracking the S&P500

This figure plots the value of the assets of index mutual funds and index ETFs tracking the S&P500 over the period June 2000 to June 2019. The red line represents the ratio of fund net assets to index value (left y-axis). The blue bars represent index fund net assets in millions of dollars (right y-axis). The data come from CRSP and ICI.

The descriptive statistics in Panel B of Table 7 concern aggregate variables, sourced at a quarterly frequency. The variables are: index fund holdings, index fund flows, and the concentration of index weights. We define index fund holdings at the end of quarter $t$ as the ratio of the value of index fund net assets to the value of the S&P500 index (i.e., the combined value of the S&P500
firms):

\[ \text{IndexFund}_t = \frac{\$\text{IndexAssets}_t}{\$\text{SP500}_t}. \]

We use two measures of index fund flows. The first is the change in index fund holdings between the end of quarters \( t - 1 \) and \( t \):

\[ \text{Flow}_{1,t} = \text{IndexFund}_t - \text{IndexFund}_{t-1}. \]

The second is dollar flows in quarter \( t \) divided by the index value at the end of that quarter:

\[ \text{Flow}_{2,t} = \frac{\$\text{Flow}_t}{\$\text{SP500}_t}. \]

ICI reports dollar flows into index mutual funds, so we use that direct measure. ICI does not report dollar flows into ETFs, so we infer these indirectly from CRSP using the change in ETF net assets and the ETF return as:

\[ \$\text{ETFFlow}_t = \$\text{ETFAssets}_t - \$\text{ETFAssets}_{t-1} \times (1 + \text{ETFRet}_t). \]

We characterize the effect of passive flows on large versus small stocks using two measures. The first is the return of a small minus big (SMB) portfolio. We construct that portfolio in each quarter by forming decile portfolios based on size, with Decile 1 containing the smallest stocks in the S&P500 index (i.e., the stocks with the smallest index weights), and Decile 10 containing the largest stocks (i.e., the stocks with the largest index weights). The SMB S&P500 portfolio consists of a long position in the stocks in Decile 1 combined with an equal short position in the stocks in Decile 10. We construct that portfolio in both equally-weighted and value-weighted terms.

The second measure is the concentration of index weights. Concentration reflects the extent to which the weight of large-capitalization index assets exceeds that of small-capitalization ones. We characterize concentration of index weights at the end of quarter \( t \) by two sub-measures: the cross-sectional standard deviation of index weights (\( \text{Dispersion}_t \)) and the Herfindahl-Hirschman
6.2 Results

6.2.1 SMB Return

We use the regression specification:

$$SMB_{SP_{i,t}} = \alpha_{i,j} + \gamma_{i,j,contemp} \times Flow_{j,contemp,t} + \gamma_{i,j,past} \times Flow_{j,past,t} + \epsilon_{i,j,t},$$

where $SMB_{SP_{i,t}}$ is the return of the SMB S&P500 portfolio in quarter $t$, with $i = ew$ if the return is computed in equal-weighted terms, and $i = vw$ if the return is computed in value-weighted terms; $Flow_{j,contemp,t}$ are contemporaneous index fund flows, with $j = 1, 2$ corresponding to the two measures of flows; and $Flow_{j,past,t}$ are past index fund flows. Our model implies $\gamma_{i,j,contemp} < 0$: passive flows raise disproportionately the prices of the largest stocks in the index.

We define contemporaneous flows in quarter $t$ to also include flows in the previous quarter:

$$Flow_{j,contemp,t} = Flow_{j,t} + Flow_{j,t-1}.$$

We define past flows to include flows in more distant quarters going back to one year and a half:

$$Flow_{j,past,t} = \sum_{i=2}^{6} Flow_{j,t-i}.$$

We include the previous quarter into contemporaneous flows to account for lags between flows and trade. For example, when ETF sponsors accept cash from authorized participants (APs) to create new ETF shares, they may not purchase the constituent securities immediately. Similarly, when APs redeem ETF shares, ETF sponsors return the constituent securities to APs but APs may not sell the securities immediately.

Table 8 shows the regression results. The coefficient $\gamma_{i,j,contemp}$ on contemporaneous flows is negative across the four specifications in Columns 1–4. The effects are statistically significant and economically large. For example, a one standard deviation increase in $Flow_{1,contemp,t}$ is associated with a contemporaneous decline in the return of the equally weighted SMB S&P500 portfolio by

26
3.56% per quarter. Hence, flows into passive funds tracking the S&P500 index tend to drive up the prices of large stocks in the index by more than the prices of small stocks in the index.

It is, of course, possible that a negative $\gamma_{i,j,contemp}$ reflects reverse causality. Suppose that large stocks in the S&P500 index perform well. Since they are the main driver of the index, the index performs well too. If investors are performance-chasers, then they invest more in the index. This gives rise to a negative relationship between index fund flows and the return of the SMB S&P500 portfolio. To partly address this concern, we include the index return in the regressions. The regression results remain similar.

The coefficient $\gamma_{i,j,past}$ on past flows is positive across the four specifications in Columns 1–4. It is statistically significant, however, only for $\text{Flow}_{2,past,t}$ (Columns 3–4). That effect is economically large. For example, a one standard deviation increase in $\text{Flow}_{2,past,t}$ predicts an increase in the future return of the equally weighted SMB portfolio by 2.82% per quarter.\footnote{A one standard deviation increase in $\text{Flow}_{1,contemp,t}$ is 0.14% ($= 0.10\% \times \sqrt{2}$). Multiplied by the slope coefficient -25.47, it yields an effect of -3.56%.
A one standard deviation increase in $\text{Flow}_{2,past,t}$ is 0.20% ($=0.09\% \times \sqrt{5}$). Multiplied by the slope coefficient 14.01, it yields and effect of 2.82%.
}

In Columns 5–8 of Table 8, we perform the same regressions as in Columns 1–4 for quarters when VIX is above average. In all specifications, the results are statistically significant despite a smaller sample. Moreover, their economic significance strengthens. In particular, the coefficient $\gamma_{i,j,past}$ on past flows is four times as large for $\text{Flow}_{1,contemp,t}$ and twice as large for $\text{Flow}_{2,past,t}$.

### 6.2.2 Concentration of Index Weights

We use the regression specification

$$\Delta\text{Concentration}_t = \alpha_j + \gamma_j \times \text{Flow}_{j,contemp,t} + \epsilon_{j,t},$$

where $\text{Concentration}_t$ is the concentration of index weights at the end of quarter $t$, measured by $\text{Dispersion}_t$ or $\text{HHI}_t$, and $\text{Flow}_{j,contemp,t}$ are contemporaneous index fund flows, with $j = 1, 2$. The regression also includes the lagged equal- or value-weighted return of the SMB S&P500 portfolio, to control for momentum. Our model implies $\gamma_j > 0$. Table 9 shows the regression results. The coefficient $\gamma_j$ is positive and statistically significant across all four specifications.
This table shows the dynamic relationship between S&P500 index fund flows and the return of the SMB S&P500 portfolio, computed in equal- and value-weighted terms. $Flow_{1,\text{contemp}}$ and $Flow_{1,\text{past}}$ are the contemporaneous and past index fund flows measured as changes in index fund holdings. $Flow_{2,\text{contemp}}$ and $Flow_{2,\text{past}}$ are the contemporaneous and past index fund flows measured as dollar flows divided by index value. $SMB_{SPew}$ and $SMB_{SPvw}$ are the return of the SMB S&P500 portfolio, computed in equal- and value-weighted terms, respectively. Columns 1–4 use the full sample, and Columns 5–8 use the sample when the VIX is above the sample mean.
Table 9: Passive Flows and Concentration in Index Weights

Panel A: $Flow_{1,contemp}$

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta Dispersion$</td>
<td>0.0282</td>
<td>0.104</td>
<td>0.0271</td>
<td>0.100</td>
<td>0.0268</td>
<td>0.0992</td>
</tr>
<tr>
<td>$\Delta HHI$</td>
<td>(2.97)</td>
<td>(2.97)</td>
<td>(2.89)</td>
<td>(2.96)</td>
<td>(2.84)</td>
<td>(2.91)</td>
</tr>
<tr>
<td>$L.SMB_{SPew}$</td>
<td>-2.29e-05</td>
<td>-2.19e-05</td>
<td>(-0.18)</td>
<td>(-0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L.SMB_{SPvw}$</td>
<td>-4.55e-05</td>
<td>-0.000108</td>
<td>(-0.35)</td>
<td>(-0.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-3.99e-05</td>
<td>-0.000151</td>
<td>-3.46e-05</td>
<td>-0.000130</td>
<td>-3.37e-05</td>
<td>-0.000127</td>
</tr>
<tr>
<td>Observations</td>
<td>76</td>
<td>76</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.106</td>
<td>0.112</td>
<td>0.107</td>
<td>0.110</td>
<td>0.109</td>
<td>0.112</td>
</tr>
</tbody>
</table>

Panel B: $Flow_{2,contemp}$

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta Dispersion$</td>
<td>0.0209</td>
<td>0.0758</td>
<td>0.0226</td>
<td>0.0836</td>
<td>0.0225</td>
<td>0.0831</td>
</tr>
<tr>
<td>$\Delta HHI$</td>
<td>(1.99)</td>
<td>(1.95)</td>
<td>(2.23)</td>
<td>(2.28)</td>
<td>(2.22)</td>
<td>(2.26)</td>
</tr>
<tr>
<td>$L.SMB_{SPew}$</td>
<td>-7.75e-05</td>
<td>-0.000224</td>
<td>(-0.61)</td>
<td>(-0.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L.SMB_{SPvw}$</td>
<td>-0.000105</td>
<td>-0.000330</td>
<td>(-0.80)</td>
<td>(-0.69)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-2.25e-05</td>
<td>-8.60e-05</td>
<td>-1.77e-05</td>
<td>-6.79e-05</td>
<td>-1.70e-05</td>
<td>-6.52e-05</td>
</tr>
<tr>
<td>Observations</td>
<td>76</td>
<td>76</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.051</td>
<td>0.049</td>
<td>0.070</td>
<td>0.071</td>
<td>0.073</td>
<td>0.074</td>
</tr>
</tbody>
</table>

This table shows the relationship between S&P500 index fund flows and changes in the concentration of index weights. Panel A shows the results for $Flow_{1,contemp}$; Panel B for $Flow_{2,contemp}$. We use two measures of concentration of index weights: the cross-sectional standard deviation ($\Delta Dispersion$) and the Herfindahl-Hirschman Index ($\Delta HHI$). We use the lagged return of the small-minus-big index portfolio as a control variable.

7 Conclusion

We show that flows into passive funds raise disproportionately the prices of the largest stocks in the index, while also making them more volatile. If, in addition, stocks are mispriced because of noise traders, then passive flows raise disproportionately the prices of the overvalued stocks among the index’s largest. Passive flows generally drive the aggregate market up even when they are
entirely due to a switch from active to passive. Underlying these effects is that passive flows make prices more sensitive to news about idiosyncratic future cashflows. We provide empirical evidence in support of our model’s main predictions.
References

