Index-Linked Trading and Stock Returns*†

Shaun William Davies‡

First Draft: March 13, 2021
This Draft: January 31, 2022

Comments and Suggestions Welcome

Abstract

I consider a model of index-linked trading in which a fraction of investors trade an index product that holds the market portfolio (e.g., an ETF). Other investors build portfolios by evaluating stocks individually. Investors are equally informed and choose portfolios to maximize their expected utility. In equilibrium, price impact from trading the index product is not equal across stocks. Index-linked trade generates cross sectional differences in returns and volatilities. Furthermore, uncertainty about indexing demand generates risk premiums in expected returns and their magnitudes depend on firm fundamentals. The findings lend a theoretical foundation to existing studies and I provide empirical support for new predictions.

*Keywords: Index investing, price impact, cross section of returns

†I would like to thank David C. Brown, Andrea Buffa, Tony Cookson, Andrew Detzel, Diego Garcia, Chris Leach, Zhengzi (Sophia) Li, Matthew Ringgenberg, Charles Robbins, Shri Santosh, Michael Stutzer, Xingtan Zhang and seminar participants at the University of Colorado, Boulder, Clemson University, Villanova University, 2021 Colorado/Texas Research Idea Conference, INQUIRE Residential Seminar 2021, and Paris December 2021 Finance Meeting. All errors are my own.

‡Leeds School of Business, University of Colorado, Boulder, Campus Box 419, Boulder, CO 80309, shaun.davies@colorado.edu.

At the end of the day, I'm against ETFs because they often create enormous distortions that can obliterate even the best of stocks.

—Jim Cramer, March 21, 2017
CNBC's Mad Money

1 Introduction

There has been an explosion of index-linked trading over the last several decades.[1] A large segment of investors has changed from a behavior of investing with actively managed mutual funds to a behavior of investing in passively managed index funds. Moreover, many day traders have moved beyond stock picking and now trade entire markets or sectors via exchange-traded funds (ETFs).[2] Market participants, like Jim Cramer, have attributed these trends to causing stock price distortions and academic researchers have taken notice.[3] There is now a burgeoning literature examining the asset pricing implications of investors buying and selling baskets of securities. Recent studies have convincingly shown that index products, like ETFs, indeed affect the cross section of returns: they introduce comovement in returns (Da and Shive (2018)), they transmit volatility to the individual stock shares (Ben-David, Franzoni, and Moussawi (2018); Krause, Ehsani, and Lien (2013)), and stocks commonly held in ETF portfolios earn a substantial risk premium (Ben-David et al. (2018); Jiang, Vayanos, and Zheng (2020)). Using a rational and informationally efficient market as the benchmark, are these findings *anomalous*? It is difficult to say as standard models do not consider the impact of index-linked trading. In this paper, I consider a simple model of index-linked trading in which there is uncertainty about both stocks' future cash flows and future index-product demand. I argue that existing empirical work is not only consistent with the appropriate null hypotheses but, even in an informationally efficient market, it would be anomalous to not find the documented effects.

My model of index-linked trade starts with the five assumptions underpinning standard asset

[1]Kenneth French's American Finance Association Presidential Address, titled "The Cost of Active Investing," includes extensive data regarding the trends in index investing and passive investing. See French (2008).
[2]As an example, see https://www.wsj.com/articles/how-etfs-swallowed-the-stock-market-11567162801 .
[3]See Wurgler (2010) for an early survey on the market impacts of index investing.

pricing models and modern portfolio theory: (i) investors have the same opinions about the possibilities of various end-of-period payoffs for all stocks, (ii) the common probability distribution describing the possible payoffs is well-behaved (e.g., jointly normal), (iii) investors are risk averse and choose portfolios that maximize their expected terminal utility, (iv) investors may take a long or short position of any size in any asset, including a risk-free asset and (v) markets are efficient and absorb information quickly and perfectly. In addition to these standard assumptions, my model adds three ingredients to study the equilibrium effects of index-linked trade. First, I assume that risky assets are in fixed supply and investors are infinitesimal. As such, while stocks' dividends are normally distributed and exogenous, prices (and expected returns) are endogenously determined via market clearing. Moreover, investors are price-takers and do not internalize any impact of their portfolio choices on equilibrium prices. Second, I assume that, for a fraction of investors, it is prohibitively costly to trade individual stocks. Instead, these investors, which I refer to as indexers, trade a basket security (e.g., an ETF) which holds the market portfolio. Third, I introduce uncertainty regarding the demand of investors that utilize the index product in the future. With these three additional assumptions, I find that index-linked trading is associated with comovement in stock returns, cross sectional differences in stock returns and volatilities, and an "index-linked trade risk premium" which differs across stocks based on their individual fundamentals.

The model's main insights center on two key observations: index investors have no discretion on stocks' individual portfolio positions (i.e., the underlying portfolio is passively managed and essentially fixed) and, while subtle, equilibrium price impact from trading the index fund is *not* equal across stocks. Thus, when investors enter or exit the index fund, the resulting price impact differs across the index fund portfolio; said differently, not only do demand curves for stock shares slope down (Shleifer (1986)), the slopes are also different. The heterogeneous price impact across stocks is driven by firm fundamentals with riskier stocks being characterized by greater price impact. Following the logic of Berk (1995), this also implies a negative correlation between price impact and market capitalization.[4]

[4]Berk (1995) argues that the negative correlation between average returns and firm size (i.e., the "size anomaly") is not anomalous and instead should be expected. The basic intuition for the argument is that riskier firms have lower market prices, all else equal.

Several practical implications stem from these two key observations and lend support to existing empirical findings. First, as index funds hold representative sets of the cross section (e.g., the S&P 500 often proxies for "the market" and is also a popular benchmark in index funds and ETFs), index investor demand should mechanically lead to return comovement. While stock returns are differentially affected by index investor demand, the directions are largely the same and return correlations are generally higher. Second, variation in index investor demand manifests itself in measurable volatility differences across stocks. Obviously, stocks commonly held by index funds and ETFs experience more return volatility from index-linked trading than stocks that are not commonly held. Not so obviously, even if all stocks are equally held by index funds and ETFs, index investor trades have stronger effects in some stocks as compared to others. Third, to the extent that index investor demand is uncertain, stocks have differential exposures to that uncertainty. As index investors trade broad portfolios, the demand uncertainty is essentially undiversifiable and individual stocks command different risk premiums based on their exposures. Thus, stock returns reflect "index-linked trade risk." These two observations also suggest that stock return variation is driven in part by index-linked trade which provides new understanding to studies that show that the rise of index investing has obfuscated firm-specific information in prices (Israeli, Lee, and Sridharan (2017); Brogaard, Ringgenberg, and Sovich (2019)).

While the model provides a new lens to view existing research through, it also provides several novel empirical predictions. First, the model predicts that there is a negative correlation between market capitalization and price impact. To support the first prediction, Figure 1 provides evidence from S&P 500 constituent stocks. In the figure, each month's S&P 500 constituent stocks are sorted into ventiles (i.e., twenty different bins) by their market capitalizations during the sample period of January 2005 through May 2020 (the smallest stocks are placed in ventile 1 and the largest stocks are placed in ventile 20). Average monthly bid-ask spreads (as percentages) are computed for each stock and the figure displays a scatter plot of the equally-weighted average monthly bid-ask spread versus ventile bin. Figure 1 shows a negative relation between market capitalization ventiles and price impact; stocks with relatively smaller market capitalizations are characterized by larger bid-ask spreads and stocks with larger market capitalizations are characterized by smaller bid-ask

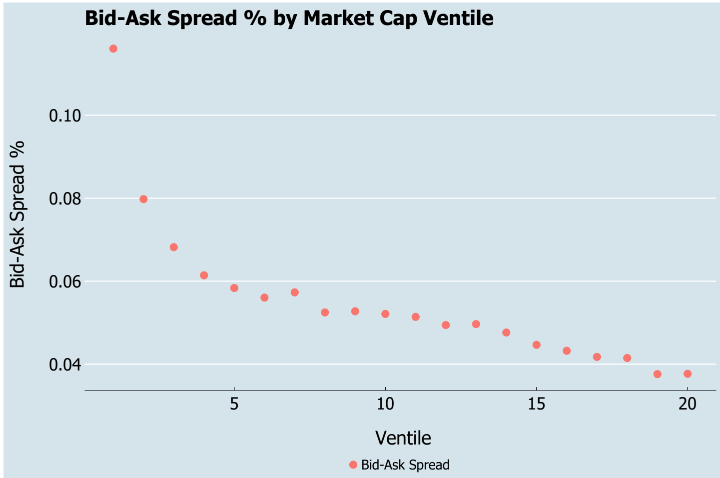


Figure 1: Bid-Ask Spread Versus Market Capitalization. Each month's S&P 500 constituent stocks are sorted into ventiles by their market capitalizations (the smallest stocks are placed in ventile 1 and the largest stocks are placed in ventile 20). With the sorted stocks, equally-weighted average monthly bid-ask spreads (as percentages) are computed and the figure displays a scatter plot of the average bid-ask spreads versus ventile bins. The sample period is January 2005 through May 2020.

spreads. Thus, when an investor purchases an index product with market-based weights (which is equivalent to purchasing the same fraction in each stock's shares outstanding), the price impact will be more severe in smaller stocks versus larger stocks.[5]

Second, the model predicts that index-linked trade generates cross sectional return volatility differences which are also negatively correlated with market capitalizations. To support the second prediction, Figure 2 examines the relations between stocks' market capitalizations, their return volatilities and S&P 500 ETF flows.[6] First, daily ETF flow volatilities for the three largest S&P 500 ETFs (SPY, IVV, and VOO) are calculated at the monthly frequency and sorted into terciles with tercile 1 including months with the lowest daily ETF flow volatility and tercile 3 including months with the highest daily ETF flow volatility.[7] Next, within each tercile-month, S&P 500 constituent stocks are sorted into ventiles with the smallest stocks being placed into ventile 1 and

[5]Bid-ask spread is just one measure of price impact. The advantages of bid-ask spread are the simplicity of the measure and its ability to be compared across stocks (i.e., it simply the round-trip cost, in percentage terms, of buying at the ask and selling at the bid). Using other measures of price impact, such as Amihud Illiquidity (Amihud (2002)), yields a nearly identical looking figure.

[6]ETF flows are investments into and out of an ETF wrapper via the ETF primary market. See Brown, Davies, and Ringgenberg (2021) for additional details.

[7]SPY is State Street Global Advisors' S&P 500 ETF, IVV is BlackRock's S&P 500 ETF, and VOO is Vanguard's S&P 500 ETF. As of July 2021, SPY, IVV and VOO were ranked as the three of the top four largest ETFs by assets under management (AUM). Collectively, as of July 2021, the three ETFs hold over $900 billion in S&P 500 stocks.

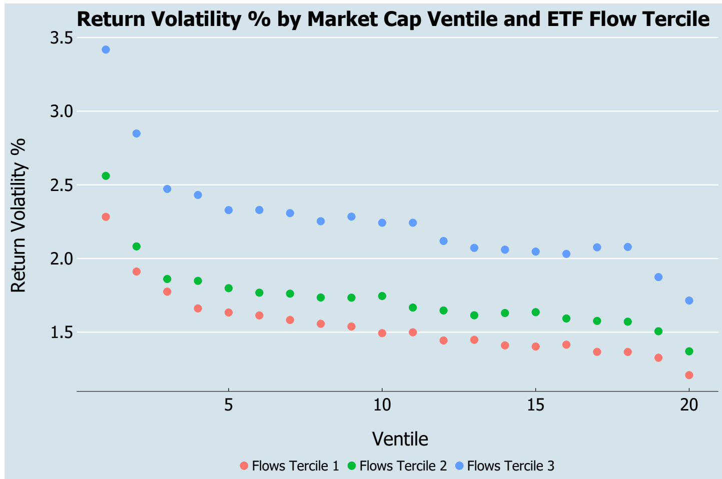


Figure 2: Return Volatility Versus Market Capitalization and ETF Flow Volatility. Months are sorted into terciles by daily S&P 500 ETF flow volatility with the lowest daily S&P 500 ETF flow volatilities appearing in tercile 1 and the highest daily S&P 500 ETF flow volatilities appearing in tercile 3. Within each S&P 500 ETF flow volatility tercile, each month's S&P 500 constituent stocks are sorted into ventiles by their market capitalizations (the smallest stocks are placed in ventile 1 and the largest stocks are placed in ventile 20). With the dual sorted stocks, equally-weighted average daily return volatilities (as percentages) are computed and the figure displays a scatter plot of the average return volatilities versus ventile bins. The sample period is January 2005 through May 2020.

the largest stocks being placed in ventile 20. Finally, within each tercile-ventile sort, the stocks' equally-weighted average daily return volatility is computed. For each ETF flow volatility tercile, Figure 2 shows a negative relation between market capitalization ventiles and average daily return volatility; stocks with relatively smaller market capitalizations are characterized by greater return volatility as compared to those with larger market capitalizations. Importantly, across ETF flow terciles, months in which ETF flows are more volatile, return volatility is higher for all ventiles and the difference between low ventiles and high ventiles increases. In other words, S&P 500 ETF flows exacerbate stocks' return volatilities across the entire S&P 500 universe and in a disproportional manner such that smaller stocks are more heavily influenced.

While Figure 1 and Figure 2 lend credence to the insights that (i) equilibrium price impacts negatively correlate with market capitalizations and (ii) index-linked trade generates cross sectional return volatility differences which are also negatively correlated with market capitalizations, later in this paper I provide formal empirical evidence to support my findings. First, I perform panel regressions in which the dependent variable is price impact (measured by either bid-ask spread or Amihud Illiquidity) and the independent variable is the natural log of market capitalization. Consistent with Figure 1, I find a statistically significant and economically meaningful negative relation between log market capitalization values and price-impact. Second, I perform additional panel regressions in which the dependent variable is return volatility and the independent variables are the natural log of market capitalization values and S&P 500 ETF flow volatility. Consistent with Figure 2, I find a statistically significant and an economically strong negative relation between log market capitalization values and return volatility. That is, smaller capitalization stocks exhibit greater return volatility. More importantly and also consistent with Figure 2, the interaction term for log market capitalization values and index flow volatility carries a statistically significant negative coefficient implying that the effect is exacerbated by index flows.

The main contribution of this paper is in providing a baseline framework to evaluate the relation between expected stock returns and index-linked trading without usual confounding tensions like asymmetric information or agency conflicts. That is, in the spirit of early capital market equilibrium models that provide predictions for expected returns and exposures to market risk (e.g., Sharpe (1964)), my model provides several insights to identify the appropriate null hypotheses relating to index-linked trade when there is uncertainty about both cash flows and indexer demand. Additionally, my findings also address a common misconception among many that market capitalization weights somehow equate price impacts across a portfolio, which they do not.[8]

While my model assumes that investors are equally informed, several studies focus on the role of index-linked trade in markets with private information. Subrahmanyam (1991) shows that uninformed investors prefer index products (e.g., stock index futures) to avoid adverse selection in individual stocks. More recently, Bhattacharya and O'Hara (2018) and Cong and Xu (2020) study the advent of ETFs and their effects on price informativeness and the allocation of liquidity (i.e., noise) traders across the securities market.[9] Similarly, Bond and Garcia (2020) considers a rational expectations model in which index investing serves as intermediate solution between active portfolio management and not participating in financial markets altogether. The study shows that, as the

[8] Admittedly, I long held the naïve belief that value-weighting a portfolio served as a sort of panacea against differential price impact.

[9] See also Stambaugh (2014) and Baruch and Zhang (2018).

cost of index investing falls, index investing becomes relatively more attractive to both marginal active managers and non-participants.[10] Thus, a decline in the cost of index investing leads to less information in prices, more noise trade, and less adverse selection.

In an independent, contemporaneous, and complementary study, Jiang et al. (2020) also studies the asset pricing implications of index-linked trade.[11] Jiang et al. (2020) shows that increases in index investing lead to higher expected returns for small capitalization stocks. The study supports this theoretical prediction by showing that a small-minus-large portfolio of S&P 500 stocks earns ten percent per year during the sample period of 2000-2019. Importantly, while both that study and my model predict higher returns for small capitalization stocks, the mechanism is different. In Jiang et al. (2020), persistent inflows to index products disproportionately increase the prices (and lower future returns) of large capitalization stocks relative to small capitalization stocks. This generates the small-minus-large return prediction in that model. Conversely, in my model, both increases and decreases in index-linked trading disproportionately affect stocks with the most price impact which are also, on average, smaller firms. Thus, to the extent that there is uncertainty about future index investor demand, the stocks with the greater price impact measures (which are also likely small capitalization stocks) carry larger index investor risk premiums and higher expected returns. Ultimately, which theoretical explanation is correct is an empirical question.

Finally, several studies have recently shown that ETF flows predict returns. Ben-David et al. (2018) documents that transitory demand shocks for ETFs are transmitted to the underlying stocks via the ETF primary market, leading to short-term return reversals at the stock-level. In a similar vein, Brown et al. (2021) shows theoretically and empirically that ETF flows signal non-fundamental demand shocks that slowly revert giving rise to medium-term return reversals at the ETF-level. Admittedly, my model does not feature return predictability from indexer's demand as shocks are i.i.d. across periods, the shocks are permanent, and the realization of shocks are immediately incorporated into prices. In this light, my model should not be viewed as a complete

[10] In related work, Brown and Davies (2017) shows that a decline in fees for passively managed index products decreases active managers' incentives to exert effort and search for mispriced assets.

[11] Jiang et al. (2020) builds from the framework of Buffa, Vayanos, and Woolley (2020), which examines active portfolio managers' deviations from market indices and the impact on prices. See also Stutzer (2018) for a related study on the equilibrium implications of style investing (i.e., the practice in which investors try to either meet or beat returns from style-specific benchmark portfolios).

characterization of the stock market and index-linked trade. Instead, it serves as the rational and informationally efficient benchmark. Incorporating additional frictions such as myopic traders, slow moving arbitrage capital, or sentiment traders may provide incremental insights beyond those studied in this paper.

2 Model

Consider a two-period stock market in which investors trade shares of multiple firms and an index fund. Investors purchase shares in period $t = 1$ and shares pay random dividends in period $t = 2$. There is no asymmetric information among investors and the only uncertainty investors face is in regards to the dividends produced by each firm (in the next subsection, I introduce uncertainty about future index fund demand in a multi-period extension). The supplies of each firm's shares are fixed and prices are endogenously determined via market clearing similar to the framework in Merton (1987).

The stock shares traded by investors come from $N \geq 2$ firms. Each firm has a unit-length measure of publicly traded shares (equivalently, one could think of these shares as percentages of the firm since they add to one). I denote the share price of firm j as P_j and the stock dividend as D_j. The vector of firm dividends are distributed according to a multivariate normal distribution in which the $N \times 1$ vector of expected dividends is denoted μ and the $N \times N$ covariance matrix of dividends is denoted Ω. I denote element j of μ as μ_j. The index fund traded by investors holds the N stocks and is market-capitalization-weighted (i.e., the market portfolio). I assume that the index fund has an expense ratio of zero and that stock shares may be bundled into or out of the fund without frictions. In addition to the firms' stock shares and the index fund, a riskless asset in perfectly elastic supply generates a gross rate of return normalized to one.

There are three types of investors: portfolio-builders, indexers, and block-holders. Portfolio-builders and indexers are risk-averse with exponential utility and CARA parameter $\gamma > 0$. Both portfolio-builders and indexers are each born endowed with wealth $W > 0$. I assume there is a continuum of length $b > 0$ of portfolio-builders and a continuum of length $m > 0$ of indexers. Portfolio-builders evaluate the merits of each firm individually and portfolio-builder k chooses to hold $s_{k,j}$ shares in firm j to maximize her expected terminal utility. Indexers are prohibited from trading individual firm shares and, instead, indexer i chooses to hold f_i shares of the index fund to maximize her terminal utility.[12] Furthermore, I allow both portfolio-builders and indexers to hold short positions.

Unlike portfolio-builders and indexers, block-holders are the simplest of investors considered. I assume that block-holders hold a measure $h_j \in [0, 1]$ in stock j and h_j is exogenously given at the beginning of the game. Denote the $N \times 1$ vector of block-holder holdings as $\mathbf{h}$. Block-holders may be thought of as insiders (e.g., a founder) that hold large undiversified positions in firm j's stock. Alternatively, block-holders may be thought of as a group of investors that have particular preferences for shares unrelated to a risk-return tradeoff (e.g., investors that purchase TSLA shares because they like Tesla). While the presence of block-holders provides additional insights and predictions, their key importance in the model is to enable there to be an observable distinction between portfolio-builders and indexers. That is, as long as $h_j \neq h_z$ for at least one pair of j and z, then portfolio-builders and indexers hold different portfolios in equilibrium; portfolio-builders hold the mean-variance efficient (MVE) portfolio and indexers hold the market portfolio. Conversely, if $h_j = h_z$ for all j and z, then, in equilibrium, portfolio-builders and indexers hold identical portfolios and there is no distinction as the MVE portfolio and market portfolio are the same.

Additional discussion regarding the role of block-holders is helpful in understanding the model. First, as just mentioned, the presence of block-holders allows for an observable distinction between the portfolios of portfolio-builders and indexers. I note that this distinction is important for the sake of intuition but not for the model's main insights. In fact, I show that the main results hold even in the limit as $h_j \to 0$ for all j. Second, while I use block-holders to generate a wedge between the MVE portfolio and the market portfolio, one could accomplish the same objective differently. For example, one could assume that the index product traded by indexers is a subset of stocks (e.g., the S&P 500 is a subset of all publicly traded stocks). With such an assumption, there would be

[12]Note, in equilibrium, the expected utility of portfolio-builders weakly exceeds that of indexers. Thus, while I assume that the measures of portfolio-builders and indexers are exogenous in the base model, I make the measures endogenous in Section OA.1 of the Internet Appendix. Specifically, by forcing investors to incur an attention cost to become a portfolio-builder (i.e., an incremental cost for portfolio building as compared to index investing), the fractions of indexers and portfolio-builders are endogenously determined.

three relevant portfolios: the market portfolio, the index fund portfolio, and the MVE portfolio.[13] Nevertheless, the main insights from that framework would be analogous to the setup outlined here.

To solve the model, I begin with the observation that a share in the index fund must have the price $\sum_{j=1}^{N} P_j$, that is, purchasing a percentage of the index fund is equivalent to purchasing that same percentage in each firm. Likewise, the dividend paid to a share of the index fund is equal to $\sum_{j=1}^{N} D_j$. As such, indexer i's terminal wealth (at $t = 2$) is

$$W_i^T = W + f_i \left(\sum_{j=1}^{N} D_j - \sum_{j=1}^{N} P_j \right), \tag{1}$$

and indexer i chooses f_i to maximize her expected utility at $t = 1$. Indexer i's optimization at $t = 1$ is,

$$\max_{f_i \in \mathbb{R}} E \left[1 - e^{-\gamma \left(W + f_i \left(\sum_{j=1}^{N} D_j - \sum_{j=1}^{N} P_j \right) \right)} \right]. \tag{2}$$

Next, consider portfolio-builders. Portfolio-builder k's terminal wealth (at $t = 2$) is,

$$W_k^T = W + \sum_{j=1}^{N} s_{k,j} \left(D_j - P_j \right), \tag{3}$$

and portfolio-builder k chooses $\mathbf{s}_k$, which is an $N \times 1$ vector representing investor k's share purchases in the N firms and element j is $s_{k,j}$, to maximize her expected utility at $t = 1$. Portfolio-builder k's optimization at $t = 1$ is,

$$\max_{\mathbf{s}_k \in \mathbb{R}^N} E \left[1 - e^{-\gamma \left(W + \sum_{j=1}^{N} s_{k,j} (D_j - P_j) \right)} \right]. \tag{4}$$

The equilibrium price for each stock is determined by market clearing; the price is set such that aggregation of investors' demands for stock j equals the supply of shares. To provide the

[13] Merton (1987) generates similar wedges between investor portfolios, the market portfolio, and the MVE portfolio by considering a setup in which investors may only trade a subset of stocks for which they have knowledge of.

most natural and intuitive explanation of equilibrium prices, hereafter I introduce new shorthand notation. Denote the market dividend as,

$$D_M \equiv \sum_{j=1}^{N} D_j, \tag{5}$$

the variance of the market dividend as,

$$\sigma_M^2 \equiv \mathbf{e}'\Omega\mathbf{e}, \tag{6}$$

in which $\mathbf{e}$ is an $N \times 1$ vector of ones. Furthermore, denote the covariance between firm j's dividend D_j and the market dividend D_M as,

$$\sigma_{j,M} \equiv \mathrm{COV}(D_j, D_M) = \sum_{y=1}^{N} \sigma_{j,y}, \tag{7}$$

in which $\sigma_{j,y}$ is the covariance between firm j's dividend and firm y's dividend.

Lemma 1. *The equilibrium price for firm j's stock is,*

$$P_j^* = \mu_j - \frac{\gamma\sigma_{j,M}}{b}\left(\sum_{z=1}^{N} \chi_{j,z}(1-h_z) - \frac{m}{m+b}\sum_{z=1}^{N} \phi_z(1-h_z)\right). \tag{8}$$

in which $\phi_z = \frac{\sigma_{z,M}}{\sigma_M^2}$, $\sum_{z=1}^{N} \phi_z = 1$, $\chi_{j,z} = \frac{\sigma_{j,z}}{\sigma_{j,M}}$ *and* $\sum_{z=1}^{N} \chi_z = 1$. *In the limit, as* $h_j \to 0$ *for all* j, *the price for firm j's stock is,*

$$P_j^* = \mu_j - \frac{\gamma\sigma_{j,M}}{m+b}. \tag{9}$$

To provide intuition for P_j^*, first consider the case in (9) in which $h_j \to 0$ for all j. In this special case, the first part of the price is the expected dividend μ_j and the second part is the risk premium $\frac{\gamma\sigma_{j,M}}{m+b}$. The risk premium contains the familiar term $\gamma\sigma_{j,M}$ which is the CARA parameter multiplied by the covariance of the firm dividend with the market dividend. The risk premium also contains the term $\left(\frac{1}{m+b}\right)$ which has a numerator that represents the total measure of shares that must be held by either portfolio-builders or indexers and the denominator is the total measure of these two investor groups. In other words, the price of stock j is its expected dividend net of a risk premium that reflects dividend risk and risk-sharing.

Next, consider the case in (8) in which $h_j \neq h_z$ for at least one pair j and z. The first part of the price is again the expected dividend μ_j. The second part is the risk premium,

$$\frac{\gamma \sigma_{j,M}}{b}\left(\sum_{z=1}^{N} \chi_{j,z}(1-h_z) - \frac{m}{m+b}\sum_{z=1}^{N}\phi_z(1-h_z)\right),$$

which contains the familiar term $\gamma\sigma_{j,M}$ but also the term $\frac{1}{b}\left(\sum\limits_{z=1}^{N} \chi_{j,z}(1-h_z) - \frac{m}{m+b}\sum\limits_{z=1}^{N}\phi_z(1-h_z)\right)$ which requires some explanation and guiding intuition.

To understand the risk premium expression, first notice that $\sum\limits_{z=1}^{N}\phi_z(1-h_z)$ is a pseudo-weighted average of the shares that must be held by either indexers or portfolio-builders across all firms. In the pseudo-weighted average, the weight for firm j is its dividend covariance with the market dividend divided by the aggregate dividend variance (i.e., the firm's dividend beta).[14] In this sense, since portfolio-builders and indexers are risk averse, $\sum\limits_{z=1}^{N}\phi_z(1-h_z)$ represents the average fraction of shares that the investors must hold across all firms on a risk-adjusted basis. For example, a firm with little covariance with the market dividend will not affect the average much but a firm with considerable covariance will. I refer to $\sum\limits_{z=1}^{N}\phi_z(1-h_z)$ as the *risk-adjusted market free float.*

Next, note that risk-adjusted market free float is multiplied by $\frac{m}{m+b}$ which represents the fraction of indexers versus portfolio-builders; if $\frac{m}{m+b}$ is large, there are relatively more indexers than portfolio-builders and vice versa when the fraction is small. As such, the product of $\sum\limits_{z=1}^{N}\phi_z(1-h_z)$ and $\frac{m}{m+b}$ may be thought of as the portion of the risk-adjusted market free float that will be held by indexers. In fact, the equilibrium demand of each indexer is $f^* = \frac{1}{m+b}\sum\limits_{z=1}^{N}\phi_z(1-h_z)$ and indexers' aggregate demand is $\int\limits_0^m \frac{1}{m+b}\sum\limits_{z=1}^{N}\phi_z(1-h_z)\,\mathrm{d}x = \frac{m}{m+b}\sum\limits_{z=1}^{N}\phi_z(1-h_z)$.

Next, consider the term $\sum\limits_{z=1}^{N}\chi_{j,z}(1-h_z)$ which is also a pseudo-weighted average. Unlike the risk-adjusted market free float which considers each firm's float and covariance with the market dividend, $\sum\limits_{z=1}^{N}\chi_{j,z}(1-h_z)$ considers each firm's float and covariance with only firm j's dividend.

[14] I use the term "pseudo-weighted average" as ϕ_j may be negative valued for some firm j (i.e., the firm's dividend has negative covariance with the market dividend).

In this sense, $\sum_{z=1}^{N} \chi_{j,z}(1-h_z)$ represents the average measure of shares that must be held on a risk-adjusted basis for stock j only. Suppose there are two firms, j and z, that have perfectly correlated dividends and the variances of their dividends is identical. Moreover, assume that these two firms' dividends are uncorrelated with all other firms'. Then, firm j and firm z are effectively identical in terms of risk and only differ with respect to the quantities of this risk, that is, $1-h_j$ and $1-h_z$. Thus, in the computation of $\sum_{z=1}^{N} \chi_{j,z}(1-h_z)$, the two firm's free floats are averaged.[15] As portfolio-builders evaluate each stock individually in the context of their portfolios, they are not concerned with the explicit supply of a firm's shares but how that supply covaries with the risk of supplies in other stocks that must also be held.

Having discussed that $\sum_{z=1}^{N} \chi_{j,z}(1-h_z)$ represents the risk-adjusted supply of firm j's stock and $\frac{m}{m+b}\sum_{z=1}^{N} \phi_z(1-h_z)$ represents the demand of indexers, $\left(\sum_{z=1}^{N} \chi_{j,z}(1-h_z) - \frac{m}{m+b}\sum_{z=1}^{N} \phi_z(1-h_z)\right)$ represents the risk-adjusted supply of firm j's shares that must be held by portfolio-builders. The risk-adjusted supply of firm j's shares per portfolio-builder is $\frac{1}{b}\left(\sum_{z=1}^{N} \chi_{j,z}(1-h_z) - \frac{m}{m+b}\sum_{z=1}^{N} \phi_z(1-h_z)\right)$. Therefore, similar to the price in (9), the price in (8) also includes stock j's expected dividend net of a risk premium that reflects dividend risk and risk-sharing.

Having solved for equilibrium prices, I now consider the price impact of index-linked trading. For analytic tractability and for the benefit of economic intuition, I make one simplifying assumption.

Assumption 1. *The risk-adjusted market free float is positive valued, that is,* $\sum_{z=1}^{N} \phi_z(1-h_z) \geq 0$.

Assumption 1 is natural as it requires that, on a risk-adjusted basis, portfolio-builders and indexers must be collectively net long. However, without Assumption 1, one could choose a covariance matrix Ω and vector of block-holder holdings $\mathbf{h}$ such that the risk-adjusted market free float is negative valued. While possible, I view this as unlikely and instead focus exclusively on the case in which the risk-adjusted market free float is positive valued. With Assumption 1, one may now consider the price impact across the cross section of stocks if an investor were to purchase an additional share of the index fund.[16] To do so, one can take the comparative static of firm j's share

[15] Another example is one in which firm j and firm z are perfectly negatively correlated and are uncorrelated with the rest of the $N-2$ firms. It is straightforward that a fraction of both $(1-h_j)$ and $(1-h_z)$ are effectively canceled out as the two stocks' dividends are perfect hedges.

[16] I note that with a market-capitalization-weighted index fund, it is equivalent to consider how investing one more

price with respect to m.

Proposition 1. *The price impact on firm j from index-linked trading is,*

$$\frac{\mathrm{d}P_j}{\mathrm{d}m} = \frac{\gamma \sigma_{j,M}}{(m+b)^2} \sum_{z=1}^{N} \phi_z (1-h_z), \tag{10}$$

which is the product of a market-wide component, $\frac{\gamma}{(m+b)^2} \sum_{z=1}^{N} \phi_z(1-h_z)$, and a firm-specific component, $\sigma_{j,M}$. The sign on $\frac{\mathrm{d}P_j}{\mathrm{d}m}$ is determined by the firm-specific component $\sigma_{j,M}$.

Corollary 1.1. *In the limit, as $h_j \to 0$ for all j, the price impact on firm j from index-linked trading is,*

$$\frac{\mathrm{d}P_j}{\mathrm{d}m} = \frac{\gamma \sigma_{j,M}}{(m+b)^2}, \tag{11}$$

which contains a market-wide component and a firm-specific component. The sign on $\frac{\mathrm{d}P_j}{\mathrm{d}m}$ is determined by the firm-specific component $\sigma_{j,M}$.

Proposition 1 and Corollary 1.1 contain the key insight of the base model; the price impact of index-linked trading is, in general, not equal across firms.[17,18] While Proposition 1 and Corollary 1.1 are from the perspective of price change, it is straightforward that the insight also holds for expected returns. Namely, define $E[r_j]$ as the expected net return of firm j, which is given explicitly by,

$$E[r_j] = \frac{\mu_j}{P_j(m)} - 1, \tag{12}$$

in which $P_j(m)$ highlights that the equilibrium price is a function of m. Using (12), the change in

dollar into the fund would affect the cross section of stock shares. However, considering price impact from dollar trade, rather than share trade, requires additional attention to how a dollar would be divided among firms based on their respective weights.

[17] Technically speaking, the price impact is equal across firms if all firms have the same dividend covariance with the market dividend.

[18] The result of Proposition 1 echoes an insight from Merton (1987). Specifically, Merton (1987) examines the equilibrium effect of expanding a firm's investor base (i.e., the number of investors that may include the firm's stock in their portfolios), which is analogous to my measure of price impact from index-linked trade. The study shows that increasing the firm's investor base increases firm value (i.e., the comparative static is positive valued) and the magnitude will be greatest for firms with large firm-specific variances.

expected returns is given by,

$$\frac{\mathrm{d}E[r_j]}{\mathrm{d}m} = -\frac{\mu_j}{P_j(m)^2}\left(\frac{\gamma\sigma_{j,M}}{(m+b)^2}\sum_{z=1}^{N}\phi_z(1-h_z)\right), \tag{13}$$

and, as $h_j \to 0$ for all j, is,

$$\frac{\mathrm{d}E[r_j]}{\mathrm{d}m} = -\frac{\mu_j}{P_j(m)^2}\frac{\gamma\sigma_{j,M}}{(m+b)^2}. \tag{14}$$

It is clear from both (13) and (14) that the impact of index-linked trading on expected returns is not, in general, equal across the cross section. In particular, stocks in which $\sigma_{j,M}$ is large (i.e., the firm specific component of the expressions) are associated with greater return volatility from index-linked trade, all else equal. Furthermore, appealing to the intuition from Berk (1995), cross sectional return volatility effects should be larger in small capitalization stocks as compared to large capitalization stocks (i.e., riskier firms are smaller, all else equal). These results imply that existing empirical findings showing a positive relation between ETF stock ownership and return volatility, for example those in Ben-David et al. (2018), are consistent with theory. These results also imply a new empirical prediction: within market-valued index funds, index-linked trade leads to greater volatility in the stocks with smaller index weights as compared to those with larger weights. In Section 3, I provide empirical support for this novel prediction.

Corollary 1.2. *The return impact from increased index-linked trade is negative for all firms that have a positive covariance with the market dividend. The return impact from increased index-linked trade is positive for all firms that have a negative covariance with the market dividend.*

The result in Corollary 1.2 states that the direction of the effect from index-linked trade depends on a firm's dividend covariance with the market dividend. As such, to the extent that there is a significant macro component to firm dividends, most firms should have a positive covariance. Thus, increases in index-linked trading should increase return comovements as measured by return correlations, which is consisted with the findings documented in Da and Shive (2018).

The results in this section relate to the observation that equilibrium price impact is not equal across stocks and price impact is driven by a stock's dividend covariance with the market dividend. This implies that, on average, price impact negatively relates to market capitalization following the logic of Berk (1995). It is important to acknowledge that notions of price impact differ in other models. For example, in Kyle (1985) price impact is a market maker's response to order flow in the presence of an informed trader (see also Easley and O'Hara (1987) and Glosten and Milgrom (1985)). To the extent that firms differ in regards to their fundamentals (i.e., the distribution of their cash flows) and to stock specific frictions (e.g., information asymmetry), more general "price impact" will also differ across stocks and index-linked trading will generate cross sectional differences in returns and return volatilities. Whether or not other measures of price impact should, on average, be negatively related to market capitalization depends on whether or not the measures' drivers are also reflected in discount rates.

2.1 Index Investor Rotations

Price impact in Proposition 1 relates to an investor moving between the index fund (i.e., the market portfolio) and cash. This is intentional to focus on index-linked trade exclusively. However, it is also important to consider the impact as it relates to an investor moving between the MVE portfolio held by portfolio-builders and the index fund. I refer to this type of trading as an index investor rotation since it is requires selling one stock portfolio and purchasing another. Thus, the pricing implications of a rotation reflect the net impact from both a purchase and sale.

To study the pricing consequences of rotations, consider a slight variant to the model in which $m \in (0,1)$ and $b = 1 - m$. In this variant, the combined measure of portfolio-builders and indexers is fixed and the price of firm j's stock is given by,

$$P_j^* = \mu_j - \frac{\gamma \sigma_{j,j}}{1-m} \left(\sum_{z=1}^{N} \chi_{j,z}(1-h_z) - m \sum_{z=1}^{N} \phi_z (1-h_z) \right). \tag{15}$$

An increase in m in (15) measures the impact from a rotation of selling the portfolio-builder portfolio and purchasing the index fund. To provide the most natural and intuitive explanation of index investor rotations, I focus on the special case in which all firms have positive covariance with the market dividend, that is, $\sigma_{j,M} \geq 0$ for all j.[19]

[19] Allowing for negative values of $\sigma_{j,M}$ introduces unnecessary nuance in the statement of Corollary 1.3.

Corollary 1.3. *The impact of index investor rotations is,*

$$-\frac{\gamma\sigma_{j,M}}{(1-m)^2}\left(\sum_{z=1}^{N}\chi_{j,z}(1-h_z)-\sum_{z=1}^{N}\phi_z(1-h_z)\right), \tag{16}$$

which is negative valued for firms in which $\sum_{z=1}^{N}\chi_{j,z}(1-h_z)$ *is larger than the risk-adjusted market free float and is positive valued for firms in which* $\sum_{z=1}^{N}\chi_{j,z}(1-h_z)$ *is smaller than the risk-adjusted market free float. In the limit, as* $h_j \to q$ *for some arbitrary constant* $q \in [0,1]$ *for all* j*, the impact of index investor rotations is zero for all firms.*

According to Corollary 1.3, if $h_j = h_z$ for all j and z, then there are no pricing implications from index investor rotations. The reason is that the market portfolio is the MVE portfolio and indexers and portfolio-builders have identical holdings. Therefore, when a portfolio-builder sells her portfolio to buy the index fund, the net price impact for all stocks is nil.

If, however, $h_j \neq h_z$ for at least one pair of j and z, there are pricing effects from index investor rotations. Firms in which $\sum_{z=1}^{N}\chi_{j,z}(1-h_z)$ exceeds the risk-adjusted market free float experience a price drop and feature higher expected returns when portfolio-builders move to being indexers. The intuition is that these firms are characterized by an excess, risk-adjusted, supply of shares and this excess supply must be held collectively by portfolio-builders. As such, when m increases and b decreases, there are less portfolio-builders to hold the excess supply and prices must fall to clear the market. Conversely, firms in which $\sum_{z=1}^{N}\chi_{j,z}(1-h_z)$ is smaller than the risk-adjusted market free float are relatively scarce and experience a price increase (lower expected return) when m increases and b decreases.

If firms are heterogeneous in terms of their float, increases in m distort aggregate risk sharing as indexers are compelled to hold the market portfolio which is not mean-variance efficient. In other words, float matters. Notably, several major stock indices have moved away from market-capitalization-based weights to float-adjusted-market-capitalization-weights. For example, Morgan Stanley Capital International (MSCI) transitioned their indices in 2002 and S&P followed with a similar transition in 2005.[20,21] Corollary 1.3 suggests that the moves by MSCI and S&P may have improved aggregate risk sharing by making the wedge between the MVE portfolio and the index fund (which is no longer truly the "market portfolio" due to the float adjustment) smaller.

In the next subsection, I consider a multi-period extension in which the measure of future indexers is random. I show that uncertainty regarding index investor demand is priced in shares and leads to an additional risk premium in expected returns. This "index-linked trade risk premium" differs across stocks based on their individual fundamentals.

2.2 Uncertainty Regarding Index-Linked Trading

To examine the price implications for uncertainty with respect to future index investor demand, I extend the model to include two periods of stock trade. Specifically, consider an overlapping generation (OLG) economy in which each generation lives two-periods and the game takes place over three periods ($t \in \{1, 2, 3\}$). The first generation (denoted by I) lives in periods $t = 1$ and $t = 2$ and the second generation (denoted by II) lives in periods $t = 2$ and $t = 3$. Each generation is constituted by a continuum of indexers and a continuum of portfolio-builders. While the measures of generation I indexers ($m_I > 0$), generation I portfolio-builders ($b_I \geq 0$), and generation II portfolio-builders ($b_{II} \geq 0$) are deterministic, the length of the generation II indexer continuum is random. That is, future indexer demand is uncertain and generation II's continuum-length is a random variable m_{II} in which $\frac{1}{m_{II}} = |\mathcal{I}|$ in which $\mathcal{I} \sim \Phi(0, \sigma_{\mathcal{I}}^2)$.[22] The setup implies larger realizations of $\mathcal{I}$ correspond to smaller values of m_{II}. I assume that $\mathcal{I}$ is common knowledge when generation II investors are born. To allow for tractable closed-form solutions, I initially focus on

[20] The MSCI announcement and discussion of its potential impact on index constituents may be found here: https://www.wsj.com/articles/SB976476678643622201. Likewise, the S&P announcement and discussion of its impact may be found here: https://www.marketwatch.com/story/sp-move-to-float-adjusted-indexes-will-create-turnover.

[21] S&P explains its float adjustment methodology and motivation as "The majority of S&P Dow Jones Indices' market capitalization-weighted indices are float-adjusted. Under float adjustment, the share counts used in calculating the indices reflect only those shares available to investors rather than all of a company's outstanding shares. Float adjustment excludes shares that are closely held by control groups, other publicly traded companies, government agencies, or other long-term strategic shareholders. With a float-adjusted index, the value of the index reflects the value available to investors in the public markets. Reducing the relative investment index investors have in stocks with limited float should enhance the investability of the index." See S&P Global (2020) for additional details.

[22] Note, $|\mathcal{I}|$ is a half-normal distributed variable (a special case of the folded normal distributions) and $E[|\mathcal{I}|] = \sigma_{\mathcal{I}}\sqrt{\frac{2}{\pi}}$.

the special case in which $b_{II} = 0$ and I assume that $h_j = 0$ for all j. Later in this analysis, I relax these requirements and solve a more general framework numerically. It is worth emphasizing that the purpose of utilizing an OLG framework is to ensure trade in future periods and to study resale price uncertainty. As such, the OLG framework should not be taken literally; the purpose of this exercise is not to study the impact of buy-and-hold investing over an investor's entire life. Instead, the purpose is to understand expected returns when future resale prices are determined in part by uncertain indexer demand.

As in the base model, all investors are risk-averse with exponential utility and each investor is born endowed with wealth $W > 0$. Furthermore, each investor may purchase assets (i.e., shares of individual firms or shares of the index fund) in the first period of her life. In the second period of an investor's life, her assets pay random dividends, she may sell her assets to the next generation (if she's not in the final generation), and she consumes everything she has (i.e., the residual of her initial wealth, the realized dividends, and any proceeds from selling her assets).

The assets traded by investors are again the shares of the N firms and the index fund. However, in this extension, trade occurs at two periods and at endogenously determined prices; at $t = 1$, generation I investors purchase the stock of firm j at endogenous price $P_{j,1}$ and, at $t = 2$, generation I investors sell the stock of firm j to generation II investors at endogenous price $P_{j,2}$. Shares generate random dividends at $t = 2$ and $t = 3$ and I denote the dividend of firm j at time t as $D_{j,t}$. In period t, the N firms' stock dividends are distributed according to a multivariate normal distribution in which the $N \times 1$ vector of expected dividends is denoted μ and the $N \times N$ covariance matrix of dividends is denoted Ω. While dividends in $t = 2$ and $t = 3$ share the same multivariate normal distribution, $t = 2$ dividends are uncorrelated with $t = 3$ dividends. Moreover, dividends are uncorrelated with $\mathcal{I}$.

To summarize, this extension's timing is as follows. At $t = 1$, generation I is born, and generation I investors purchase the shares of the N firms (and the index fund) to maximize their expected utility. At $t = 2$, shares realize their first dividends, m_{II} is realized, generation II is born, generation I investors sell their shares (and the index fund) to generation II investors, and generation I investors consume their wealth and die. At $t = 3$, shares realize their second and final dividends and generation II investors consume their wealth and die.

I solve the model by backwards induction. At $t = 3$, there are no portfolio-builders (since $b_{II} = 0$ by assumption) and indexers have terminal wealth quantities similar to those studied in Section 2. Consequently, generation II indexers' expected terminal utilities at $t = 2$ are also redolent to their expected utilities in Section 2 and equilibrium prices mirror those in Lemma 2; the equilibrium $t = 2$ price for firm j's stock is,

$$P^*_{j,2} = \mu_j - \frac{\gamma\sigma_{j,M}}{m_{II}}, \tag{17}$$

and the vector of $t = 2$ prices is denoted $\mathbf{P}^*_2$.[23]

Given the equilibrium stock prices at $t = 2$, I now consider generation I investors' terminal wealth at $t = 2$. Portfolio-builder k's terminal wealth at $t = 2$ is

$$W^T_{k,2} = W + \sum_{j=1}^{N} s_{k,j,1}\left(D_{j,2} + P_{j,2} - P_{j,1}\right), \tag{18}$$

in which $s_{k,j,1}$ is the number of shares investor k purchased in firm j at $t = 1$. Similarly, indexer i's terminal wealth at $t = 2$ is

$$W^T_{i,2} = W + f_{i,1}\left(\sum_{j=1}^{N} D_{j,2} + P_{j,2} - \sum_{j=1}^{N} P_{j,1}\right), \tag{19}$$

in which $f_{i,1}$ is the number of index fund shares investor i purchased at $t = 1$. Using generation I investors' terminal wealth quantities, at $t = 1$, generation I investors maximize their expected utilities. That is, generation I portfolio-builders solve,

$$\max_{\mathbf{s}_{k,1}\in\mathbb{R}^N} E\left[1 - e^{-\gamma\left(W + \sum_{j=1}^{N} s_{k,j,1}(D_{j,2}+P_{j,2}-P_{j,1})\right)}\right], \tag{20}$$

[23] Technically, there are an infinite number of equilibrium price vectors when $b_{II} = 0$. Specifically, since indexers only trade the index fund, any price vector $\hat{\mathbf{P}} \in \mathbb{R}^N$ that satisfies $\mathbf{e}'\hat{\mathbf{P}} = \sum_{j=1}^{N} \mu_j - \frac{\gamma\sigma_{j,M}}{m_{II}}$ is an equilibrium price vector since indexers care only about the price of the total index and not the prices of individual stocks. I eliminate the arbitrariness of the price vector by choosing the one that would also satisfy portfolio-builder's first-order conditions. Later in the analysis, I numerically solve a multi-period setting in which $h_j \neq h_z$ for at least one pair j and z and $b_{II} > 0$. In that setting, there exist a unique equilibrium price vector.

in which $\mathbf{s}_{k,1}$ is an $N \times 1$ vector representing investor k's share purchases in the N firms and element j is $s_{i,j,1}$. Similarly, generation I indexers solve,

$$\max_{f_{i,1} \in \mathbb{R}} E\left[1 - e^{-\gamma\left(W + f_{i,1}\left(\sum_{j=1}^{N} D_{j,2} + \sum_{j=1}^{N} P_{j,2} - \sum_{j=1}^{N} P_{j,1}\right)\right)}\right], \tag{21}$$

in which $f_{i,1}$ is the number of index fund shares investor i purchased at $t = 1$. Furthermore, note that there is uncertainty regarding both the $t = 2$ dividends *and* the $t = 2$ share prices since m_{II} is an unknown random variable. The equilibrium price for each stock is determined by market clearing; the price is set such that aggregation of investors' demands for stock j equals the supply of shares.

Lemma 2. *The equilibrium $t = 1$ price for firm j's stock is,*

$$P_{j,1}^{*} = 2\mu_j - \frac{\gamma\sigma_{j,M}}{m_I + b_I} - \gamma\sigma_{j,M}\sigma_{\mathcal{I}}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I + b_I} + \frac{\phi\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)}{\mathcal{N}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)}\right), \tag{22}$$

in which $\mathcal{N}(\cdot)$ denotes the standard normal distribution cumulative density function (CDF) and $\phi(\cdot)$ denotes the standard normal distribution probability density function (PDF).

To understand the different risk premiums in the $t = 1$ price, it is helpful to establish the expected resale price (i.e., $E[P_{j,2}]$) at $t = 1$.

Lemma 3. *At $t = 1$, the expected $t = 2$ price is,*

$$E[P_{j,2}] = \mu_j - \frac{\gamma\sigma_{j,M}\sigma_{\mathcal{I}}\sqrt{2}}{\sqrt{\pi}}. \tag{23}$$

Using Lemma 2 and Lemma 3, the $t = 1$ price for firm j's stock may be rewritten as,

$$P_{j,1}^{*} = \mu_j + E[P_{j,2}] - \frac{\gamma\sigma_{j,M}}{m_I + b_I} - \gamma\sigma_{j,M}\sigma_{\mathcal{I}}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I + b_I} + \frac{\phi\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)}{\mathcal{N}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)} - \sqrt{\frac{2}{\pi}}\right). \tag{24}$$

To understand the price in (24), first note that the expected payoff from holding a share in firm j is the sum of the expected $t = 2$ dividend μ_j and the expected resale price $E[P_{j,2}]$. The $t = 2$ dividend is risky and requires a risk premium of $\frac{\gamma\sigma_{j,M}}{m_I+b_I}$ which is a function of investors' CARA parameter γ, the dividend's covariance with the market dividend $\sigma_{j,M}$, and the measure of investors over which risk is shared $m_I + b_I$. The $t = 2$ resale price is also risky because there is uncertainty regarding indexer demand. As such, a risk premium relating to that uncertainty is also required in the amount of $\gamma\sigma_{j,M}\sigma_{\mathcal{I}}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I} + \frac{\phi\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)}{\mathcal{N}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)} - \sqrt{\frac{2}{\pi}}\right)$. Before proceeding, it is helpful to provide a supporting result.

Lemma 4. *The function,*

$$V(x) \equiv x + \frac{\phi(x)}{\mathcal{N}(x)}, \tag{25}$$

is increasing in x for all $x \in \mathbb{R}^+$. In the limit, as $x \to 0$, $V(x)$ is equal to $\sqrt{\frac{2}{\pi}}$.

Lemma 4 implies that the quantity $\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I} + \frac{\phi\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)}{\mathcal{N}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)} - \sqrt{\frac{2}{\pi}}\right)$ is weakly positive valued and increasing in $\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}$. Using Lemma 4, the proceeding proposition characterizes the risk premium relating to indexer demand uncertainty.

Proposition 2. *Uncertainty regarding the realization m_{II} introduces a $t = 1$ risk premium of,*

$$\gamma\sigma_{j,M}\sigma_{\mathcal{I}}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I} + \frac{\phi\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)}{\mathcal{N}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)} - \sqrt{\frac{2}{\pi}}\right). \tag{26}$$

The risk premium is positive valued, is increasing in $\sigma_{\mathcal{I}}$, γ, $\sigma_{j,M}$, and σ_M^2, and is decreasing in m_I and b_I.

Proposition 2 provides a novel prediction regarding the cross section of returns: index investor demand uncertainty leads to a risk premium in stocks and the stocks with the largest loadings on the risk are those with the most price impact from index-linked trading. Moreover, to the extent that price impact negatively correlates with market capitalization, the result in Proposition 2 provides a new lens to view existing empirical work through. First, it is well-established that firm size, measured by market capitalization, performs strongly in explaining the cross-section of average returns (e.g., see Banz (1981)). In fact, size is commonly included as a risk factor in asset pricing models (e.g., the three factor model in Fama and French (1993)). Proposition 2 suggests that estimated loadings on the size factor may be exacerbated as the factor may also proxy for exposure to the index-linked trade risk premium (which also implies that estimated models' factor loadings may be biased). The result also helps understand a portfolio excess return, which is related to the size effect, recently documented in Jiang et al. (2020). Specifically, Jiang et al. (2020) documents that a small-minus-large portfolio of S&P 500 stocks earns ten percent per year during the sample period of 2000-2019. While that study attributes the excess return to index flows disproportionately increasing the prices of large capitalization stocks and lowering their future returns, my explanation is different. Namely, my model suggests that the outperformance of small stocks in the S&P 500 is due to their greater exposure to index-linked trade risk.

Corollary 2.1. *Index-linked trade risk premiums are exacerbated by firm-level and market-wide dividend uncertainty.*

Corollary 2.1 says in words that the cross derivatives of the expression in (26) with respect to $\sigma_{\mathcal{I}}$ and $\sigma_{j,M}$ and with respect to $\sigma_{\mathcal{I}}$ and σ_M^2 are positive valued. In other words, an increase in future indexer demand uncertainty (i.e., an increase in $\sigma_{\mathcal{I}}$) increases index-linked trade risk premiums and the increases are even more pronounced for stocks with large dividend covariances (large $\sigma_{j,M}$) and for all stocks when market uncertainty is high (large σ_M^2). The result of Corollary 2.1 provides the novel empirical prediction that index-linked trade risk premiums are higher in times of individual stock volatility and market volatility.

While it is not analytically tractable to solve for closed-form $t = 1$ prices when $b_{II} \neq 0$ or $h_j \neq h_z$ for at least one pair of j and z, it is possible to study the effects via numerical estimation.[24] I consider a setting in which there are $N = 3$ firms and all firms have the same expected dividend, that is, $\mu_1 = \mu_2 = \mu_3 = 1$. The three firms differ in regards to their dividend risk and their available

[24] In the Online Appendix, I detail the methods used for numerical estimation of the model. See Section OA.2. Additionally, if $h_j \neq h_z$ for at least one pair of j and z, then b_{II} must be strictly greater than zero for an equilibrium to exist.

float. The covariance matrix Ω is given by,

$$\Omega = \begin{pmatrix} 0.3 & 0.1 & 0.1 \\ 0.1 & 0.5 & 0.1 \\ 0.1 & 0.1 & 0.7 \end{pmatrix}, \tag{27}$$

such that firm 1's dividend is the safest, firm 3's dividend is the riskiest and the covariance between all firm dividends is the same. The block-holders vector is given by,

$$\mathbf{h} = \begin{pmatrix} 0.0 \\ 0.2 \\ 0.0 \end{pmatrix}, \tag{28}$$

such that firm 2 has relatively less float and its shares are relatively more scarce. I also assume that $m_I = \sqrt{\frac{2}{\pi}}$, $b_I = 1 - \sqrt{\frac{2}{\pi}}$, $\gamma = 0.1$, and $W = 1$. To study the comparative statics related to $\sigma_{\mathcal{I}}^2$, I solve for equilibrium $t = 1$ prices, dividend risk premiums, and index-linked trade risk premiums allowing $\sigma_{\mathcal{I}}^2$ to take values in the interval $(0, 1.7)$. The two panels in Figure 3 depict the dividend risk premiums and index-linked trade risk premiums (measured in dollars) as functions of $\sigma_{\mathcal{I}}^2$; the solid line represents firm 1 which has the smallest dividend variance, the dashed line represents firm 2 which has the intermediate level of dividend variance and the least amount of float, and the dashed-dotted line represents firm 3 which has the greatest dividend variance. In both panels, the horizontal axis represents $\sigma_{\mathcal{I}}^2$ and the vertical axis represents the level of the relevant risk premium. The dotted horizontal line denotes a risk premium of zero.

Panel (a) in Figure 3 shows that changes in $\sigma_{\mathcal{I}}^2$ do not affect the risk premium related to dividend risk as all three lines for the firms are flat. Consistent with intuition, firm 3, which has the riskiest dividend, commands the largest dividend risk premium. Firm 2, which has the intermediate level of dividend risk actually commands the *smallest* dividend risk premium. This is because $h_2 = 0.2$ which implies that 20% of shares are held by block-holders and only 80% of shares remain for indexers and portfolio-builders, making the stock relatively scarce as compared to other firms.

The second panel, Panel (b) in Figure 3, shows that changes in $\sigma_{\mathcal{I}}^2$ affect stocks differentially

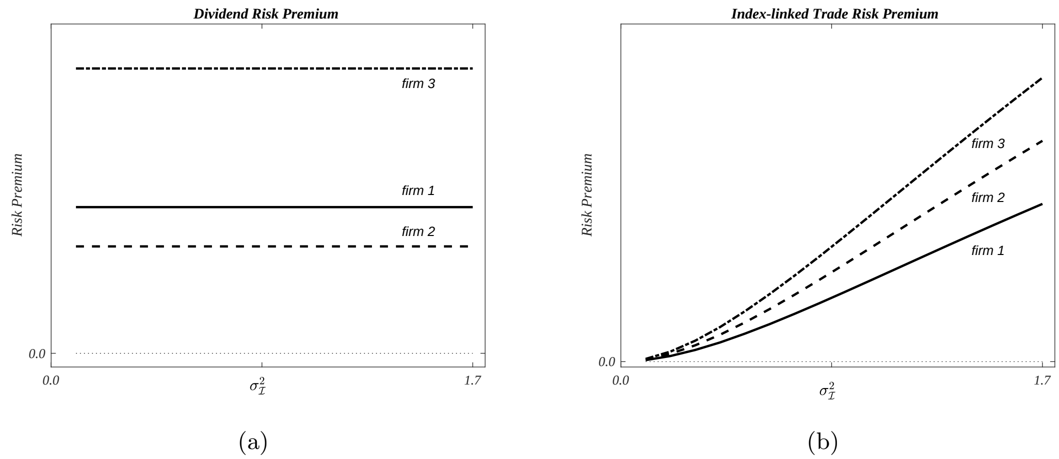


Figure 3: Comparative statics of $\sigma_{\mathcal{I}}^2$. The first panel depicts the relation between $\sigma_{\mathcal{I}}^2$ and the equilibrium dividend risk premium for each firm. The second panel depicts the relation between $\sigma_{\mathcal{I}}^2$ and the equilibrium index-linked trade risk premium for each firm. Risk premiums are measured in dollars and are located on the vertical axis. In both panels, the solid line represents firm 1, the dashed line represents firm 2, the dashed-dotted line represents firm 3, and the dotted horizontal line denotes risk premium of zero.

based on their individual fundamentals. Across all levels of $\sigma_{\mathcal{I}}^2$, index-linked trade risk premiums are increasing and firm 3 commands the highest premium and firm 1 commands the lowest premium (which is consistent with the ordering of their respective dividend risks). As such, while dividend risk itself is distinct from index-linked trade risk, stocks' price impacts reflect dividend risk and, accordingly, determine the level of their index-linked trade risk premiums.

2.2.1 Index Investor Rotation Uncertainty

The risk premium related to index-linked trade, characterized in Proposition 2, occurs because investors do not know the measure of index investors in the subsequent period; future uncertainty about indexers' movements between the index product and cash induces risk in a stock's resale price and affects that stock's value today. It is also important to consider risk as it relates to investors moving between the index fund and the MVE portfolio held by portfolio-builders. In a similar vein to the analysis in Section 2.1, I examine the implications stemming from index investor rotation risk, that is, future uncertainty regarding investors movements into the index product and out of the MVE portfolio and vice versa.

To study rotation uncertainty, I adopt the multiple period framework from Section 2.2 with a few adjustments. First, I assume that,

$$b_I + m_I = b_{II} + m_{II} = 1, \tag{29}$$

such that the total measure of investors is constant across periods and the measure equals one. Second, I introduce uncertainty regarding rotations by having $m_{II} = \mathcal{N}(\mathcal{R})$ and $b_{II} = 1 - \mathcal{N}(\mathcal{R})$ in which $\mathcal{R} \sim \Phi(0, \sigma^2_{\mathcal{R}})$ and $\mathcal{N}(\cdot)$ denotes the standard normal distribution CDF. Third, I relax the requirement that $h_j = h_z$ for all j and z. Instead, I allow for firms to differ in regards to their available float. Intuitively, if $h_j = h_z$ for all j and z, there is no risk premium related to index investor rotation uncertainty since both portfolio-builders and indexers hold the same portfolios in equilibrium.

The equilibrium $t = 2$ price of stock j, after the realization of $\mathcal{R}$, is given by,

$$P^*_{j,2} = \mu_j - \frac{\gamma \sigma_{j,j}}{1 - \mathcal{N}(\mathcal{R})} \left(\sum_{z=1}^{N} \chi_{j,z}(1 - h_z) - \mathcal{N}(\mathcal{R}) \sum_{z=1}^{N} \phi_z (1 - h_z) \right). \tag{30}$$

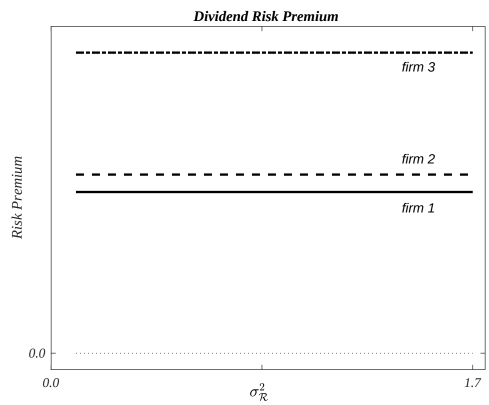
As such, both generation I portfolio-builders and indexers maximize their expected terminal utility at $t = 1$ by taking into consideration both the dividend risk and resale price risk associated with each stock. Solving for $t = 1$ prices in this extension of the model is analytically intractable and numerical estimation is necessary. I consider a setting in which there are $N = 3$ firms and all firms have the same expected dividend, that is, $\mu_1 = \mu_2 = \mu_3 = 1$. The three firms differ in regards to their dividend risk and their available float. The covariance matrix Ω is the same as the one given in (27), such that firm 1's dividend is the safest, firm 3's dividend is the riskiest and the covariance between all firm dividends is the same. The block-holders vector is the same as the one given in (28), such that firm 2 has relatively less float and its shares are relatively more scarce. I also assume that $m_I = b_I = 0.5$, $\gamma = 0.1$, and $W = 1$. To study the comparative statics related to $\sigma^2_{\mathcal{R}}$, I solve for equilibrium $t = 1$ prices, dividend risk premiums, and investor rotation risk premiums allowing $\sigma^2_{\mathcal{R}}$ to take values in the interval $(0, 1.7)$. The two panels in Figure 4 depict the dividend risk premiums and investor rotation risk premiums (measured in dollars) as functions of $\sigma^2_{\mathcal{R}}$; the solid line represents firm 1 which has the smallest dividend variance, the dashed line represents firm 2 which has the intermediate level of dividend variance and the least amount of float, and the dashed-dotted line represents firm 3 which has the greatest dividend variance. In both panels, the horizontal axis represents $\sigma^2_{\mathcal{R}}$ and the vertical axis represents the level of the relevant risk premium. The dotted horizontal line denotes a risk premium of zero.

Panel (a) in Figure 4 shows that changes in $\sigma^2_{\mathcal{R}}$ do not affect the risk premium related to dividend risk as all three lines for the firms are flat. Consistent with intuition, firm 3, which has the riskiest dividend, commands the largest dividend risk premium and firm 1, which has the least risky dividend, commands the smallest dividend risk premium. Notably, while not explicitly shown, firm 2's risk premium is lower than it would be in the case in which $h_2 = 0$; as 20% of shares are held by block-holders and only 80% of shares must be held by indexers and portfolio-builders, the equilibrium dividend risk premium is lower.[25]

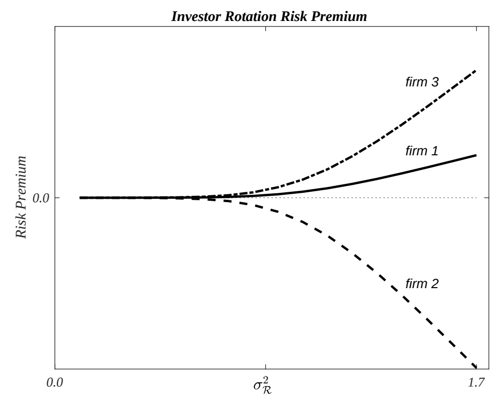
The second panel, Panel (b) in Figure 4, shows that changes in $\sigma^2_{\mathcal{R}}$ affect stocks differentially based on their individual fundamentals and their available free floats. First, note that firm 2 commands a *negative* investor rotation premium and its magnitude is increasing with $\sigma^2_{\mathcal{R}}$. To understand why the sign is negative, recall that indexers' portfolios are fixed. Thus, on a relative basis, firm 2's shares are scarce and moving from the MVE portfolio to the market portfolio increases firm 2's share price. Importantly, if the realization of $\mathcal{R}$ is large and $\mathcal{N}(\mathcal{R})$ is near one, then almost all investors are indexers. In this situation portfolio-builders hold extremely concentrated positions in firm 1 and firm 3 shares, and perhaps short positions in firm 2 shares to satisfy indexers' demand. This generates a convex-like payoff for firm 2's shares and explains why the investor rotation risk premium for firm 2 is not only negative valued but also increasing in magnitude with $\sigma^2_{\mathcal{R}}$.

The second panel also shows that firm fundamentals interact with the investor rotation risk premium. Firm 1 and firm 3 have the same available float (i.e., $h_1 = h_3 = 0$), but differ in regards to their dividend variance with firm 3 having a riskier dividend. As can be seen in the figure, firm 3 commands a larger investor rotation risk premium and its magnitude is increasing with $\sigma^2_{\mathcal{R}}$. This

[25] I note that the rank order of dividend risk premiums differs from those depicted in Figure 3. The difference in rank orderings is due to different measures of initial investors; in Figure 3 the total measure is $m_I + b_I = \sqrt{\frac{2}{\pi}}$ and in Figure 4 the total measure is $m_I + b_I = 1$.



(a)



(b)

Figure 4: Comparative statics of $\sigma^2_{\mathcal{R}}$. The first panel depicts the relation between $\sigma^2_{\mathcal{R}}$ and the equilibrium dividend risk premium for each firm. The second panel depicts the relation between $\sigma^2_{\mathcal{R}}$ and the equilibrium investor rotation risk premium for each firm. Risk premiums are measured in dollars and are located on the vertical axis. In both panels, the solid line represents firm 1, the dashed line represents firm 2, the dashed-dotted line represents firm 3, and the dotted horizontal line denotes risk premium of zero.

highlights that while dividend risk is distinct from investor rotation risk, dividend risk affects price impact which feeds back into investor rotation risk premiums. In other words, firm fundamentals determine a stock's loading on investor rotation risk.

3 Supporting Empirical Analysis

To support the model and its findings, I perform a series of tests using a universe of stocks commonly held in index funds. The tests are by no means exhaustive and, instead, they motivate future studies. First, I examine the relation between stocks' price impacts (measured by bid-ask spread and Amihud Illiquidity) and their size measured by the natural log of their market capitalization values. I find that stocks with small market capitalizations are characterized by larger price impacts. Second, I examine the relation between stocks' return volatilities, their size and index-linked trading. I find that small capitalization stocks are characterized by greater return volatility and the effect is exacerbated by index flows. The evidence is consistent with the model's insights that price impact differs across the cross section and that index flows lead to cross sectional variation in return volatilities.

3.1 Data

I combine data from Bloomberg and the Center for Research in Security Prices (CRSP). From Bloomberg, I retrieve data on State Street Global Advisors's ETF SPY, BlackRock's ETF IVV, and Vanguard's ETF VOO which all provide exposure to the S&P 500 and are physically replicated (i.e., shares of the ETF are backed by physical shares of the companies in the S&P 500). The data I retrieve for SPY, IVV and VOO are daily shares outstanding and share prices which are used to calculate daily ETF flows.[26] From CRSP, I retrieve daily data on individual stock returns, trade volume, share prices, and shares outstanding. From CRSP, I also retrieve monthly data on S&P 500 index constituents to identify which stocks are included in the S&P 500 index.

As a proxy for index-linked trade I utilize ETF flows for SPY, IVV and VOO. First, to calculate daily ETF flows for each ETF, I perform the calculation,

$$flow_{\tau,j} = \frac{SO_{\tau,j}}{SO_{\tau-1,j}} - 1, \tag{31}$$

in which SO_τ is ETF j's shares outstanding on day τ and $\tau - 1$ denotes the previous trading day. $flow_{\tau,j}$ can be negative valued (ETF shares are redeemed in net) or $flow_{\tau,j}$ can be positive valued (ETF shares are created in net). Both negative and positive values of $flow_{\tau,j}$ imply an ETF flow, with the sign on $flow_{\tau,j}$ providing the direction. Second, I combine ETF flows across ETFs using the three funds' asset weights,

$$flow_\tau = \frac{\sum_{j=1}^{3} flow_{\tau,j} \times AUM_{\tau,j}}{\sum_{j=1}^{3} AUM_{\tau,j}}, \tag{32}$$

[26]See Brown et al. (2021) for an in-depth discussion of ETF flows and the ETF primary market that facilitates the flows (i.e., the ETF arbitrage mechanism).

in which $AUM_{\tau,j}$ is the AUM of ETF j on day τ and $AUM_{\tau,j}$ is calculated as,

$$\frac{SharesOutstanding_{\tau,j} + SharesOutstanding_{\tau-1,j}}{2} \times \frac{price_{\tau,j} + price_{\tau-1,j}}{2}.$$

Positive values of $flow_\tau$ imply net inflows to S&P 500 ETFs and negative values of $flow_\tau$ imply net outflows to S&P 500 ETFs.

As I am interested in the cross sectional impact on return volatilities from index-linked trades, I calculate the standard deviation of daily flows in each month t ($volFlow_t$). While SPY, IVV and VOO are just three ETFs among many, they are consistently some of the most heavily traded ETFs and they serve as a bellwether for index investor demand. As such, I interpret $volFlow_t$ as a proxy for indexer trade volatility in S&P 500 stocks.

In each month t, I take the S&P 500 constituents as my main sample of stocks. I also take the largest 500 NYSE-traded stocks (with a share price greater than \$5) that are not included in the S&P 500 that month as a control group. For each stock j in month t, I calculate the natural log of its market capitalization value LN($mktCap_{j,t}$) in which $mktCap_{j,t}$ is the market capitalization of firm j in month t.[27] To measure price impact, I use both bid-ask spread and the Amihud Illiquidity measure (Amihud (2002)). I calculate each stock's average monthly bid-ask spread (as a percentage) using bid prices, ask prices, and their midpoint,

$$bidAsk_{j,t} = \sum_{\tau=1}^{T} \left(\frac{2(Ask_{j,\tau} - Bid_{j,\tau})}{Ask_{j,\tau} + Bid_{j,\tau}} \right) \Big/ T, \tag{33}$$

in which month t consists of T trading days. I calculate each stock's monthly Amihud Illiquidity measure as,

$$illiq_{j,t} = \sum_{\tau=1}^{T} \left(\frac{|r_{j,\tau}|}{\$VOL_{j,\tau}} \right) \Big/ T, \tag{34}$$

in which month t consists of T trading days, $|r_{j,\tau}|$ is the absolute value of stock j's return on trading day τ and $\$VOL_{j,\tau}$ is the dollar volume of stock j on trading day τ. For purposes of scale, $illiq_{j,t}$

[27] I note that outright market capitalizations are not utilized in the S&P 500 to calculate portfolio weights due to the index provider's methodology to account for free float. Nevertheless, there is good reason to believe that portfolio weights calculated from outright market capitalizations are highly correlated with actual S&P 500 portfolio weights.

Table 1: Summary statistics

Summary statistics are reported for the data in the sample period from January 1, 2005 through May 31, 2020. *S&P 500 Constituents* denotes the sample of stocks held monthly in the S&P 500 throughout the sample period. *Control Group* denotes the sample of the 500 largest NYSE traded stocks each month that are not S&P 500 constituents. *volFlow* is standardized to have a mean of zero and a standard deviation of one.

	S&P 500 Constituents				Control Group			
Variable	Mean	Median	Std. Dev.	# Obs.	Mean	Median	Std. Dev.	# Obs.
bidAsk	0.05%	0.03%	0.08%	93,058	0.08%	0.05%	0.12%	92,500
illiq	0.02	0.01	0.03	93,058	0.10	0.04	0.77	92,500
volRet	1.85%	1.47%	1.44%	93,058	1.99%	1.62%	1.40%	92,500
LN(*mktCap*)	23.48	23.38	1.09	93,058	22.27	22.10	0.77	92,500
volFlow	0.00%	-0.29%	1.00%	185	0.00%	-0.29%	1.00%	185

is multiplied by 100,000,000.

I also calculate the standard deviation of daily stock returns for each stock j in month t using daily returns ($volRet_t$). Table 1 reports the summary statistics for the variables used in the empirical tests for both S&P 500 constituent stocks and my control group. I note that all measures, other than the natural log of market capitalization values and Amihud Illiquidity values, are reported as percentages and volatility measures are reported at the daily frequency (one can multiply by $\sqrt{252}$ to annualize the values). Furthermore, I standardize *volFlow* so that it has a mean of zero and a standard deviation of one. Table 1 shows that S&P 500 constituent stocks generally have slightly smaller bid-ask spreads, are more liquid, have slightly less return volatility, and are slightly larger as compared to stocks in the control group.[28] Nevertheless, the control group stocks provide a useful sample for a placebo test later in the empirical analysis.

3.2 Results

Motivated by Panel (a) of Figure 1, I first explore whether there is a relation between market capitalization and price impact. Using insights from the model, in particular the result in Proposition 1, the null hypothesis is that there is a negative relation between market capitalization and

[28] I note that the number of observations between S&P 500 stocks and the control group differs because there are times at which the S&P 500 holds multiple share classes of the same index constituent.

price impact.[29] I perform the following regression:

$$PriceImpact_{j,t} = \beta \text{LN}(mktCap_{j,t}) + \Psi_t + \epsilon_{j,t}, \quad (35)$$

in which $PriceImpact_{j,t}$ is stock j's measure of price impact in month t, LN($mktCap_{j,t}$) is the natural log of stock j's market capitalization in month t, β is the estimated coefficient on LN($mktCap_{j,t}$), Ψ_t a year-month fixed effect, and $\epsilon_{j,t}$ is the error term. The results for $bidAsk_{j,t}$ (Panel A) and $illiq_{j,t}$ (Panel B) are reported in Column (1) of Table 2. In Column (2), the same regression is performed, but with the addition of stock fixed effects. In both columns, standard errors are clustered by year-month.

In Panel A, the estimated coefficient β in both (1) and (2) is negative valued and statistically significant at a 1% p-value threshold. In terms of economic magnitude, the coefficients are -1 basis points (bps) and -3 bps respectively which implies that a 1% increase in market capitalization relates to a 1 bps decline in spreads across stocks and a 3 bps decline in spreads within stocks (since the independent variable is in natural logs). In Panel B, the results echo the same sentiment; the estimated coefficient on β in (1) is -0.01, the estimated coefficient in (2) is -0.02, and both are statistically significant at a 1% p-value threshold. In other words, smaller market capitalization stocks are less liquid than larger market capitalization stocks.

I next explore whether there is a relation between firm sizes (measured by the natural log of market capitalizations) and return volatilities. Using insights from the model, the null hypothesis is that there is a negative relation between firm size and return volatilities and that the effect is amplified by index-linked trade. First, I perform the same regressions reported in Table 2 but replace the dependent variable $bidAsk_{j,t}$ with the daily volatility of returns $volRet_{j,t}$. The results are reported in Column (1) and Column (3) of Panel A in Table 3. Column (1) and Column (3) of Panel A show a statistically significant negative relation between firm size and return volatility, which is consistent with my model as well as many existing studies including Merton (1987) and Berk (1995). In terms of economic magnitude, across stocks, a 1% increase in firm size relates to a

[29] I note that while the null hypothesis is that there is a negative relation, I assess statistical significance using two-tailed tests as is standard in the literature.

Table 2: Price impact and market capitalization

The table reports the results from OLS regressions for each month's S&P 500 constituent stocks to examine associations between stocks' price impacts and firm size (measured by the natural log of market capitalization). $bidAsk_{j,t}$ (%) is the average monthly bid-ask spread of stock j in month t reported as a percentage. LN($mktCap_{j,t}$) is the natural log of firm j's market capitalization in month t. t-statistics that are clustered by year-month to account for cross-sectional correlation are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Panel A: Bid-Ask Spreads

	$bidAsk_{j,t}$ (%)	
	(1)	(2)
LN($mktCap_{j,t}$)	-0.01***	-0.03***
	(-13.90)	(-10.46)
Observations	93,058	93,058
Adjusted R^2	0.37	0.48
Year-month fixed effects	Yes	Yes
Stock fixed effects	No	Yes

Panel B: Amihud Illiquidity

	$illiq_{j,t}$	
	(1)	(2)
LN($mktCap_{j,t}$)	-0.01***	-0.02***
	(-13.66)	(-7.98)
Observations	93,058	93,058
Adjusted R^2	0.36	0.54
Year-month fixed effects	Yes	Yes
Stock fixed effects	No	Yes

decrease in daily return volatility of 24 bps (or, about 3.8% annually). Within stocks, a 1% increase in firm size relates to a decrease in daily return volatility of 71 bps (or, about 11.3% annually).

To further relate the empirical analysis to the model, I perform the following regression:

$$bidAsk_{j,t} = \beta\text{LN}(mktCap_{j,t}) + \theta\text{LN}(mktCap_{j,t}) \times volFlow_t + \Psi_t + \epsilon_{j,t}, \tag{36}$$

in which $volFlow_t$ is the daily volatility of ETF flows and θ is the estimated coefficient on the interaction of LN($mktCap_{j,t}$) $\times$ $volFlow_t$. I note that the estimated coefficient on $volFlow_t$ is absorbed by the year-month fixed effect. Column (2) reports the results without stock fixed effects and Column (4) reports the results with stock fixed effects. The model shows that ETF flows exacerbate cross sectional differences in return volatilities and, as such, the null hypothesis is that θ is negative valued. In both Column (2) and Column (4) the estimated coefficient β are essentially identical to those reported in Column (1) and Column (3). In both columns, the interaction term θ is estimated as a statistically significant negative coefficient as the theory suggests. In Column (2) of Panel A, the estimate is -3 bps which is economically significant. To see this, consider rewriting

Table 3: Market capitalization, return volatilities, and index-linked trade

The table reports the results from OLS regressions for each month's S&P 500 constituent stocks and a control group of stocks to examine associations between stocks' return volatilities and firm size (measured by the natural log of market capitalization) interacted with index-linked trade volatility. $volRet_{j,t}$ (%) is the daily return volatility from returns for firm j in month t. LN($mktCap_{j,t}$) is the natural log of firm j's market capitalization in month t. $volFlow_t$ is the daily ETF flow volatility of S&P 500 ETFs. In Panel A, results for S&P 500 stocks are reported. In Panel B, results for a control group of each month's 500 largest NYSE-traded stocks not included in the S&P 500. t-statistics that are clustered by year-month to account for cross-sectional correlation are reported in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Panel A: S&P 500 Stocks

	$volRet_{j,t}$ (%)			
	(1)	(2)	(3)	(4)
LN($mktCap_{j,t}$)	-0.24***	-0.24***	-0.71***	-0.71***
	(-16.25)	(-16.63)	(-10.54)	(-10.67)
LN($mktCap_{j,t}$) × $volFlow_t$		-0.03*		-0.03**
		(-1.82)		(-2.36)
Observations	93,058	93,058	93,058	93,058
Adjusted R^2	0.50	0.50	0.63	0.63
Year-month fixed effects	Yes	Yes	Yes	Yes
Stock fixed effects	No	No	Yes	Yes

Panel B: Control Group

	$volRet_{j,t}$ (%)			
	(1)	(2)	(3)	(4)
LN($mktCap_{j,t}$)	-0.09***	-0.09***	-0.41***	-0.41***
	(-8.91)	(-8.99)	(-12.20)	(-12.19)
LN($mktCap_{j,t}$) × $volFlow_t$		0.01		0.00
		(1.27)		(0.00)
Observations	92,500	92,500	92,500	92,500
Adjusted R^2	0.48	0.48	0.65	0.65
Year-month fixed effects	Yes	Yes	Yes	Yes
Stock fixed effects	No	No	Yes	Yes

the regression as,

$$volRet_{j,t} = (\beta + \theta volFlow_t)\,\mathrm{LN}(mktCap_{j,t}) + \Psi_t + \epsilon_{j,t}, \tag{37}$$

and note that a standard deviation of $volFlow_t$ is equal to one (which can be seen in Table 1). With this rewriting, a one standard deviation increase in ETF flow volatility amplifies the impact of log market capitalization by 12.5%.[30] In Column (4), the estimated coefficient on the interaction term is essentially the same as the one in Column (2), but its economic importance, while still significant, is dampened. The dampened economic significance is due to the estimates of β being larger which implies the amplification via ETF flows is smaller.

The results in Panels A provide support for the model by showing that return volatility differences are amplified by index-linked trade, precisely as the model predicts. Moreover, it is important to note that this effect is driven by trade itself increasing return volatility, rather than by market conditions. Specifically, all regressions include year-month fixed effects and Column (3) and Column (4) include stock fixed effects. To provide additional evidence that ETF flows amplify cross sectional differences in return volatilities, Panel B of Table 3 performs the same regressions but with the control group rather than the sample of S&P 500 constituent stocks. As can be seen in Panel B, there is a statistically significant negative relation between firm size and return impact as theory would suggest. Importantly, however, S&P 500 ETF flows have no effect on non-S&P 500 stocks' return volatilities as the estimated coefficients on the interaction term are near zero and are statistically insignificant.

4 Conclusion

This paper considers a simple modeling framework to analyze the impact of index-linked trading on stock returns. The model succinctly shows that price impact is not equal across the cross section of stocks. As such, trades into and out of index products generate cross sectional differences in returns and return volatilities. The model's insights suggest that existing empirical findings are

[30]To arrive at 12.5%, the following calculation is performed, $\frac{(-0.24-0.03\times 1.00)}{-0.24} - 1 = 12.5\%$.

not only consistent with the appropriate null hypotheses, but it would be anomalous to *not* find the documented effects. The model also shows that cross sectional differences in return volatility differences are exacerbated by index-linked trade volatility which is a novel prediction. The empirical analysis shows that S&P 500 ETF flows exacerbate return volatility differences, confirming the model's prediction. Naturally, this implies that, to the extent that index-linked trade is uncertain, different stocks have different exposures to that uncertainty and command different risk premiums.

The results also provide a word of warning for future studies. There may be a temptation by researchers to use stocks' heterogeneous price impact from index-linked trading (e.g., ETF flows) to identify exogenous variation in stock prices (i.e., non-fundamental price variation). Such temptations are likely as several previous studies have similarly used mutual fund outflows in an attempt to also isolate non-fundamental price variation.[31] Moreover, the temptation is likely to be especially strong since index funds and ETFs are passively managed and there are not concerns about mangers choosing which assets to trade or not trade. Despite these strong considerations, researchers must recognize that, even if index flows are random, the price variation may not be. As my model shows, price impact is driven in part by firm fundamentals. As such, future studies measuring non-fundamental price variation from index-linked trade may be inadvertently measuring information about firm fundamentals. Thus, careful attention must be given to any identification strategy that uses index-linked trades.

[31] Notable examples using mutual fund outflows to study stock prices include Coval and Stafford (2007) and Edmans, Goldstein, and Jiang (2012). See also Berger (2019) and Wardlaw (2020) for discussions relating to the strengths and weaknesses of using mutual fund flows to identify exogenous variation.

References

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18.

Baruch, S., & Zhang, X. (2018). Is index trading benign. *Working Paper*.

Ben-David, I., Franzoni, F., & Moussawi, R. (2018). Do ETFs increase volatility? *The Journal of Finance*, *73*(6), 2471–2535.

Berger, E. (2019). Selection bias in mutual fund fire sale. *Working Paper*.

Berk, J. B. (1995). A critique of size-related anomalies. *The Review of Financial Studies*, *8*(2), 275–286.

Bhattacharya, A., & O'Hara, M. (2018). Can ETFs increase market fragility? Effect of information linkages in ETF markets. *Working Paper*.

Bond, P., & Garcia, D. (2020). The equilibrium consequences of indexing. *Working Paper*.

Brogaard, J., Ringgenberg, M. C., & Sovich, D. (2019). The economic impact of index investing. *The Review of Financial Studies*, *32*(9), 3461–3499.

Brown, D. C., & Davies, S. W. (2017). Moral hazard in active asset management. *Journal of Financial Economics*, *125*(2), 311–325.

Brown, D. C., Davies, S. W., & Ringgenberg, M. C. (2021). ETF arbitrage, non-fundamental demand, and return predictability. *Review of Finance*, *25*(4), 937–972.

Buffa, A. M., Vayanos, D., & Woolley, P. (2020). Asset management contracts and equilibrium prices. *Working Paper*.

Cong, L. W., & Xu, D. (2020). Rise of factor investing: Security design and asset pricing implications. *Working Paper*.

Coval, J., & Stafford, E. (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, *86*(2), 479–512.

Da, Z., & Shive, S. (2018). Exchange traded funds and asset return correlations. *European Financial Management*, *24*(1), 136–168.

Easley, D., & O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics*, *19*(1), 69–90.

Edmans, A., Goldstein, I., & Jiang, W. (2012). The real effects of financial markets: The impact of prices on takeovers. *The Journal of Finance*, *67*(3), 933–971.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*, 3-56.

French, K. R. (2008). Presidential address: The cost of active investing. *The Journal of Finance*, *63*(4), 1537–1573.

Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, *14*(1), 71–100.

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

Jiang, H., Vayanos, D., & Zheng, L. (2020). Tracking biased weights: Asset pricing implications of value-weighted indexing. *Working Paper*.

Krause, T., Ehsani, S., & Lien, D. (2013). Exchange traded funds, liquidity, and market volatility. *Working Paper*.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, *53*(6), 1315–1335.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, *42*(3), 483–510.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425–442.

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

S&P Global. (2020). Float adjustment methodology. *S&P Dow Jones Indices, A Division of S&P Global*.

Stambaugh, R. F. (2014). Presidential address: Investment noise and trends. *The Journal of Finance*, *69*(4), 1415–1453.

Stutzer, M. (2018). Style investing and the ICAPM. *Quantitative Finance and Economics*, *2*(3), 702–716.

Subrahmanyam, A. (1991). A theory of trading in stock index futures. *The Review of Financial Studies*, *4*(1), 17–51.

Wardlaw, M. (2020). Measuring mutual fund flow pressure as shock to stock returns. *The Journal of Finance*, *75*(6), 3221–3243.

Wurgler, J. (2010). On the economic consequences of index-linked investing. *Working Paper*.

Appendix

Proof of Lemma 1: With exponential utility and normally distributed dividends, indexer i's maximization in (2) is equivalent to,

$$\max_{f_i \in \mathbb{R}} f_i(\mathbf{e}'\mu - \mathbf{e}'\mathbf{P}) - \frac{\gamma f_i^2}{2}\mathbf{e}'\Omega\mathbf{e}, \tag{A1}$$

in which $\mathbf{e}$ is an $N \times 1$ vector of ones and $\mathbf{P}$ is the $N \times 1$ vector of stock prices. The first-order condition is necessary and sufficient to solve the optimization and the first-order condition is given by,

$$0 = (\mathbf{e}'\mu - \mathbf{e}'\mathbf{P}) - \gamma f_i \mathbf{e}'\Omega\mathbf{e}. \tag{A2}$$

Because all indexers are homogenous, are infinitesimal, are price-takers and face the same concave optimization, their demands are homogenous. As such, I drop the subscript i hereafter. An indexer's optimal demand in the index fund is given by,

$$f^* = \frac{\mathbf{e}'\mu - \mathbf{e}'\mathbf{P}}{\gamma\mathbf{e}'\Omega\mathbf{e}}, \tag{A3}$$

and the aggregate demand of all indexers is given by,

$$\int_0^m \frac{\mathbf{e}'\mu - \mathbf{e}'\mathbf{P}}{\gamma\mathbf{e}'\Omega\mathbf{e}}\,\mathrm{d}x = m\left(\frac{\mathbf{e}'\mu - \mathbf{e}'\mathbf{P}}{\gamma\mathbf{e}'\Omega\mathbf{e}}\right). \tag{A4}$$

Portfolio-builder k's maximization in (4) is equivalent to,

$$\max_{\mathbf{s}_k \in \mathbb{R}^N} \mathbf{s}_k'(\mu - \mathbf{P}) - \frac{\gamma}{2}\mathbf{s}_k'\Omega\mathbf{s}_k. \tag{A5}$$

First-order conditions are necessary and sufficient to solve the optimization and the first-order conditions (in matrix form) are given by,

$$\vec{0} = (\mu - \mathbf{P}) - \gamma\Omega\mathbf{s}_k, \tag{A6}$$

in which $\vec{0}$ is an $N \times 1$ vector of zeros. Because all portfolio-builders are homogenous, are infinitesimal, are price-takers and face the same concave optimization, their demands are homogenous. As such, I drop the subscript k hereafter. In equilibrium, it must be the case that investor k's demand for firm j is $s_j^* = \frac{1-h_j-m\left(\frac{\mathbf{e}'\mu-\mathbf{e}'\mathbf{P}}{\gamma\mathbf{e}'\Omega\mathbf{e}}\right)}{b}$ for markets to clear. That is,

$$1 = \underbrace{h_j + m\left(\frac{\mathbf{e}'\mu - \mathbf{e}'\mathbf{P}}{\gamma\mathbf{e}'\Omega\mathbf{e}}\right)}_{\substack{\text{Demand from block-holders} \\ \text{and indexers}}} + \int_0^b s_j^* \, \mathrm{d}x, \tag{A7}$$

which implies,

$$s_j^* = \frac{1 - h_j - m\left(\frac{\mathbf{e}'\mu-\mathbf{e}'\mathbf{P}}{\gamma\mathbf{e}'\Omega\mathbf{e}}\right)}{b}, \tag{A8}$$

for element j in $\mathbf{s}$. Consequently there are N unknowns (the prices of each firms' shares) and there are N equations from (A6). Solving the system yields,

$$P_j = \mu_j - \frac{\gamma}{b}\left(\sum_{y=1}^{N}\sigma_{j,y}\right)\left(\sum_{z=1}^{N}\chi_{j,z}(1-h_z) - \frac{m}{m+b}\sum_{z=1}^{N}\phi_z(1-h_z)\right), \tag{A9}$$

in which $\sigma_{j,y}$ is the covariance between firm j's dividend and firm y's dividend (implying $\sigma_{j,j}$ is the variance of firm j's dividend) and in which,

$$\phi_z \equiv \frac{\sum_{x=1}^{N}\sigma_{z,x}}{\mathbf{e}'\Omega\mathbf{e}} = \frac{\sigma_{z,M}}{\sigma_M^2}, \tag{A10}$$

with,

$$\sum_{z=1}^{N}\phi_z = 1, \tag{A11}$$

and in which

$$\chi_{j,z} \equiv \frac{\sigma_{j,z}}{\sum_{x=1}^{N}\sigma_{j,x}} = \frac{\sigma_{j,z}}{\sigma_{j,M}}, \tag{A12}$$

with,

$$\sum_{z=1}^{N}\chi_{j,z} = 1. \tag{A13}$$

Using the shorthand notation from the main prose, (A9) simplifies to,

$$P_j = \mu_j - \frac{\gamma \sigma_{j,M}}{b} \left(\sum_{z=1}^{N} \chi_{j,z}(1-h_z) - \frac{m}{m+b} \sum_{z=1}^{N} \phi_z (1-h_z) \right), \tag{A14}$$

In the limit, as $h_j \to 0$ for all j, the price of firm j in (A9) is given by,

$$P_j = \mu_j - \frac{\gamma}{m+b} \left(\sum_{y=1}^{N} \sigma_{j,y} \right), \tag{A15}$$

and, using the shorthand notation from the main prose, may be rewritten as,

$$P_j = \mu_j - \frac{\gamma \sigma_{j,M}}{m+b}. \tag{A16}$$

■

Proof of Proposition 1: The proof is performed by taking the derivative of (A14) with respect to m,

$$\frac{\mathrm{d}P_j}{\mathrm{d}m} = \frac{\gamma \sigma_{j,M}}{(m+b)^2} \sum_{z=1}^{N} \phi_z (1-h_z), \tag{A17}$$

which is the product of a market-wide component,

$$\frac{\gamma}{(m+b)^2} \sum_{z=1}^{N} \phi_z (1-h_z), \tag{A18}$$

and a firm-specific component,

$$\sigma_{j,M}. \tag{A19}$$

■

Proof of Corollary 1.1:

The result follows directly from evaluating (A17) at $h_j = 0$ for all j.

■

Proof of Corollary 1.2:

The result follows directly from Proposition 1, Corollary 1.1, and Assumption 1.

■

Proof of Corollary 1.3: In the case in which there is a constant measure of non-block-holders, which is normalized to one, (A9) may be rewritten as,

$$P_j = \mu_j - \frac{\gamma\sigma_{j,M}}{1-m}\left(\sum_{z=1}^{N}\chi_{j,z}(1-h_z) - m\sum_{z=1}^{N}\phi_z(1-h_z)\right), \tag{A20}$$

and the derivative of P_j with respect to m is,

$$\frac{\mathrm{d}P_j}{\mathrm{d}m} = -\frac{\gamma\sigma_{j,M}}{(1-m)^2}\left(\sum_{z=1}^{N}\chi_{j,z}(1-h_z) - \sum_{z=1}^{N}\phi_z(1-h_z)\right), \tag{A21}$$

which is a combination of market-wide and firm-specific components.

■

Proof of Lemma 2:

The proceeding auxiliary lemmas provide useful results for the proof.

Lemma A1. *In the limit, as $b_{II} \to 0$, the equilibrium $t = 2$ price for firm j's stock is,*

$$P^*_{j,2} = \mu_j - \frac{\gamma\sigma_{j,M}}{m_{II}}. \tag{A22}$$

Proof of Lemma A1:

The proof follows the proof of Lemma 1 and is omitted for the sake of brevity.

■

Lemma A2. *If x is distributed normally with mean μ and variance σ^2 and a is a real constant, then*

$$E[e^{a|x|}] = e^{-a\mu+\frac{a^2\sigma^2}{2}}\mathcal{N}\left(-\frac{(\mu-a\sigma^2)}{\sigma}\right) + e^{a\mu+\frac{a^2\sigma^2}{2}}\mathcal{N}\left(\frac{(\mu+a\sigma^2)}{\sigma}\right). \tag{A23}$$

in which $\mathcal{N}(z)$ denotes the integral of a standard normal distribution probability density function (PDF) from $-\infty$ to z, that is,

$$\mathcal{N}(z) \equiv \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\,\mathrm{d}y. \tag{A24}$$

In the special case in which $\mu = 0$, $E[e^{a|x|}]$ is given by,

$$E[e^{a|x|}] = 2e^{\frac{a^2\sigma^2}{2}}\mathcal{N}(a\sigma). \tag{A25}$$

Proof of Lemma A2: Consider $E[e^{a|x|}]$, in which x is distributed normally with mean μ and variance σ^2 and a is a constant. Then, $E[e^{a|x|}]$ may be written explicitly as,

$$E[e^{a|x|}] = \int_{-\infty}^{\infty} e^{a|x|} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \, \mathrm{d}x, \tag{A26}$$

which is equivalent to,

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2+2a|x|\sigma^2}{2\sigma^2}} \, \mathrm{d}x. \tag{A27}$$

One can separate the preceding expression into two integrals, one for the negative domain of x and one for the positive domain of x, and eliminate the absolute value function from x,

$$\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2-2ax\sigma^2}{2\sigma^2}} \, \mathrm{d}x + \int_{0}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2+2ax\sigma^2}{2\sigma^2}} \, \mathrm{d}x. \tag{A28}$$

Completing the square to isolate x in each of the two integrals yields,

$$\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-(\mu-a\sigma^2))^2-2a\mu\sigma^2+a^2\sigma^4}{2\sigma^2}} \, \mathrm{d}x + \int_{0}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-(\mu+a\sigma^2))^2+2a\mu\sigma^2+a^2\sigma^4}{2\sigma^2}} \, \mathrm{d}x, \tag{A29}$$

and simplifies to,

$$e^{-a\mu+\frac{a^2\sigma^2}{2}} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-(\mu-a\sigma^2))^2}{2\sigma^2}} \, \mathrm{d}x + e^{a\mu+\frac{a^2\sigma^2}{2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-(\mu+a\sigma^2))^2}{2\sigma^2}} \, \mathrm{d}x. \tag{A30}$$

Denote the integral of a standard normal distribution PDF from $-\infty$ to an arbitrary constant z as $\mathcal{N}(z)$, that is,

$$\mathcal{N}(z) \equiv \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \, \mathrm{d}y. \tag{A31}$$

Using the function $\mathcal{N}(z)$, (A30) may be rewritten as,

$$e^{-a\mu+\frac{a^2\sigma^2}{2}} \mathcal{N}\left(-\frac{(\mu-a\sigma^2)}{\sigma}\right) + e^{a\mu+\frac{a^2\sigma^2}{2}} \left(1 - \mathcal{N}\left(-\frac{(\mu+a\sigma^2)}{\sigma}\right)\right), \tag{A32}$$

and using the symmetry of the standard normal distribution around zero, further rewritten as,

$$e^{-a\mu+\frac{a^2\sigma^2}{2}}\mathcal{N}\left(-\frac{(\mu-a\sigma^2)}{\sigma}\right)+e^{a\mu+\frac{a^2\sigma^2}{2}}\mathcal{N}\left(\frac{(\mu+a\sigma^2)}{\sigma}\right). \tag{A33}$$

Finally, if $\mu = 0$, then (A33) simplifies to,

$$2e^{\frac{a^2\sigma^2}{2}}\mathcal{N}\left(a\sigma\right). \tag{A34}$$

■

Lemma A3. *Generation I portfolio-builder k's objective function is,*

$$\mathbf{s}'_{k,1}(2\mu-\mathbf{P}_1)-\frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1}-\frac{\gamma\left(\gamma\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)\left(\gamma\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)\sigma_{\mathcal{I}}^2}{2}-\frac{1}{\gamma}\ln\left(\mathcal{N}\left(\gamma^2\sigma_{\mathcal{I}}\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)\right). \tag{A35}$$

Generation I indexer i's objective function is,

$$f_{i,1}\mathbf{e}'(2\mu-\mathbf{P}_1)-\frac{\gamma f_{i,1}^2}{2}\mathbf{e}'\Omega\mathbf{e}-\frac{\gamma\left(\gamma f_{i,1}\mathbf{e}'\Omega\mathbf{e}\right)\left(\gamma f_{i,1}\mathbf{e}'\Omega\mathbf{e}\right)\sigma_{\mathcal{I}}^2}{2}-\frac{1}{\gamma}\ln\left(\mathcal{N}\left(\gamma^2\sigma_{\mathcal{I}}f_{i,1}\mathbf{e}'\Omega\mathbf{e}\right)\right). \tag{A36}$$

Proof of Lemma A3:

With exponential utility and normally distributed dividends, portfolio-builder k's objective function from the maximization in (20) is equivalent to,

$$1-\int_{-\infty}^{\infty}e^{-\gamma\left(\mathbf{s}'_{k,1}(\mu+(\mu-\gamma|\mathcal{I}|\Omega\mathbf{e})-\mathbf{P}_1)-\frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1}\right)}\phi(\mathcal{I})\,\mathrm{d}\mathcal{I}, \tag{A37}$$

in which $\mathcal{I}$ is a zero-mean normally distributed variable with variance $\sigma_{\mathcal{I}}^2$ and $\phi(\mathcal{I})$ is the PDF for $\mathcal{I}$. The equation in (A37) may be rewritten as,

$$1-e^{-\gamma\left(\mathbf{s}'_{k,1}(2\mu-\mathbf{P}_1)-\frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1}\right)}\int_{-\infty}^{\infty}e^{\gamma^2|\mathcal{I}|\mathbf{s}'_{k,1}\Omega\mathbf{e}}\phi(\mathcal{I})\,\mathrm{d}\mathcal{I}. \tag{A38}$$

Define,

$$\nu=\gamma\mathbf{s}'_{k,1}\Omega\mathbf{e}, \tag{A39}$$

which is a constant scalar value. Using ν, (A38) may be rewritten.

$$1 - e^{-\gamma\left(\mathbf{s}'_{k,1}(2\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1}\right)} \int_{-\infty}^{\infty} e^{\gamma\nu|\mathcal{I}|}\phi(\mathcal{I})\,\mathrm{d}\mathcal{I}, \tag{A40}$$

and using Lemma A2, (A40) becomes,

$$1 - 2e^{-\gamma\left(\mathbf{s}'_{k,1}(2\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1}\right)} e^{\frac{\gamma^2\nu^2\sigma_{\mathcal{I}}^2}{2}} \mathcal{N}\left(\gamma\nu\sigma_{\mathcal{I}}\right), \tag{A41}$$

and further simplifies to,

$$1 - 2e^{-\gamma\left(\mathbf{s}'_{k,1}(2\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1} - \frac{\gamma\nu^2\sigma_{\mathcal{I}}^2}{2}\right)} \mathcal{N}\left(\gamma\nu\sigma_{\mathcal{I}}\right). \tag{A42}$$

Maximizing the objective function in (A42) is equivalent to maximizing,

$$-e^{-\gamma\left(\mathbf{s}'_{k,1}(2\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1} - \frac{\gamma\nu^2\sigma_{\mathcal{I}}^2}{2}\right)} \mathcal{N}\left(\gamma\nu\sigma_{\mathcal{I}}\right), \tag{A43}$$

or, equivalently, the natural log transformation of (A43),

$$\gamma\left(\mathbf{s}'_{k,1}(2\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1} - \frac{\gamma\nu^2\sigma_{\mathcal{I}}^2}{2}\right) - \ln\left(\mathcal{N}\left(\gamma\nu\sigma_{\mathcal{I}}\right)\right), \tag{A44}$$

which, is also equivalent to maximizing,

$$\mathbf{s}'_{k,1}(2\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1} - \frac{\gamma\nu^2\sigma_{\mathcal{I}}^2}{2} - \frac{1}{\gamma}\ln\left(\mathcal{N}\left(\gamma\nu\sigma_{\mathcal{I}}\right)\right). \tag{A45}$$

The final step is to substitute the explicit form of ν into (A46),

$$\mathbf{s}'_{k,1}(2\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1} - \frac{\gamma\left(\gamma\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)\left(\gamma\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)\sigma_{\mathcal{I}}^2}{2} - \frac{1}{\gamma}\ln\left(\mathcal{N}\left(\gamma^2\sigma_{\mathcal{I}}\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)\right). \tag{A46}$$

Constructing generation I indexer i's objective function follows immediately and is omitted for the sake of brevity.

■

Using Lemma A3, portfolio-builder k's $t = 1$ maximization problem is,

$$\max_{\mathbf{s}_{k,1}\in\mathbb{R}_+^N} \mathbf{s}'_{k,1}(2\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1} - \frac{\gamma\left(\gamma\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)\left(\gamma\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)\sigma_{\mathcal{I}}^2}{2}$$
$$- \frac{1}{\gamma}\ln\left(\mathcal{N}\left(\gamma^2\sigma_{\mathcal{I}}\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)\right). \tag{A47}$$

First-order conditions are necessary and sufficient to solve the optimization and the first-order conditions (in matrix form) are given by,

$$\vec{0} = (2\mu - \mathbf{P}_1) - \gamma\Omega\mathbf{s}_{k,1} - \gamma^3\sigma_{\mathcal{I}}^2\left(\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)(\Omega\mathbf{e}) - \frac{\phi\left(\gamma^2\sigma_{\mathcal{I}}\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)}{\mathcal{N}\left(\gamma^2\sigma_{\mathcal{I}}\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)}\gamma\sigma_{\mathcal{I}}\Omega\mathbf{e}, \tag{A48}$$

in which $\phi\left(\gamma^2\sigma_{\mathcal{I}}\mathbf{s}'_{k,1}\Omega\mathbf{e}\right)$ represents the standard normal distribution PDF evaluated at $\gamma^2\sigma_{\mathcal{I}}\mathbf{s}'_{k,1}\Omega\mathbf{e}$. Because all portfolio-builders are homogenous, are infinitesimal, are price-takers and face the same concave optimization, their demands are homogenous. As such, I drop the subscript k for generation I portfolio-builders hereafter.

Also using Lemma A3, indexer i's $t = 1$ maximization problem is,

$$\max_{\mathbf{f}_{i,1}\in\mathbb{R}_+} f_{i,1}\mathbf{e}'(2\mu - \mathbf{P}_1) - \frac{\gamma f_{i,1}^2}{2}\mathbf{e}'\Omega\mathbf{e} - \frac{\gamma\left(\gamma f_{i,1}\mathbf{e}'\Omega\mathbf{e}\right)\left(\gamma f_{i,1}\mathbf{e}'\Omega\mathbf{e}\right)\sigma_{\mathcal{I}}^2}{2}$$
$$- \frac{1}{\gamma}\ln\left(\mathcal{N}\left(\gamma^2\sigma_{\mathcal{I}}f_{i,1}\mathbf{e}'\Omega\mathbf{e}\right)\right). \tag{A49}$$

The first-order condition is necessary and sufficient to solve the optimization and the first-order conditions is given by,

$$0 = \mathbf{e}'(2\mu - \mathbf{P}_1) - \gamma f_{i,1}\mathbf{e}'\Omega\mathbf{e} - \gamma\left(\gamma\mathbf{e}'\Omega\mathbf{e}\right)\left(\gamma f_{i,1}\mathbf{e}'\Omega\mathbf{e}\right)\sigma_{\mathcal{I}}^2 - \frac{1}{\gamma}\frac{\phi\left(\gamma^2\sigma_{\mathcal{I}}f_{i,1}\mathbf{e}'\Omega\mathbf{e}\right)}{\mathcal{N}\left(\gamma^2\sigma_{\mathcal{I}}f_{i,1}\mathbf{e}'\Omega\mathbf{e}\right)}\gamma^2\sigma_{\mathcal{I}}\mathbf{e}'\Omega\mathbf{e}. \tag{A50}$$

Because all indexers are homogenous, are infinitesimal, are price-takers and face the same concave optimization, their demands are homogenous. As such, I drop the subscript i for generation I indexers hereafter.

In equilibrium, it must be the case that portfolio-builder k's demand for firm j at $t = 1$ is

$s^*_{j,1} = \frac{1-f^*_1}{b_I}$ for markets to clear. That is,

$$1 - f^*_1 = \int_0^{b_I} s^*_{j,1}\ \mathrm{d}x \Rightarrow s^*_{j,1} = \frac{1-f^*_1}{b_I}. \tag{A51}$$

This also implies that $s^*_{j,1} = s^*_{z,1}$ for all firms j and z, meaning that portfolio-builders hold an equal fraction of all N firms. Thus, portfolio-builders hold the equivalent of an index fund. With this insight, the N equations in (A48) may be summarized by a single equation of the form,

$$0 = \mathbf{e}'(2\mu - \mathbf{P}_1) - \gamma s_1 \mathbf{e}'\Omega\mathbf{e} - \gamma\left(\gamma\mathbf{e}'\Omega\mathbf{e}\right)\left(\gamma s_1\mathbf{e}'\Omega\mathbf{e}\right)\sigma^2_{\mathcal{I}} - \frac{1}{\gamma}\frac{\phi\left(\gamma^2\sigma_{\mathcal{I}} s_1\mathbf{e}'\Omega\mathbf{e}\right)}{\mathcal{N}\left(\gamma^2\sigma_{\mathcal{I}} s_1\mathbf{e}'\Omega\mathbf{e}\right)}\gamma^2\sigma_{\mathcal{I}}\mathbf{e}'\Omega\mathbf{e}, \tag{A52}$$

in which s_1 is the quantity that a portfolio-builder holds in all firms. Note, (A50) and (A52) are identical. As such, the solution to the portfolio-builder optimization and the indexer optimization must be equal, that is, $s^*_1 = f^*_1$. Moreover, because of market clearing, it must be the case that $s^*_1 = f^*_1 = \frac{1}{m_I + b_I}$. As such, (A48) may be rewritten as,

$$\vec{0} = (2\mu - \mathbf{P}_1) - \frac{\gamma}{m_I + b_I}\Omega\mathbf{e} - \frac{\gamma^3\sigma^2_{\mathcal{I}}}{m_I + b_I}\left(\mathbf{e}\Omega\mathbf{e}\right)\left(\Omega\mathbf{e}\right) - \frac{\phi\left(\frac{\gamma^2\sigma_{\mathcal{I}}}{m_I+b_I}\mathbf{e}\Omega\mathbf{e}\right)}{\mathcal{N}\left(\frac{\gamma^2\sigma_{\mathcal{I}}}{m_I+b_I}\mathbf{e}\Omega\mathbf{e}\right)}\gamma\sigma_{\mathcal{I}}\Omega\mathbf{e}. \tag{A53}$$

Consequently, firm j's $t = 1$ price is given by,

$$P^*_{j,1} = 2\mu_j - \frac{\gamma}{m_I + b_I}\left(\sum_{k=1}^{N}\sigma_{j,k}\right) - \gamma\sigma_{\mathcal{I}}\left(\frac{\gamma^2\sigma_{\mathcal{I}}}{m_I + b_I}\left(\mathbf{e}\Omega\mathbf{e}\right) + \frac{\phi\left(\frac{\gamma^2\sigma_{\mathcal{I}}}{m_I+b_I}\mathbf{e}\Omega\mathbf{e}\right)}{\mathcal{N}\left(\frac{\gamma^2\sigma_{\mathcal{I}}}{m_I+b_I}\mathbf{e}\Omega\mathbf{e}\right)}\right)\left(\sum_{k=1}^{N}\sigma_{j,k}\right). \tag{A54}$$

Using the short-hand notation introduced in the main prose, (A54) becomes,

$$P^*_{j,1} = 2\mu_j - \frac{\gamma\sigma_{j,M}}{m_I + b_I} - \gamma\sigma_{\mathcal{I}}\sigma_{j,M}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma^2_M}{m_I + b_I} + \frac{\phi\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma^2_M}{m_I+b_I}\right)}{\mathcal{N}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma^2_M}{m_I+b_I}\right)}\right). \tag{A55}$$

■

Proof of Lemma 3:

The result follows directly from the fact that $E\left[|\mathcal{I}|\right] = \sigma_{\mathcal{I}}\sqrt{\frac{2}{\pi}}$.

■

Proof of Lemma 4:

The result is easily confirmed by evaluating $x + \frac{\phi(x)}{\mathcal{N}(x)}$ for $x \in \mathbb{R}^+$.

■

Proof of Proposition 2:

The results follow directly from Lemma 3 and Lemma 4.

■

Proof of Corollary 2.1: The result follows from the following two cross derivative calculations;

$$\frac{\mathrm{d}^2\left(\gamma\sigma_{j,M}\sigma_{\mathcal{I}}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}+\frac{\phi\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)}{\mathcal{N}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)}-\sqrt{\frac{2}{\pi}}\right)\right)}{\mathrm{d}\sigma_{j,M}\,\mathrm{d}\sigma_{\mathcal{I}}}>0 \tag{A56}$$

$$\frac{\mathrm{d}^2\left(\gamma\sigma_{j,M}\sigma_{\mathcal{I}}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}+\frac{\phi\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)}{\mathcal{N}\left(\frac{\gamma^2\sigma_{\mathcal{I}}\sigma_M^2}{m_I+b_I}\right)}-\sqrt{\frac{2}{\pi}}\right)\right)}{\mathrm{d}\sigma_M^2\,\mathrm{d}\sigma_{\mathcal{I}}}>0. \tag{A57}$$

■

Online Appendix for "Index-Linked Trading and Stock Returns"

DAVIES, SHAUN WILLIAM[32]

OA.1 Endogenous Measure of Portfolio-Builders

In the base model, it is clear that portfolio-builders' expected utility weakly exceeds that of indexers. In this extension, I adopt the baseline model framework from Section 2 but I allow m and b to be endogenously determined via an incremental attention cost that is incurred when becoming a portfolio-builder. The incremental attention cost, denoted $c > 0$, may be thought of the resources required to evaluate each stock individually as compared to evaluating a position in only the index product. For simplicity, I assume that the combined measure of indexers and portfolio-builders equals one.

With this framework, each investor i plays a mixed strategy $\rho \in [0, 1]$; with probability $1 - \rho$ investor i pays c and becomes a portfolio-builder and with probability ρ investor i does not incur the fixed cost c and becomes an indexer. For simplicity, I assume that investors play symmetric strategies. As such, by the strong law of large numbers, a measure ρ of investors are indexers and a measure $1 - \rho$ are portfolio-builders.

To solve for the equilibrium mixed strategy, note that if ρ is internally valued then investors are indifferent between becoming an indexer or a portfolio-builder in equilibrium. If investors strictly prefer being an indexer in equilibrium, then $\rho = 1$. If investors strictly prefer begin a portfolio-builder in equilibrium then $\rho = 0$. Using the certainty equivalents of indexers' expected utility and portfolio-builders' expected utility, the equilibrium value of ρ is determined by the expression,

$$\mathbf{s}'(\mu - \mathbf{P}) - \frac{\gamma}{2}\mathbf{s}'\Omega\mathbf{s} - c - \left(f(\mathbf{e}'\mu - \mathbf{e}'\mathbf{P}) - \frac{\gamma f^2}{2}\sigma_M^2\right), \tag{OA1}$$

[32] Citation format: Davies, Shaun William Davies, Online Appendix for "Index-Linked Trading and Stock Returns," 2021, Working Paper.

in which element j of $\mathbf{P}$ is given by,

$$P_j = \mu_j - \frac{\gamma \sigma_{j,M}}{1-\rho} \left(\sum_{z=1}^{N} \chi_{j,z}(1-h_z) - \rho \sum_{z=1}^{N} \phi_z(1-h_z) \right), \tag{OA2}$$

element j of $\mathbf{s}$ is given by,

$$s_j = \frac{1}{1-\rho} \left(1 - h_j - \rho \sum_{z=1}^{N} \phi_z(1-h_z) \right) \tag{OA3}$$

and f is given by,

$$f = \sum_{z=1}^{N} \phi_z(1-h_z). \tag{OA4}$$

In (OA1), the expression in the large set of parenthesis represents indexers' certainty equivalent and the other terms represent portfolio-builders' certainty equivalent net of the cost c. If the expression in (OA1) is negative valued for all $\rho \in [0,1]$, then the only equilibrium is one in which all investors are indexers. Conversely, if the expression in (OA1) is positive valued for all $\rho \in [0,1]$, then the only equilibrium is one in which all investors are portfolio-builders. If (OA1) equals zero for some internal value of ρ, then investors mix between being portfolio-builders and indexers. If a separating equilibrium exists, the value of ρ is given by,

$$\rho^* = 1 - \sqrt{\frac{\gamma \left(\sum_{j=1}^{N} \sum_{y=j+1}^{N} \left(-\sigma_{j,y}^2 + (\sigma_{j,M} - \sigma_{j,y})(\sigma_{y,M} - \sigma_{j,y}) - \sum_{z \neq j, z \neq y}^{N} \sigma_{j,y}\sigma_{z,M} \right) (h_j - h_y)^2 \right)}{2c\sigma_M^2}}. \tag{OA5}$$

ρ^* is increasing in c; when the costs of becoming a portfolio-builder are higher, more investors become indexers. ρ^* is decreasing in γ; when investors have greater coefficients of risk-aversion, it is more valuable to hold the MVE portfolio and more investors become portfolio-builders. The specific comparative statics relating to the covariance matrix are more nuanced, however, it can be seen that as there are greater wedges of available float, that is, more non-zero values of $(h_j - h_y)$, ρ^* is smaller. This is due to there being a greater wedge between the market portfolio and the MVE portfolio.

OA.2 Numerical Estimation

In this section, I outline the numerical estimation methods used in Section 2.2 when closed-form solutions for prices are not obtainable. First, note that with exponential utility and normally distributed dividends (which are independent of $\mathcal{I}$), portfolio-builder k's objective function from the maximization in (20) is equivalent to,

$$1 - \int_{-\infty}^{\infty} e^{-\gamma\left(W + \mathbf{s}'_{k,1}(\mu + \mathbf{P}_2(\mathcal{I}) - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1}\right)} \phi(\mathcal{I}) \, \mathrm{d}\mathcal{I}, \tag{OA6}$$

in which $\mathcal{I}$ is a zero-mean normally distributed variable with variance $\sigma^2_{\mathcal{I}}$ and $\phi(\mathcal{I})$ is the PDF for $\mathcal{I}$. The equation in (OA6) may be rewritten as,

$$1 - e^{-\gamma\left(W + \mathbf{s}'_{k,1}(\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1}\right)} \int_{-\infty}^{\infty} e^{-\gamma\mathbf{s}'_{k,1}\mathbf{P}_2(\mathcal{I})} \phi(\mathcal{I}) \, \mathrm{d}\mathcal{I}. \tag{OA7}$$

The first-order condition with respect to $s_{k,j,1}$ is given by,

$$\begin{aligned} 0 = {} & \gamma \left((\mu_j - P_{j,1}) - \gamma \sum_{z=1}^{N} s_{k,z,1}\sigma_{j,z} \right) e^{-\gamma\left(W + \mathbf{s}'_{k,1}(\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1}\right)} \int_{-\infty}^{\infty} e^{-\gamma\mathbf{s}'_{k,1}\mathbf{P}_2(\mathcal{I})} \phi(\mathcal{I}) \, \mathrm{d}\mathcal{I} \\ & - e^{-\gamma\left(W + \mathbf{s}'_{k,1}(\mu - \mathbf{P}_1) - \frac{\gamma}{2}\mathbf{s}'_{k,1}\Omega\mathbf{s}_{k,1}\right)} \left(\frac{\mathrm{d}}{\mathrm{d}s_{k,j,1}} \int_{-\infty}^{\infty} e^{-\gamma\mathbf{s}'_{k,1}\mathbf{P}_2(\mathcal{I})} \phi(\mathcal{I}) \, \mathrm{d}\mathcal{I} \right), \end{aligned} \tag{OA8}$$

As all portfolio-builders are price-takers and face the same concave optimization I drop the subscript k hereafter and (OA8) simplifies to,

$$0 = (\mu_j - P_{j,1}) - \gamma \sum_{z=1}^{N} s_{z,1}\sigma_{j,z} - \frac{\left(\frac{\mathrm{d}}{\mathrm{d}s_{j,1}} \int_{-\infty}^{\infty} e^{-\gamma\mathbf{s}'_1\mathbf{P}_2(\mathcal{I})} \phi(\mathcal{I}) \, \mathrm{d}\mathcal{I} \right)}{\gamma \int_{-\infty}^{\infty} e^{-\gamma\mathbf{s}'_1\mathbf{P}_2(\mathcal{I})} \phi(\mathcal{I}) \, \mathrm{d}\mathcal{I}}. \tag{OA9}$$

Next, indexer i's objective function from the maximization in (21) is equivalent to,

$$1 - \int_{-\infty}^{\infty} e^{-\gamma\left(W + f_{i,1}\mathbf{e}'(\mu + \mathbf{P}_2(\mathcal{I}) - \mathbf{P}_1) - \frac{f_{i,1}^2\gamma}{2}\mathbf{e}'\Omega\mathbf{e}\right)}\phi(\mathcal{I})\ \mathrm{d}\mathcal{I}, \tag{OA10}$$

The equation in (OA10) may be rewritten as,

$$1 - e^{-\gamma\left(W + f_{i,1}\mathbf{e}'(\mu - \mathbf{P}_1) - \frac{f_{i,1}^2\gamma}{2}\mathbf{e}'\Omega\mathbf{e}\right)} \int_{-\infty}^{\infty} e^{-\gamma f_{i,1}\mathbf{e}'\mathbf{P}_2(\mathcal{I})}\phi(\mathcal{I})\ \mathrm{d}\mathcal{I}. \tag{OA11}$$

The first-order condition with respect to $f_{i,1}$ is given by,

$$\begin{aligned} 0 = \gamma\left(\mathbf{e}'(\mu - \mathbf{P}_1) - f_{i,1}\gamma\mathbf{e}'\Omega\mathbf{e}\right) e^{-\gamma\left(W + f_{i,1}\mathbf{e}'(\mu - \mathbf{P}_1) - \frac{f_{i,1}^2\gamma}{2}\mathbf{e}'\Omega\mathbf{e}\right)} \int_{-\infty}^{\infty} e^{-\gamma f_{i,1}\mathbf{e}'\mathbf{P}_2(\mathcal{I})}\phi(\mathcal{I})\ \mathrm{d}\mathcal{I} \\ - e^{-\gamma\left(W + f_{i,1}\mathbf{e}'(\mu - \mathbf{P}_1) - \frac{f_{i,1}^2\gamma}{2}\mathbf{e}'\Omega\mathbf{e}\right)} \left(\frac{\mathrm{d}}{\mathrm{d}f_{i,1}} \int_{-\infty}^{\infty} e^{-\gamma f_{i,1}\mathbf{e}'\mathbf{P}_2(\mathcal{I})}\phi(\mathcal{I})\ \mathrm{d}\mathcal{I}\right). \end{aligned} \tag{OA12}$$

As all indexers are price-takers and face the same concave optimization I drop the subscript i hereafter and (OA12) simplifies to,

$$0 = \mathbf{e}'(\mu - \mathbf{P}_1) - f_1\gamma\mathbf{e}'\Omega\mathbf{e} - \frac{\left(\frac{\mathrm{d}}{\mathrm{d}f_1}\int_{-\infty}^{\infty} e^{-\gamma f_1\mathbf{e}'\mathbf{P}_2(\mathcal{I})}\phi(\mathcal{I})\ \mathrm{d}\mathcal{I}\right)}{\gamma\int_{-\infty}^{\infty} e^{-\gamma f_1\mathbf{e}'\mathbf{P}_2(\mathcal{I})}\phi(\mathcal{I})\ \mathrm{d}\mathcal{I}}. \tag{OA13}$$

The equilibrium investor demands must satisfy market-clearing, that is, for all j, the following equality must hold,

$$1 = h_j + s_{j,1} + f_1. \tag{OA14}$$

There are $2N + 1$ unknown variables: the prices of the N firms, portfolio-builders' demand for the N firms, and indexers' demand for the index fund. There are also $2N + 1$ unique equations: (OA9) implies N equations, (OA13) implies one equation, and (OA14) implies N equations. Thus, there are $2N + 1$ unknowns and $2N + 1$ unique equations. Numerical solutions to the system of equations are obtained using MATLAB. The numerical solutions are obtained similarly for the analysis in Section 2.2.1 in which the total measure of investors is fixed but there is uncertainty about the future fraction that will be indexers.