# Mapping Causes to Consequences: The Impact of Indexing[*]

Hong Liu[†] Yajun Wang [‡]

November 13, 2023

## Abstract

We analyze the impact of indexing when investors can endogenously acquire information. We find that the impact critically depends on what drives the rise of indexing. If indexing is driven by reduced participation costs or decreased liquidity trading in the non-index market, then: (1) price informativeness and relative price efficiency increase, (2) the welfare of non-liquidity traders decreases, and (3) both stock correlations and the expected market capitalization increase. The opposite is true if indexing is driven by increased liquidity trading in the index market. Our analysis highlights the importance of identifying the driving forces behind the rise of indexing.

*JEL* Classification Codes: G11, G12, G14, D82.

Keywords: Indexing, Information Acquisition, Risk Premium, Capital Allocation, Correlations, Welfare.

[*]We thank Philip Bond, Phil Dybvig, Pete Kyle, Guofu Zhou and seminar participants at Fudan University, SUFE, SWUFE, Washington University in St. Louis, University of Maryland, CICF, and MFA for helpful comments.

[†]Olin Business School, Washington University in St. Louis and CHIEF, liuh@wustl.edu.

[‡]Zicklin School of Business, Baruch College, yajun.wang@baruch.cuny.edu.

# Mapping Causes to Consequences: The Impact of Indexing


## Abstract

We analyze the impact of indexing when investors can endogenously acquire information. We find that the impact critically depends on what drives the rise of indexing. If indexing is driven by reduced participation costs or decreased liquidity trading in the non-index market, then: (1) price informativeness and relative price efficiency increase, (2) the welfare of non-liquidity traders decreases, and (3) both stock correlations and the expected market capitalization increase. The opposite is true if indexing is driven by increased liquidity trading in the index market. Our analysis highlights the importance of identifying the driving forces behind the rise of indexing.



*JEL* Classification Codes: G11, G12, G14, D82.

Keywords: Indexing, Information Acquisition, Risk Premium, Capital Allocation, Correlations, Welfare.

# 1 Introduction

Index investing has grown substantially during the past few decades.[1] Indexers pay lower management fees and passive funds save the cost of stock picking. However, with this continued growth in recent years, some major concerns arise in the industry, media, and academia. For example, does a significant rise of index investing reduce information acquisition, thereby impairing market efficiency? Will this rise reduce the welfare of market participants, exacerbate stock price comovements, and thus amplify systemic risks? In this paper, we develop a rational expectations equilibrium model to study the consequences of indexing. We find that the effects of index investing depend critically on the causes of its rise.

The rise of indexing can be caused by many factors. In our analysis, we focus on three of them: (1) a decline in the index participation cost; (2) a decline in liquidity trading in the non-index market; and (3) an increase in liquidity trading in the index market. A lowered cost to participate in the index market draws savers or discretionary investors toward indexing. A decrease in the liquidity trading in the non-index market reduces the trading profitability for informed investors and thus more discretionary traders become indexers and save the participation cost of the non-index market. Likewise, an increase in liquidity trading in the index market makes trading in the index market more profitable and thus draws more discretionary traders toward indexing.

We find that if the rise of indexing results from either reduced participation costs or a decline in liquidity trading in the non-index market, then: (1) Price informativeness in both the index and non-index markets increases; (2) Stock correlations intensify; (3) The expected market capitalization grows; (4) The magnitude of index return reversal decreases; and (5) The welfare of non-liquidity traders decreases. In contrast, if the rise of indexing is driven by increased liquidity trading in the index market, its impact is reversed.

Our findings thus highlight the importance of identifying the underlying driving force behind the rise of indexing in understanding its economic impact. In addition, if multiple causes

[1] During 2016, actively-managed funds experienced \$285 billion of outflows while passive funds attracted \$429 billion of inflows. The proliferation of ETFs is now approaching 2,000 funds and nearly \$3.0 trillion of asset under management.

are at play in practice, the overall effects of indexing might be insignificant as the impact due to different causes may offset one another. This aligns with empirical findings such as those by Coles, Heath, and Ringgenberg (2022), who report that passive investing does not compromise price efficiency.

In our model, investors can trade two risky securities— the index and the non-index— in addition to a risk-free asset. These securities replicate the trading of two stocks. While trading the risk-free asset is cost-free for all investors, trading the index and non-index may involve participation costs (e.g., learning costs). We consider four types of investors: (i) active investors who have no participation costs for either the index or the non-index securities, and thus always trade both risky assets; (ii) discretionary investors who incur participation costs for trading either the index or the non-index. They endogenously choose to be indexers or active investors; (iii) savers who are only endowed with the risk-free asset incur a participation cost for trading the index but do not trade the non-index at all; and (iv) liquidity traders who exogenously trade both risky assets to hedge against endowment risks. Only active traders and discretionary investors may acquire private information about the fundamental value of the risky assets by paying additional information acquisition costs. We focus on the case where acquiring information in one market increases the information acquisition cost in the other, capturing the qualitative feature of limited information processing capacity.

When indexing rises due to either decreased participation costs in the index market or reduced liquidity trading within the non-index market, the price informativeness of both index and non-index market increases. This increase can be attributed to some discretionary traders potentially shifting towards indexing, driven by either lowered costs associated with indexing or by decreased profitability in trading in the non-index market. These discretionary indexers tend to acquire more precise private information compared to their active trader counterparts. The reason is simple: active traders split their focus between index and non-index markets, which in turn increases their information acquisition costs in the index market. As a result, with an influx of these well-informed discretionary investors into the index market, its price informativeness increases. At the same time, as these traders migrate away from the non-index market, the remaining participants find greater value in obtaining more accurate information about the non-index, leading to an increase in its price informativeness.

The impact of indexing on important economic measures like market risk premium, market capitalization, stock price correlations, index return reversal, and the welfare of market participants is intrinsically tied to its effect on price informativeness. Specifically, as price informativeness increases, there is a corresponding decrease in aggregate uncertainty. This in turn results in a decline in the market risk premium and a rise in asset prices. Furthermore, a reduction in the index participation cost or reduced liquidity trading (and thus lower profitability) in the non-index market prompts either more savers or discretionary traders to turn to indexing, they infuse more capital into the stock market and boost the expected market capitalization.

The rise in indexing has sparked concerns about a potential increase in the comovement of asset prices, especially as more investors trade stocks bundled within indices. We find that, when indexing rises due to either decreased index participation costs or reduced liquidity trading within the non-index market, asset prices comovements do amplify. This amplification can be attributed to a greater number of investors acquiring and acting on their private information about the fundamental value of the index. As prices become more informative and responsive to this fundamental value, stocks within the index tend to move more synchronously.

Within the framework of rational expectations equilibrium models, the index exhibits return reversal mainly because liquidity trading drives equilibrium prices away from the fundamental values. As prices become more informative, the magnitude of return reversal decreases since prices stray less from the index's fundamental value. The diminished index return reversal implies reduced profits for non-liquidity traders who trade against liquidity traders, leading to a decline in their welfare. In contrast, the welfare of liquidity traders increases.

The impact of indexing on various market facets such as price informativeness, market risk premium, market capitalization, stock price correlations, and the welfare of market participants is reversed when the rise of indexing is driven by increased liquidity trading in the index market. An increase in liquidity trading in the index market enhances its profitability, attracting more discretionary traders to switch to indexing. While these discretionary traders typically acquire more precise private information due to the elevated marginal value of private information, the effect is dominated by the increased liquidity trading and as a result, the price informativeness of the index decreases. Since active investors acquire more precise information in the index market, which increases their information acquisition cost in the non-index market, active investors choose to acquire less precise information about the non-index market and thus the price informativeness of the non-index market also decreases.

It follows that, when indexing rises due to increased liquidity trading in the index market, the market risk premium increases and the expected market capitalization decreases. Asset prices comovements attenuate because prices are less responsive to fundamental values since there are more liquidity trading (e.g., non-fundamentals trading) in the index market. As prices become less informative, the magnitude of return reversal intensifies as prices deviate more from the index's fundamental value, implying that non-liquidity traders benefit more from trading with liquidity traders. Consequently, non-liquidity investors' welfare increases while liquidity investors' welfare decreases.

Our paper highlights the importance of discerning the mechanisms driving the rise of index investing. The observed empirical relationship between the surge in index investment and price informativeness can guide researchers in pinpointing the likely causes behind increased indexing. Specifically, if prices grow more informative alongside a rise in indexing, it suggests that reduced index participation costs or decreased liquidity trading in the non-index market may be the drivers. On the other hand, if prices become less informative, it hints at increased liquidity trading in the index market as the main driving force. If price informativeness remains unchanged, it suggests that multiple causes may be at play in practice, resulting in an insignificant overall effect.

Our paper is closely related to Bond and Garcia (2022). They find that an increase in indexing results in decreased price informativeness, heightened market risk premium, magnified index return reversals, and increased welfare of market participants. They focus on the rise of indexing attributed to a reduced index participation cost under the assumption of exogenously endowed information. In their model, a reduction in the index participation cost attracts more traders who are less informed and have endowment shocks, thus injecting more noise into the index market.[2] In contrast, in our model, discretionary investors can optimally choose the precision of private information they acquire. When discretionary investors switch to indexing—motivated by either lowered index participation cost or diminished liquidity trading in the non-index market—they tend to acquire more precise information compared to active traders and thus the price informativeness may increase when indexing rises. As a result, in our analysis, if the rise of indexing is driven by reduced participation costs or decreased liquidity trading in the non-index market, then the price informativeness increases, the magnitude of index return reversal decreases, and the welfare of non-liquidity traders decreases. In essence, our paper not only complements Bond and Garcia (2022) but also underlines the necessity of discerning the underlying driving force of the growth of index investing to fully grasp its implications.

Our model builds upon the foundational framework of Grossman and Stiglitz (1980), adapting it to a multi-asset environment to study the impact of indexing. Unlike Diamond and Verrecchia (1981), Ganguli and Yang (2009), and Bond and Garcia (2022), but in line with Admati (1985), we treat liquidity trades as entirely exogenous. Nonetheless, this assumption is unlikely to critically influence our findings, as it mirrors a scenario with endogenous liquidity trades when risk aversion approaches infinity. Van Nieuwerburgh and Veldkamp (2009) study information acquisition in multi-asset markets. They use information acquisition capacity constraints to model the tension between acquiring information in different markets. The negative information acquisition externality assumption in our model echoes their stance: acquiring more information in one market increases the cost of acquiring information in another, consequently decreasing information acquisition. Baruch and Zhang (2022) find that while indexing does not impact the validity of the CAPM risk-return relation, it does reduce the price efficiency of individual stocks. In their study, investors observe the realizations of private signals without any cost. In contrast, we focus on how indexing affects price informativeness and market risk premium through its impact on endogenous information acquisition and the associated trade-off between costs and benefits. As a result, indexing may increase market efficiency if it is driven

[2]Their findings align with our results when the rise of indexing is due to increased liquidity trading in the index market.

by reduced participation costs or decreased liquidity trading in the non-index market.

Benchmarking to an index is qualitatively similar to indexing. Breugem and Buss (2018) show that benchmarking to an index reduces price informativeness. In contrast, Chen, Hu, and Wang (2022) show that price informativeness can increase with investors' benchmarking concerns about an asset when traders employ an integrative learning technology by observing a private signal about a linear combination of asset payoffs. Unlike our model, a change in benchmarking is exogenous in both Breugem and Buss (2018) and Chen et al. (2022). This leads to their results different from ours in several key aspects.

The empirical findings regarding the effects of indexing on price informativeness are mixed. Coles et al. (2022) use Russell Index reconstitutions as a source of exogenous variation in passive investing and find that the rise of index investing does not significantly affect market efficiency. Israeli, Lee, and Sridharan (2017) find that an increase in ETF ownership is associated with lower future earnings response coefficients and less informative asset prices. Brogaard, Ringgenberg, and Sovich (2019) find a decline in investment efficiency for index commodity firms. In contrast, Glosten, Nallareddy, and Zou (2021) find that ETF activity increases informational efficiency by improving the link between short-run fundamentals and stock prices. Antoniou, Li, Liu, Subrahmanyam, and Sun (2023) find that higher ETF ownership is associated with an increased sensitivity of real investment to Tobin's q and improved predictive power of stock returns for future earnings. Farboodi, Matray, Venkateswaran, and Veldkamp (2022) find that while price informativeness increases for the S&P 500 firms, it declines for the broader market. These mixed empirical findings align well with the main point of our paper: the effects of indexing critically depend on the underlying factors that drive its rise.

The remainder of the paper proceeds as follows. In Section 2, we present the model. Section 3 derives the equilibrium. Section 4 examines the equilibrium effects associated with the rise of indexing. We conclude in Section 5. All proofs are provided in the Appendix.

## 2 The Model

**The Asset Market.** We consider a one-period model where a continuum of investors can trade one risk-free asset and two stocks at time 0 to maximize their expected utility at time 1. The net supply of the risk-free asset is zero. Each stock $i$ ($i = 1, 2$) has a supply of one share and a final payoff of $V_i$ with distribution $N(\mu_i, 1/\tau)$, where $V_1$ and $V_2$ are correlated with a correlation coefficient $\rho \in (-1, 1)$.

As pointed out by Bond and Garcia (2022), trading the two stocks is equivalent to trading the market portfolio $m$ ( also referred to as the "index portfolio") and the spread portfolio $s$ (referred to as "non-index portfolio"), where the payoffs of the index $m$ and the non-index $s$ are given as

$$V_m = \tfrac{V_1+V_2}{2} \quad \text{and} \quad V_s = \tfrac{V_1-V_2}{2}, \tag{1}$$

respectively. It can be shown that $V_m$ and $V_s$ are independently distributed as $N(\mu_{vm}, 1/\tau_{vm})$ and $N(\mu_{vs}, 1/\tau_{vs})$, where $\mu_{vm} = \frac{\mu_1+\mu_2}{2}$, $\tau_{vm} = \frac{2}{1+\rho}\tau$, $\mu_{vs} = \frac{\mu_1-\mu_2}{2}$, and $\tau_{vs} = \frac{2}{1-\rho}\tau$.[3]

Since each stock has a supply of one share and thus the total payoff of the two stocks is $V_1 + V_2$, this implies that the total supply of the index portfolio is $s_m = 2$ and that of the non-index portfolio is $s_s = 0$. Because trading the two portfolios is equivalent to directly trading the two stocks, and the independence of the index and non-index portfolios payoffs considerably simplifies the analysis, in subsequent analysis, we assume that investors trade the two portfolios directly and determine their equilibrium prices. We will then derive the equilibrium prices for the stocks, based on the relationship between the payoffs of the stocks and the portfolios.

**Participation Costs and Types of Investors.** There is no participation cost in investing in the risk-free asset, but some investors incur participation costs in terms of utility loss (e.g., from

[3]The equal variance of $V_1$ and $V_2$ is assumed so that $\mathrm{Cov}(V_m, V_s) = 0$. Given that $V_m$ and $V_s$ are normally distributed, this implies that $V_m$ and $V_s$ are independent of each other, substantially simplifying subsequent analysis. Without this assumption, solving the model would involve solving for 10 coefficients as opposed to 5, and it appears that closed-form solutions would not be readily obtainable, if possible at all. On the other hand, one can always vary the supply of a stock to change the variance of its payoff per share and solve for the equilibrium using the above approach. Therefore, the assumption of equal variance is purely for analytical simplicity.

time and attention consumption) before trading the index or the non-index portfolios.[4]

A mass $\lambda_A$ of investors are exogenous active investors, denoted as "A" investors, who incur no participation costs for trading in either the index or the non-index market, thereby always engaging in trading both risky assets. They can also acquire private information about these two assets by paying information acquisition costs, which will be specified subsequently. Each A investor is endowed with 2 shares of the index portfolio and $-\beta < 0$ units of the risk-free asset. These active investors may represent funds that have a relatively low cost in stock-picking.

A mass $\lambda_D = 1 - \lambda_A$ of investors are discretionary traders, referred to as "D" investors. These investors need to pay $k_{Dm} > 0$ to participate in the index and $k_s > 0$ to participate in the non-index market. Like an A investor, a D investor can also acquire private information about these two risky assets after incurring information acquisition costs. Each D investor is endowed with 2 shares of the index portfolio and $-\beta$ units of the risk-free asset. D investors may represent funds that experience relatively high costs in stock picking and might opt to invest solely in the index market.

A mass $\lambda_S$ of investors, referred to as "S" investors, are initially savers endowed with $\frac{1}{\lambda_S}\beta$ units of the risk-free asset but no risky assets. These S investors incur a participation cost of $k_m > 0$ when trading in the index market and they do not trade the non-index portfolio due to high participation costs. In addition, acquiring private information is not feasible for them due to their high information acquisition costs. S investors might represent institutional investors who begin with cash and allocate investments between an index fund and a money market fund.

A mass $\lambda_L$ of investors are liquidity traders, denoted as "L" investors, whose trading in both risky portfolios is exogenous and subject to a participation cost of $k_m$ and $k_s$ respectively. These L investors have zero endowment of both the risky portfolios and the risk-free asset. Instead, each liquidity trader $i$ has a random endowment $e_{ji} = Z_j + u_{ji}$ shares of a non-traded asset $j$ (e.g., two streams of labor income) for $j \in \{m, s\}$, with each share of the non-traded asset $j$ paying the same amount at time 1 as $V_j$, where $Z_j$ and $u_{ji}$ are independently distributed as

[4] The case where the participation cost is monetary and is paid to the active investors yields the same qualitative results if the cost is exogenous (e.g., due to competition).

$N(0, 1/\tau_{zj})$ and $N(0, 1/\tau_{uj})$ respectively, and $\tau_{zj}$ and $\tau_{uj}$ are constants.[5] $Z_j$ (resp. $u_{ji}$) can be interpreted as an aggregate (resp. idiosyncratic) shock in the endowment. Liquidity traders have mean-variance preferences over the final wealth and an infinite risk aversion.[6] As a result, liquidity traders have to perfectly hedge their endowment risk, which leads them to take the opposite positions in markets $m$ and $s$ to their endowment of the non-traded assets, incurring a participation cost of $k_m + k_s$. After hedging, since they have no risk exposure, they have no incentive to acquire any information.

All non-liquidity-traders have constant absolute risk averse (CARA) preferences with a risk aversion coefficient of $\gamma > 0$. As in Grossman and Stiglitz (1980), the presence of liquidity traders is necessary for the existence of an equilibrium in our model. The liquidity traders are similar to noise traders commonly assumed in the literature. The main difference is that we can measure liquidity traders' utility by the expected terminal wealth which depends on market prices. Given this setup, we can shed light on how indexing affects liquidity traders' welfare.

In actual financial markets, hedge funds can be viewed as "A" investors who do not passively invest only in index funds. Mutual funds can be considered as "D" investors in our model. Some of them may choose to become active mutual funds ("DA" investors), while others might choose to be indexers ("DI" investors). Institutional investors that allocate capital between index funds and money market funds are considered as "S" investors who do not select stocks and do not acquire private information about assets. Hedgers, who trade assets primarily for hedging or rebalancing reasons and not based on the fundamental value of assets, can be categorized as "L" investors.

**Information Acquisition.** At a certain cost, each investor $i$ of type $t \in \{A, D\}$ can observe independent private signals $Y_{tji}$ at time 0 about the payoff of the risky asset $j$, where

$$Y_{tji} = V_j + \varepsilon_{tji}, \quad j \in \{m, s\},$$

[5] It is sufficient to assume that the payoff of the non-traded asset $j$ is perfectly correlated with $V_j$ so that hedging motive is present. This is equivalent to assuming that liquidity traders are noise traders who have exogenous trading demand.

[6] With finite risk aversion, the derivation is more complicated, but the qualitative results are the same.

and the noise term $\varepsilon_{tji}$ is independently distributed as $N(0,1/\tau_{tji})$. The cost of acquiring private information with precisions $\tau_{tm}$ and $\tau_{ts}$ is $C_t(\tau_{tm},\tau_{ts})$ for $t \in \{A,D\}$. We only consider symmetric equilibria where investors of the same type make the same trading and information acquisition decision. As a result, the precision choices are the same across investors of the same type, and thus we omit the $i$ index in the precision variables.

**Investors' Problems.** Let $P_m$ and $P_s$ be the equilibrium prices of the index portfolio and the non-index portfolio respectively, $I_{ti}$ be time 0 information set of investor $i$ of type $t$, and $\Theta_{tji}$ be the number of shares of the $j$ portfolio bought by investor $i$ of type $t$ at time 0, for $j \in \{m,s\}$ and $t \in \{A,D,S,L\}$. We denote D (S, resp.) investors who choose to be indexers as DI (SI, resp.) investors and D investors who choose to trade in both markets as DA investors. Investor $i$ of type $t$ ($t \in \{A,D,SI\}$) chooses $(\Theta_{tmi},\Theta_{tsi})$ to solve

$$\max \mathrm{E}\left[-\mathrm{e}^{-\gamma(\tilde{W}_{ti}-k_{Dm}\ 1_{\{t=DI\}}-k_m\ 1_{\{t=SI\}}-(k_{Dm}+k_s)\ 1_{\{t=DA\}})}\middle| I_{ti}\right], \tag{2}$$

subject to the budget constraint

$$\tilde{W}_{ti}=(2V_m-\beta)\ 1_{\{t=A,D\}}+\frac{1}{\lambda_S}\beta\ 1_{\{t=S\}}+\Theta_{tmi}(V_m-P_m)+\Theta_{tsi}(V_s-P_s)-C_t(\tau_{tm},\tau_{ts})\ 1_{\{t=A,D\}}, \tag{3}$$

where $t=DI$ if and only if $\Theta_{Dmi}\neq 0$ and $\Theta_{Dsi}=0$, $t=DA$ if and only if $\Theta_{Dmi}\neq 0$ and $\Theta_{Dsi}\neq 0$, $t=SI$ if and only if $\Theta_{Smi}\neq 0$ and $\Theta_{Ssi}=0$.

The information set of investor $i$ of type $A, DA, DI$, and $SI$ is $I_{ti}=(Y_{tmi},Y_{tsi},P_m,P_s)$, where the precision of $Y_{tmi}$ and $Y_{tsi}$ is zero for $SI$ investors, and the precision of $Y_{tsi}$ is zero for $DI$ investors.

For liquidity traders, $\Theta_{Lmi}=-(Z_m+u_{mi})$ and $\Theta_{Lsi}=-(Z_s+u_{si})$.

**Market-Clearing Condition.** The time 0 equilibrium is $\{P_j,\Theta_{tji}\}$, for $j \in \{m,s\}$, $t \in \{A,D,S,L\}$ such that $\Theta_{tji}$ solves the above problem for investor $i$ (equation (2)) of type $t \in \{A,D,S\}$ and the following market clearing condition is satisfied:

$$\sum_{t\in\{A,D,S,L\}} \int_i \Theta_{tji} di = 0, \quad j \in \{m, s\}. \tag{4}$$

**Choice of Indexing and Precisions of Signals.** Just before time 0, type $A$ investors choose their optimal precisions of signals about the payoffs of both the index and non-index portfolios. Type D investors decide whether to become indexers, become active investors, or simply retain their endowment; they also decide on their optimal precisions of signals about the payoffs. Type S investors choose either to become indexers or to continue investing solely in the risk-free asset.

In equilibrium, depending on the participation costs $k_{Dm}$, $k_m$, and $k_s$, we have the following possible cases: (1) S (resp. D) investors do not trade in the index market if and only if the participation cost $k_m$ (resp. $k_{Dm}$) is sufficiently large; and (2) D investors do not trade in the non-index market if and only if $k_s$ is sufficiently large.

# 3 The Equilibrium

The solution of the optimization problem for investor $i$, as described by equation (2), yields the optimal number of shares for both the index and non-index portfolios for investor $i$ of type $t$:

$$\begin{aligned} \Theta_{tmi} &= \frac{\mathrm{E}[V_m|I_{ti}] - P_m}{\gamma \mathrm{Var}[V_m|I_{ti}]} - s_m 1_{\{t=A,D\}}, \quad t = A, DA, DI, SI, \\ \Theta_{tsi} &= \frac{\mathrm{E}[V_s|I_{ti}] - P_s}{\gamma \mathrm{Var}[V_s|I_{ti}]} - s_s, \quad t = A, DA, \end{aligned} \tag{5}$$

where the total supplies for the market and non-index portfolios are $s_m = 2$ and $s_s = 0$, respectively. The information set of investor $i$ of type $t$ is $I_{ti} = (Y_{tmi}, Y_{tsi}, P_m, P_s)$.

In a linear equilibrium, we conjecture and later verify that

$$P_j = a_j + b_j V_j - d_j Z_j, \quad j \in \{m, s\}, \tag{6}$$

where $a_j$, $b_j$, and $d_j$ are constants that will be determined in Theorem 1.

By direct computation, we obtain the following expressions for the conditional expectation and variance of portfolio $j$'s payoff for investor $i$ of type $t = A, D$, and for $j = m, s$:

$$\begin{aligned}
\mathrm{E}[V_j|I_{ti}] &= \frac{\mu_{vj}\tau_{vj} + Y_{tji}\tau_{tji} + \frac{b_j^2}{d_j^2}\left(V_j - \frac{d_j}{b_j}Z_j\right)\tau_{zj}}{\tau_{vj} + \tau_{tji} + \frac{b_j^2}{d_j^2}\tau_{zj}}, \\
\mathrm{Var}[V_j|I_{ti}] &= \left(\tau_{vj} + \tau_{tji} + \frac{b_j^2}{d_j^2}\tau_{zj}\right)^{-1}.
\end{aligned} \tag{7}$$

In our analysis, we focus on the linear symmetric equilibrium, where investors of the same type adopt identical trading strategies and select the same information precision. We can simplify our notation by omitting the $i$ index in the precision variables. Hence, the precision of signals for investor $i$ of type $t$ will be denoted by $\tau_{tj}$, rather than $\tau_{tji}$.

For $j \in \{m, s\}$, let $\tau_j$ denote the total precision of private information and $\rho_j$ denote the price informativeness about asset $j$, we have

$$\tau_j := \lambda_A \tau_{Aj} + \lambda_{DA}\tau_{DAj} + \lambda_{DI}\,\tau_{DIj} 1_{j=m}, \tag{8}$$

$$\rho_j := \left(Var[V_j|P_j]\right)^{-1} = \tau_{vj} + \frac{\tau_j^2}{\gamma^2\lambda_L^2}\tau_{zj}. \tag{9}$$

Equation (9) implies that, for given values of $\tau_{vj}$ and $\tau_{zj}$, the price informativeness in market $j$ increases with the total precision $\tau_j$ of private information in market $j$.

Given the signal precisions $\tau_{tj}$, $t \in \{A, DI, DA\}$, $j \in \{m, s\}$, there is a unique linear symmetric equilibrium. The following theorem presents the equilibrium prices.

**Theorem 1** *For $j \in \{m, s\}$, the equilibrium price is given in equation (6) and coefficients $a_m$, $b_m$, and $d_m$ are given as*

$$\begin{aligned}
a_m &= \frac{(\lambda_A + \lambda_{DA} + \lambda_{DI} + \lambda_{SI})\mu_{vm}\tau_{vm} - 2\gamma}{\tau_m + (\lambda_A + \lambda_{DA} + \lambda_{DI} + \lambda_{SI})\rho_m}, \\
b_m &= 1 - \frac{(\lambda_A + \lambda_{DA} + \lambda_{DI} + \lambda_{SI})\tau_{vm}}{\tau_m + (\lambda_A + \lambda_{DA} + \lambda_{DI} + \lambda_{SI})\rho_m}, \quad d_m = \frac{\gamma\lambda_L b_m}{\tau_m},
\end{aligned} \tag{10}$$

*and the coefficients $a_s$, $b_s$, and $d_s$ are given as*

$$a_s = \frac{(\lambda_A + \lambda_{DA})\mu_{vs}\tau_{vs}}{\tau_s + (\lambda_A + \lambda_{DA})\rho_s}, \quad b_s = 1 - \frac{(\lambda_A + \lambda_{DA})\tau_{vs}}{\tau_s + (\lambda_A + \lambda_{DA})\rho_s}, \quad d_s = \frac{\gamma\lambda_L b_s}{\tau_s}, \tag{11}$$

*with $\tau_j$ and $\rho_j$ given by equations (8) and (9) and the mass of DI, DA and SI investors determined endogenously.*

Theorem 1 implies that the risk premium of the market portfolio $m$ is equal to

$$\mathrm{E}[V_m - P_m] = \frac{2\gamma}{\tau_m + (\lambda_A + \lambda_{DA} + \lambda_{DI} + \lambda_{SI})\rho_m}. \tag{12}$$

Equation (12) implies that the market risk premium and the price informativeness $\rho_m$ tend to move in opposite directions, *ceteris paribus*. This is because as price informativeness increases, the aggregate uncertainty in the market reduces. In contrast, the risk premium of the non-index portfolio is zero because the aggregate supply of the non-index portfolio is zero and thus there is no aggregate risk. Therefore, the expected price of the non-index is

$$\mathrm{E}[P_s] = \mu_{vs}. \tag{13}$$

Just before time 0, investors optimally choose the precisions of their private signals. Let $C_{tm}$, $C_{ts}$, $C_{tmm}$, and $C_{tss}$ denote respectively the first and the second derivative of the information acquisition cost function $C_t(\tau_{tm}, \tau_{ts})$ with respect to $\tau_{tm}$ and $\tau_{ts}$, and $C_{tms}$ denote the cross derivative, for $t \in \{A, D\}$. To ensure the existence and the uniqueness of equilibrium, we make the following assumption:

**Assumption 1** *$C_{tj} \geq 0$ for $j \in \{m, s\}$ with equality only at $\tau_{tm} = \tau_{ts} = 0$. $C_{tmm} > 0$, $C_{tss} > 0$, and $C_{tmm}C_{tss} - C_{tms}^2 \geq 0$ for $t \in \{A, D\}$.*

Assumption 1 ensures the convexity of the cost function. The information acquisition externality (IAE) for type $t$ ($t = A, D$) investors can be measured by

$$\varphi_t := -C_{tms} = -\frac{\partial^2 C_t(\tau_{tm}, \tau_{ts})}{\partial \tau_{tm} \partial \tau_{ts}}. \tag{14}$$

A negative $\varphi_t$ indicates that increasing information acquisition in one market elevates the information acquisition cost in the other market. This negative externality could be interpreted as an investor having a fixed information acquisition (or processing) capacity; hence, gathering more information in one market reduces the capacity for acquiring information in another. A positive $\varphi_t$ suggests that acquiring information in one market can decrease the acquisition cost in another, pointing to a positive IAE. This might be understood as the experience in one market aiding the efficiency of information gathering in another, consequently lowering its cost.

Given the optimal signal precisions and trading strategies of other investors, D investors decide whether to pay the cost $k_{Dm}$ to trade the index and $k_s$ for the non-index portfolio. S investors determine whether to pay the cost $k_m$ to participate in the index market.

**Theorem 2** *Assume $k_{Dm} = 0$.*[7] *Under Assumption 1, there exists a linear symmetric equilibrium where the equilibrium information precisions solve the following five equations:*

$$2\gamma C_{tj}(\tau_{tm}, \tau_{ts}) = (\tau_{tj} + \rho_j)^{-1}, \quad t \in \{A, DA\}, j \in \{m, s\}, \tag{15}$$

*and*

$$2\gamma C_{DIm}(\tau_{DIm}, 0) = (\tau_{DIm} + \rho_m)^{-1}, \tag{16}$$

*the proportion $\lambda_{D0} = 0$ and the equilibrium masses $\lambda_{SI}$ and $\lambda_{DI}$ are such that (1) either all S investors strictly prefer investing only in the risk-free asset (i.e., $\lambda_{SI} = 0$) or all S investors strictly prefer indexing (i.e., $\lambda_{SI} = \lambda_S$) or each S investor is indifferent between pure savers and indexers (i.e., $\lambda_S > \lambda_{SI} > 0$); and (2) either all D investors strictly prefer investing only in the index (i.e., $\lambda_{DI} = \lambda_D$) or all D investors strictly prefer being active (i.e., $\lambda_{DA} = \lambda_D$) or each D investor is indifferent between investing only in the index and being active (i.e., $\lambda_D > \lambda_{DI} > 0$).*

[7] Because of the complexity of the endogenous choices of D investors, we are only able to prove the existence and uniqueness of a linear symmetric equilibrium when D investors do not pay the participation cost $k_{Dm}$. However, since all endogenous choices are continuous with respect to $k_{Dm}$ due to the continuity of the functions involved, our primary qualitative findings remain valid at least for small values of $k_{Dm}$.

Equation (15) represents the first-order conditions for A and DA investors regarding their choice of precisions in both the index and non-index markets. Equation (16) is the first-order condition for DI investors when choosing the precision in the index market.

Intuitively, as trading in the index market becomes less costly, more savers choose to invest in the index. Similarly, when the participation cost for the non-index market rises, more discretionary investors tend to invest exclusively in the index. We outline these findings in the subsequent proposition.

**Proposition 1** *1. As the participation cost $k_m$ for the index market decreases, the proportion of savers opting to invest in the index ($\lambda_{SI}$) increases.*

*2. As the participation cost $k_s$ for the non-index market increases, the proportion of discretionary traders choosing to invest exclusively in the index ($\lambda_{DI}$) also rises.*

# 4 Model Implications

Due to the complexity arising from endogenous information acquisition, analytical solutions to the information acquisition problems are unlikely available. In this section, we undertake a numerical analysis to understand the equilibrium effects associated with the rise of indexing. We primarily explore three potential causes behind the surge in index investing:

1. A decrease in the participation cost, $k_m$, in the index market;
2. A decrease in liquidity trading in the non-index market measured by $\tau_{zs}^{-1/2}$, which is proportional to the expected liquidity trading volume; and
3. An increase in liquidity trading in the index market measured by $\tau_{zm}^{-1/2}$.[8]

The relationship between the rise of indexing and key variables such as price informativeness, market risk premium, and welfare depends critically on the driving forces behind this rise.

[8] For $j \in \{m, s\}$, the expected liquidity trading volume in market $j$ is equal to $\mathrm{E}\left|\int_i e_{ji} di\right| = \mathrm{E}\left|\int_i (Z_j + u_{ji}) di\right| = \sqrt{\frac{2}{\pi}}\,\tau_{zj}^{-1/2}$.

In the subsequent analysis, we focus on the negative information acquisition externality case because it captures the effect of limited information processing capacity which has better empirical support. We assume that S and D investors have the same participation costs in the index market, i.e., $k_{Dm} = k_m$. We use the following default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_{Dm} = k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

We should note that we do not attempt to calibrate our model to match empirical data. The parameters are selected to ensure that slight adjustments of certain parameter values can illustrate all main qualitative findings.

For clarity, in this numerical section, we assume identical information cost parameters for both A and D investors. The only distinction between the two is the market participation costs. Thus, A and DA investors always choose the same optimal information precisions in both markets.[9]

## 4.1 Driving Forces behind the Rise of Indexing

As a result of the growing number of ETFs and low-cost index funds, coupled with intensified competition among them, the cost of investing in an index has likely decreased. This reduction in cost could be a driving force behind the increasing popularity of indexing. As illustrated in the top-left panel of Figure 1, as the participation cost of the index market ($k_m$) decreases, the equilibrium mass of indexers ($\lambda_I$) increases. As depicted in Figure 1, the rise of indexing can be attributed to a rise in savers, discretionary investors turning to indexing, or a combination of both. As the indexing cost $k_m$ reduces, more savers transition to indexing from solely investing in risk-free assets, and more discretionary investors do the same, moving away from merely holding the initial endowment.

Another potential driving force behind the rise of indexing could be changes in liquidity

[9] We also conduct similar analysis when A and D investors have different information cost parameter values. The main qualitative results are the same.

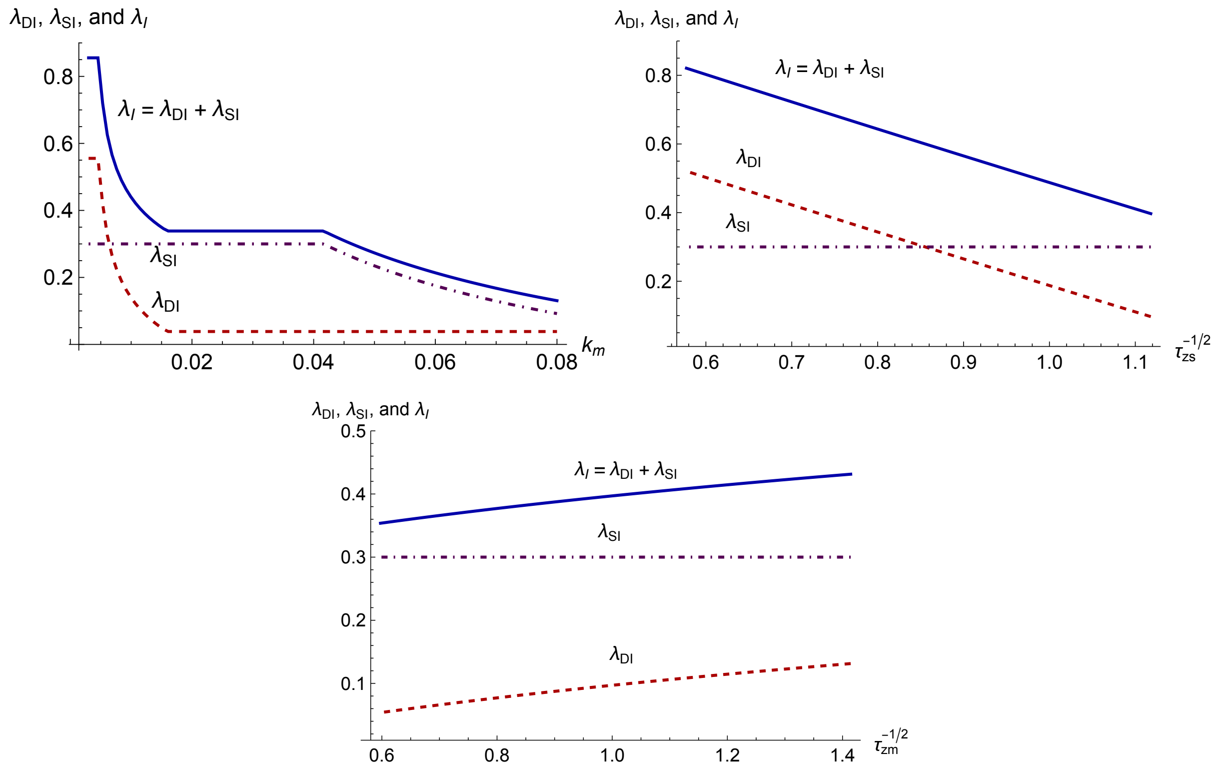


**Figure 1.** The equilibrium mass of endogenous indexers. Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

trading. When the average liquidity trading volume in the non-index market (measured by $\tau_{zs}^{-1/2}$) decreases, trading in the non-index market becomes less profitable for informed investors, *ceteris paribus*. Essentially, profiting from private information speculation becomes more challenging in the absence of liquidity traders who transact based on non-fundamental reasons. Consequently, more discretionary traders may stay away from the non-index market to save the participation cost $k_s$. The top right panel of Figure 1 confirms this intuition. As $\tau_{zs}^{-1/2}$ decreases, the average liquidity trading volume in the non-index market decreases and as a result $\lambda_{DI}$ increases.

Likewise, if the liquidity trading (measured by $\tau_{zm}^{-1/2}$) in the index market increases, the index market's profitability for discretionary traders grows, leading more discretionary traders to gravitate towards indexing.

In summary, Figure 1 indicates that a rise in indexing can result from a reduction in the index market participation cost, a decline in liquidity trading in the non-index market, or a rise in liquidity trading in the index market. Next, we show that increases in indexing due to different causes can lead to different impacts on key economic variables.

## 4.2 The Effects on Price Informativeness

We begin with analyzing the relationship between the rise of indexing and price informativeness. The top-left panel of Figure 2 suggests that when the increase in indexing is caused by a reduction in the index market participation cost $k_m$, the price informativeness in both the index and non-index markets does not change initially. This is because, as $k_m$ drops, the initial increase in indexing is due to an increase in the savers who switch to indexers. Savers do not acquire private information and the amount of information acquired by A investors, DA, and DI investors remains unchanged (see the top-left panel of Figure 3). Therefore, the initial rise in indexing does not affect the price informativeness for either the index ($\rho_m^*$) or the non-index ($\rho_s^*$) markets.[10] This is in line with some empirical findings (e.g., Coles et al. (2022)).

However, when $k_m$ reduces substantially, some discretionary traders who just hold their initial endowment (i.e., D0 investors) opt to become indexers. These new discretionary indexers then choose to acquire private information, which, in turn, increases the overall price informativeness in the index market. This occurs despite the fact that DI and DA investors acquire less precise private information about the index market (see the top-left panel of Figure 3) due to the decline of the marginal value of acquiring private information. Furthermore, because DA investors acquire less precise information in the index market, their information acquisition cost in the non-index market decreases. As a result, they acquire more precise private information about the non-index market (see the top-right panel of Figure 3), leading to a rise in the price informativeness of the non-index market.

[10] In all the graphical illustrations presented in this paper, we represent key economic variables as functions of $k_m$, $\tau_{zs}$, or $\tau_{zm}$. This is to highlight the effects of indexing stemming from various drivers: a decrease in the index market participation cost ($k_m$), a drop in liquidity trading in the non-index market ($\tau_{zs}^{-1/2}$), or an increase in liquidity trading in the index market ($\tau_{zm}^{-1/2}$).

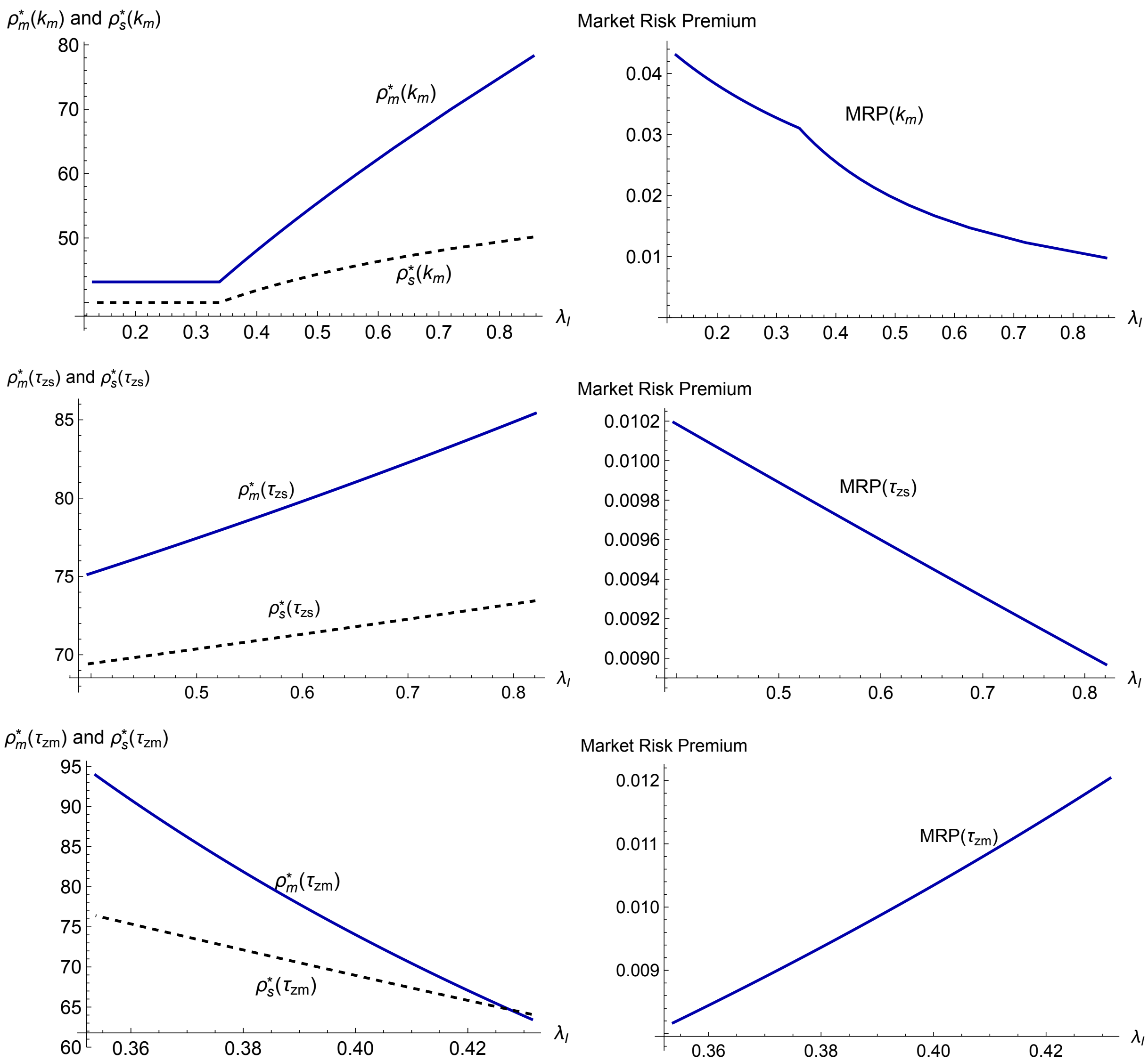


**Figure 2.** The price informativeness and market risk premium. Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

When the rise of indexing is driven by a decrease in the liquidity trading in the non-index market, as shown in the middle-left panel of Figure 2, the price informativeness in both markets increases. Intuitively, as $1/\tau_{zs}$ decreases, the non-index market's profitability decreases, and thus more discretionary investors become indexers to save the non-index market participation cost $k_s$. As shown in the middle-left panel of Figure 3, DI investors acquire more precise private information about the index than DA traders for any given $\tau_{zs}$, because DA investors acquire information in both index and non-index markets and there is negative information acquisition externality. The influx of DI traders who acquire more precise private information in the index market increases its price informativeness. Concurrently, as more discretionary traders migrate away from the non-index market, the marginal value of information acquisition in the non-index market rises. Consequently, as depicted in the middle-right panel of Figure 3, DA traders tend to acquire more precise signals about the non-index market, leading to enhanced price informativeness of the non-index market.

In contrast, as shown in the bottom-left panel of Figure 2, when the rise of indexing is driven by an influx of liquidity traders into the index market, the price informativeness for both markets decreases. As the volume of liquidity trading in the index market increases, the profitability in the index market increases and more discretionary traders become indexers. The bottom-left panel of Figure 3 indicates that for any given $\tau_{zm}$, DI traders obtain more precise private information about the index than DA traders. Furthermore, as $\tau_{zm}$ drops, liquidity trading volume increases on average, the precision of private information for both DI and DA traders also increases due to the elevated marginal value of information acquisition. Despite this, the index's price informativeness drops. This is because the surge in the liquidity trading reduces the price informativeness and this force dominates. Given that DA investors acquire more precise information in the index market, which increases their information acquisition cost in the non-index market, DA traders choose to gather less precise information about the non-index market. This trend, depicted in the bottom-right panel of Figure 3, leads to a decline in the price informativeness of the non-index market.

The right panels of Figure 2 show that as the price informativeness of the index increases, the market risk premium declines, aligning with equation (12). There exists an inverse relationship between the market risk premium and the index's price informativeness, $\rho_m$. This is because, a rise in price informativeness corresponds to decreased aggregate uncertainty in the market.

Recall that when the growth in indexing is due to a reduction in the index market participation cost, the price informativeness initially remains unchanged. This is because the initial growth is driven solely by savers transitioning to indexers. Despite this, the market risk premium still experiences a decline within this range. The reason is that more savers are now sharing the same level of aggregate risk within the index market.

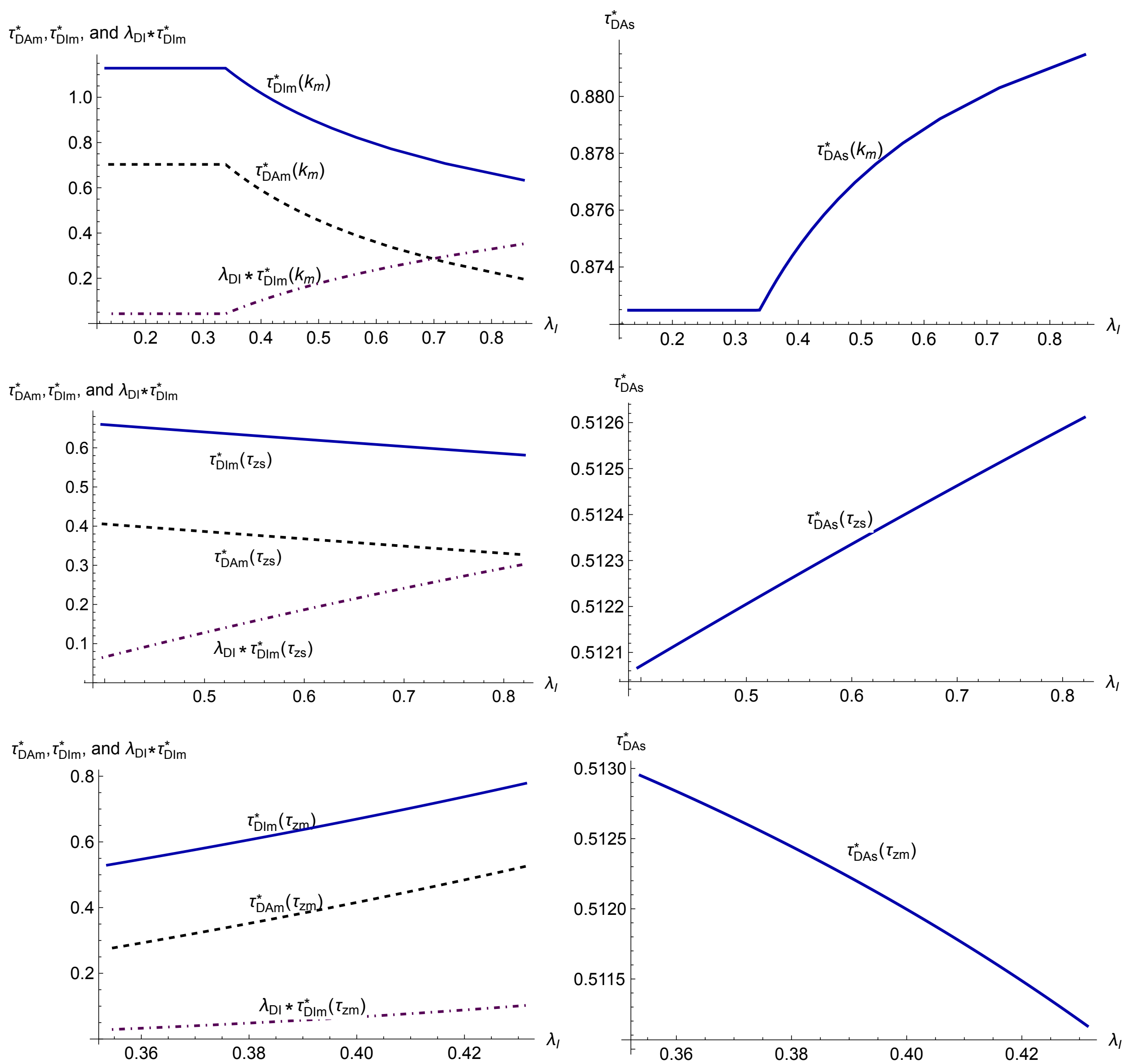


**Figure 3.** The precision of traders' private information against the rise in indexing. Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

When the rise of indexing is caused by a decline in the participation cost in the index market or a decline in liquidity trading in the non-index market, our findings regarding the price informativeness and market risk premium are opposite to the conclusions drawn by Bond and Garcia (2022). They find that an increase in indexing leads to a decline in price informativeness and a surge in the market risk premium. They focus on the rise of indexing attributed to a reduced index participation cost. In their model, a reduction in the index participation cost attracts more traders who are less informed and have endowment shocks, thereby introducing greater noise into the index market. Different from our model, information acquisition is exogenous in Bond and Garcia (2022) and thus switching to be an indexer does not change information acquisition. In contrast, in our model, when discretionary investors switch to indexing—prompted either by a lowered index participation cost or by diminished liquidity trading in the non-index market—they tend to acquire more precise information compared to their active trader counterparts and thus the price informativeness may increase with indexing. In addition, in our model, a decrease in the index participation cost may attract more savers to become indexers and thus increase risk sharing in the index market and subsequently decrease market risk premium even when the price informativeness of the index market remains unchanged.

A common concern regarding the rapid growth of index investment is the potential for indexers to "free-ride" on the information acquisition efforts of others, leading to a decline in market efficiency. However, Figure 3 suggests the opposite. In fact, certain indexers may gather information with greater precision compared to their active counterparts, leading to a scenario where active investors might actually "free-ride" on some indexers. We formalize this observation with the following proposition:

**Proposition 2** *Assuming the information acquisition externality for D investors is negative (i.e., $\varphi_D < 0$), then $\tau_{DAm} < \tau_{DIm}$.*

Proposition 2 highlights that discretionary indexers can achieve greater precision in information acquisition. This is attributable to the negative externality in information acquisition: active investors, who participate in and source information from both markets, face heightened costs for gathering information in the index market. This leads to a reduction in information acquisition by these active investors.

We proceed to assess the influence of indexing on other crucial economic indicators, including market capitalization, stock price correlations, index return reversal, and the welfare of market participants. These effects are inherently linked to its impact on price informativeness.

### 4.3 The Effects on Market Capitalization

Given equation (1), the implied prices of Stocks 1 and 2, $P_1$ and $P_2$ respectively, can be expressed as:

$$P_1 = P_m + P_s, \quad P_2 = P_m - P_s. \tag{17}$$

The expected market capitalizations of Stocks 1 and 2 are

$$\begin{aligned} EMC_1 &:= \mathrm{E}[P_1] = \mu_{vm} + \mu_{vs} - \mathrm{E}[V_m - P_m], \\ EMC_2 &:= \mathrm{E}[P_2] = \mu_{vm} - \mu_{vs} - \mathrm{E}[V_m - P_m], \end{aligned} \tag{18}$$

where $\mathrm{E}[V_m - P_m]$ is the market risk premium as shown in (12). Equation (18) suggests that the expected capitalizations for both stocks decrease with the market risk premium, which decreases in the price informativeness $\rho_m$. It's noteworthy that the expected capitalization is not influenced by the price informativeness in the non-index market. This is because the total supply of the non-index portfolio is zero, which means that the total capital allocated to the non-index portfolio is also zero.

Because the rise in indexing impacts the market risk premium, it also influences the market capitalization of Stocks 1 and 2. The top two panels of Figure 4 depict that as indexing grows, due to either a reduced participation cost in the index market or a decline in liquidity trading in the non-index market, there's a larger average capital allocation to stocks. A reduction in the index participation cost encourages either savers or discretionary traders to become indexers, infusing more capital into the stock market. Similarly, diminished liquidity trading in the non-index market, resulting in lower profitability. This reduction in profitability prompts more discretionary investors to shift to the index market. They acquire more precise private information compared to their active trader counterparts and invest more extensively in the index

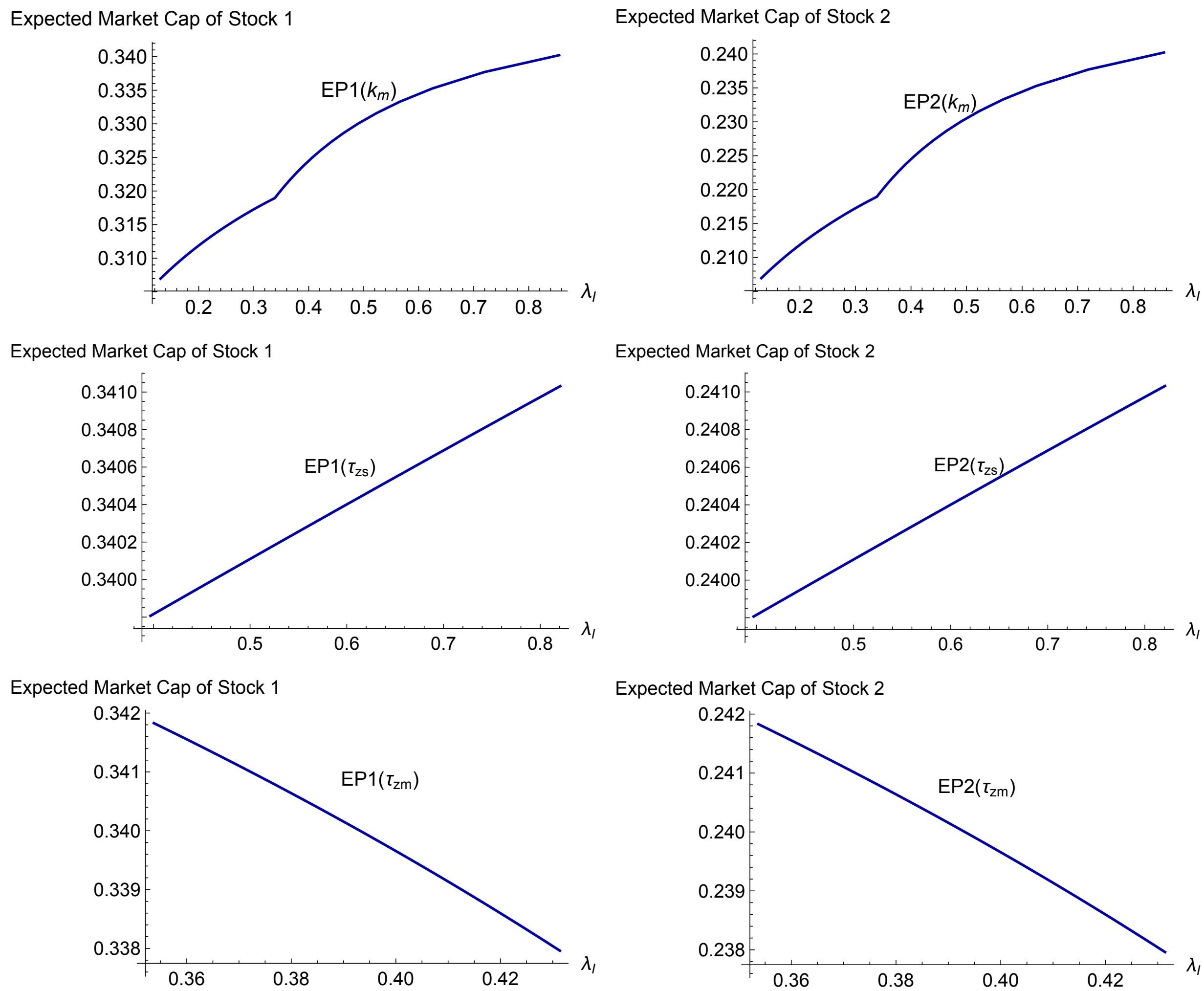


**Figure 4.** The equilibrium expected market capitalizations against the proportion of indexers $\lambda_I$. Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

market, thus bringing additional capital into the stock market.

Conversely, if the increase in indexing stems from a boost in liquidity trading in the index market, attracting certain discretionary traders to exclusively invest there, the reduced price informativeness of this market means that risk-averse investors demand a higher risk premium and a lower stock price to participate. Consequently, the expected capitalization of both stocks decreases, as shown in the bottom panels of Figure 4.

### 4.4 The Effects on Comovement

From equation (17), we can calculate the correlation and variance of stock prices $P_1$ and $P_2$ as

$$\begin{aligned} \mathrm{Var}(P_1) = \mathrm{Var}(P_2) &= \frac{b_m^2}{\tau_{vm}} + \frac{b_s^2}{\tau_{vs}} + \frac{d_m^2}{\tau_{zm}} + \frac{d_s^2}{\tau_{zs}}, \\ \mathrm{Corr}(P_1, P_2) &= \frac{b_m^2/\tau_{vm} - b_s^2/\tau_{vs} + d_m^2/\tau_{zm} - d_s^2/\tau_{zs}}{b_m^2/\tau_{vm} + b_s^2/\tau_{vs} + d_m^2/\tau_{zm} + d_s^2/\tau_{zs}}. \end{aligned} \tag{19}$$

The rise of indexing has sparked concerns about a potential increase in the comovement of asset prices, given that more investors are trading stocks bundled within the same index. This aligns with the illustration of the top two panels of Figure 5.

If the growth in indexing results from either a decrease in the index market's participation cost or a decline in liquidity trading in the non-index market, the correlation between stock prices, $P_1$ and $P_2$, tends to increase. Lower index participation costs draw in both savers and discretionary traders to adopt indexing. As discussed previously, with an increasing number of discretionary traders becoming indexers, price informativeness across both index and non-index markets increases. As a result, the prices of Stocks 1 and 2 are more attuned to their payoffs. We show numerically that the coefficients on $V_m$ and $V_s$ in prices $P_m$ and $P_s$ (i.e., $b_m$ and $b_s$) increase with $\lambda_I$ under these circumstances (Figure A-1 in the Appendix), leading to an increased correlation and variance in stock prices as the indexer proportion rises.

In contrast, if the indexing surge is attributed to amplified liquidity trading in the index market, which causes more discretionary traders to transition to indexing, as discussed previously, price informativeness of both index and non-index markets decreases. In this case, both the correlation and variance of stock prices tend to decrease as the indexer proportion increases, as depicted in the bottom panels of Figure 5. The rationale is that heightened exogenous liquidity trading in the index market makes prices less responsive to asset payoffs. Numerically, we show that the coefficients $b_m$ and $b_s$ for $V_m$ and $V_s$ in prices $P_m$ and $P_s$ respectively, decrease in $\lambda_I$ when the rise in indexing is driven by increased liquidity trading in the index market. Thus, both correlation and variance may decrease as indexing rises.

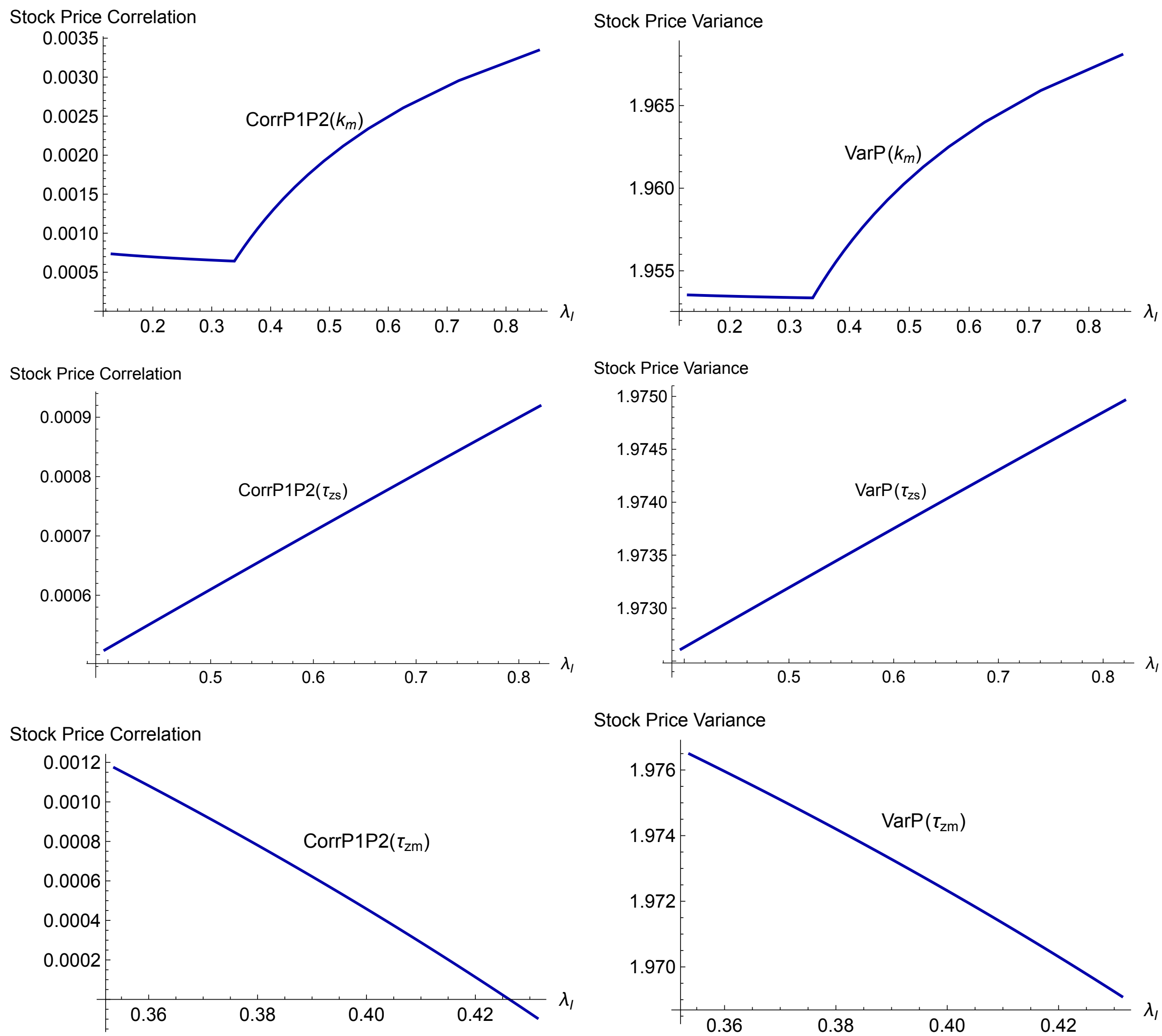


**Figure 5.** Price correlations, and stock variances against indexers mass $\lambda_I$. Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

## 4.5 The Effects on Index Level Reversal

Following Bond and Garcia (2022), we define the relative price efficiency of stocks 1 and 2 as

$$RPE := \text{Var}[V_1 - V_2 | P_1 - P_2]^{-1} = \text{Var}[2V_s | P_s]^{-1} = \tfrac{1}{4}\rho_s. \tag{20}$$

As in REE models, the index exhibits a return reversal, specifically $\text{E}[V_m - P_m | P_m]$ decreases in $P_m$. Following Bond and Garcia (2022), we measure the strength of reversal by the steepness of the negative slope of $\mathrm{E}[V_m - P_m|P_m]$ with respect to $P_m$. Using Projection Theorem,

$$
\begin{aligned}
IR := \frac{\partial \mathrm{E}[V_m - P_m|P_m]}{\partial P_m} &= \frac{\mathrm{Cov}[V_m - P_m, P_m]}{\mathrm{Var}[P_m]} \\
&= -\frac{\gamma\lambda_L \tau_{vm}}{\left(\tau_{vm} + \frac{\tau_m^2}{\gamma^2\lambda_L^2}\tau_{zm}\right)\left(\gamma\lambda_L + \frac{\tau_m}{\gamma\lambda_L}(1-\lambda_{D0}+\lambda_{SI})\tau_{zm}\right)}.
\end{aligned}
\tag{21}
$$

Equation (21) suggests that a lower total precision $\tau_m$ of private information leads to a larger magnitude of reversal. This is intuitive: with lower information precision, the price informativeness of the index decreases, indicating more noise in the index market. This pushes the price further away from the fundamental value of the index at time 0, but at time 1, it must align with the index's fundamental value $V_m$. Consequently, the extent of the return reversal becomes more pronounced.

Bond and Garcia (2022) find that as indexing increases, both the relative price efficiency and index price reversal tend to increase. Our analysis, however, suggests that the effects of the rise in indexing on both relative price efficiency and index price reversal hinge on the specific causes behind this increase. Specifically, the relative price efficiency—proportional to the price informativeness of the non-index market— tends to increase when the rise of indexing is attributed to a reduction in the index market's participation cost or a reduction in liquidity trading in the non-index market. In contrast, it tends to decrease if the rise of indexing is due to increased liquidity trading in the index market, as discussed in Section 4.2.

The magnitude of index reversal is inversely related to the total precision of private information in the index market and, consequently, decreases with the price informativeness of the index market. Therefore, the magnitude of price reversal (represented by the absolute value of the negative slope) decreases when the rise of indexing results from a reduced participation cost in the index market or from declining liquidity trading in the non-index market. On the other hand, it increases when the rise of indexing is driven by increased liquidity trading in the index market. These results are consistent with the numerical illustrations in Figure 6.

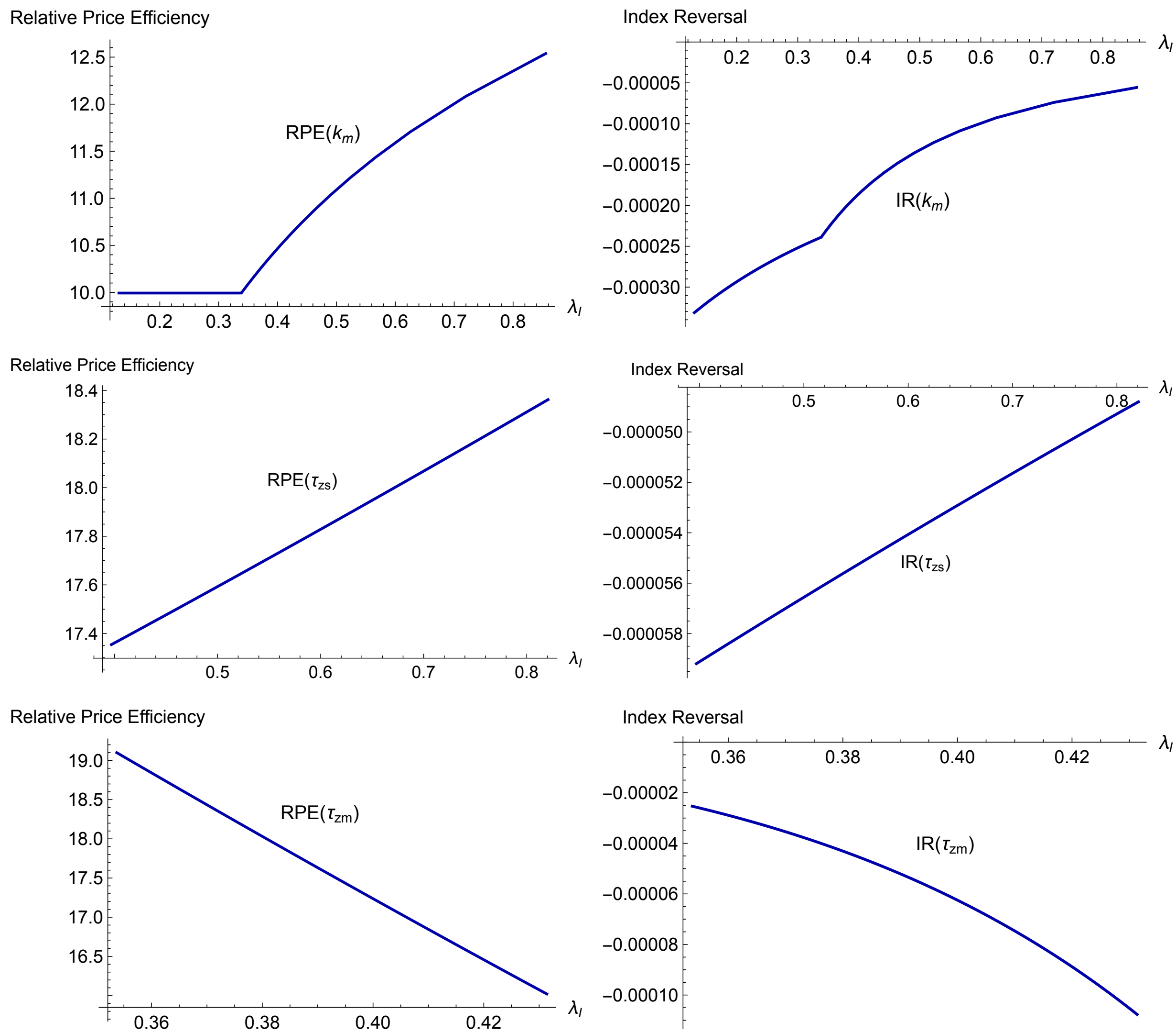


**Figure 6.** The relative price efficiency and index level reversals against indexers mass $\lambda_I$. Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

## 4.6 The Effects on Welfare

The influence of the rise of indexing on the welfare of non-liquidity traders also depends on the underlying causes of this rise. As Figure 7 indicates, liquidity traders' welfare increases but non-liquidity traders' welfare decreases if the growth in indexing arises from either a decrease in the participation cost of the index market or a reduction in liquidity trading in the non-index market. In contrast, the results reverse if increased liquidity trading in the index market is the primary driver behind the rise of indexing.

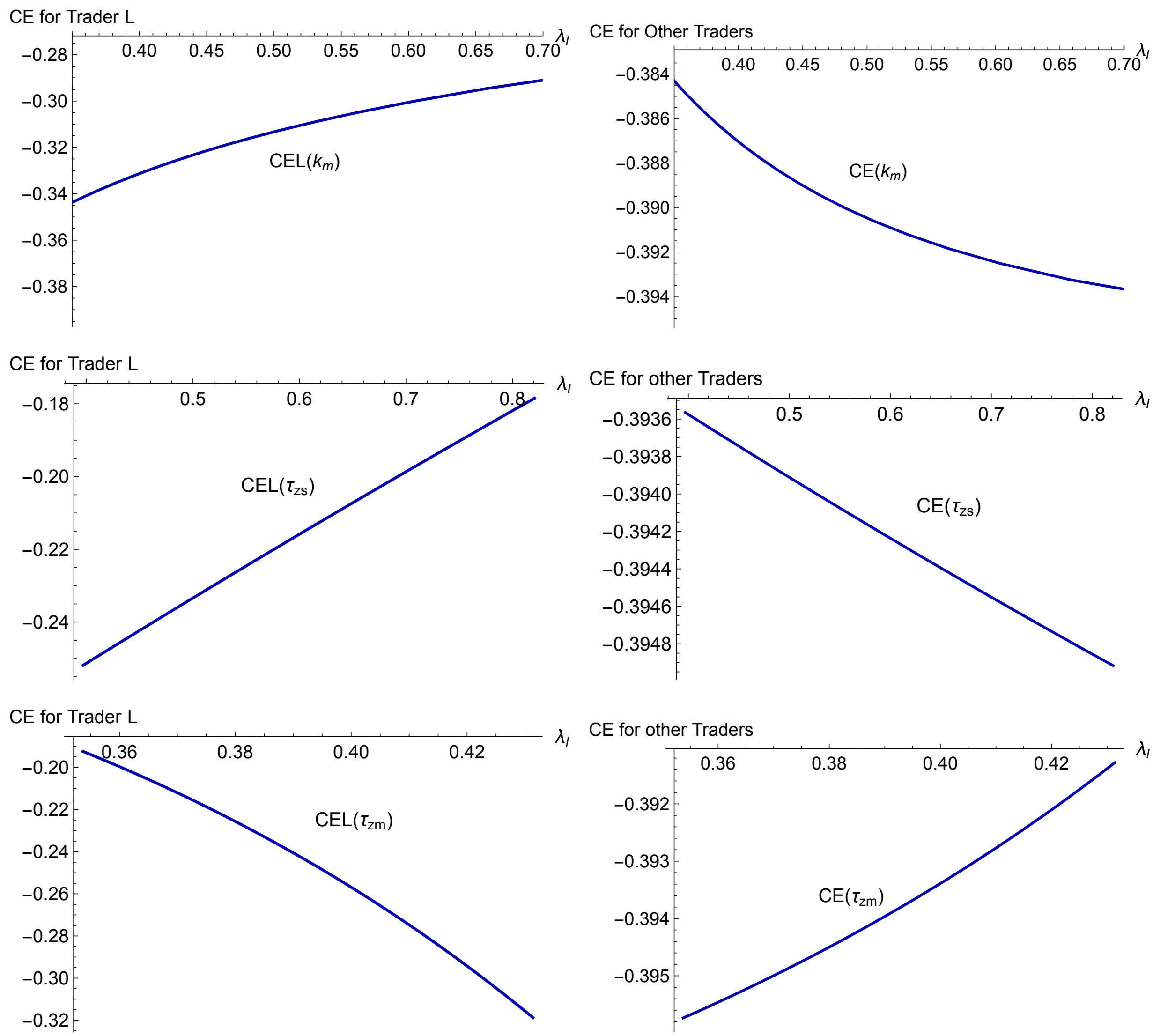


**Figure 7.** The welfare of liquidity traders and non-liquidity traders against indexers mass $\lambda_I$. Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

As the index participation cost decreases, more savers start to invest in the index market. Consistent with Bond and Garcia (2022), we find liquidity traders benefit from reduced costs and consequently achieve greater welfare. On the other hand, the welfare of active and discretionary investors decrease, primarily because enhanced risk sharing in the index market leads to a reduced market risk premium, as illustrated in Figure 7. The top-left panel of Figure 8 suggests that the welfare loss of non-liquidity traders can dominate the welfare gain of the liquidity traders so that the social welfare decreases.

When the rise of indexing is attributed to a decline in liquidity trading in the non-index market, similar to the above case, the middle panels of Figure 7 suggest that the welfare of liquidity traders also increases and that of non-liquidity traders also decreases. However, a notable point captured by the top-right section of Figure 8 is that, while the welfare drop for active and discretionary traders is relatively small, liquidity traders experience large gains, due to increased price informativeness in both markets, resulting in prices that better reflect underlying values. As a result, the social welfare increases.

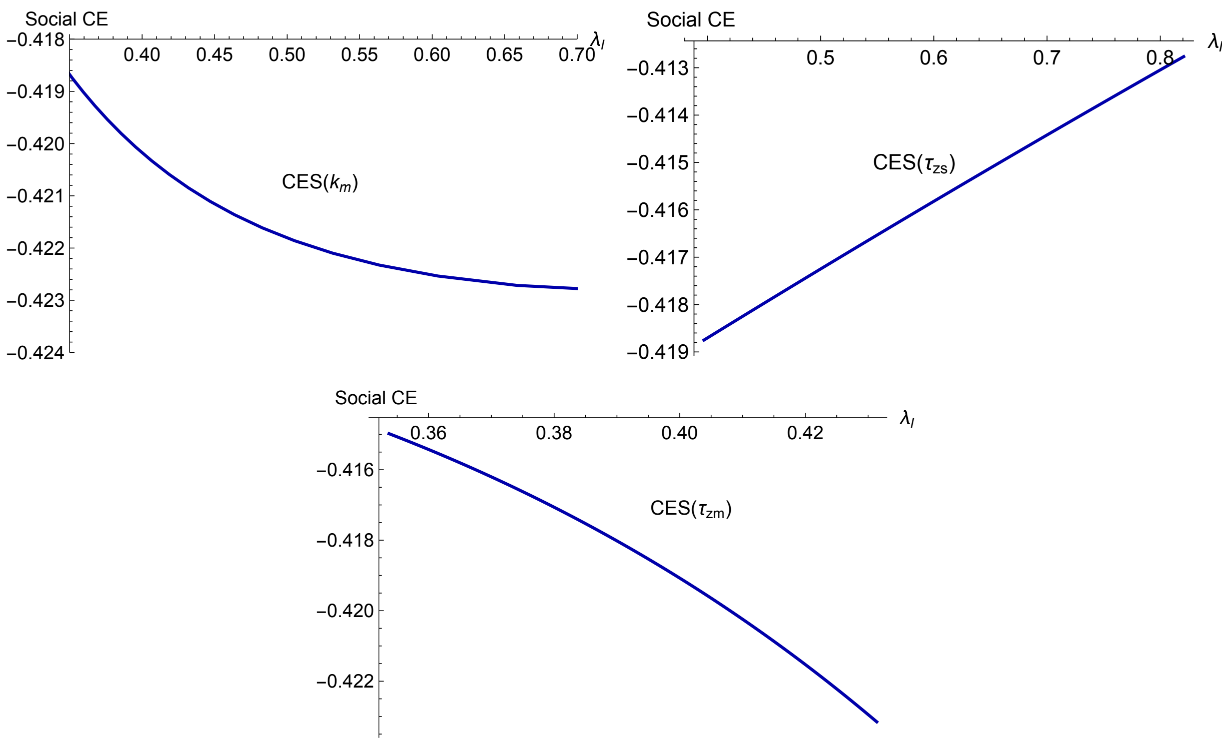


**Figure 8.** The social welfare against indexers mass $\lambda_I$. Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

On the other hand, if amplified liquidity trading in the index market propels the rise in indexing, the resulting decline in price informativeness in both markets hurts liquidity traders but benefits non-liquidity traders. Non-liquidity traders capitalize on increased trading opportunities with liquidity traders, as illustrated in the lower sections of Figure 7. Consequently, the welfare of the liquidity traders can decrease due to a decline in price informativeness across both markets. In addition, the bottom section of Figure 8 shows that the overall welfare might decrease, as the benefits reaped by non-liquidity traders are overshadowed by the losses incurred by liquidity traders.

### 4.7 Net Effects

Our analysis highlights the importance of identifying the underlying driving force of the growth of index investing in understanding its implications. As shown above, the rise of indexing can be attributed to reduced liquidity trading in the non-index market and increased liquidity trading in the index market, which often produce opposite effects. As an illustrative example, we examine the net effects of increased liquidity trading in the index market and reduced liquidity trading in the non-index market.

As previously discussed, the proportion of indexing, denoted by $\lambda_I$, increases either due to enhanced liquidity trading in the index market (i.e., a higher value of $\tau_{zm}^{-1/2}$) or due to reduced liquidity trading in the non-index market (i.e., a lower value of $\tau_{zs}^{-1/2}$). Figures 9 and 10 illustrate the consequences of these changes in indexing on several measures of interest: the price informativeness of the index, the market risk premium, the expected market capitalization, the correlation and variance of stock prices, index return reversal, the welfare of liquidity and non-liquidity traders, as well as the overall social welfare, under scenarios where liquidity trading increases in the index market and decreases in the non-index market.[11]

When there is decreased liquidity trading in the non-index market (i.e., a decrease in $\tau_{zs}^{-1/2}$), we observe the following outcomes: an increase in price informativeness; higher variances and correlations among stock prices; an increased expected market capitalization; a reduced magnitude of index return reversal; a decrease in the welfare of non-liquidity traders; an increase in the welfare of liquidity traders; and an increase in overall social welfare.

[11] Initially, we set $\tau_{zm}^{-1/2} = \tau_{zs}^{-1/2} = 1$. For our analysis, we examine the effects of variations where $\tau_{zm}^{-1/2}$ is increased (i.e., $\tau_{zm}^{-1/2} \geq 1$) to represent higher liquidity trading in the index market, and $\tau_{zs}^{-1/2}$ is decreased (i.e., $\tau_{zs}^{-1/2} \leq 1$) to indicate lower liquidity trading in the non-index market.

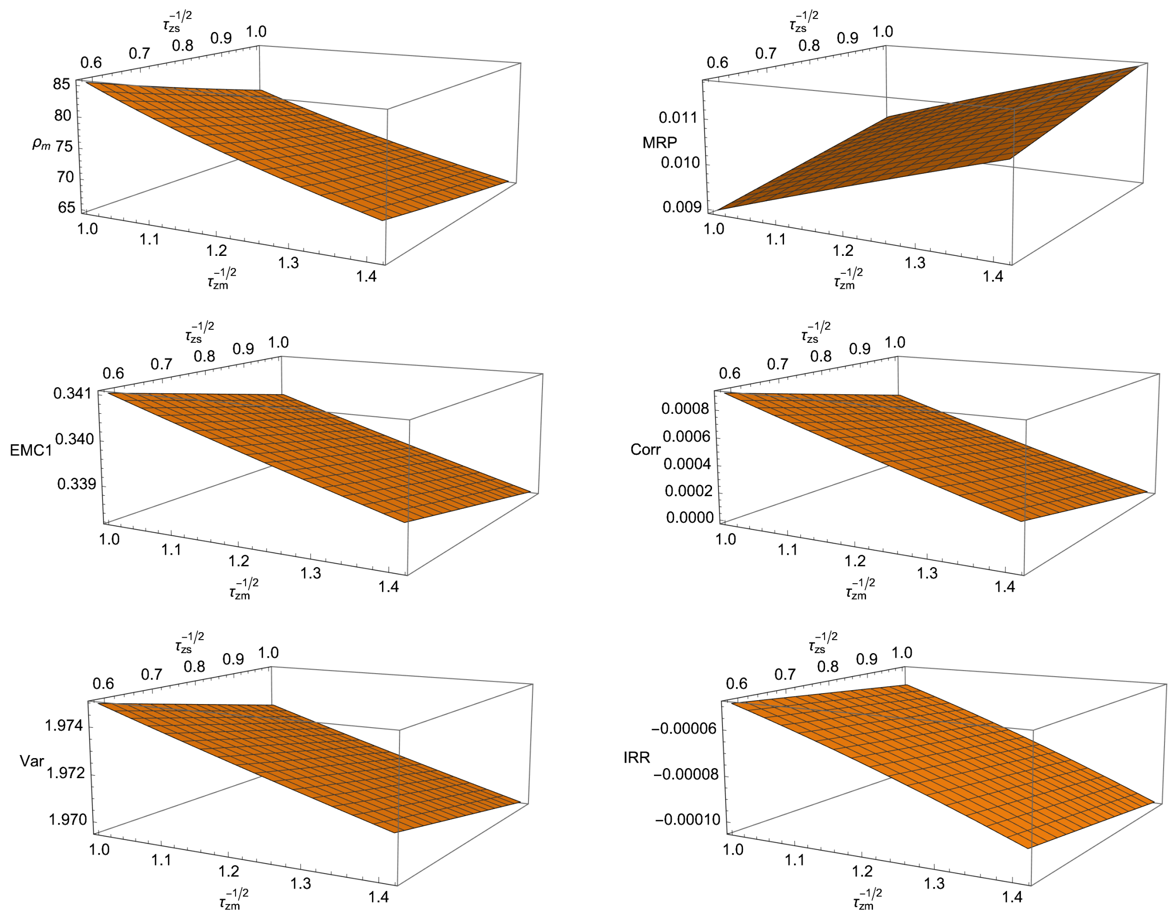


**Figure 9.** The net effects on the price informativeness of the index (top left), the market risk premium (top right), the expected market capitalization (middle left), the correlation and variance of prices (middle right and bottom left, respectively), and the index return reversal (bottom right). Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

Conversely, with increased liquidity trading in the index market (i.e., an increase in $\tau_{zm}^{-1/2}$), the opposite effects are noted: price informativeness decreases; variances and correlations among stock prices reduce; the expected market capitalization decreases; the magnitude of index return reversal increases; and there is an increase in the welfare of non-liquidity traders, but a decrease in the welfare of liquidity traders and in social welfare.

As depicted in Figures 9 and 10, when multiple factors are influencing the market, the net effects of indexing may become less pronounced, potentially canceling each other out.

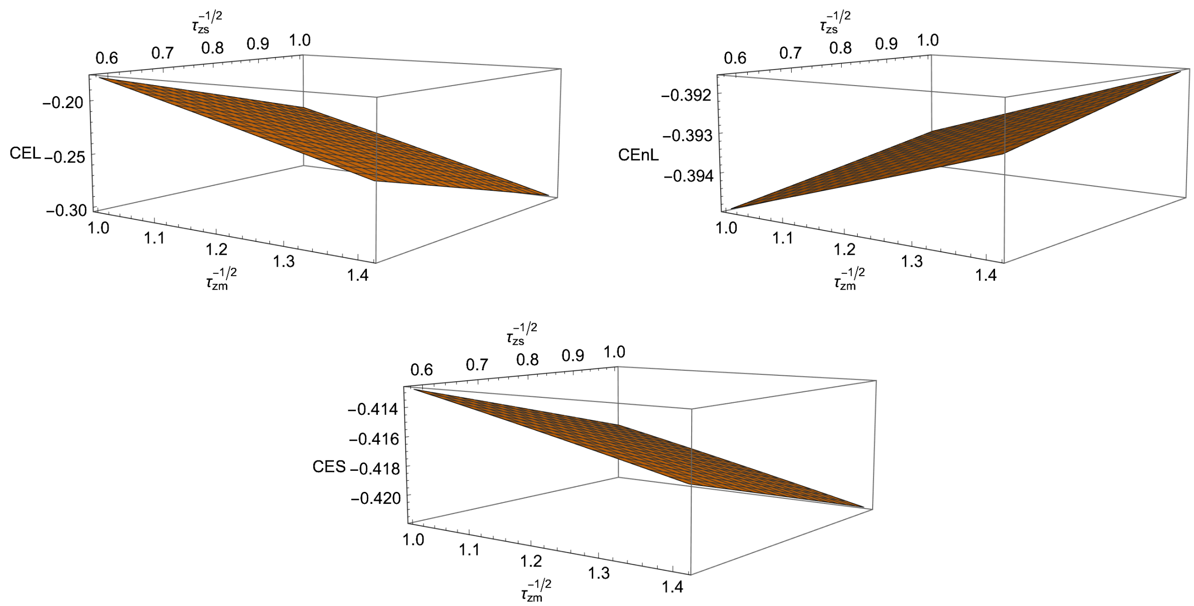


**Figure 10.** The net effects on welfare of liquidity traders (top left), welfare of non-liquidity traders (top right), and social welfare (bottom). Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

# 5 Conclusion

We use a rational expectations equilibrium model with endogenous information acquisition to study the consequences of indexing. We show that the effects of index investing critically depend on the causes of the rise of indexing. Specifically, if the rise of indexing results from either reduced participation costs or a decline in liquidity trading in the non-index market, then (1) Price informativeness in both the index and non-index markets increases; (2) Stock correlations intensify; (3) The expected market capitalization grows; (4) The magnitude of index return reversal decreases; (5) The welfare of non-liquidity traders decreases. In contrast, if the rise of indexing is driven by increased liquidity trading in the index market, the opposite effects take place.

Our paper underscores the importance of understanding the underlying factors driving the rise of index investing. By examining the relationship between the growth in index investment and changes in economic measures such as price informativeness, market capitalizations, and stock correlations, researchers can infer potential reasons behind this rise in indexing. If there is an observed increase in price informativeness concurrent with a rise in indexing, it may indicate that factors such as reduced index participation costs or a decline in liquidity trading in the non-index market are at play. On the other hand, a decrease in price informativeness might suggest an increase in liquidity trading within the index market. If price informativeness remains unchanged, it suggests that multiple causes may be at play in practice, the net effects of indexing might be insignificant. Our research thus emphasizes the importance of identifying the main drivers behind the rise of index investing to fully understand its broader consequences.

## References


Admati, A. R. 1985. A Noisy Rational Expectations Equilibrium for Multi-asset Securities Markets. *Econometrica* 53:629–657.

Antoniou, C., F. W. Li, X. Liu, A. Subrahmanyam, and C. Sun. 2023. Exchange-Traded Funds and Real Investment. *Review of Financial Studies* 36:1043–1093.

Baruch, S., and X. Zhang. 2022. The Distortion in Prices Due to Passive Investing. *Management Science* 68:6219–6234.

Bond, P., and D. Garcia. 2022. The Equilibrium Consequences of Indexing. *Review of Financial Studies* 35:3175–3230.

Breugem, M., and A. Buss. 2018. Institutional Investors and Information Acquisition: Implications for Asset Prices and Informational Efficiency. *Review of Financial Studies* 32:2260–2301.

Brogaard, J., M. Ringgenberg, and D. Sovich. 2019. The Economic Impact of Index Investing. *Review of Financial Studies* 32:3461–99.

Chen, W., B. Hu, and Y. Wang. 2022. How Does Benchmarking Affect Market Efficiency? — The Role of Learning Technology. Working paper.

Coles, J. L., D. Heath, and M. C. Ringgenberg. 2022. On Index Investing. *Journal of Financial Economics* 145:665–683.

Diamond, D. W., and R. E. Verrecchia. 1981. Information Aggregation in a Noisy Rational Expectations Economy. *Journal of Financial Economics* 9:221–235.

Farboodi, M., A. Matray, V. Venkateswaran, and L. Veldkamp. 2022. The Economic Impact of Index Investing. *Review of Financial Studies* 35:3101–3138.

Ganguli, J., and L. Yang. 2009. Complementarities, Multiplicity, and Supply Information. *Journal of the European Economic Association* 7:90–115.

Glosten, L., S. Nallareddy, and Y. Zou. 2021. ETF Activity and Informational Efficiency of Underlying Securities. *Management Science* 67:22–47.

Grossman, S., and J. Stiglitz. 1980. On the Impossibility of Informationally Efficient Markets. *American Economic Review* 70:393–408.

Israeli, D., C. Lee, and S. Sridharan. 2017. Is There a Dark Side to Exchange Traded Funds? An Information Perspective. *Review of Accounting Studies* 22:1048–1083.

Van Nieuwerburgh, S., and L. Veldkamp. 2009. Information Immobility and the Home Bias Puzzle. *Journal of Finance* 64:1187–1215.

# Appendix A

**Proof of Theorem 1:**

The optimal number of shares for both the index and non-index portfolios for investor $i$ of type $t$ is:

$$\begin{aligned}\Theta_{tmi} &= \frac{\mathrm{E}[V_m|I_{ti}]-P_m}{\gamma \mathrm{Var}[V_m|I_{ti}]} - s_m 1_{\{t=A,D\}}, \text{ for } t = A, DA, DI, SI, \\ \Theta_{tsi} &= \frac{\mathrm{E}[V_s|I_{ti}]-P_s}{\gamma \mathrm{Var}[V_s|I_{ti}]} - s_s, \text{ for } t = A, DA,\end{aligned} \tag{A-1}$$

where $s_m = 2$ and $s_s = 0$. The information set of investor $i$ of type $A, DA$, $DI$, and $SI$ is $I_{ti} = (Y_{tmi}, Y_{tsi}, P_m, P_s)$, where the precision of $Y_{tmi}$ and $Y_{tsi}$ is zero for $SI$ and $L$ investors, and the precision of $Y_{tsi}$ is zero for $DI$ investors. Because $Y_{tmi}$ and $Y_{tsi}$ are independent, and $V_m$ and $V_s$ are independent, the conditional expectation of $V_j$ only depends on $(Y_{tji}, P_j)$. Direct computation yields that for $t = A, D$, and $j = m, s$,

$$\mathrm{E}[V_j|I_{ti}] = \frac{\mu_{vj}\tau_{vj} + Y_{tji}\tau_{tji} + \frac{b_j}{d_j^2}(-a_j + P_j)\tau_{zj}}{\tau_{vj} + \tau_{tji} + \frac{b_j^2}{d_j^2}\tau_{zj}}, \quad \mathrm{Var}[V_j|I_{ti}] = \left(\tau_{vj} + \tau_{tji} + \frac{b_j^2}{d_j^2}\tau_{zj}\right)^{-1}. \tag{A-2}$$

Using the market clearing conditions

$$\sum_{t\in\{A,DI,DA,SI\}} \int_i \Theta_{tmi}\, di - \lambda_L Z_m = 0, \quad \sum_{t\in\{A,DA\}} \int_i \Theta_{tsi}\, di - \lambda_L Z_s = 0, \tag{A-3}$$

where the integration is over all investors of the same type. We obtain

$$P_m = \frac{(1+\lambda_{SI})\tau_{vm}\mu_m - 2\gamma + \left(\tau_m + \frac{b_m^2}{d_m^2}(1+\lambda_{SI})\tau_{zm}\right)V_m - \left(\gamma\lambda_L + \frac{b_m}{d_m}(1+\lambda_{SI})\tau_{zm}\right)Z_m}{\tau_m + \left(\tau_{vm} + \frac{b_m^2}{d_m^2}\tau_{zm}\right)(1+\lambda_{SI})}, \tag{A-4}$$

setting the coefficients of $V_m$ and $Z_m$ and the constant term to be $b_m$, $-d_m$, and $a_m$. We solve the coefficients $a_m$, $b_m$, and $d_m$ as

$$a_m = \frac{(\lambda_A + \lambda_{DA} + \lambda_{DI} + \lambda_{SI})\mu_{vm}\tau_{vm} - 2\gamma}{\tau_m + (\lambda_A + \lambda_{DA} + \lambda_{DI} + \lambda_{SI})\rho_m},$$
$$b_m = 1 - \frac{(\lambda_A + \lambda_{DA} + \lambda_{DI} + \lambda_{SI})\tau_{vm}}{\tau_m + (\lambda_A + \lambda_{DA} + \lambda_{DI} + \lambda_{SI})\rho_m}, \quad d_m = \frac{\gamma\lambda_L b_m}{\tau_m},$$

Similarly, we solve the coefficients $a_s, b_s$, and $d_s$ as presented in equation (11).

**Proof of Theorem 2:**

For type $t = A, D, SI$, trader $i$'s expected utility function is

$$\mathrm{E}\left[-\mathrm{e}^{-\gamma\left((2V_m-\beta)1_{\{t=A,D\}}+\frac{1}{\lambda_S}\beta 1_{\{t=S\}}+\Theta_{tmi}(V_m-P_m)+\Theta_{tsi}(V_s-P_s)-C_t(\tau_{tm},\tau_{ts})1_{\{t=A,D\}}-k_m 1_{\{t=SI\}}-k_s 1_{\{t=DA\}}\right)}\middle| I_{ti}\right], \tag{A-5}$$

substituting equation (5) into (A-5), equation (A-5) can be written as

$$\mathrm{E}\left[-\mathrm{e}^{-\gamma\left((2P_m-\beta)1_{\{t=A,D\}}+\frac{1}{2\gamma}\frac{(\mathrm{E}[V_m|I_{ti}]-P_m)^2}{\mathrm{Var}[V_m|I_{ti}]}+\frac{1}{2\gamma}\frac{(\mathrm{E}[V_s|I_{ti}]-P_s)^2}{\mathrm{Var}[V_s|I_{ti}]}1_{\{t=A,DA\}}-C_t(\tau_{tm},\tau_{ts})1_{\{t=A,D\}}\right)}\right]$$
$$\times\, \mathrm{e}^{-\gamma\left(\frac{1}{\lambda_S}\beta 1_{\{t=S\}}-k_m 1_{\{t=SI\}}-k_s 1_{\{t=DA\}}\right)}. \tag{A-6}$$

The problem of choosing precisions $(\tau_{tmi}, \tau_{tsi})$ to maximize the expected utility (A-6) is equivalent to solving

$$\max_{\tau_{tmi},\tau_{tsi}} \mathrm{E}\left[-\exp\left(-\frac{1}{2}\frac{(\mathrm{E}[V_m|I_{ti}]-P_m)^2}{\mathrm{Var}[V_m|I_{ti}]}-\frac{1}{2}\frac{(\mathrm{E}[V_s|I_{ti}]-P_s)^2}{\mathrm{Var}[V_s|I_{ti}]}1_{\{t=A,DA\}}+\gamma C_t(\tau_{tm},\tau_{ts})1_{\{t=A,D\}}\right)\right], \tag{A-7}$$

which is equivalent to solving

$$\max_{\tau_{tmi},\tau_{tsi}} -\sqrt{\frac{\mathrm{Var}[V_m|I_{ti}]}{\mathrm{Var}[V_m|P_m]}}\exp\left(-\frac{1}{2}\frac{(\mathrm{E}[V_m|P_m]-P_m)^2}{\mathrm{Var}[V_m|P_m]}-\frac{1}{2}\frac{(\mathrm{E}[V_s|P_s]-P_s)^2}{\mathrm{Var}[V_s|P_s]}1_{\{t=A,DA\}}+\gamma C_t(\tau_{tm},\tau_{ts})1_{\{t=A,D\}}\right). \tag{A-8}$$

Therefore, type $A$ and type $DA$ investors choose the precisions $(\tau_{tmi}, \tau_{tsi})$ to maximize

$$-\gamma C_t(\tau_{tmi},\tau_{tsi}) + \frac{1}{2}\log(\tau_{tmi}+\tau_{vm}+r_m^2\tau_{zm}) + \frac{1}{2}\log(\tau_{tsi}+\tau_{vs}+r_s^2\tau_{zs}), \quad t \in \{A, DA\}, \tag{A-9}$$

where $r_j := \frac{\tau_j}{\gamma\lambda_L}$ for $j = m, s$. Type $DI$ chooses the precision $\tau_{tmi}$ to maximize

$$-\gamma C_t(\tau_{tmi}, 0) + \frac{1}{2}\log(\tau_{tmi} + \tau_{vm} + r_m^2\tau_{zm}), \quad t \in \{DI\}. \tag{A-10}$$

Under Assumption 1, it can be easily verified that the objective functions are all globally strictly concave in the choice precision variables, and therefore given $r_m$ and $r_s$, there are unique solutions. Since investors of the same type choose the same precisions, we omit the index $i$. Note that savers (type $SI$) do not acquire information. Optimization yields the following first order equations,

$$\begin{aligned} 2\gamma C_{tj}(\tau_{tm}, \tau_{ts}) &= (\tau_{tj} + \rho_j)^{-1}, \quad t \in \{A, DA\}, j \in \{m, s\}, \\ 2\gamma C_{DIm}(\tau_{DIm}, 0) &= (\tau_{DIm} + \rho_m)^{-1}. \end{aligned} \tag{A-11}$$

Define the optimal precision functions as

$$\tau_{tj}^* = f_j^t(r_m, r_s), j \in \{m, s\}, \quad t \in \{A, DA, DI\}, \tag{A-12}$$

with $f_s^{DI}(r_m, r_s) = 0$ since $DI$ investors do not acquire information in the non-index market. Taking derivatives with respect to $r_m$ in the first order conditions (A-11) yields,

$$\frac{\partial f_m^t(r_m, r_s)}{\partial r_m} < 0, \quad t \in \{A, DA, DI\}. \tag{A-13}$$

Similarly,

$$\frac{\partial f_s^t(r_m, r_s)}{\partial r_s} < 0, \quad t \in \{A, DA\}. \tag{A-14}$$

Note that in equilibrium

$$r_m = \frac{\lambda_A\tau_{Am} + \lambda_{DA}\tau_{DAm} + \lambda_{DI}\tau_{DIm}}{\gamma\lambda_L}, \quad r_s = \frac{\lambda_A\tau_{As} + \lambda_{DA}\tau_{DAs}}{\gamma\lambda_L}. \tag{A-15}$$

Therefore, we need to show that there exists a unique solution $(r_m^*, r_s^*)$ to the equations

$$f_m(r_m, r_s) \equiv \lambda_A f_m^A(r_m, r_s) + \lambda_{DA} f_m^{DA}(r_m, r_s) + \lambda_{DI} f_m^{DI}(r_m, r_s) - \gamma \lambda_L r_m = 0, \quad \text{(A-16)}$$

and

$$f_s(r_m, r_s) \equiv \lambda_A f_s^A(r_m, r_s) + \lambda_{DA} f_s^{DA}(r_m, r_s) - \gamma \lambda_L r_s = 0. \quad \text{(A-17)}$$

It is clear that for any given $r_s$, $f_m(0, r_s) > 0$ and $f_m(\infty, r_s) < 0$ since optimal precisions $f_m^A(0, r_s) \geq 0$, $f_m^{DA}(0, r_s) \geq 0$, $f_m^{DI}(0, r_s) > 0$, and $f_m^t(\infty, r_s) = 0$ for $t \in \{A, DA, DI\}$ as implied by the first order conditions. In addition, we have $\frac{\partial f_m(r_m, r_s)}{\partial r_m} < 0$ by (A-13) for any given $r_s$. Therefore, for any given $r_s$, there is a unique positive solution $r_m = g(r_s)$ such that equation (A-16) holds.

Plugging $r_m = g(r_s)$ into the second equation of (A-17), we have $f_s(g(r_s), r_s) = 0$. We have $f_s(g(0), 0) > 0$ because the precisions $f_s^A(r_m, r_s)$ and $f_s^{DA}(r_m, r_s)$ are all positive for any $r_s$ and $r_m$, and $f_s(g(\infty), \infty) < 0$ because $f_s^A(r_m, \infty) = 0$ and $f_s^{DA}(r_m, \infty) = 0$ for any $r_m$. By continuity, there must exist a solution $r_s^* > 0$ to $f_s(g(r_s), r_s) = 0$, which implies that there exists $r_s^* > 0$ and $r_m^* = g(r_s^*) > 0$ that solve $f_m(r_m, r_s) = 0$ and $f_s(r_m, r_s) = 0$. Therefore, there exists an equilibrium for a given pair $(\lambda_{DI}, \lambda_{SI})$.

Next we show that there exists $(\lambda_{DI}, \lambda_{SI})$ that solves the investors' optimal participation problem. Note that when the participation cost $k_s = 0$, it is always better for D investors to invest in both of the risky assets because of the diversification effect, so $\lambda_{DI} = 0$, while when $k_s = \infty$, it is always better to invest only in the market portfolio, and so $\lambda_{DI} = \lambda_D$. Similarly, when the participation cost $k_m = 0$, it is always better for S investors to invest in the index and the risk-free asset, so $\lambda_{SI} = \lambda_S$, while when $k_m = \infty$, it is always better for savers to invest only in the risk-free asset, and so $\lambda_{SI} = 0$. By continuity, there exists an equilibrium.

**Proof of Proposition 1:**

Part 1. Suppose given $k_m = k_{m1}$ the equilibrium endogenous proportion of savers who opt to indexing is $\lambda_{SI1}^*$. When $k_m$ is decreased to $k_{m2}$, the utility of SI investors becomes greater than that of S0 investors because SI investors now pay a smaller participation cost of $k_m$. Therefore, some S0 investors will switch to be SI investors and thus the new equilibrium endogenous indexing $\lambda^*_{SI2}$ must be larger than $\lambda^*_{SI1}$.

Part 2. Similarly, suppose given $k_s = k_{s1}$ the equilibrium endogenous indexing is $\lambda^*_{DI1}$. When $k_s$ is increased to $k_{s2}$, the utility of DA investors becomes smaller than that of DI investors because DI investors do not pay the participation cost of $k_s$. Therefore, some DA investors must switch to be DI investors and thus the new equilibrium endogenous indexing $\lambda^*_{DI2}$ must be greater than $\lambda^*_{DI1}$.

**Proof of Proposition 2:**

Note that the information acquisition cost function, $C(\tau_{tm}, \tau_{ts})$, is identical for both DA and DI investors. The value of $\tau_{DAm}$ is derived from the first-order condition of the optimization problem for DA investors:

$$2\gamma C_{Dm}(\tau_{DAm}, \tau_{DAs}) = \frac{1}{\tau_{DAm} + \rho_m}. \tag{A-18}$$

If $\varphi_D < 0$, then we have

$$2\gamma C_{Dm}(\tau_{DAm}, 0) < 2\gamma C_{Dm}(\tau_{DAm}, \tau_{DAs}) = \frac{1}{\tau_{DAm} + \rho_m}. \tag{A-19}$$

Equation (A-19) implies that $\tau_{DAm} < \tau_{DIm}$.

To demonstrate this, let's consider the contrary: suppose $\tau_{DAm} \geq \tau_{DIm}$. From equation (A-19), we obtain:

$$2\gamma C_{Dm}(\tau_{DAm}, 0) < \frac{1}{\tau_{DAm} + \rho_m} \leq \frac{1}{\tau_{DIm} + \rho_m}, \tag{A-20}$$

since $\tau_{DIm}$ solves

$$2\gamma C_{Dm}(\tau_{DIm}, 0) = \frac{1}{\tau_{DIm} + \rho_m}. \tag{A-21}$$

Equations (A-20) and (A-21) imply that $\tau_{DIm} > \tau_{DAm}$ since $C_{Dm}(\tau_{DAm}, 0)$ increases in $\tau_{DAm}$. This leads to a contradiction.

**The Coefficients $b_m$ and $b_s$:**

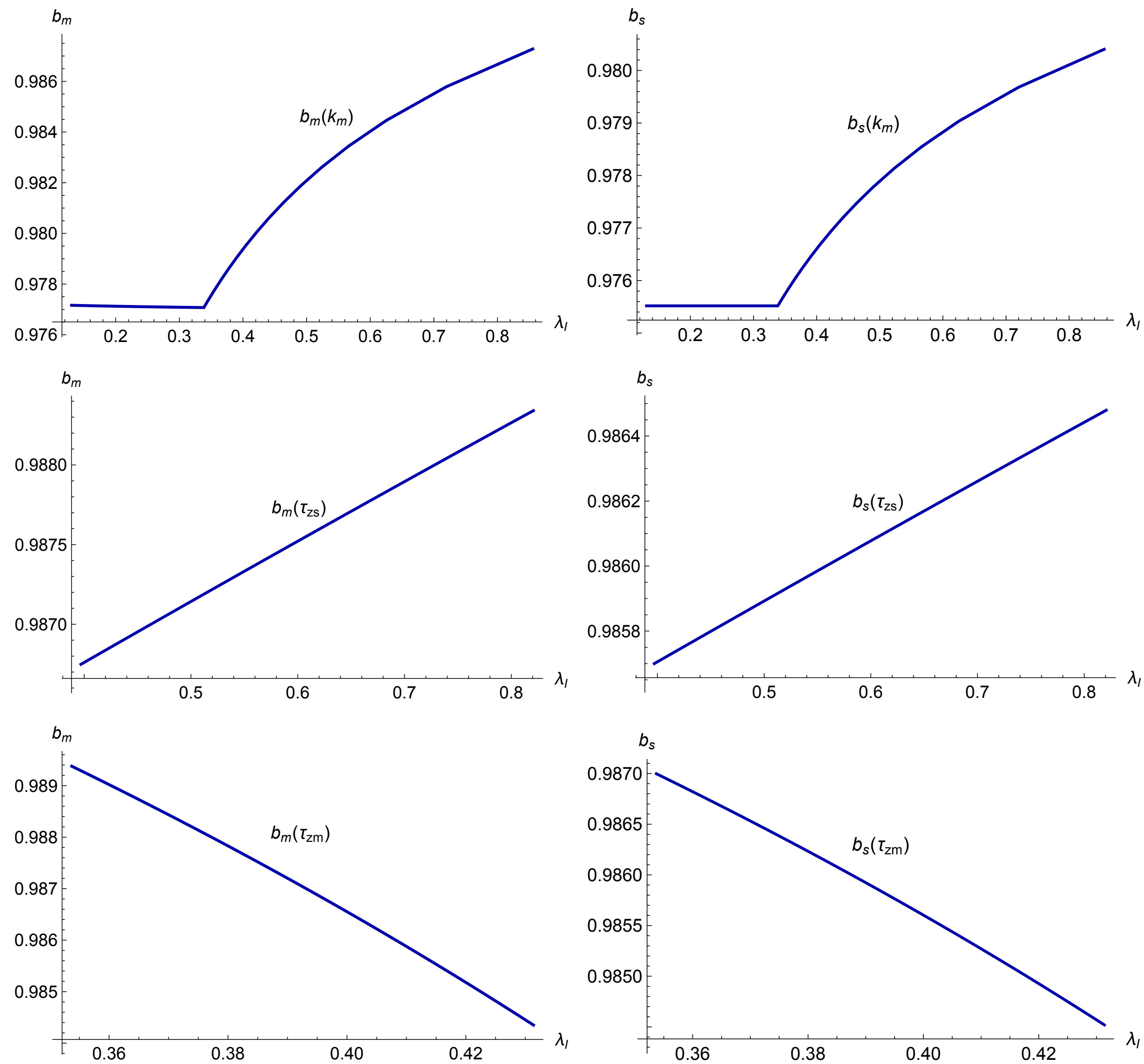


**Figure A-1.** The coefficients $b_m$ and $b_s$ for $V_m$ and $V_s$ in prices $P_m$ and $P_s$ respectively, against indexers mass $\lambda_I$. Default parameter values: $\tau_{vm} = \tau_{vs} = \tau_{zm} = \tau_{zs} = 1$, $\beta = 1$, $\gamma = 0.5$, $\mu_{vm} = 0.3$, $\mu_{vs} = 0.05$, $\lambda_A = 0.4$, $\lambda_D = 0.6$, $\lambda_S = 0.3$, $\lambda_L = 0.1$, $k_m = 0.002$, $k_s = 0.002$, and $C_t(\tau_{tm}, \tau_{ts}) = c_{tm}\tau_{tm}^2 + c_{ts}\tau_{ts}^2 + c_{tms}\tau_{tm}\tau_{ts}$ with $c_{tm} = c_{ts} = 0.01$, and $c_{tms} = 0.01$ for $t = A, DI, DA$.

As illustrated in Figure A-1, the coefficients on $V_m$ and $V_s$ in prices $P_m$ and $P_s$ (i.e., $b_m$ and $b_s$) increase with $\lambda_I$ if the rise of indexing is driven by reduced index participation costs or decreased liquidity trading in the non-index market. In contrast, the coefficients $b_m$ and $b_s$ decrease in $\lambda_I$ when the rise in indexing is driven by increased liquidity trading in the index market.