The equilibrium consequences of indexing*

Philip Bond† Diego García‡

First draft: September 2016
This draft: April 2021

Abstract

We develop a benchmark model to study the equilibrium consequences of indexing in a standard rational expectations setting. Individuals incur costs to participate in financial markets, and these costs are lower for individuals who restrict themselves to indexing. A decline in indexing costs directly increases the prevalence of indexing, thereby reducing the price efficiency of the index and augmenting relative price efficiency. In equilibrium, these changes in price efficiency in turn further increase indexing, and raise the welfare of uninformed traders. For well-informed traders, the share of trading gains stemming from market timing increases relative to stock selection trades.

JEL classification: D82, G14.
Keywords: indexing, welfare.

*We thank seminar audiences at Baruch, BI Oslo, Boston College, the Hanqing Institute of Renmin University, Hong Kong University, Insead, the ISB CAF conference, LBS, LSE, MIT, Michigan State, Notre Dame, NYU, Stanford, the SUFE-SIIFE Conference on Financial Markets and Corporate Finance, the University of Alberta, UC San Diego, the University of Chicago, the University of Maryland, the University of North Carolina, the University of Rochester, UT Austin, the University of Utah, and University of Virginia, and Yale for helpful comments. Any remaining errors are our own.

†Philip Bond, University of Washington, Email: `apbond@uw.edu`; Webpage: `http://faculty.washington.edu/apbond/`.

‡Diego García, University of Colorado Boulder, Email: `diego.garcia@colorado.edu`; Webpage: `http://leeds-faculty.colorado.edu/garcia/`.

1 Introduction

The standard investment recommendation that financial economists offer to retail investors it to purchase a low-fee index mutual fund or exchange-traded fund (ETF), a strategy often described as "index investing," or simply "indexing." An increasing number of indexing products are available, and are increasingly inexpensive and accessible, and more and more investors follow this advice.[1] In this paper, we develop a benchmark model to analyze the equilibrium consequences of a decrease in indexing costs, paying particular attention to participation and welfare.

Our model has the following three features, all of which are necessary for the topic at hand. First, and because much of the discussion of indexing revolves around retail investors, some agents in our model have much less information about future asset cash flows than others. Second, while agents seek trading profits, they also have other trading motives;[2] we model these as stemming from a desire to reallocate risk from non-financial income, with the effect on agents' desired trades resembling discount rate shocks. Third, there are costs to participating in financial markets, which are potentially lower for investors who index. We characterize what happens in such a model as indexing costs fall.

The direct consequence of a fall in indexing costs is, naturally, to draw investors into indexing, and away from both more active trading strategies and non-participation in financial markets. The marginal investor who switches from non-participation to indexing is relatively uninformed. Similarly, the marginal investor who switches out of active trading into indexing is someone who was less informed than other active traders. So the direct consequence of falling indexing costs is to reduce the price efficiency of the index and introduce a common "noise" factor to stock prices of firms covered by the index, while simultaneously reducing "liquidity" in individual stocks.[3]

The central questions we ask in this paper are: What equilibrium effects follow from these direct consequences of a fall in indexing costs? And what is the net effect on investors' welfare? In particular, do equilibrium forces dampen the direct consequences, or are the direct consequences instead self-reinforcing?

We find that the direct consequences are self-reinforcing. Although the reduction in price efficiency associated with indexing may sound undesirable, investor welfare increases, even for the least informed investors, as we discuss below. The increase in the welfare of indexing investors draws in more relatively uninformed investors, amplifying the original effect.

Conversely, the consequence of relatively uninformed investors switching from active trad-

[1] See, for example, French (2008) and Stambaugh (2014).
[2] Absent such motives, the no-trade theorem would apply.
[3] See evidence in, for example, Ben-David, Franzoni, and Moussawi (2018).

ing to index investing is to increase the price efficiency of individual stock prices. Although this sounds desirable, it reduces investor welfare. This induces yet more investors to abandon active trading, again amplifying the original effect.

As the above discussion suggests, our key analytical result is that increases in participation in the market for a financial asset raise the welfare of those already participating. Put differently, participation decisions are strategic complements.[4] The underlying economic force is that an individual investor prefers to trade in a market in which the average investor is relatively uninformed; since the marginal investor is less informed than the average investor, this generates strategic complementarity. Although it may seem intuitive that an investor prefers his average counterparty to be uninformed, this same force leads to prices that are more divorced from cash flows, i.e., lower price efficiency, which is often interpreted as being *undesirable*. Concretely, lower price efficiency in our framework means that prices are more exposed to fluctuations in expected non-financial income that investors wish to partially hedge, i.e., fluctuations in discount rates. As such, establishing an investor's preference for less informed counterparties entails establishing that the benefits of a less informed average counterparty exceed the costs of lower price efficiency.[5] Perhaps surprisingly, and despite the fact that we work with a canonical model of the type introduced by Diamond and Verrecchia (1981), this analysis has not been conducted in the existing literature.

The central empirical predictions of our analysis are about price efficiency. As indexing costs fall, and indexing increases, price efficiency of the index as a whole falls, while the relative price efficiency of individual stocks increases. Moreover, price efficiency is lower for stocks covered by the index than for those outside. It further follows that index reversals become more pronounced, and a greater fraction of the trading gains of relatively informed investors are attributable to "market timing" strategies as opposed to "stock selection" strategies. Section 5 reviews empirical support for these findings.

Asides from its implications for the equilibrium effects of indexing, our paper also speaks to the wider question of whether and how the financial sector contributes to social welfare (see, e.g, Baumol (1965)). In particular, we work with a canonical model in which a financial market exists because it facilitates risk-sharing, and show that informed trading generally worsens this risk-sharing function, while uninformed trading improves it. In Section 7 we also explore an extension of our baseline model in which the information produced by financial

[4]Grossman and Stiglitz (1980) analyze traders' decisions to become informed, taking the set of trading agents as given, and show that information acquisition decisions are strategic substitutes: the incremental trading profits from private information decline as more traders become informed. In contrast, we consider a setting without exogenous noise traders, and analyze traders' participation decisions.

[5]This result is related to the so-called "Hirshleifer effect" (Hirshleifer (1971)), but does not follow directly from it; see subsection 3.1 below, and also related discussions in Marín and Rahi (1999) and Dow and Rahi (2003).

markets guides resource allocation decisions (see Bond, Goldstein, and Edmans (2012) for a survey), and in doing so formalize a positive welfare effect of the increase in the relative price efficiency that follows from lower indexing costs.

Related literature: In its general theme, our analysis is related to papers such as Subrahmanyam (1991), Cong and Xu (2019), Stambaugh (2014), and Bhattacharya and O'Hara (2018). Subrahmanyam (1991) models the introduction of index futures, while Cong and Xu (2019) and Bhattacharya and O'Hara (2018) model the introduction of ETFs. An important assumption in all three papers is how the introduction of a new financial product affects the allocation of "noise" or liquidity traders across assets.[6] Stambaugh (2014) analyzes the effects of an exogenous decline in noise traders for financial markets, and in particular, for actively managed investment funds. Relative to all these papers, we model the behavior of all financial market participants, and in particular, analyze how a fall in indexing costs affects participation decisions and welfare. Like us, Peress (2005) analyzes participation, in his case in an economy with a single risky asset that is initially held by noise traders. In his analysis, the direct effect of a fall in participation costs is to increase participation, leading to smaller positions for rational traders. In contrast to our analysis, these smaller positions in turn reduce the gains to participation, dampening the direct effect of lower costs.[7]

In an independent, contemporaneous, and complementary paper, Baruch and Zhang (2018) likewise study the equilibrium consequence of indexing, though from a very different perspective. They consider a multi-asset version of Grossman (1976), so that without indexing prices fully reveal agents' private signals. In this setting they show that an exogenous increase in indexing reduces the amount of information prices contain about individual assets, while the amount of information prices contain about aggregates is unaffected.

We have deliberately based our analysis on the canonical model of financial markets of Diamond and Verrecchia (1981). Like Grossman and Stiglitz (1980), Hellwig (1980), and Admati (1985), these authors analyze trade between differentially informed agents, but different from these papers, there are no exogenous "noise" or "liquidity" trades. Instead, agents have heterogeneous and privately observed exposures to risk. Consequently, financial markets hold the potential to increase welfare by allowing agents to redistribute risk. Perhaps surprisingly, and although a model of this type has been analyzed by a significant number of authors, results on welfare are scarce.[8] A significant algebraic complication in characterizing

[6]Subrahmanyam (1991), and Cong and Xu (2019) allow for optimization by a subset of these traders.

[7]In Peress (2005), the smaller positions that stem from lower participation costs also reduce the amount of information collected by "active" traders. As in our analysis, this reduces the information content of prices, though via a distinct channel. Also related is Gârleanu and Pedersen (2018), who among other things consider the choice between actively and passively managed funds in a financial market with a single risky asset and noise traders. Gârleanu and Pedersen (2020) extend their analysis to a multi-asset model.

[8]For results on the effect of information on welfare in different equilibrium models of financial markets,

welfare is that, when combined with the asset price, each agent's private exposure shock contains information about expected asset payoffs. To avoid this complication, Verrecchia (1982) and Diamond (1985) consider sequences of economies in which the variance of each individual's exposure shock grows with the number of agents, and directly study the limit of this sequence. In the limit economy, each agent's exposure shock has infinite variance, and so expected utility prior to the realization of the exposure shock is undefined, in turn preventing the analysis of participation decisions prior to the realization of exposure shocks.[9]

In an independent, contemporaneous, and complementary paper, Kawakami (2017) also makes progress in characterizing welfare in a Diamond and Verrecchia (1981) model. Whereas we study an economy with a continuum of agents and allow for heterogeneity in the precision of signals about cash flows that agents observe, thereby allowing us to consider the effect of an increase in participation by relatively uninformed agents, Kawakami instead considers a finite-agent economy with homogeneous signal precisions, in which an increase in the size of the market is associated with better diversification of individual exposure shocks. Analytically, we make more explicit use than Kawakami of market-clearing conditions, which allows us to incorporate heterogeneity in signal precisions in a tractable way.

Marín and Rahi (1999) obtain welfare results in a relatively specialized setting: there are two classes of agents, one class of which sees identical signals about asset payoffs and private endowments, and another class of completely uninformed agents. Moreover, the traded asset is in zero net supply. Dow and Rahi (2003) also analyze welfare, and obtain some tractability by inserting a risk-neutral market maker into the economy, which reduces the applicability of the model for analyzing aggregate financial markets. In closely related settings, Medrano and Vives (2004) argue that "the expressions for the expected utility of a hedger ... are complicated," whereas Kurlat and Veldkamp (2015) write that "there is no closed-form expression for investor welfare." The complications, common to our model as well, stem from the role of exposure shocks as signals about asset cash flows, on top of the standard risk sharing role that motivates trade.

see, for example, Schlee (2001) and Kurlat (2019). The former paper analyzes the value of public signals in a setting in which individual endowments are public once they are realized, and so trades can be conditioned on them, and characterizes conditions in which improvements in public information cause a Pareto deterioration. In contrast, in our setting (inherited from Diamond and Verrecchia (1981)), individual signals are private, and individual endowments are likewise private, even after they are realized. The latter paper analyzes what is essentially an origination market in which sellers are strictly better informed than buyers, and characterizes the ratio of the social to private value of buyer information. In this setting, improvements in seller information reduce adverse selection, and so the social value of information is positive. In contrast, in our setting improvements in trader information do not necessarily reduce adverse selection, and our results imply that improvements in information of a positive measure of agents reduce the welfare of all counterparties.

[9] If instead one modeled participation decisions as being made after the exposure shock, then almost all agents would participate, since their exposure shocks are so large.

The strategic complementarity of participation decisions is related to results in Admati and Pfleiderer (1988) and Chowdhry and Nanda (1991) on the incentive for liquidity to traders to trade in the same periods and locations, respectively, as other liquidity traders. Different from in our analysis, in these papers liquidity traders trade for unspecified exogenous reasons, so there is no channel via which price efficiency can benefit such traders.

Finally, we emphasize that we examine "indexing" in the sense of a "passive" investment strategy based on an index. Other authors have analyzed the distinct topic of the role of indices as benchmarks that affect the compensation of fund managers: see, e.g., Admati and Pfleiderer (1997), Basak and Pavlova (2013), Breugem and Buss (2018), Buffa, Vayanos, and Woolley (2019).

2 The model

2.1 Preferences, assets, endowments, information

We work with a version of Diamond and Verrecchia (1981) in which there is a unit interval of agents (see Ganguli and Yang (2009) and Manzano and Vives (2011)) and multiple assets. We let $i \in [0,1]$ index agents. We emphasize that this is a canonical setting, in which risk-sharing benefits lead to gains from trade, which in turn allows for informed trading. Indeed, we are unable to think of a simpler or more standard framework that still contains the three key features discussed in the introduction.

Each agent i has preferences over terminal wealth W_i with a constant absolute risk aversion (CARA) of γ. There are m risky assets available for trading; and also a risk-free asset that is in perfectly elastic supply and has a net return that we normalize to 0.[10] Each asset $k \in \{1,\ldots,m\}$ produces a payoff given by random variable $\tilde{X}_k$, with identical mean, and identical variance τ_X^{-1}. All exogenous random variables in the model are normally distributed and mutually independent. Each agent's initial endowment of each asset is $\tilde{S}$. We characterize the competitive equilibrium of the economy, where agents are small relative to the market, and act as price-takers. The equilibrium price of asset k is $\tilde{P}_k$.

In addition, agents have other sources of income (e.g., labor income, non-traded capital income) that are correlated with the cash flows of the risky assets. For simplicity, we assume the correlation is perfect. Specifically, let $\tilde{Z}_k$ and $\tilde{u}_{ik}$ be random variables, and define agent

[10] Alternatively, one can assume the risk-free asset is in zero net supply, and interpret the prices of risky assets as forward prices (i.e., the amount of certain wealth that will be exchanged for the risky asset's payoff at the terminal date). In this case, the market for the risk-free asset clears by Walras's Law.

i's income from sources other than the risky assets by

$$\sum_{k=1}^{m} \left(\tilde{Z}_k + \tilde{u}_{ik} \right) \tilde{X}_k. \tag{1}$$

Here, $\tilde{Z}_k + \tilde{u}_{ik}$ represents agent i's non-financial exposure to the cash flow risk $\tilde{X}_k$. Agent i privately observes the sum $\tilde{Z}_k + \tilde{u}_{ik}$, but not its individual components $\tilde{Z}_k$ and $\tilde{u}_{ik}$. So agents know their own income exposures $\tilde{Z}_k + \tilde{u}_{ik}$, but remain uncertain about the aggregate component of other agents' exposures, $\tilde{Z}_k$. The variances of $\tilde{Z}_k$ and $\tilde{u}_{ik}$ are τ_Z^{-1} and τ_u^{-1} respectively, and $\mathbb{E}\left[\tilde{Z}_k\right] = \mathbb{E}\left[\tilde{u}_{ik}\right] = 0$.

Our environment is symmetric across assets; this considerably simplifies the analysis because it makes possible the change of basis that we describe in subsection 2.3. Moreover, agents have the same risk aversion, initial asset endowments, and ex ante exposures to other sources of risk, though they differ in their access to information, as described below. These agent-symmetry properties are important in allowing us to tractably characterize the expected utility from participation in financial markets, and hence participation decisions. That said, the agent-symmetry assumptions can be at least slightly relaxed. For example, it is straightforward to instead assume that some agents have greater initial endowments of financial assets but smaller ex ante exposures to non-financial risk, provided the combined ex ante financial and non-financial exposure remains constant across agents. Similarly, we could allow some agents to have greater initial endowments of financial assets but lower risk aversions, provided the two sources of variation offset each other.

Gains from trade stem from agents' differential and privately observed exposures $\tilde{Z}_k + \tilde{u}_{ik}$. In equilibrium, fluctuations in aggregate exposures $\tilde{Z}_k$ affect prices, and so resemble discount rate fluctuations. It is also worth noting that one can give a more behavioral interpretation to $\tilde{Z}_k + \tilde{u}_{ik}$: looking ahead to agents' trades (11), $\tilde{Z}_k + \tilde{u}_{ik}$ can be interpreted simply as a shock to agent i's desired holding of asset k, independent of the source of this shock.

The terminal wealth of agent i is determined by the combination of trading profits, payoffs from initial asset endowments, and other income (1). For notational convenience, define

$$\tilde{e}_{ik} = \tilde{S} + \tilde{Z}_k + \tilde{u}_{ik}$$

to represent agent i's net exposure to cash flow $\tilde{X}_k$, stemming from the combination of initial asset holdings and non-financial exposure.

We denote by $\tilde{\theta}_{ik}$ agent i's trade of asset k. The terminal wealth of an agent who makes the vector of trades $\tilde{\theta}_i$ is

$$W\left(\tilde{\theta}_i, \tilde{e}_i\right) \equiv \sum_{k=1}^{m} \tilde{\theta}_{ik}\left(\tilde{X}_k - \tilde{P}_k\right) + \tilde{S}\tilde{X}_{ik} + \left(\tilde{Z}_k + \tilde{u}_{ik}\right)\tilde{X}_{ik} = \sum_{k=1}^{m}\left(\tilde{\theta}_{ik} + \tilde{e}_{ik}\right)\left(\tilde{X}_k - \tilde{P}_k\right) + \tilde{e}_{ik}\tilde{P}_k. \tag{2}$$

Prior to trading, each agent i observes private signals of the form

$$\tilde{y}_{ik} = \tilde{X}_k + \tilde{\epsilon}_{ik},$$

where $\tilde{\epsilon}_{ik}$ is a random variable with mean 0 and variance τ_i^{-1}.[11]

The precisions τ_i^{-1} of private signals are heterogeneous across agents, so that some agents are more informed than others. Without loss, we order agents so that signal precision τ_i is decreasing in i; and for simplicity, we assume τ_i is strictly decreasing.[12]

An agent's information set at the time of trading is hence the triple of m-vectors $\left(\tilde{y}_i, \tilde{e}_i, \tilde{P}\right)$, which consists of signals about cash flows $\tilde{y}_i$; exposures $\tilde{e}_i$; and prices $\tilde{P}$.

2.2 Indexing and participation

Agents incur a cost $\kappa > 0$ of fully participating in financial markets, reflecting a combination of information collection and processing costs, psychic costs, expected trading costs, and the cost of potentially trading in a less than optimal way. Agents make participation decisions prior to observing any of $\left(\tilde{y}_i, \tilde{e}_i, \tilde{P}\right)$. This timing assumption for the participation decision is important for tractability, since it ensures that all random variables are normally distributed at the trading stage.

In addition to fully participating in financial markets, agents have the option of participating only via trading an "index" asset. The index covers the first $l \leq m$ of the m assets, where we assume that l is a power of 2 (this greatly enhances tractability, as will become clear in the next subsection). Since all assets have the same supply $\tilde{S}$, the index asset likewise

[11]Note that an agent i has the same quality signal about all assets. We leave the interaction of heterogeneity of signal precision across assets with heterogeneity of signal precisions across agents for future research.

[12]Formally, our model is one in which all agents either invest directly in individual stocks, or else invest via passive index funds (see below). An alternative interpretation is that agents with a low i index (and hence precise signals) are relatively good at identifying skilled mutual and hedge funds (Gârleanu and Pedersen (2018)), and the direct investments in the model are made through such intermediaries. Looking ahead, and as one would expect, agents' desired trades depend on their exposure realizations. So in the intermediated investment interpretation just described, agents would also pay attention to general "styles" of funds, in addition to the skill of managers. See also García and Vanden (2009).

has an equal holding of each asset $k \leq l$ covered, and produces a cash flow

$$X_1 \equiv l^{-\frac{1}{2}} \sum_{k=1}^{l} \tilde{X}_k, \tag{3}$$

where $l^{\frac{1}{2}}$ is an index divisor, set so that $\mathrm{var}(X_1) = \tau_X^{-1}$. Because of equal supply $\tilde{S}$ of the underlying assets, the index can be viewed as either equal- or value-weighted, since the value share of an asset $k \leq l$ in an index trade and in the market for assets $j \leq l$ is simply $\frac{\tilde{P}_k}{\sum_{j=1}^{l} \tilde{P}_j}$.

Indexing by agent i corresponds to buying or selling equal amounts of all assets in the index, and zero units of assets outside the index, i.e., trade vectors $\tilde{\theta}_i$ such that $\tilde{\theta}_{ij} = \tilde{\theta}_{ik}$ for any $j, k \in \{1, \ldots, l\}$ and $\tilde{\theta}_{ik} = 0$ for $k > l$.

The advantage of participating in financial markets only via indexing is that the participation cost is lower, which we denote by $\kappa_1 \in (0, \kappa)$. The lower participation cost of indexing reflects lower trading costs, because of the availability of low cost index mutual funds and exchange traded funds (ETFs); lower cognitive demands and attention costs; and lower information costs, since as our formal analysis will show, a sufficient statistic for agent i's private information under indexing is the average signal for assets in the index, $\frac{1}{l} \sum_{k=1}^{l} \tilde{y}_{ik}$, which can be interpreted as agent i restricting attention to broad economic aggregates.

Looking ahead, the main comparative static we will be interested in is a fall in the cost of indexing κ_1. This corresponds to falling fees, greater availability, and greater awareness of products such as low-cost index funds and ETFs. It may also reflect an increase in public awareness of the standard advice given by finance academics.

Finally, individuals i who do not participate in financial markets at all pay no participation cost, but do not trade, i.e., $\tilde{\theta}_{ik} = 0$ for all assets k.

The definition of an equilibrium in terms of pricing, participation decisions, and trading strategies is standard. However, we postpone a formal statement of equilibrium conditions until subsection 2.4. This allows us to give the definition directly in terms of a spanning set of synthetic assets, which we introduce next, and use to conduct our analysis.

2.3 Disentangling markets

Although the fundamentals of different assets are independent in all dimensions, the presence of indexing agents links the prices of distinct assets. For example, if j and k are two distinct assets covered by the index, then an indexing agent's exposure $\tilde{e}_{ij}$ to cash flow risk $\tilde{X}_j$ affects the agent's desired trade of asset k as well as of asset j.

Because of the entanglement that indexing produces between different assets covered by the index, it is analytically very convenient to change basis and study the economy in terms of a set of synthetic assets that are mutually independent even in the presence of indexers.[13]

The case of two assets ($l = m = 2$) is simple. The first synthetic asset is the index portfolio, $X_1 = \frac{1}{\sqrt{2}}\left(\tilde{X}_1 + \tilde{X}_2\right)$. The second synthetic asset is $X_2 = \frac{1}{\sqrt{2}}\left(\tilde{X}_1 - \tilde{X}_2\right)$, which can be labeled a "spread" asset, as it allows agents to trade on the relative mispricing between assets 1 and 2. We next generalize this construction to $l > 2$ assets in the index. We first give a representative example, and then formalize the change of basis.

Example: Suppose there are $m = 5$ assets and the index covers the first $l = 4$. Then consider the following set of 5 synthetic assets, where the 1st synthetic asset is the index asset, and pays X_1 as defined in (3); the 5th synthetic asset coincides with the underlying asset 5, i.e., it pays $X_5 = \tilde{X}_5$; and the remaining 3 synthetic assets are long-short positions in assets covered by the index, and pay X_2, X_3, X_4 defined by

$$\begin{aligned} X_2 &= \frac{1}{2}\left(\tilde{X}_1 + \tilde{X}_3 - \tilde{X}_2 - \tilde{X}_4\right), \\ X_3 &= \frac{1}{2}\left(\tilde{X}_1 + \tilde{X}_2 - \tilde{X}_3 - \tilde{X}_4\right), \\ X_4 &= \frac{1}{2}\left(\tilde{X}_1 + \tilde{X}_4 - \tilde{X}_2 - \tilde{X}_3\right). \end{aligned}$$

The five synthetic assets span the underlying assets $\tilde{X}_k$. Moreover, they are uncorrelated (i.e., $\mathrm{cov}\left(X_j, X_k\right) = 0$ for all j, k), and each has variance τ_X^{-1}. Note that $\mathbb{E}\left[X_1\right] = \sqrt{4}\mathbb{E}\left[\tilde{X}_1\right]$, while $\mathbb{E}\left[X_j\right] = 0$ for synthetic assets $j = 2, 3, 4$. Similarly the per-agent endowment of the index synthetic asset 1 is $\sqrt{4}\tilde{S}$, while synthetic assets $j = 2, 3, 4$ are in zero net supply.

Everything about the above example extends to the general case (formally, see Lemmas A-1 and A-2 in appendix). Specifically, we construct l synthetic assets to span assets $1, \ldots, l$. Synthetic asset 1 is the index asset that pays X_1 as defined in (3), with $\mathbb{E}\left[X_1\right] = \sqrt{l}\mathbb{E}\left[\tilde{X}_1\right]$ and per-agent endowment $S_1 \equiv \sqrt{l}\tilde{S}$. Synthetic assets $j = 2, \ldots, l$ have payoffs denoted X_j, with $\mathbb{E}\left[X_j\right] = 0$ and per-capita endowment $S_j \equiv 0$. For assets $j > l$ outside the index we simply define $X_j = \tilde{X}_j$, i.e., the synthetic and fundamental asset coincide. The correlation of any pair of synthetic assets is 0. The variance of each synthetic asset is τ_X^{-1}.

In general, we use tildes to denote quantities related to fundamental assets, and the absence of a tilde to denote quantities related to synthetic assets. For example, P is the vector of prices of synthetic assets, and θ_i is the vector of agent i's trades of synthetic assets.

[13]This change of basis is not essential to solve for equilibrium prices; see Admati (1985)'s analysis of a multi-asset version of Hellwig (1980).

The terminal wealth of agent i given the vector of synthetic exposures e_i and trades θ_i is

$$W(\theta_i, e_i) = \sum_{k=1}^{m} (\theta_{ik} + e_{ik})(X_k - P_k) + e_{ik} P_k.$$

To solve for equilibrium prices and welfare, we work directly with the synthetic assets described above. By construction, the synthetic assets are independent of each other in all respects. Moreover, and importantly, indexing corresponds simply to the constraint that an agent can trade only synthetic asset 1, with no trade of the other synthetic assets. This means we can analyze the equilibrium in the market for each synthetic asset in isolation (given CARA preferences). Formally, we denote the set of trades available to agents who pay the reduced participation cost κ_1 by $\Theta_1 \equiv \{\theta_i \in \mathbb{R}^m : \theta_{ik} = 0 \text{ if } k \neq 1\}$.

2.4 Equilibrium

The equilibrium definition is a straightforward extension of that used in competitive rational expectations models (Grossman and Stiglitz (1980), Hellwig (1980)), with the participation decision incorporated. To ease the formal statement of participation decisions, we first define the expected utilities $\mathcal{U}_i^A(e_i)$, $\mathcal{U}_i^I(e_i)$, $\mathcal{U}_i^0(e_i)$ associated with full participation, or "active" trading; with indexing, or "passive" investing; and with non-participation. For consistency with subsequent notation, we define these objects conditional on the vector of exposure realizations, e_i, and exclusive of the participation costs κ and κ_1. In particular, it is unnecessary for our analysis to explicitly integrate out uncertainty over exposures e_i.

$$\begin{aligned}
\mathcal{U}_i^A(e_i) &\equiv \mathbb{E}\left[\max_{\theta_i} \mathbb{E}\left[u\left(W(\theta_i, e_i)\right) | y_i, e_i, P\right] | e_i\right], \\
\mathcal{U}_i^I(e_i) &\equiv \mathbb{E}\left[\max_{\theta_i \in \Theta_1} \mathbb{E}\left[u\left(W(\theta_i, e_i)\right) | y_i, e_i, P\right] | e_i\right], \\
\mathcal{U}_i^0(e_i) &\equiv \mathbb{E}\left[u\left(\sum_{k=1}^{m} e_{ik} X_k\right) | e_i\right].
\end{aligned}$$

Definition 1 *A rational expectations equilibrium consists of non-overlapping sets of agents who fully participate, N, and who index, N_1; trading strategies $\{\theta_i(y_i, e_i, P)\}_{i \in [0,1]}$; and price function $P(X, Z)$. The equilibrium conditions are that markets clear almost surely,*

$$\int_0^1 \theta_i(y_i, e_i, P)\, di = 0; \tag{4}$$

each agent i's trading strategy is optimal given his participation decision and prices,

$$\theta_i(y_i, e_i, P) \in \arg\max_{\hat{\theta}_i} \mathbb{E}\left[u\left(W\left(\hat{\theta}_i, e_i\right)\right)|y_i, e_i, P\right] \text{ if } i \in N, \tag{5}$$

$$\theta_i(y_i, e_i, P) \in \arg\max_{\hat{\theta}_i \in \Theta_1} \mathbb{E}\left[u\left(W\left(\hat{\theta}_i, e_i\right)\right)|y_i, e_i, P\right] \text{ if } i \in N_1, \tag{6}$$

$$\theta_{ik}(y_i, e_i, P) = 0 \text{ for all assets } k \text{ if } i \notin N \cup N_1; \tag{7}$$

and participation decisions are optimal, i.e.,

$$\begin{aligned}
\mathbb{E}\left[\mathcal{U}_i^A(e_i)\exp(\gamma\kappa)\right] &\geq \max\left\{\mathbb{E}\left[\mathcal{U}_i^I(e_i)\exp(\gamma\kappa_1)\right], \mathbb{E}\left[\mathcal{U}_i^0(e_i)\right]\right\} \text{ if } i \in N, \\
\mathbb{E}\left[\mathcal{U}_i^I(e_i)\exp(\gamma\kappa_1)\right] &\geq \max\left\{\mathbb{E}\left[\mathcal{U}_i^A(e_i)\exp(\gamma\kappa)\right], \mathbb{E}\left[\mathcal{U}_i^0(e_i)\right]\right\} \text{ if } i \in N_1, \\
\mathbb{E}\left[\mathcal{U}_i^0(e_i)\right] &\geq \max\left\{\mathbb{E}\left[\mathcal{U}_i^A(e_i)\exp(\gamma\kappa)\right], \mathbb{E}\left[\mathcal{U}_i^I(e_i)\exp(\gamma\kappa_1)\right]\right\} \text{ if } i \notin N \cup N_1.
\end{aligned}$$

Throughout, we assume

$$4\gamma^2\left(\tau_Z^{-1} + \tau_u^{-1}\right) < \tau_X \tag{8}$$

$$\gamma^2 > 4\tau_0\tau_u, \tag{9}$$

where τ_0 is the precision of agent 0's information, i.e., the highest precision in the population. Condition (8) ensures that expected utility is well-defined for non-participating agents; without this condition, such agents are exposed to so much risk that their expected utilities are infinitely low. Condition (9) ensures that an equilibrium exists at the trading stage (see Lemma 3 below). Loosely speaking, without this condition there is too much trading based on information relative to trading based on risk-sharing; Ganguli and Yang (2009) impose essentially the same condition.[14]

3 Informed trading and welfare in each asset market

Given our construction of synthetic assets as being independent from each other in all respects, in this section we analyze the equilibrium of the market for an arbitrary synthetic asset k in isolation. Likewise, we evaluate the expected utility associated with trading asset k in isolation. For clarity, we retain the asset subscript k in the main text, while generally omitting it in proofs in the appendix.

[14] The main extension in Manzano and Vives (2011) relative to Ganguli and Yang (2009) is that they allow for the error terms in the trader's signals to be correlated. Non-zero correlation eliminates the need for condition (9). Since our focus is on welfare, we choose to study the slightly more tractable model with conditionally independent estimation errors.

3.1 A public information benchmark

To gain intuition for our subsequent results on the relation between price efficiency and the value of participation, it is useful to consider a benchmark case in which agents' private signals y_{ik} are replaced with a finite number of public signals about X_k, with all other aspects of the model left unchanged (in particular, exposures e_{ik} are private information, and trade occurs only after agents observe these exposures). In this case, all agents have the same posterior of X_k at the trading stage, and so by market clearing (4) and the expression for the optimal trade θ_{ik} in (11) below,

$$\theta_{ik} + e_{ik} = S_k + Z_k.$$

So each agent's terminal wealth is

$$W_{ik} = (S_k + Z_k)(X_k - P_k) + e_{ik}P_k = (S_k + Z_k)X_k + u_{ik}P_k. \tag{10}$$

The first term in (10) corresponds to simply dividing the economy's aggregate cash flow equally among agents, and is unaffected by signal precision; indeed, it corresponds to the (symmetric)[15] unconstrained solution to the social planner's problem. In contrast, the second term depends on the price P_k, which is certainly affected by signal precision.

Expression (10) indicates that, ceteris paribus, agents dislike variance of the price P_k. In turn, the variance of P_k is determined by its dependence on the cash flow X_k and the aggregate exposure Z_k. Higher-precision signals about X_k increase P_k's dependence on X_k, and hence, ceteris paribus, the variance of P_k (corresponding to the Hirshleifer effect); but they tend to decrease P_k's dependence on Z_k, and hence, ceteris paribus, the variance of P_k. As such, the overall effect is unclear.[16]

3.2 Equilibrium at the trading stage

We first characterize the equilibrium outcome of the trading stage, taking participation decisions as given. It is immediate that agents with lower signal precision enjoy lower utility

[15]In non-symmetric solutions, each agent has terminal wealth $W_i = (S_k + Z_k)X_k + K_i$, where K_i is a constant, and $\int K_i di = 0$.

[16]It is worth noting that the limiting case of perfect information about X_k is straightforward. In this case, the price P_k simply equals X_k, and so (10) reduces to $W_{ik} = e_{ik}X_k$, which is the autarchy outcome. Hence welfare is minimized by perfect information about X_k, since in this case the financial market cannot provide any risk sharing (the Hirshleifer effect). Our analysis below concerns the more relevant non-limit case. Moreover, note that Diamond (1985) characterizes how welfare changes as the precision of public information changes, though with the mathematical compromises discussed earlier. In Appendix C, we show that welfare in this benchmark case indeed monotonically declines in the precision of public information, though the proof is non-trivial, consistent with the discussion in the main text.

from trading. Hence there is a cutoff agent n_k such that all better-informed agents $i \leq n_k$ participate, and all worse-informed agents $i > n_k$ do not participate. As standard, we focus on linear equilibria, i.e., equilibria in which the price of asset k is a linear function of the cash flow X_k and the aggregate exposure Z_k. We characterize equilibria by conjecturing that the price has this linear form, and then verifying (Lemma 3).

We relegate much of this analysis to the appendix. In particular, Lemmas A-4, A-5, A-6 establish some basic properties that we use in subsequent results. In the main text, we focus on two results—Lemmas 1 and 2—which turn out to be critical for Propositions 1 and 2 covering participation decisions. Both Lemmas 1 and 2 hold generally in Diamond-Verrecchia economies—and indeed, Lemma 1 holds in a wide class of CARA-normal settings. Nonetheless, we are unaware of previous statements of either result.

In a linear equilibrium, each agent's optimal trade has the standard mean-variance form,

$$\theta_{ik} + e_{ik} = \frac{1}{\gamma} \frac{\mathbb{E}\left[X_k | y_{ik}, e_{ik}, P_k\right] - P_k}{\operatorname{var}\left(X_k | y_{ik}, e_{ik}, P_k\right)}. \tag{11}$$

As typical for this class of models, an important equilibrium quantity is the relative sensitivity of price P_k to the true cash flow X_k and aggregate exposure Z_k, which we denote by ρ_k:

$$\rho_k \equiv -\frac{\frac{\partial P_k}{\partial X_k}}{\frac{\partial P_k}{\partial Z_k}}.$$

We refer to ρ_k as the *price efficiency* of the risky asset, since

$$\operatorname{var}(X_k | P_k)^{-1} = \tau_X + \rho_k^2 \tau_Z, \tag{12}$$

$$\operatorname{var}(X_k | y_{ik}, e_{ik}, P_k)^{-1} = \tau_X + \rho_k^2 (\tau_Z + \tau_u) + \tau_i. \tag{13}$$

These expressions (derived in the proof of Lemma A-5) measure the ability of an outside observer and agent i, respectively, to forecast the cash flow X_k.

The conditional variance $\operatorname{var}\left(X_k | y_{ik}, e_{ik}, P_k\right)$ is an important determinant of agent i's trades. Lemma 1 establishes that the harmonic mean of the conditional variance of participating agents equals the covariance of returns $X_k - P_k$ and cash flows X_k. We stress that the proof is very concise, and makes use only of the market clearing condition (4), the general form of demand (11), and the assumption that random variables are distributed normally.

Lemma 1 *In a linear equilibrium,*

$$\frac{1}{n_k} \int_0^{n_k} \frac{1}{\operatorname{var}\left(X_k | y_{ik}, e_{ik}, P_k\right)} di = \frac{1}{\operatorname{cov}\left(X_k - P_k, X_k\right)}. \tag{14}$$

Note that Lemma 1 nests the special case in which all agents are completely uninformed about the cash flow X_k, so that the price is unrelated to X_k, and so for any agent i, $\text{var}\left(X_k|y_{ik}, e_{ik}, P_k\right) = \text{var}\left(X_k\right) = \text{cov}\left(X_k - P_k, X_k\right)$.

Among other things, we use Lemma 1 to characterize the equilibrium risk premium $\mathbb{E}\left[X_k - P_k\right]$. Taking the unconditional expectation of (11) gives

$$\mathbb{E}\left[\theta_{ik}\right] + S_k = \frac{1}{\gamma}\frac{\mathbb{E}\left[X_k - P_k\right]}{\text{var}\left(X_k|y_{ik}, e_{ik}, P_k\right)}.$$

Combined with market clearing (4) (specifically, $\int_0^{n_k} \mathbb{E}\left[\theta_{ik}\right] di = 0$), we obtain:

Corollary 1 *In a linear equilibrium,*

$$\mathbb{E}\left[X_k - P_k\right] = \gamma S_k \text{cov}\left(X_k - P_k, X_k\right).$$

As for Lemma 1, it may help to note that Corollary 1 nests the special case in which no agent has any information, and so $\mathbb{E}[X_k - P_k] = \gamma S_k \text{var}(X_k)$.

Both prices and exposure shocks play two distinct roles in determining an agent's demand: they directly affect demand, and separately, they also affect an agent's beliefs about the cash flow X, thereby indirectly affecting demand. To clarify this dual role, we write $\theta_{ik}\left(y_{ik}, e_{ik}, \hat{e}_{ik}, P_k, \hat{P}_k\right)$ for the demand of an agent who has exposure e_{ik} and can trade at price P_k, but whose posterior of X_k is formed using the information set $\left(y_{ik}, \hat{e}_{ik}, \hat{P}_k\right)$. Even though $\hat{e}_{ik} = e_{ik}$ and $\hat{P}_k = P_k$, keeping separate track of the two roles of prices and exposure shocks is conceptually useful. In particular, this separation allows the following result, which says that prices contain more information about cash flows than exposures do. The intuition is that exposures are only informative about cash flows because they provide information on whether a high price is due to a high cash flow X_k or a low aggregate exposure shock Z_k. That is, the information provided by exposures is subsidiary to that provided by prices.

Lemma 2 *In a linear equilibrium, the ratio of the informational to non-informational effect of prices on demand exceeds the ratio of the informational to non-informational effect of exposures on demand,*

$$\frac{\left|\int_0^{n_k} \frac{\partial \theta_{ik}}{\partial \hat{P}_k} di\right|}{\left|\int_0^{n_k} \frac{\partial \theta_{ik}}{\partial P_k} di\right|} > \frac{\left|\int_0^{n_k} \frac{\partial \theta_{ik}}{\partial \hat{e}_{ik}} di\right|}{\left|\int_0^{n_k} \frac{\partial \theta_{ik}}{\partial e_{ik}} di\right|}. \tag{15}$$

The only departure of our model's trading stage relative to Ganguli and Yang (2009) and Manzano and Vives (2011) is that agents have heterogeneous signal precisions. As such, our next result closely follows these previous papers. Moreover, as in these previous analyses, our trading stage features two equilibria. We follow Manzano and Vives (2011) and focus on the stable equilibrium, which is the one with lower price efficiency.[17]

Lemma 3 *Given participation n_k, there is unique stable linear equilibrium, in which price efficiency ρ_k is the smaller root of the quadratic*

$$\rho_k^2 \tau_u - \gamma \rho_k + \frac{1}{n_k} \int_0^{n_k} \tau_i di = 0. \tag{16}$$

Price efficiency ρ_k is decreasing in participation n_k.

From Lemma 3, price efficiency is determined by the average information precision of agents who actively trade, given by the term $\frac{1}{n_k} \int_0^{n_k} \tau_i di$. As participation increases, newly participating agents lower this average. Even though these agents bring more information to the market, they also bring more trade motivated by risk-sharing concerns, which functions in the same way as noise. Consequently, the net effect is to reduce price efficiency.[18]

3.3 Expected utility from participation

We next turn to agents' participation decisions. To do so, we first characterize an agent's expected utility from participation. As noted in the introduction, a concise representation of expected utility in economies of this type has proved challenging to obtain in related work. The key to a concise representation is Corollary 1's link between the risk premium $\mathbb{E}[X_k - P_k]$, the aggregate amount of risk to share, S_k, and the endogenous covariance between returns $X_k - P_k$ and asset cash flows X_k. Looking ahead, a concise representation is important in order to establish the strategic complementarity of participation decisions (Proposition 2), which in turn allows us to take comparative statics in participation costs.

[17] Manzano and Vives (2011) give a mathematical definition of stability. One way to think about stability is in terms of condition (B-7) in the proof of Lemma 3. The right hand side (RHS) describes agents' demands, which in turn depend on price efficiency (this can be seen explicitly from (B-8)). The left hand side (LHS) of (B-7) describes how prices must behave to clear the market, given agents' demands on the RHS. Equilibrium price efficiency is a fixed point of this relation. Moreover, the RHS is increasing in ρ_k, at least in the neighborhood of any solution. If the RHS crosses the 45^o line from below, the corresponding equilibrium is unstable in the following sense: A small upwards perturbation in agents' beliefs about price efficiency affects agents' demands and increases the RHS. To preserve market clearing, this then pushes ρ_k up, and precisely because the RHS crosses the 45^o line from below, the change in ρ_k is greater than the original perturbation in agents' beliefs about ρ_k, i.e., instability.

[18] More generally, the comparative static in Lemma 3 would hold even if agents have ex ante heterogeneous hedging needs, providing that the marginal participating agent adds more trading due to hedging than trading due to information, relative to the average participating agent. See also subsection 8.3.

We define the single-asset analogues of $\mathcal{U}_i^A(e_i)$ and $\mathcal{U}_i^0(e_i)$ by

$$\begin{aligned}
\mathcal{U}_{ik}^A(e_{ik}) &\equiv \mathbb{E}\left[\max_{\theta_{ik}} \mathbb{E}\left[u\left((\theta_{ik}+e_{ik})(X_k-P_k)+e_{ik}P_k\right)|y_i,e_{ik},P_k\right]|e_{ik}\right], \\
\mathcal{U}_{ik}^0(e_{ik}) &\equiv \mathbb{E}\left[u\left(e_{ik}X_k\right)|e_{ik}\right] = -\exp\left(-\gamma e_{ik}\mathbb{E}[X_k] + \frac{\gamma^2 e_{ik}^2}{2\tau_X}\right), \qquad (17)
\end{aligned}$$

i.e., the expected utilities from participation in the market for asset k (active trading), and from not participating in the market for asset k. As before, we write both quantities conditional on the exposure realization e_{ik}, and write $\mathcal{U}_{ik}^A$ exclusive of the participation cost. Non-participation utility $\mathcal{U}_{ik}^0(e_{ik})$ follows from the standard certainty equivalence formula.

Proposition 1 *In a linear equilibrium with price efficiency ρ_k,*

$$\mathcal{U}_{ik}^A(e_{ik}) = \left(d_{ik}(\rho_k)\, D_k(\rho_k)\right)^{-1/2} \exp\left(-\frac{1}{2}\Lambda_k(\rho_k)\left(e_{ik}-S_k\right)^2\right)\mathcal{U}_{ik}^0(e_{ik}), \qquad (18)$$

where

$$\begin{aligned}
d_{ik}(\rho_k) &\equiv \frac{\text{var}\left(X_k|e_{ik},P_k\right)}{\text{var}(X_k|y_i,e_{ik},P_k)}, \qquad (19)\\
D_k(\rho_k) &\equiv \frac{\text{var}\left(X_k-P_k|e_{ik}\right)}{\text{var}\left(X_k|e_{ik},P_k\right)}, \qquad (20)\\
\Lambda_k(\rho_k) &\equiv \frac{\left(\frac{\text{cov}(P_k,e_{ik})}{\text{var}(e_{ik})} + \gamma\text{cov}(X_k-P_k,X_k)\right)^2}{\text{var}(X_k-P_k|e_{ik})}. \qquad (21)
\end{aligned}$$

Moreover, participation n_k affects $d_{ik}(\rho_k)$, $D_k(\rho_k)$, and $\Lambda_k(\rho_k)$ only via price efficiency ρ_k.

In light of Proposition 1, we often write $\mathcal{U}_{ik}^A(e_{ik};\rho_k)$ to make explicit its dependence on price efficiency ρ_k. A participating agent's gain relative to non-participation utility $\mathcal{U}_{ik}^0$ is represented by the benefits Λ_k and D_k, which stem from risk-sharing and are the same for all agents, no matter how precise or imprecise their private information; and d_{ik}, which stems from the the advantages of more precise private information.

We note that the risk-sharing gains are increasing in the absolute difference of an agent's exposure shock e_{ik} relative to the average endowment in the economy, as such agents have more to gain from trade. These risk-sharing gains are also increasing in Λ_k, defined in (21), and a composite of three terms. To interpret these terms, note first that $\text{cov}(X_k - P_k, X_k)$ is positive, since prices do not fully reflect future cash flows (formally, Lemma 1); that $\text{cov}(P_k, e_{ik})$ is negative since prices are decreasing in the aggregate exposure Z_k (formally, Lemma A-6); and that Lemma 2 implies that the combined numerator term $\frac{\text{cov}(P_k,e_{ik})}{\text{var}(e_{ik})}$ +

$\gamma \text{cov}(X_k - P_k, X_k)$ is positive.[19] The three terms in Λ_k have the following interpretation. First, agents' final wealths depend on their exposure shocks e_{ik} only to the extent that they cannot hedge these at the trading stage: from (11), agents undo their risk exposures by trading against them. Thus, they prefer prices to covary as little as possible with their exposures, i.e., for $|\text{cov}(e_{ik}, P_k)|$ to be small. Second, welfare is decreasing in $\text{cov}(P_k, X_k)$, which is closely related to price efficiency ρ_k (in particular, by Lemma 1 it is increasing in price efficiency), capturing the Hirshleifer effect that risk-sharing is hampered when agents have accurate information at the time of trading. Third, welfare is decreasing in the variance of trading profits, $\text{var}(X_k - P_k|e_{ik})$, as one would expect.

Turning to D_k, the denominator $\text{var}(X_k|e_{ik}, P_k)$ is decreasing in price efficiency ρ_k. Loosely speaking, one would also expect the numerator $\text{var}(X_k - P_k|e_{ik})$ to be decreasing in price efficiency, since price efficiency makes P_k more closely related to X_k. The proof of Proposition 2 establishes that the numerator indeed falls, and moreover is the dominant effect, so D_k is decreasing in price efficiency, and hence increasing in participation. The key step in the proof follows from the fact that the demand curve slopes down, even in spite of the informational role of higher prices forecasting higher cash flows (formally, Lemma A-6).

The gains from trading on private information are captured by d_{ik}, which has a form familiar from existing literature. In particular, d_{ik} measures the extent to which observing the private signal y_{ik} improves agent i's forecast of the cash flow X_k, relative to a forecast based only on the price P_k and the agent's private exposure e_{ik}. As one would expect, d_{ik} and hence utility is increasing in the precision of an agent's information, τ_i.

3.4 Strategic complementarity in participation decisions

Next, we show that agents' individual participation decisions exhibit strategic complementarity. As discussed in the introduction, this is the key analytical result in the paper. And as noted earlier, the key step in the proof is Lemma 2's implication that the information in prices affects demand more than the information in exposure shocks.

Proposition 2 *As participation n_k increases, each individual agent's utility from participation $\mathcal{U}_{ik}^A(e_{ik}; \rho_k)$ increases.*

Economically, the key driving force behind strategic complementarity is that, as participation n_k increases, price efficiency ρ_k drops (Lemma 3). Loosely speaking, lower price efficiency increases the amount of risk-sharing that the financial market enables. Specifically, the risk sharing function of the financial market is to enable agents with high idiosyncratic exposures u_{ik} to share cash flow risk $u_{ik}X_k$ with agents with low idiosyncratic exposures.

[19]This last implication is established in (B-25) in the proof of Proposition 2.

Lower price efficiency corresponds to agents having less information about the cash flow X_k, making risk sharing easier to sustain as in Hirshleifer (1971). However, and as discussed in the context of the public information benchmark of subsection 3.1, the risk sharing benefits of less efficient prices must be compared to the potential costs of more volatile prices, since if prices are less efficient, they are relatively more exposed to the aggregate exposure shock Z_k (essentially, the discount rate), and this can easily lead to greater volatility. Proposition 2 establishes that the benefits of lower price efficiency always dominate the potential costs.

4 The effect of declining indexing costs

We are now in a position to address our main question: How does a decline in the cost of indexing, as represented by the parameter κ_1, affect equilibrium outcomes?

Since the same number of agents participate in all the non-index synthetic assets $k \neq 1$, by Lemma 3, price efficiency is the same for all such assets. We write ρ_{-1} for this common level of price efficiency, along with ρ_1 for the price efficiency of the index asset.

We start by explicitly writing the expected utilities for agents who fully participate, who index, and who do not participate. Using the asset-by-asset utility for agents who do not participate in a given asset market, the expected utility of agents who do not participate is:

$$\mathcal{U}_i^0(e_i) = -\left|\prod_{k=1}^{m} \mathcal{U}_{ik}^0(e_{ik})\right|. \tag{22}$$

Similarly, the expected utility of agents who fully participate in financial markets is:

$$\mathcal{U}_i^A(e_i; \rho_1, \rho_{-1}) = -\left|\mathcal{U}_{i1}^A(e_{i1}; \rho_1) \prod_{k=2}^{m} \mathcal{U}_{ik}^A(e_{ik}; \rho_{-1})\right|. \tag{23}$$

The expected utility of indexers is a mixture of these two cases:

$$\mathcal{U}_i^I(e_i; \rho_1) = -\left|\mathcal{U}_{i1}^A(e_{i1}; \rho_1) \prod_{k=2}^{m} \mathcal{U}_{ik}^0(e_{ik})\right|. \tag{24}$$

Relative to "active traders," these agents miss out on the gains from trading assets outside the index ($k = l + 1, \ldots, m$), as well as from trading assets covered by the index in different proportions to their index weights (as represented by non-zero positions in assets $k = 2, \ldots, l$). On the other hand, indexing agents benefit from lower participation costs, $\kappa_1 < \kappa$. (Recall that $\mathcal{U}_i^A$ and $\mathcal{U}_i^I$ are defined as exclusive of participation costs κ and κ_1.)

As we noted earlier, the sets of agents who participate fully, N, and who participate either fully or via indexing, $N \cup N_1$, must both consist of all agents with precision levels below some cutoff. Accordingly, define

$$\begin{aligned} n_1 &\equiv \sup N \cup N_1, \\ n_{-1} &\equiv \sup N. \end{aligned}$$

Hence n_1 is the number of agents who trade the index asset 1, while n_{-1} is the number of agents trading all the remaining assets. Note that certainly $n_1 \geq n_{-1}$, i.e., more agents trade the index asset than any other asset, because all agents who participate either fully or via indexing trade the index asset. The number of agents who index, in the sense of participating *only* via indexing, is $n_1 - n_{-1}$.

In particular, there are two distinct possible types of equilibrium. An *indexing equilibrium* is one in which $n_1 > n_{-1}$; and a *no-indexing equilibrium* is one in which $n_1 = n_{-1}$.

Given these observations, an equilibrium is fully characterized by the values of n_1 and n_{-1}, i.e., by the marginal agents who trade the index asset 1 and other assets $k \neq 1$. Accordingly, we frequently denote a specific equilibrium by (n_1, n_{-1}).

Because of the strategic complementarities established in Proposition 2, participation is self-reinforcing, and so there may simultaneously exist equilibria with high participation levels, and equilibria with low participation levels. Whether such multiplicity in fact arises depends on the distribution of information precisions τ_i, on which we have imposed almost no assumptions. We state our results below allowing for such equilibrium multiplicity.

4.1 The prevalence of indexing

With these preliminaries in place, we can state our main result on how indexing costs affect participation decisions. In particular, we consider what happens as the indexing cost κ_1 falls. This corresponds to falling fees, greater availability, and greater awareness of products such as low-cost index funds and ETFs. It may also reflect an increase in public awareness of the standard advice given by finance academics.

We take this comparative static while leaving the cost of full participation, κ, unchanged; however, our results remain qualitatively unchanged if κ also falls, but by less than κ_1.

As indexing costs κ_1 fall, indexing equilibria are easier to support, and feature more agents indexing and fewer agents fully participating. Conversely, no-indexing equilibria are harder to support.

Proposition 3

(A) For any indexing cost κ_1, at least one equilibrium exists.

(B) As the indexing cost falls, indexing equilibria are easier to support, and feature more indexing. That is, for indexing costs κ_1, κ_1' such that $\kappa_1 < \kappa_1'$:
(i) If an indexing equilibrium exists at κ_1', an indexing equilibrium exists at κ_1 also. Moreover, indexing equilibria at κ_1 feature more total participation, i.e., higher values of n_1.
(ii) If an indexing equilibrium exists at κ_1', the maximum amount of indexing in equilibria at κ_1 exceeds the maximum amount of indexing in equilibria at κ_1'.
(C) As the indexing cost falls, no-indexing equilibria are harder to support. That is, for indexing costs κ_1, κ_1' such that $\kappa_1 < \kappa_1'$, if a no-indexing equilibrium exists at κ_1, a no-indexing equilibrium exists at κ_1' also.

Note that the statement in B(i) should be interpreted in the sense of Milgrom and Roberts (1994): if multiple indexing equilibria exist, the maximum and minimum values of n_1 across such equilibria are higher at the lower index participation cost κ_1.

The economics behind Proposition 3 are relatively straightforward given the strategic complementarity of participation decisions (Proposition 2). As indexing costs κ_1 fall, this directly increases the gain to participation-via-indexing. This in turn raises the gain to other agents of participation-via-indexing, amplifying the original effect. At the same time, the fall in indexing costs κ_1 *raises* the marginal cost $\kappa - \kappa_1$ of full participation, with analogous effects: there is both a direct fall in full participation, which reduces the gain to other agents of fully participating, in turn further reducing full participation. Combining, the amount of indexing increases, with entry at both margins—some people who did not previously participate start trading the index, and some people who previously traded non-index assets switch to trading just the index.

Because agents who switch from full participation to indexing are better informed than the average indexer, some readers may speculate that these marginal indexers would raise the price efficiency of the index and reduce the attractiveness of indexing, contrary to our argument. The reason that this does not happen is that the agents who switch from full participation to indexing were already trading the index under full participation, and so their switch does not affect trade in the index. Instead, the amount of trade in the index is determined entirely by the some-participation/no-participation margin, i.e., n_1.[20]

Given the strategic complementarity of participation decisions, Proposition 3 is largely an application of monotone comparative statics (Milgrom and Roberts (1994)). The main

[20] In a model with wealth effects, it is possible that agents who switch from full participation to indexing would trade the index more aggressively after the switch because they are no longer exposed to the risks associated with trading non-index assets. On the other hand, the inability of these same agents to hedge non-financial exposures that are uncorrelated with the index pushes in the opposite direction. Neither effect arises in our model with CARA utility.

difficulties, which are handled in the formal proof, lie in simultaneously allowing for the possibility of indexing and no-indexing equilibria, and in the fact that a fall in κ_1 simultaneously makes index participation more attractive but full participation less attractive.

4.2 Indexing and price efficiency

As discussed, reductions in indexing costs affect agents' trading decisions both directly, and indirectly because of changes in price efficiency. Indeed, by Proposition 1 the spillover effect of other agents' trading decision can be summarized entirely by price efficiency.

Here, we collect our analysis's implications for how reductions in indexing costs affect price efficiency. A necessary preliminary is to relate the price efficiency of synthetic assets, which is what our analysis makes direct predictions about, to the price efficiency of actual assets. To do so, it is in turn useful to define the *relative price efficiency* of assets j, k by

$$\mathrm{var}(\tilde{X}_j - \tilde{X}_k | \tilde{P}_j - \tilde{P}_k)^{-1}.$$

That is, the relative price efficiency of assets j and k is the extent to which the relative price of assets j and k forecasts the relative future cash flows of this same pair of assets. Empirical papers such as Bai, Philippon, and Savov (2016) estimate relative price efficiency because they include time fixed effects.

Lemma 4 *The relative price efficiency of any pair of assets in the index, and also of any pair of assets outside the index, is measured by ρ_{-1}.*

Index price efficiency, i.e., $\mathrm{var}(X_1|P_1)^{-1}$, is directly measured by ρ_1 (recall (12)). For a broad-based index, such as the S&P 500, index price-efficiency is close to market price efficiency.

The following is then immediate from Lemma 3 and Proposition 3:

Corollary 2 *Let κ_1 be an indexing cost such that an indexing equilibrium exists. If the indexing cost falls, then index price efficiency falls, while relative price efficiency of assets both inside and outside the index rises.*

In the case of multiplicity, Corollary 2 should be interpreted as in B(i) and B(ii) of Proposition 3: across indexing equilibria, the maximum and minimum levels of index price efficiency are lower, and the maximum level of relative price efficiency is higher, for lower values of κ_1.

A separate and basic prediction of our analysis is that the index asset has lower price efficiency than all other synthetic assets, i.e., $\rho_1 < \rho_{-1}$. This prediction maps to a statement about actual assets. Assets covered by the index are linear combinations of the index asset with other synthetic assets. In contrast, assets outside the index coincide with synthetic assets. Intuitively, an "averaging" argument then suggests that assets covered by the index have lower price efficiency. The following result makes this intuition precise.

Lemma 5 *In any indexing equilibrium, the price efficiency of assets covered by the index is strictly lower than the price efficiency of assets not covered by the index, i.e., for j, k such that $j \leq l < k$,* $\operatorname{var}\left(\tilde{X}_j|\tilde{P}_j\right)^{-1} < \operatorname{var}\left(\tilde{X}_k|\tilde{P}_k\right)^{-1}$.

Index inclusion introduces a common component to stock prices. Specifically: The proof of Lemma 5 establishes that the price of assets $k > l$ outside the index takes the form

$$\tilde{P}_k = \text{constant} + b_m \rho_{-1} \tilde{X}_k - b_m \tilde{Z}_k,$$

for some constant b_m, while the price of assets $j \leq l$ covered by the index takes the form

$$\tilde{P}_j = \text{constant} + b_l \tilde{\rho} \tilde{X}_j - b_l \tilde{Z}_j + X_0 + Z_0, \tag{25}$$

for some constant $\tilde{b}_l$, with $\tilde{\rho} < \rho_{-1}$, and random variables X_0 and Z_0 that are independent of both $\tilde{X}_j$ and $\tilde{Z}_j$. The random variables X_0 and Z_0 are the common components that indexing introduces. They are linear combinations of, respectively, cash flows of stocks in the index, $\left\{\tilde{X}_{k'}\right\}_{k'=1}^{l}$, and exposure shocks related to stocks in the index, $\left\{\tilde{Z}_{k'}\right\}_{k'=1}^{l}$.

As discussed in the introduction, one might have conjectured that the fall in the price efficiency of the index associated with lower indexing costs would make indexing *less* attractive for uninformed investors, thereby generating a countervailing force. Instead, our analysis shows that the entry of uninformed investors is self-reinforcing: as more such investors enter, and price efficiency drops, indexing becomes more attractive rather than less, attracting still more uninformed investors.

Conversely, one might have conjectured that the rise in relative price efficiency of assets covered by the index would make individual trades more attractive. Again, this is not the case: this increase makes trading individual assets less attractive for uninformed investors, and so is again self-reinforcing.

4.3 Indexing, market timing, and stock selection

The comparative statics for price efficiency (Corollary 2) in turn imply:

Corollary 3 *Let κ_1 be an indexing cost such that an indexing equilibrium exists. If the indexing cost falls, the expected utility of indexing agents increases. For agents who fully participate, the share of trading gains stemming from trading the index asset increases.*

In the case of multiplicity, Corollary 2 should be interpreted as in B(i) and B(ii) of Proposition 3: across indexing equilibria, the maximum and minimum expected utility of indexing agents is higher for lower values of κ_1, as is the maximum share of trading gains stemming from trading the index asset for fully participating agents.

As the cost κ_1 of indexing falls, the utility of indexing agents increases, reflecting both the direct benefit of the lower cost, and the equilibrium benefit of lower index price efficiency.

In contrast, better informed agents who fully participate are subject to conflicting forces. On the one hand, the decline in index price efficiency makes it easier to make trading profits from the index, via "market timing" trades (see also next subsection, along with Section 5 for empirical references). On the other hand, the increase in relative price efficiency makes it harder to profit from "stock selection" trades; this formalizes the idea that as indexing becomes cheaper, trading individual stocks is increasingly the preserve of relatively informed traders, who consequently struggle to profitably trade against each other.

4.4 Indexing, reversals, and informed trading

A direct implication of market-clearing (4) and agents' trading decisions (11) is that

$$\frac{1}{\gamma}\mathbb{E}\left[X_1 - P_1|P_1\right]\frac{1}{n_1}\int_0^{n_1}\frac{1}{\operatorname{var}\left(X_1|y_{i1}, e_{i1}, P_1\right)} = S_1 + \mathbb{E}\left[Z_1|P_1\right], \tag{26}$$

with analogous identities for other assets. Moreover, $\mathbb{E}\left[Z_1|P_1\right]$ is decreasing in the price P_1, since P_1 is negatively correlated with exposure Z_1 (Lemma A-6). Consequently, our setting exhibits price reversals, with high prices today associated with lower expected returns.[21]

The strength of reversals is captured by the steepness of the negative slope of $\mathbb{E}\left[X_1 - P_1|P_1\right]$, i.e., when this relation is strongly negative, the expected returns following high prices are much lower than following low prices. This is determined by price efficiency, and hence in turn by participation decisions, as shown in the next lemma.

Lemma 6 $\frac{\partial}{\partial P_1}\mathbb{E}\left[X_1 - P_1|P_1\right]$ *is negative, and decreases (i.e., becomes further from* 0*) as price efficiency ρ_1 declines, and hence as the cost of index participation κ_1 decreases.*

[21] Consequently, if investors observe a low price for an asset, and have no exposure to economic shocks, they should take long positions in the asset, since its conditional expected return is high. That is, investors can profit from buying "value" stocks. Although this point is often overlooked, it is nonetheless a standard implication of models of the type we consider here (see, e.g., Biais, Bossaerts, and Spatt, 2010).

High prices are more likely when the average exposure Z_1 is high. Consequently, agents who observe a high asset price are unable to fully infer whether the high price indicates a high future cash flow, or a high value of Z_1 (i.e., high aggregate unwillingness to buy the asset). Although all agents face this inference problem, agents with more precise private signals are better able to resolve it, and to shy away from the asset when future cash flows are in fact low. To express this formally, fix an arbitrary $\hat{n} \in (0, n_1)$, so that agents $[0, \hat{n}]$ correspond to relatively well-informed investors, while agents $[\hat{n}, n_1]$ correspond to relatively uninformed investors. From agents' trades (11), the difference in the average position of well-informed to uninformed investors is given by

$$\frac{1}{\gamma}\left(\frac{1}{\hat{n}}\int_0^{\hat{n}} \tau_i di - \frac{1}{n_1 - \hat{n}}\int_{\hat{n}}^{n_1} \tau_i di\right)(X_1 - P_1). \tag{27}$$

That is, informed investors own a high share of the asset precisely when returns are high, and a low share precisely when returns are low.

Conversely, relatively uninformed investors own a high share of the asset precisely when returns are low. By Lemma 6, it further follows that relatively uninformed investors own a high share when current prices are high. Hence relatively uninformed investors engage in behavior that resembles "trend chasing," and experience lower average returns.

5 Empirical implications

The sharpest predictions of our model concern price efficiency (subsection 4.2). To recap, our analysis predicts that (i) as indexing becomes easier, relative price efficiency rises (Corollary 2), (ii) price efficiency is lower for stocks covered by the index than for those outside it (Lemma 5), and (iii) as indexing becomes easier, the price efficiency of the index as a whole decreases (Corollary 2).

A number of recent empirical papers have studied related predictions, especially in regard to ETFs. With regard to (i), Bai, Philippon, and Savov (2016) and Farboodi, Matray, and Veldkamp (2018) find that relative price efficiency has trended upwards over approximately the last 50 years, broadly the same time period in which indexing has become more prevalent. Over a more recent period, Glosten, Nallareddy, and Zou (forthcoming) document that relative price efficiency increases precisely as ETF ownership of the underlying shares increases.[22]Antoniou et al. (2019) show that an increase in ETF ownership is associated

[22]We also note that Farboodi, Matray, and Veldkamp (2018) additionally show that relative price efficiency has declined for stocks outside the S&P 500, a finding that is inconsistent with our model. Among other things, these authors emphasize the importance of accounting for changes in firm size, which is outside the scope of our analysis. Israeli, Lee, and Sridharan (2017) estimate similar empirical specifications to Glosten,

with a strengthening of the link between a firm's investment and its own stock price; since their empirical analysis includes time fixed effects, this suggests that managers believe that increased ETF ownership increases relative price efficiency, consistent with our model.[23]

With regard to (ii), Coles, Heath, and Ringgenberg (2020) and Bennett, Stulz, and Wang (2020) find that inclusion in the Russell 2000 and SP 500 indices, respectively, is associated with a decline in price efficiency. Farboodi, Matray, and Veldkamp (2018) show that for stocks that have been included in the S&P 500 at some point in time, price efficiency is lower during periods in which they are included. Qin and Singal (2015) find that price efficiency of a stock decreases in a measure of its "indexed ownership." Ben-David, Franzoni, and Moussawi (2018) document that ETF ownership increases stock volatility, consistent with the hypothesis that ETFs enable "liquidity shocks [to] propagate." Also consistent with our analysis, they provide evidence that the increase in volatility is *not* due to increased price efficiency. Similarly, Brogaard, Ringgenberg, and Sovich (forthcoming) show that firms that use commodities covered by leading commodity indices make less efficient production decisions than firms that use commodities outside these indices.

While prediction (iii) is harder to directly test, a closely related prediction is that, as indexing increases, informed trading profits will stem increasingly from "timing" strategies based on the entire index, rather than individual asset trades. (See Corollary 3, with the caveat that expected utilities are distinct objects from expected profits.) There is at least some empirical evidence supporting this prediction. AQR document that the correlation between hedge fund returns and market returns has risen from 0.6 to 0.9 over the last two decades.[24] Related, Stambaugh (2014) documents a decline in asset-selection strategies by active mutual funds over the same period. Also related, and using data since 2000, Gerakos, Linnainmaa, and Morse (2019) show that a significant fraction of returns generated by active mutual funds stem from market timing strategies.

Our model can also be used to study the effect of index inclusion on return variances and covariances. For example, and as one would expect, index inclusion introduces a common component to stock prices; see discussion immediately following Lemma 5. This is consistent with evidence in Vijh (1994), Barberis, Shleifer, and Wurgler (2005), Greenwood (2008), and Bennett, Stulz, and Wang (2020); and related, Da and Shive (2018) and Leippold, Su,

Nallareddy, and Zou (forthcoming), but use lagged changes in ETF ownership, and show that these are associated with decreases rather than increases in relative price efficiency.

[23]Section 7 below formally incorporates managerial learning from prices. Antoniou et al. (2019) further decompose stock prices into systematic and idiosyncratic components, where the systematic component captures differences explained by time and industry effects. They find that increased ETF ownership has a greater effect on the relation between investment and the systematic component of cash flows. This is consistent with an asset k in our model corresponding to an industry as opposed to an individual firm.

[24]See "Hedge fund correlation risk alarms investors," Financial Times, June 29th, 2014.

and Ziegler (2016) present evidence that individual stock correlations have risen with ETF ownership. As noted, Ben-David, Franzoni, and Moussawi (2018) present evidence that ETF ownership raises the volatility of individual stock prices.

The analysis in subsection 4.4 predicts that relatively uninformed ownership of an asset increases when prices are high, and that this is followed by low subsequent returns. This is consistent with the empirical evidence in Ben-Rephael, Kandel, and Wohl (2012) for the market as a whole (i.e., the index asset), and in Jiang, Verbeek, and Wan (2017) and Grinblatt et al. (2020) in the cross-section.[25] Furthermore, Lemma 6 is consistent with the empirical findings of Baltussen, van Bekkum, and Da (2019), who use international data to document that negative serial correlation in index returns is associated with greater indexing, and with Ben-David, Franzoni, and Moussawi (2018) who find that "ETF ownership increases the negative autocorrelation in stock prices."

Finally, it is worth noting that our model does *not* generate a price premium for index inclusion, contrary to at least some empirical evidence. The reason is that changes in participation in our model are always accompanied by changes in the supply of the asset being traded, since we assume all agents have equal initial endowments.[26] Because of this, index inclusion affects prices only via its effect on price efficiency; and since price efficiency falls, prices fall also. In contrast, if one were to relax the assumption of equal endowments, the effect on prices would be determined by the relative strength of changes in price efficiency and changes in the aggregate risk-sharing capacity of participating agents. In general, we emphasize again that our model's primary empirical predictions operate via price efficiency.

6 Direct versus equilibrium effects

The consequences of a decline in indexing costs stem from a combination of direct effects (e.g., more people index because of lower costs) and equilibrium effects (e.g., the direct increase in indexing in turn makes indexing even more attractive). In this section, we discuss the relative contributions of the direct and equilibrium effects.

[25]Less directly, this prediction is also consistent with the finding in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014) that the fraction of mutual fund returns stemming from timing strategies is greater in recessions. Specifically, in our setting informed investors should shift into the index when index prices are low, and out of the index when index prices are high. To the extent to which the second half of this strategy is constrained by difficulties shorting the index, this generates the findings of Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014). (We should also note that the same authors suggest a distinct explanation in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016).)

[26]Recall, in turn, that we use this assumption in the proof of Proposition 1. The empirical evidence on the effects of index inclusion is mixed; while older studies found positive price reactions, Bennett, Stulz, and Wang (2020) find that in more recent data "the positive announcement effect on the stock price of [S&P 500] index inclusion has disappeared."

6.1 Spillover effects of changes in informedness

We have focused on the comparative static associated with indexing costs κ_1. Naturally, however, there are other interesting questions that one could pursue using our framework. Here, we examine the consequences of an increase in the informedness of the most informed agents in the economy. The idea that the most informed agents have grown more informed has been suggested by a number of authors (e.g., Glode, Green, and Lowery (2012), Fishman and Parker (2015), Bolton, Santos, and Scheinkman (2016); closely related is the idea that "big data" and data processing have grown in importance, e.g., Farboodi, Matray, and Veldkamp (2018)). This exercise has the virtue of highlighting the role of equilibrium effects; in particular, since participation costs are unchanged, there is no direct effect.

Concretely, we consider a parameterization of our economy in which an indexing equilibrium (n_1, n_{-1}) exists, with $n_{-1} > 0$ so that some agents fully participate and actively trade assets other than the index. We then consider the consequences of increasing the precision of signals of the most informed agents $[0, \hat{n}]$, where $\hat{n} < n_{-1}$.

This increase in the informedness of the most informed agents raises the price efficiency of both index and non-index assets. Consequently, fewer agents fully participate, and fewer agents participate in financial markets (either fully or via indexing). In other words, even though an increase in signal precisions makes prices more informative, it also makes financial markets less beneficial. Hence our analysis formalizes the channel via which an increase in sophistication of financial specialists places "ordinary" investors at an increasing disadvantage (see references above). It also speaks to the older question of whether insider trading is socially costly (e.g., Fishman and Hagerty (1992), Khanna, Slezak, and Bradley (1994)).

6.2 Numerical decomposition of direct versus equilibrium effects

We now return to our main comparative static of considering a drop in indexing costs κ_1. We again start from an indexing equilibrium (n_1, n_{-1}) with $n_{-1} > 0$, and associated equilibrium price efficiency levels ρ_1 and ρ_{-1}. A fall in κ_1 leads to a new equilibrium with more total participation (n_1 higher) and less full participation (n_{-1} lower); see Proposition 3. This equilibrium effect can be decomposed into the *direct* effect associated with the change in cost κ_1, which we define as agents' participation decisions under the new costs, but assuming that price efficiency remains fixed at the original level (ρ_1, ρ_{-1}); and the remaining change, which by definition is the *equilibrium* effect.

We illustrate the relative size of these effects via a numerical example. We emphasize that the example is for illustrative purposes only; a realistic calibration is beyond the scope of this paper, and in any case, for such a calibration one would want to consider a dynamic version

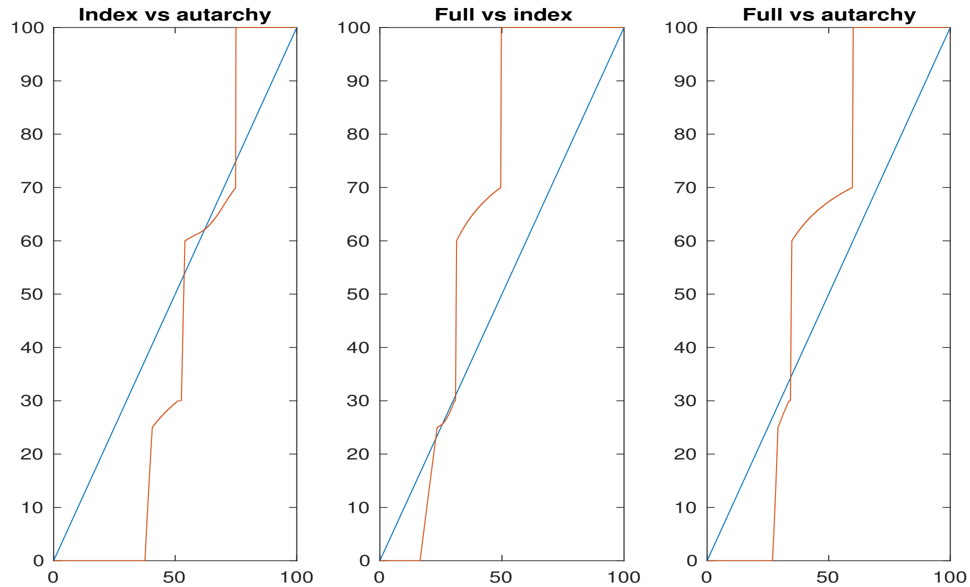

Figure 1: Equilibrium participation

of our model. Appendix D details the parameter values used. We have chosen values—especially for signal precisions τ_i and participation costs κ and κ_1—so that equilibria with interior levels of participation exist. In particular, the example has three groups of agents, with minimal intra-group variation: agents with high precision signals, agents with medium precision signals, and almost-uninformed agents. This simple setting delivers equilibria in which strict subsets of agents fully participate, index, and do not participate.

As a first step, Figure 1 shows the equilibria of our economy. The lefthand panel plots the following: for a candidate fraction of agents who trade the index asset (the horizontal axis), find the price efficiency of the index asset (Lemma 3), and then find the fraction of agents who prefer indexing to autarchy. Stable equilibria correspond to points where this function crosses the 45^o line from above. The middle and righthand plots show the corresponding functions for the choice of full participation versus indexing, and the choice of full participation versus autarchy.

One point that is immediately apparent from Figure 1 is that the strategic complementarity effect is strong, and multiple equilibria exist. This arises in all numerical examples that we examined. Specifically, there are four stable equilibria: No-one participates ($n_1 = n_{-1} = 0$); everyone fully participates ($n_1 = n_{-1} = 1$); $n_1 = 62.14\%$ index and no-one fully participates ($n_{-1} = 0$); and $n_1 = 62.14\%$ participate with $n_{-1} = 25.65\%$ fully participating, so that 36.49% index. We focus on the last of these equilibria. (We do not see a good theoretical argument for selecting among these equilibria; and the one we focus on matches the empirical fact that some agents trade actively, some index, and some do not participate.)

	Participation under initial costs κ, κ_1	Participation given reduced κ_1	Change due to direct effect	Change due to equilibrium effect
n_1	62.14%	63.44%	+0.77%	+0.53%
n_{-1}	25.65%	24.13%	−0.30%	−0.22%

Table 1: Decomposition into direct and equiulibrium effects.

Next, we consider a small 1% decrease in the cost of indexing κ_1. Table 1 displays the results of the decomposition described above. The decomposition shows that the direct and equilibrium effects are of the same order of magnitude, and that both are sizable (the total elasticity with respect to cost changes is approximately 2).

7 Indexing and firm performance: Feedback effects

The rise of index investing has generated a variety of concerns about the effect of indexing on firm performance, including, for example, the fears that indexing investors spend less effort on firm governance, or that indexing results in extensive common ownership, reducing competition between firms, thereby reducing consumer surplus.[27] We have deliberately focused our analysis on indexing's effects in an endowment economy, in order to clearly delineate what we believe is an important channel.

The central economic force in our model is that indexing costs affect participation, which affects price efficiency, which in turn affects participation. As such, the link between indexing and firm performance that our analysis can best speak to is the possible effect of prices on firms' real decisions, a channel often referred to as the "feedback effect" (see Bond, Goldstein, and Edmans (2012) for a survey). To this end, in this section we extend our baseline model to allow for feedback effects. The extension is rich enough to embody two key effects, yet simple enough to remain tractable.[28]

[27]For a discussion of the first point, along with some evidence against, see, for example, Appel, Gormley, and Keim (2016). For a discussion of the second point, see, for example, Azar, Schmalz, and Tecu (2018) and Schmalz (2018).

[28]A standard tractability hurdle in feedback models is to make sure that the distribution of the endogenous cash flow is such that agents' portfolio decisions generate prices that allow for a tractable treatment of agents' updating from prices. Papers such as Sockin and Xiong (2015), Bond and Goldstein (2015), Sockin (2018), and Goldstein and Yang (2019) all contain competitive models of financial markets with asymmetric information in which economic agents extract information from financial prices and use this information to affect firm cash flows. On top of this standard hurdle, we also need to ensure that it remains feasible to characterize participation decisions (Proposition 2, in turn building on Proposition 1). The issue here is that Corollary 1—which, as we have noted, is key to characterizing expected utilities and hence participation—does not generalize to settings with feedback effects. The reason that our analysis in this section is nonetheless tractable is that it is one in which the feedback effects shift index cash flows only by a constant; the effects on long-short portfolios (e.g., synthetic asset 2) are more complicated, but these assets are in zero net supply,

7.1 Extended model

There are two assets covered by the index, $m = l = 2$. (These two assets can be interpreted as covering broad sectors of the economy.) We denote by $\tilde{V}_k$ the cash flow that is produced by each of assets $k = 1, 2$. This cash flow is determined by a mixture of exogenous and endogenous factors. Specifically, let $\eta \geq 0$ be a constant; the endogenous components of cash flows, which we denote by $\tilde{\Xi}_1$ and $\tilde{\Xi}_2$, are determined by

$$\tilde{\Xi}_1 = \arg\max_{\xi} \mathbb{E}\left[\xi - \exp\left(\xi - \eta\tilde{X}_1 + \eta\tilde{X}_2\right) | \tilde{P}_1, \tilde{P}_2\right] \quad (28)$$

$$\tilde{\Xi}_2 = \arg\max_{\xi} \mathbb{E}\left[\xi - \exp\left(\xi - \eta\tilde{X}_2 + \eta\tilde{X}_1\right) | \tilde{P}_1, \tilde{P}_2\right], \quad (29)$$

where $\tilde{X}_k$ continues to denote an exogenous and randomly distributed random variable; and the total cash flow of each asset $k = 1, 2$ is

$$\tilde{V}_k = \tilde{X}_k + \tilde{\Xi}_k.$$

In words: $\tilde{X}_k$ is an exogenous shock to the business environment of firm (or sector) k, and $\tilde{\Xi}_k$ is the effect of the actions of firm k's executives on cash flows. These actions—for example, firm expansion, or the consumption of private benefits—are chosen prior to the realization of exogenous business environment shocks $\tilde{X}_k$; but the realized net benefit or cost of these actions to executives depends on the realization of the business environment. In particular, expansion of firm 1 has a lower marginal cost for firm 1's executives when firm 1's business environment is more favorable, or when firm 2's business environment is less favorable.[29]

Feedback effects arise in this setting because firms' executives seek to learn about the future business environment, $\tilde{X}_1$, $\tilde{X}_2$, from the financial asset prices $\tilde{P}_1$, $\tilde{P}_2$. Note that by setting the parameter η to 0, executives' decisions are independent of expectations about $\tilde{X}_1, \tilde{X}_2$, and hence independent of prices $\tilde{P}_1$, $\tilde{P}_2$, so that this framework nests our baseline model of exogenous cash flows as a special case.

Finally, an agent's non-financial income continues to be given by (1), and in particular, is unaffected by firms' endogenous decisions.

7.2 Endogenous cash flows

It is straightforward to show that there is an equilibrium in which prices are linear functions of exogenous business environment shocks $\tilde{X}_1$, $\tilde{X}_2$ and aggregate exposure shocks $\tilde{Z}_1$, $\tilde{Z}_2$.

and for zero net supply assets Corollary 1 trivially generalizes.

[29]The use of the exponential function as the cost function in (28) and (29) is not essential; see Bond and Goldstein (2015).

Given this, executives' posteriors of $\tilde{X}_1$, $\tilde{X}_2$ are normal, and (28), (29) rewrite as

$$\tilde{\Xi}_k = \arg\max_{\xi} \left(\xi - \exp\left(\xi - \eta \mathbb{E}\left[\tilde{X}_k - \tilde{X}_{-k} | \tilde{P}_1, \tilde{P}_2 \right] + \frac{1}{2}\eta^2 var\left(\tilde{X}_k - \tilde{X}_{-k} | \tilde{P}_1, \tilde{P}_2 \right) \right) \right),$$

where $\tilde{X}_{-k}$ denotes the business environment $j \neq k$. Hence

$$\tilde{\Xi}_k = \eta \mathbb{E}\left[\tilde{X}_k - \tilde{X}_{-k} | \tilde{P}_1, \tilde{P}_2 \right] - \frac{1}{2}\eta^2 var\left(\tilde{X}_k - \tilde{X}_{-k} | \tilde{P}_1, \tilde{P}_2 \right).$$

We again analyze the economy in terms of synthetic assets, defined exactly as before. Since there are two underlying assets there are two synthetic assets, which we refer to as the index and spread asset, having cash flows $V_1 = X_1 + \Xi_1$ and $V_2 = X_2 + \Xi_2$, where Ξ_1 and Ξ_2 denote the endogenous components of the synthetic asset cash flows:

$$\Xi_1 = -\eta^2 var\left(\tilde{X}_1 - \tilde{X}_2 | P_1, P_2 \right) = -2\eta^2 var\left(X_2 | P_2 \right) \quad (30)$$

$$\Xi_2 = 2\eta \mathbb{E}\left[\tilde{X}_1 - \tilde{X}_2 | P_1, P_2 \right] = 2^{\frac{3}{2}} \eta \mathbb{E}\left[X_2 | P_2 \right]. \quad (31)$$

Note that the price P_1 of the index asset contains no information about the *difference* in business environments X_2, as reflected in (30) and (31).

Feedback affects the index cash flow by a term that is constant in equilibrium, $-2\eta^2 var\left(X_2 | P_2 \right)$. This term captures the intuitive idea that cash flows are higher when price efficiency is greater, so that $var\left(X_2 | P_2 \right)$ is lower. It is the spread asset's price efficiency that matters, i.e., relative price efficiency.

In contrast, the unconditional mean of spread asset's cash flow V_2 is 0, regardless of either the strength of feedback η or of price efficiency. Instead, feedback in the spread asset implies that its cash flow responds to price realizations, increasing the volatility of the spread asset's cash flow. This effect is stronger when price efficiency is greater and hence the conditional expectation $\mathbb{E}\left[X_2 | P_2 \right]$ is more volatile.

7.3 Equilibrium

The endogenous components of the cash flows, Ξ_k, are deterministic functions of prices P_k. One can see this directly from (28), (29), without appealing to the specific expressions in (30), (31). Given this, it is convenient to analyze the economy in terms of the prices of assets that pay only the exogenous cash flows $V_k - \Xi_k = X_k$. We denote the price of these assets by P_k^X, and note that, precisely because Ξ_k is deterministic given prices,

$$P_k^X = P_k - \Xi_k.$$

Hence agent i's terminal wealth given trades θ_i is

$$W_i = \sum_{k=1}^{2} \theta_{ik} \left(X_k - P_k^X \right) + S_k \left(X_k + \Xi_k \right) + \left(Z_k + u_{ik} \right) X_k$$
$$= \sum_{k=1}^{2} \theta_{ik} \left(X_k - P_k^X \right) + e_{ik} X_k + S_k \Xi_k, \tag{32}$$

where per capita endowments are $S_1 = \sqrt{2}\tilde{S}$ and $S_2 = 0$ for the index and spread assets.

An immediate implication of these observations is that, given participation decisions, equilibrium price efficiency is exactly as before (Lemma 3).

The combination of (30), (31) and (32) means that participation decisions in the feedback economy are no harder to analyze than in our baseline case without feedback ($\eta = 0$). For the index asset, feedback adds a cash flow component Ξ_1 that is constant in equilibrium, and is received both by agents who trade the index asset and those who do not participate ($\theta_{i1} = 0$) and simply hold their original endowment. Consequently, feedback has no effect on participation decisions for the index asset.

For the spread asset, feedback adds a cash flow component Ξ_2 that fluctuates with the price of the spread asset P_2. But because the spread asset is in zero supply, $S_2 = 0$, this endogenous component does not affect participation decisions, as (32) makes clear.

7.4 Indexing costs and firm performance

With the equilibrium characterization in hand, we revisit our central topic, viz, the equilibrium effects of a falling cost of indexing. As noted, price efficiency and participation decisions are unchanged from the baseline case. Nonetheless, feedback introduces two new effects.

First, as indexing costs κ_1 fall, the increase in the price efficiency of the spread asset reduces uncertainty about business conditions $\tilde{X}_1$ and $\tilde{X}_2$ inducing executives to work harder (or consume fewer private benefits), thereby increasing the cash flow of firms, and hence of the index asset. This creates an additional welfare benefit for all agents in the economy who are endowed with a long position in the underlying assets (in our model, this is all agents).

Second, as indexing costs κ_1 fall, the increase in the price efficiency of the spread asset increases the sensitivity of cash flows to the price of the spread asset P_2, consistent with the empirical findings in Antoniou et al. (2019) and Bennett, Stulz, and Wang (2020);[30] and since the equilibrium price P_2^X of the exogenous part of cash flows coincides with the baseline case, it follows that cash flows grow more volatile. Nonetheless, this increase in cash flow

[30]Here, we are viewing empirically-measured investment as a proxy for the decisions $\tilde{\Xi}_k$; our analysis predicts that the sensitivity of the relative decision Ξ_2 to the price difference P_2 rises with indexing.

volatility leaves welfare unchanged, since returns $V_k - P_k = X_k - P_k^X$ are unaffected, and agents have no initial endowment of the spread asset (see (32)).[31]

8 Discussion

8.1 Multiple indices

So far, we have focused on the case of a single index, which we have typically interpreted as a broad-based market index such as the S&P500. In reality, a large number of indices co-exist, and investment vehicles such as ETFs are increasingly available on this assortment of indices (see, for example, Lettau and Madhaven (2018) and Easley et al. (2020)).

We can straightforwardly extend our analysis to cover multiple indices, at least when the indices are orthogonal to each other. To fix ideas, we consider here the case of two indices: the first a broad-based market index, corresponding to synthetic asset 1; and the second an index corresponding to the long-short positions associated with synthetic asset 2, which by construction is uncorrelated with synthetic asset 1. For example, the second index might be an index tracking the value factor, which is constructed as a long-short portfolio, and which indeed has low correlation with the overall market. Moreover, there are a number of ETFs that track the value factor.

The participation cost of participating via the market index is κ_1, and the participation cost of participating via the value factor is κ_2. Hence an agent now has two additional participation options—participation via just the value factor, and participation via the combination of the market index and value factor—along with the three possibilities considered so far (no participation, market index, full participation).

A fall in the cost κ_2 of participating via the value factor has effects that are analogous to those analyzed in the main body of the paper. As this cost falls, more agents trade the value factor, decreasing the price efficiency of the value factor, thereby further increasing the attractiveness of the value factor. In a typical case, the increased attractiveness of participation via value factor strategies reduces the number of agents who fully participate in financial markets, increasing the price efficiency of assets not covered by either the market index of the value factor, and also of synthetic assets other than those corresponding to the market index and value factor, and that may correspond to other factors. In particular, increased trading of the value factor is accompanied by an increase in the number of people

[31] As noted from the outset, the particular framework we have analyzed is special in significant respects, and other specifications of feedback would potentially yield different implications for a reduction in indexing costs. But the framework we use has the advantages of being highly tractable, and allowing us to highlight two specific consequences of extending our model to allow prices to affect cash flows.

who market index, in the sense of trading the market index asset but not individual assets. Similarly, a decline in the cost of market indexing typically increases the number of people trading the value factor but not individual stocks.

8.2 Intensive margin of information acquisition

While the participation decision in our model can be thought of as an information acquisition decision on the extensive margin, we have abstracted away from the intensive margin of agents' information acquisition efforts. Here, we briefly consider an extension that features an intensive margin. For simplicity, we focus on the case of two assets, $l = m = 2$, so that the synthetic assets are simply the index and spread asset.

Suppose now that, instead of being exogenously endowed with informative signals about cash flows y_{ik}, agents spend resources to collect information. Specifically, for constants $c_i > 0$ and ψ, in order to collect signals about the index asset 1 and spread asset 2 with precisions τ_{i1} and τ_{i2}, respectively, agent i incurs a cost

$$c_i \left(\tau_{i1}^2 + \tau_{i2}^2 + \psi \tau_{i1} \tau_{i2} \right). \tag{33}$$

Agent i faces convex costs of increasing precision, and these costs vary across agents. We assume that agents choose precisions at the same time that they make participation decisions.

The cost specification (33) says that agents collect two types of information: information about market aggregates, y_{i1}, and information about relative cash flows, y_{i2}. We believe this assumption is realistic; it also has the important advantage of allowing us to continue to analyze the economy in terms of the synthetic index and spread assets.[32]

The parameter ψ in (33) captures the multitasking dimension of information acquisition. In particular, if $\psi > 0$, then an increase in the precision of the signal about X_1 increases the the marginal cost of precision about X_2, and vice versa.

Absent multitasking considerations (i.e., $\psi = 0$), the above model generates qualitatively the same implications as our main model. Specifically, a fall in the cost of indexing κ_1 has the direct effect of decreasing participation in the spread asset and increasing participation in the index asset. If precision decisions are artificially held fixed, these changes in participation decrease price efficiency in the index asset and increase price efficiency in the spread asset. Consequently, in equilibrium agents who trade the spread asset reduce the precisions of their spread asset signals, while agents who trade the index asset increase the precision of their

[32] In contrast, if we instead assumed that agents can choose precisions of signals about the underlying cash flows $\tilde{X}_1$ and $\tilde{X}_2$, then the change-of-basis to synthetic assets is much less useful. In a different context, Van Nieuwerburgh and Veldkamp (2010) establish conditions under which investors would indeed want to specialize and focus their information acquisition efforts on a small number of assets.

index asset signals. These endogenous changes in precision dampen the effect of the fall in indexing costs, but do not qualitatively change the predictions, since the net effect is still that equilibrium price efficiency rises in the spread asset and falls in the index asset.[33]

Multitasking considerations—and specifically, the case in which precision of one signal increases the marginal cost of the other signal, i.e., $\psi > 0$—introduce an additional effect, as follows. To isolate the effect in a transparent way, we consider the case in which κ_1 is sufficiently low that all agents trade the index asset, and a strict subset of agents trade spread asset. In this case, a further reduction in κ_1 reduces the number of agents trading the spread asset; but does not affect the number of agents trading the index asset, since everyone already trades it. So if precision decisions are artificially held fixed, the price efficiency of the spread asset increases, while the price efficiency of the index asset is unchanged.

The new effect introduced by multitasking is that agents respond by increasing the precision of signals about the index asset. Agents who trade the spread asset respond to the increase in its price efficiency by reducing the precision of signals about the spread asset, and—because of multitasking—increasing the precision of signals about the index asset. Agents who previously traded the spread asset, but no longer do so after the fall in indexing costs, likewise respond by increasing the precision of signals about the index asset, again for multitasking reasons. Consequently, and different from our baseline model, price efficiency of the index asset rises rather than falls. In this case, the fall in the cost of indexing generates indirect effects that reduce the benefits of trading the index asset—though, of course, the fall in indexing costs still carries a direct benefit.

As we noted, this multitasking effect is starkest in the case in which a fall in indexing costs does not increase the number of agents trading the index asset. The case in which both effects operate is more complicated, and the net effect is a quantitative question that we leave for future research.

8.3 Revisiting the link between participation and price efficiency

The substantive implications of our analysis are all contingent on Lemma 3's implication that the price efficiency of an asset is *decreasing* in the level of participation in that asset. Recall that this implication arises because the marginal new participating agent is less informed than the average participating agent, but has the same non-informational trading motives. Consequently, the marginal new participating agent adds proportionately more "noise" than

[33] In particular, the comparative static of Lemma 3 continues to hold, as follows. Consider an increase in participation n_k. Suppose to the contrary that price efficiency ρ_k increases. But then agents who were already participating lower their signal precisions. So price efficiency falls, both because of the increase in participation and because of the lower signal precisions. The contradiction establishes that price efficiency is decreasing in participation.

information to the market. Moreover, the empirical evidence surveyed in Section 5 provides at least some support for the idea that price efficiency is decreasing participation.

Nonetheless, it is important to acknowledge that there are plausible economic forces that would predict that price efficiency is instead increasing in participation. We discuss two extensions of our basic model to conclude our analysis.

Heterogenous risk aversion: If agents have differing levels of risk-aversion, γ_i, it is straightforward to show that equilibrium price efficiency of asset k is given by the smaller root of the quadratic

$$\frac{\rho_k^2 \tau_u}{|N_k|} \int_{N_k} \frac{1}{\gamma_i} di - \rho_k + \frac{1}{|N_k|} \int_{N_k} \frac{\tau_i}{\gamma_i} di = 0, \tag{34}$$

where N_k is the set of agents participating in asset k. Hence a sufficient condition for price efficiency to decrease in participation (as in Lemma 3) is that both $\frac{1}{|N_k|} \int_{N_k} \frac{1}{\gamma_i} di$ and $\frac{1}{|N_k|} \int_{N_k} \frac{\tau_i}{\gamma_i} di$ decrease in participation. Similarly, a sufficient condition for price efficiency to increase in participation is that both $\frac{1}{|N_k|} \int_{N_k} \frac{1}{\gamma_i} di$ and $\frac{1}{|N_k|} \int_{N_k} \frac{\tau_i}{\gamma_i} di$ increase in participation.

In particular, if agents differ only in risk aversion, while having the same signal precisions, then price efficiency is increasing in participation if the most risk-averse agents gain most from participation. While, unfortunately, participation decisions are much harder to evaluate analytically when risk aversion is heterogeneous, it is intuitive that participation gains are indeed increasing in risk aversion, and numerical analysis supports this intuition.[34]More generally, our analysis suggests that price efficiency is decreasing in participation if the main source of heterogeneity among agents is differing signal precisions; but increasing in participation if instead the main source of heterogeneity is differing risk aversion.

Quadratic trading costs in place of fixed participation costs: Our main model assumes that agents pay fixed costs (κ and κ_1) to participate in financial markets. An alternative modeling approach would be to instead assume that agents incur costs at the time of trading. To fix ideas, we consider here the highly tractable case in which trading costs take the form $\frac{1}{2}\hat{\kappa}\theta_{ik}^2$, for some constant $\hat{\kappa} > 0$, which is potentially smaller for trades of the index asset.[35]

[34]See online appendix for expressions used in the numerical calculation. The parameter values used are the same as in subsection 6.2, and detailed in Appendix D, other than for risk aversion γ_i, which is distributed uniformly over $\left[\frac{27}{50}, \frac{33}{50}\right]$. The main impediment to an analytical evaluation of participation is that the generalization of Corollary 1 delivers $\mathbb{E}\left[X - P\right] = \frac{S\text{cov}(X-P,X)}{\frac{1}{|N|}\int_N \frac{1}{\gamma_i} di}$; it then follows that the substitution of Corollary 1 in the proof of Proposition 1 is useful only for an agent whose risk aversion γ_i matches the harmonic mean of participating agents.

[35]In an economy in which each agent can observe the number of agents with a given signal precision who buy (respectively, sell) the asset, Dávila and Parlatore (forthcoming) obtain results for linear and fixed transaction costs.

Naturally, transaction costs reduce trade sizes, with (11) generalizing to

$$\theta_{ik} = \frac{\mathbb{E}\left[X_k | y_{ik}, e_{ik}, P_k\right] - P_k - e_{ik}\gamma \mathrm{var}\left(X_k | y_{ik}, e_{ik}, P_k\right)}{\gamma \mathrm{var}\left(X_k | y_{ik}, e_{ik}, P_k\right) + \hat{\kappa}}.$$

Under this specification of costs, all agents participate in all markets, and costs affect the intensive rather than extensive margin of participation. The relation between price efficiency and the magnitude of the transaction cost parameter $\hat{\kappa}$ depends on the relative effect of the transaction cost on trade sensitivity to cash flow signals y_{ik} and exposure shocks e_{ik}, along with how this relative sensitivity varies across agents with differing signal precisions.

Dávila and Parlatore (forthcoming) analyze exactly this question in a closely related economy, and find that the relation depends on the specific form of heterogeneity across agents. In particular, they examine a model in which agents have heterogeneous prior beliefs about cash flows,[36] and also observe subsequent signals of cash flows. They find that price efficiency is decreasing in the transaction cost parameter if agents differ in their signal precisions; but is increasing if instead agents differ in the accuracy of their prior beliefs.[37]

9 Conclusion

We develop a benchmark model to study the equilibrium consequences of indexing in a standard rational expectations setting (Grossman and Stiglitz (1980); Hellwig (1980); Diamond and Verrecchia (1981)). Individuals must incur costs to participate in financial markets, and these costs are lower for individuals who restrict themselves to indexing strategies. Individuals' participation decisions exhibit strategic complementarity, and consequently, equilibrium effects reinforce the direct consequences of declining costs of indexing. As indexing becomes cheaper we find that: (1) indexing increases, while individual stock trading decreases; (2) aggregate price efficiency falls, while relative price efficiency increases; (3) the welfare of relatively uninformed traders increases; (4) for well-informed traders, the share of trading gains stemming from market timing increases, and the share of gains from stock selection decreases; (5) market-wide reversals become more pronounced. We link these predictions to existing empirical evidence.

We end on a more speculative note, by noting two ways in which a fall in indexing costs reduces inequality stemming from public financial markets in our framework.[38] First,

[36] In Dávila and Parlatore's model, agents do not have exposure to non-financial cash flows ($Z + u_i$ in our notation) and instead trade because of heterogeneous prior beliefs.

[37] For completeness, in an online appendix we show that the former result relating to signal precisions holds in our model also, at least for the case in which the population distribution of signal precisions is binary.

[38] Mihet (2020) examines how inequality is affected by changes in participation and information acquisition

the switch away from active investing decreases both the mean and variance of investment returns for the best informed agents who continue to actively invest. Second, increased participation by relatively uninformed agents reduces the income variance of these agents. Further work on how financial innovations affect inequality and welfare is a natural avenue for future research.

References

Admati, A. R. 1985. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica* 53(3):629–57.

Admati, A. R., and P. Pfleiderer. 1988. A theory of intraday patterns: Volume and price variability. *Review of Financial Studies* 1:3–40.

———. 1997. Does it all add up? Benchmarks and the compensation of active portfolio managers. *The Journal of Business* 70:323–50.

Antoniou, C., F. W. Li, X. Liu, A. Subrahmanyam, and C. Sun. 2019. The real effects of exchange-traded funds. *Working Paper* .

Appel, I., T. Gormley, and D. Keim. 2016. Passive investors, not passive owners. *Journal of Financial Economics* 111–41.

Azar, J., M. Schmalz, and I. Tecu. 2018. Anticompetitive effects of common ownership. *Journal of Finance* 73:1513–65.

Bai, J., T. Philippon, and A. Savov. 2016. Have financial markets become more informative? *Journal of Financial Economics* 122:625–54.

Baltussen, G., S. van Bekkum, and Z. Da. 2019. Indexing and stock market serial dependence around the world. *Journal of Financial Economics* 132:26–48.

Barberis, N., A. Shleifer, and J. Wurgler. 2005. Comovement. *Journal of Financial Economics* 75:283–317.

Baruch, S., and X. Zhang. 2018. Is index trading benign? *Working Paper* .

Basak, S., and A. Pavlova. 2013. Asset prices and institutional investors. *American Economic Review* 103:1728–58.

Baumol, W. J. 1965. *The stock market and economic efficiency.* New York: Fordham University Press.

Ben-David, I., F. Franzoni, and R. Moussawi. 2018. Do ETFs increase volatility? *Journal of Finance* 73:2471–535.

costs, using a model close to Peress (2005).

Ben-Rephael, A., S. Kandel, and A. Wohl. 2012. Measuring investor sentiment with mutual fund flows. *Journal of Financial Economics* 363–82.

Bennett, B., R. M. Stulz, and Z. Wang. 2020. Does joining the S&P 500 index hurt firms? *Working Paper* .

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

Biais, B., P. Bossaerts, and C. Spatt. 2010. Equilibrium asset pricing and portfolio choice under asymmetric information. *Review of Financial Studies* 23:1503–43.

Bolton, P., T. Santos, and J. A. Scheinkman. 2016. Cream skimming in financial markets. *Journal of Finance* 71:709–36.

Bond, P., and I. Goldstein. 2015. Government intervention and information aggregation by prices. *Journal of Finance* 2777–812.

Bond, P., I. Goldstein, and A. Edmans. 2012. The real effects of financial markets. *Annual Review of Financial Economics* 339–60.

Breugem, M., and A. Buss. 2018. Institutional investors and information acquisition: Implications for asset prices and informational efficiency. *Review of Financial Studies* 32:2260–301.

Brogaard, J., M. C. Ringgenberg, and D. Sovich. forthcoming. The economic impact of index investing. *Review of Financial Studies* .

Buffa, A. M., D. Vayanos, and P. Woolley. 2019. Asset management contracts and equilibrium prices. *Working Paper* .

Chowdhry, B., and V. Nanda. 1991. Multimarket trading and market liquidity. *Review of Financial Studies* 4:483–511.

Coles, J. L., D. Heath, and M. C. Ringgenberg. 2020. On index investing. *Working Paper* .

Cong, L. W., and D. X. Xu. 2019. Rise of factor investing: asset prices, informational efficiency, and security design. *Working Paper* .

Da, Z., and S. Shive. 2018. Exchange traded funds and asset return correlations. *European Financial Management* 24:136–68.

Dávila, E., and C. Parlatore. forthcoming. Trading costs and informational efficiency. *Journal of Finance* .

Diamond, D. W. 1985. Optimal release of information by firms. *Journal of Finance* 40:1071–94.

Diamond, D. W., and R. E. Verrecchia. 1981. Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics* 9:221–35.

Dow, J., and R. Rahi. 2003. Informed trading, investment, and welfare. *Journal of Business* 439–54.

Easley, D., D. Michayluk, M. O'Hara, and T. J. Putniss. 2020. The active world of passive investing. *Working Paper* .

Farboodi, M., A. Matray, and L. Veldkamp. 2018. Where has all the big data gone? *Working Paper* .

Fishman, M. J., and K. M. Hagerty. 1992. Insider trading and the efficiency of stock prices. *RAND Journal of Economics* 23:106–22.

Fishman, M. J., and J. A. Parker. 2015. Valuation, adverse selection, and market collapses. *Review of Financial Studies* 28:2575–607.

French, K. R. 2008. Presidential address: The cost of active investing. *The Journal of Finance* 63:1537–73.

Ganguli, J., and L. Yang. 2009. Complementarities, multiplicity, and supply information. *Journal of the European Economic Association* 7(1):90–115.

García, D., and J. M. Vanden. 2009. Information acquisition and mutual funds. *Journal of Economic Theory* 144:1965–95.

Gârleanu, N., and L. H. Pedersen. 2018. Efficiently inefficient markets for assets and asset management. *Journal of Finance* 73:1663–712.

———. 2020. Active and passive investing. *Working Paper* .

Gerakos, J., J. Linnainmaa, and A. Morse. 2019. Asset managers: Institutional performance and factor exposures. *Working Paper* .

Glode, V., R. C. Green, and R. Lowery. 2012. Financial expertise as an arms race. *Journal of Finance* 67:1723–59.

Glosten, L., S. Nallareddy, and Y. Zou. forthcoming. ETF activity and informational efficiency of underlying securties. *Management Science* .

Goldstein, I., and L. Yang. 2019. Commodity financialization and information transmission. *Working Paper* .

Greenwood, R. 2008. Excess comovement of stock returns: Evidence from cross-sectional variation in nikkei 225 weights. *Review of Financial Studies* 21:1153–86.

Grinblatt, M., G. Jostova, L. Petrasek, and A. Philipov. 2020. Style and skill: Hedge funds, mutual funds, and momentum. *Working Paper* .

Grossman, S. J. 1976. On the efficiency of competitive stock markets where trades have diverse information. *Journal of Finance* 31(2):573–85.

Grossman, S. J., and J. E. Stiglitz. 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70(3):393–408.

Hellwig, M. F. 1980. On the aggregation of information in competitive markets. *Journal of Economic Theory* 22:477–98.

Hirshleifer, J. 1971. The private and social value of information and the reward to inventive activity. *The American Economic Review* 61:561–74.

Israeli, D., C. Lee, and S. A. Sridharan. 2017. Is there a dark side to exchange traded funds? An information perspective. *Review of Accounting Studies* 1048–83.

Jiang, H., M. Verbeek, and Y. Wan. 2017. Information content when mutual funds deviate from benchmarks. *Management Science* 2038–53.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp. 2014. Time-varying fund manager skill. *Journal of Finance* 69.

———. 2016. A rational theory of mutual funds' attention allocation. *Econometrica* 84.

Kawakami, K. 2017. Welfare consequences of information aggregation and optimal market size. *American Economic Journal: Microeconomics* 303–23.

Khanna, N., S. L. Slezak, and M. Bradley. 1994. Insider trading, outside search, and resource allocation: why firms and society may disagree on insider trading restrictions. *Review of Financial Studies* 7:575–608.

Kurlat, P. 2019. The social value of financial expertise. *American Economic Review* 109:556–90.

Kurlat, P., and L. Veldkamp. 2015. Should we regulate financial information? *Journal of Economic Theory* 158:697–720.

Leippold, M., L. Su, and A. Ziegler. 2016. How index futures and ETFs affect stock return correlations. *Working Paper* .

Lettau, M., and A. Madhaven. 2018. Exchange-traded funds 101 for economists. *Journal of Economic Perspectives* 32:135–54.

Manzano, C., and X. Vives. 2011. Public and private learning from prices, strategic substitutability and complementarity, and equilibrium multiplicity. *Journal of Mathematical Economics* 47(3):346–69.

Marín, J. M., and R. Rahi. 1999. Speculative securities. *Economic Theory* 14:653–68.

Medrano, L. A., and X. Vives. 2004. Regulating insider trading when investment matters. *Review of Finance* 199–277.

Mihet, R. 2020. Financial technology and the inequality gap. *Working Paper* .

Peress, J. 2005. Information vs. entry costs: What explains u.s. stock market evolution? *Journal of Financial and Quantitative Analysis* 40:563–94.

Qin, N., and V. Singal. 2015. Indexing and stock price efficiency. *Financial Management* 44:875–904.

Schlee, E. E. 2001. The value of information in efficient risk-sharing arrangements. *American Economic Review* 91:509–24.

Schmalz, M. 2018. Common-ownership concentration and corporate conduct. *Annual Review of Financial Economics* 10.

Sockin, M. 2018. Financial sector transparency and real investment. *Working Paper* .

Sockin, M., and W. Xiong. 2015. Informational frictions and commodity markets. *Journal of Finance* 70:2063–98.

Stambaugh, R. F. 2014. Presidential address: Investment noise and trends. *The Journal of Finance* 69:1415–53.

Subrahmanyam, A. 1991. A theory of trading in stock index futures. *Review of Financial Studies* 17–51.

Van Nieuwerburgh, S., and L. Veldkamp. 2010. Information acquisition and under- diversification. *Review of Economic Studies* .

Verrecchia, R. E. 1982. Information acquisition in a noisy rational expectations economy. *Econometrica* 50(6):1415–30.

Vijh, A. M. 1994. S&P 500 trading strategies and stock betas. *Review of Financial Studies* 7:215–51.

A Results omitted from main text

Note: Throughout the appendix, we frequently omit asset subscripts in order to enhance notational transparency.

Lemma A-1 *For any positive integers m and $l \leq m$ such that l is a power of 2, there exists an $m \times m$ matrix A with the following properties: A is symmetric and invertible, with $A^{-1} = A$ (i.e., A is involutory); $A_{jk} = 0$ if $j \neq k$ and either $j > l$ or $k > l$; $A_{jk} = 1$ if $j = k > l$; $A_{1k} = l^{-\frac{1}{2}}$ and $|A_{jk}| = l^{-\frac{1}{2}}$ for all $j, k \leq l$; for any j, $j' \neq j$, $\sum_{k=1}^{m} A_{jk} A_{j'k} = 0$; and $\sum_{k=1}^{l} A_{jk} = 0$ for $j = 2, \ldots, l$.*

Proof of Lemma A-1: We focus on cases $l = m$, since the generalization to $l < m$ is trivial. The proof is inductive: Given the existence of an $m \times m$ matrix A with the stated properties, we construct a $2m \times 2m$ matrix B with the same properties. Specifically, define

$$B = \frac{1}{\sqrt{2}} \begin{pmatrix} A & A \\ A & -A \end{pmatrix}.$$

With the exception of the symmetry and inversion properties, it is straightforward to see that B has the desired properties. To establish that B is involutory, simply note that

$$BB = \frac{1}{2} \begin{pmatrix} A & A \\ A & -A \end{pmatrix} \begin{pmatrix} A & A \\ A & -A \end{pmatrix} = \frac{1}{2} \begin{pmatrix} AA + AA & AA - AA \\ AA - AA & AA + AA \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2I_m & 0_m \\ 0_m & 2I_m \end{pmatrix} = I_{2m},$$

where I_m denotes the $m \times m$ identity matrix and 0_m denotes the $m \times m$ matrix in which all entries are zero. To establish that B is symmetric, simply note that

$$B^\top = \frac{1}{\sqrt{2}} \begin{pmatrix} A^\top & A^\top \\ A^\top & -A^\top \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} A & A \\ A & -A \end{pmatrix} = B.$$

Finally, for the base case of $m = 1$, simply define $A = (1)$. This completes the proof.

Lemma A-2 *Let A be a matrix with the properties stated in Lemma A-1. Define synthetic assets as paying off*

$$X_k \equiv \sum_{j=1}^{m} A_{kj} \tilde{X}_j. \tag{A-1}$$

The synthetic assets have the properties stated in the main text. Moreover, the price vectors for fundamental and synthetic assets are related by $P = A\tilde{P}$ and $\tilde{P} = AP$.

Proof of Lemma A-2: The properties of the synthetic assets all follow directly from the properties of the matrix A. For the price vectors, the statement $P = A\tilde{P}$ is immediate from construction. Since A is involutory, it follows that $AP = \tilde{P}$.

Lemma A-3 *Suppose X is a normally distributed random variable, and that an information set $\mathcal{F}$ consists of a set of normally distributed random variables. Then the derivative of the conditional expectation $\mathbb{E}[X|\mathcal{F}]$ with respect to a realization $\hat{X}$ of X is*

$$\frac{\partial}{\partial \hat{X}} \mathbb{E}[X|\mathcal{F}] = 1 - \frac{\text{var}(X|\mathcal{F})}{\text{var}(X)}.$$

Proof of Lemma A-3: Let Σ_{22} be the variance matrix of the random variables in $\mathcal{F}$; and Σ_{12} be the row vector of covariances between X and the random variables in $\mathcal{F}$. The vector of coefficients from the regressing each variable in $\mathcal{F}$ on X is $\frac{\Sigma_{12}}{\mathrm{var}(X)}$. So by the properties of multivariate normality,

$$\frac{\partial}{\partial \hat{X}} \mathbb{E}\left[X|\mathcal{F}\right] = \Sigma_{12}\Sigma_{22}^{-1}\frac{\Sigma_{12}^{\top}}{\mathrm{var}\left(X\right)}.$$

Also from multivariate normality,

$$\mathrm{var}\left(X|\mathcal{F}\right) = \mathrm{var}\left(X\right) - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^{\top}.$$

Combining these two equations completes the proof.

Lemma A-4 *In a linear equilibrium, $\frac{\partial P_k}{\partial Z_k} \neq 0$, and hence ρ_k is well-defined.*

Proof of Lemma A-4: Differentiation of market clearing (4) with respect to Z yields

$$\frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial P}di + \frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial \hat{P}}di + \int_0^n \frac{\partial \theta_i}{\partial e_i}di + \int_0^n \frac{\partial \theta_i}{\partial \hat{e}_i}di = 0. \tag{A-2}$$

Suppose that, contrary to the claimed result, $\frac{\partial P}{\partial Z} = 0$. Then Z and e_i provide no information about the cash flow X, so that $\frac{\partial \theta_i}{\partial \hat{e}_i} = 0$ for all agents. In contrast, the non-informational effect of exposure shocks on the portfolio decision is certainly negative (see (11)). But then the LHS of (A-2) is strictly negative, a contradiction, completing the proof.

Lemma A-5 *If P_k is a linear function of X_k and Z_k then the effects of non-informational factors on demand θ_{ik} are given by*

$$\frac{\partial \theta_{ik}}{\partial e_{ik}} = -1; \quad \frac{\partial \theta_{ik}}{\partial P_k} = -\frac{1}{\gamma}\frac{1}{\mathrm{var}\left(X_k|y_{ik}, e_{ik}, P_k\right)}; \tag{A-3}$$

while the effects of informational factors on demand θ_{ik} satisfy

$$\frac{\partial \theta_{ik}}{\partial y_{ik}} = \frac{\tau_{ik}}{\gamma}; \quad \frac{\partial \theta_{ik}}{\partial \hat{e}_{ik}} = \frac{\rho_k}{\gamma}\tau_u; \quad \frac{\partial P_k}{\partial Z_k}\frac{\partial \theta_{ik}}{\partial \hat{P}_k} = -\frac{\rho_k}{\gamma}\left(\tau_Z + \tau_u\right); \quad \frac{\partial P_k}{\partial X_k}\frac{\partial \theta_{ik}}{\partial \hat{P}_k} = \frac{\rho_k^2}{\gamma}\left(\tau_Z + \tau_u\right). \tag{A-4}$$

Furthermore, these imply that

$$\frac{\frac{\partial P_k}{\partial Z_k}\frac{\partial \theta_{ik}}{\partial \hat{P}_k}}{\frac{\partial \theta_{ik}}{\partial \hat{e}_{ik}}} = -\frac{\tau_Z + \tau_u}{\tau_u}, \tag{A-5}$$

$$\frac{\partial P_k}{\partial Z_k}\frac{\partial \theta_{ik}}{\partial \hat{P}_k} + \frac{\partial \theta_{ik}}{\partial \hat{e}_{ik}} = -\frac{\rho_k}{\gamma}\tau_Z. \tag{A-6}$$

Proof of Lemma A-5: Consider first the case in which $\frac{\partial P}{\partial X} \neq 0$. The information content of (y_i, e_i, P) is the same as the information content of

$$\left(y_i, \frac{P - \mathbb{E}[P]}{\frac{\partial P}{\partial X}} + \mathbb{E}[X], \frac{P - \mathbb{E}[P]}{\frac{\partial P}{\partial X}} + \mathbb{E}[X] + \rho^{-1}(e_i - S)\right) = \left(X + \epsilon_i, X - \rho^{-1} Z, X + \rho^{-1}(u_i + s_i - S)\right).$$

Since ϵ_i, Z, and $u_i + s_i - S$ are independent and all have mean 0, the conditional variance expressions (12) and (13) follow by standard normal-normal updating. Using $\frac{\partial P}{\partial X} = -\rho \frac{\partial P}{\partial Z}$, the corresponding conditional expectation $\mathbb{E}[X|y_i, e_i, P]$ is given by

$$\frac{\mathbb{E}[X|y_i, e_i, P]}{\operatorname{var}(X|y_i, e_i, P)} = \tau_X \mathbb{E}[X] + \rho(\tau_Z + \tau_u)\left(\rho \mathbb{E}[X] - \frac{P - \mathbb{E}[P]}{\frac{\partial P}{\partial Z}}\right) + \rho \tau_u (e_i - S) + \tau_i y_i. \tag{A-7}$$

Finally, if $\frac{\partial P}{\partial X} = 0$ then neither the price nor the exposure shock e_i contains any information about X; and $\rho = 0$; so (12), (13), and (A-7) are all immediate.

The stated expressions are then immediate from the demand equation (11), completing the proof.

Lemma A-6 *In a linear equilibrium, the aggregate demand curve slopes down, i.e.,*

$$\int_0^{n_k} \frac{\partial \theta_{ik}}{\partial P_k} di + \int_0^{n_k} \frac{\partial \theta_{ik}}{\partial \hat{P}_k} di < 0, \tag{A-8}$$

and the price is an increasing function of X_k, and a strictly decreasing function of Z_k,

$$\frac{\partial P_k}{\partial X_k} \geq 0 \text{ and } \frac{\partial P_k}{\partial Z_k} < 0,$$

and so in particular $\rho_k \geq 0$.

Proof of Lemma A-6: From Lemma A-5, $\frac{\partial \theta_i}{\partial P} < 0$ for all agents.

If $\frac{\partial P}{\partial X} = 0$, then P contains no information about X, so $\int_0^n \frac{\partial \theta_i}{\partial \hat{P}} di = 0$, and (A-8) is then immediate.

If instead $\frac{\partial P}{\partial X} \neq 0$, then $\frac{\partial P}{\partial X} \int_0^n \frac{\partial \theta_i}{\partial \hat{P}} di > 0$ by Lemma A-5. By (B-6) and Lemma A-5,

$$\frac{\partial P}{\partial X} \int_0^n \frac{\partial \theta_i}{\partial P} di + \frac{\partial P}{\partial X} \int_0^n \frac{\partial \theta_i}{\partial \hat{P}} di = -\int_0^n \frac{\partial \theta_i}{\partial y_i} di < 0. \tag{A-9}$$

Hence $\frac{\partial P}{\partial X} > 0$, which (again using (A-9)) implies (A-8).

Note that the above arguments also establish that $\frac{\partial P}{\partial X} \geq 0$.

Lemma A-4 establishes that $\frac{\partial P}{\partial Z} \neq 0$. So to establish $\frac{\partial P}{\partial Z} < 0$, suppose to the contrary that $\frac{\partial P}{\partial Z} > 0$. Then $\rho \leq 0$, and Lemma A-5 implies $\int_0^n \frac{\partial \theta_i}{\partial e_i} di < 0$ and $\int_0^n \frac{\partial \theta_i}{\partial \hat{e}_i} di \leq 0$. Combined with (A-8), this in turn implies that the LHS of (A-2) is strictly negative. The contradiction completes the proof.

Lemma A-7 *Let $\xi \in \mathbb{R}^n$ be a normally distributed random vector with mean μ and variance-covariance matrix Σ. Let $b \in \mathbb{R}^n$ be a given vector, and $A \in \mathbb{R}^{n\times n}$ a symmetric matrix. If $I - 2\Sigma A$ is positive definite, then $\mathbb{E}\left[\exp(b^\top \xi + \xi^\top A\xi)\right]$ is well defined, and given by:*

$$\mathbb{E}\left[\exp\left(b^\top \xi + \xi^\top A\xi\right)\right] = |I - 2\Sigma A|^{-1/2} \exp\left(b^\top \mu + \mu^\top A\mu + \frac{1}{2}(b + 2A\mu)^\top (I - 2\Sigma A)^{-1}\Sigma(b + 2A\mu)\right). \tag{A-10}$$

Proof of Lemma A-7: Standard result.

B Proofs of results stated in main text

Proof of Lemma 1: For use throughout, note that the price P is normally distributed in a linear equilibrium. Differentiation of market clearing (4) with respect to X gives

$$\frac{\partial}{\partial X}\int_0^n \theta_i di = 0.$$

Substituting in the portfolio θ_i from (11); recalling the property of multivariate normality that conditional variances do not depend on the realizations of random variables; and noting that $\frac{\partial P}{\partial X} = \frac{\text{cov}(P,X)}{\text{var}(X)}$, it follows that

$$\int_0^n \frac{\frac{\partial}{\partial X}\mathbb{E}\left[X|y_i, e_i, P\right]}{\text{var}\left(X|y_i, e_i, P\right)} di = \int_0^n \frac{\frac{\text{cov}(P,X)}{\text{var}(X)}}{\text{var}\left(X|y_i, e_i, P\right)} di. \tag{B-1}$$

By Lemma A-3,

$$\frac{\partial}{\partial X}\mathbb{E}\left[X|y_i, e_i, P\right] = 1 - \frac{\text{var}\left(X|y_i, e_i, P\right)}{\text{var}\left(X\right)}. \tag{B-2}$$

(Note that the RHS of (B-2) is simply the R^2 of regressing cash flows on (y_i, e_i, P).) Substitution of (B-2) into (B-1) yields

$$\left(1 - \frac{\text{cov}\left(P, X\right)}{\text{var}\left(X\right)}\right) \int_0^n \frac{1}{\text{var}\left(X|y_i, e_i, P\right)} di = \int_0^n \frac{1}{\text{var}\left(X\right)} di,$$

which is equivalent to (14), completing the proof.

Proof of Lemma 2: At various points in the proof, we make use of $\rho < 0$ and $\frac{\partial P}{\partial Z} < 0$ (by Lemma A-6), and $\frac{\partial \theta_i}{\partial \tilde{P}} > 0$ (by Lemma A-5).

Re-arranging market-clearing (A-2) gives

$$0 = \frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial P}di + \int_0^n \frac{\partial \theta_i}{\partial e_i}di + \left(\frac{\frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial \hat{P}}di}{\int_0^n \frac{\partial \theta_i}{\partial \tilde{e}_i}di} + 1\right)\int_0^n \frac{\partial \theta_i}{\partial \hat{e}_i}di. \tag{B-3}$$

By Lemma A-6 and market-clearing (A-2), we know

$$\int_0^n \frac{\partial \theta_i}{\partial e_i}di + \int_0^n \frac{\partial \theta_i}{\partial \hat{e}_i}di < 0. \tag{B-4}$$

From (A-6) of Lemma A-5, $\frac{\partial P}{\partial Z}\frac{\partial \theta_i}{\partial \hat{P}} + \frac{\partial \theta_i}{\partial \hat{e}_i} < 0$, which together with $\frac{\partial \theta_i}{\partial \hat{e}_i} > 0$ implies

$$\frac{\frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial \hat{P}}di}{\int_0^n \frac{\partial \theta_i}{\partial \hat{e}_i}di} + 1 < 0.$$

So substituting (B-4) into (B-3) gives

$$\begin{aligned} 0 &> \frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial P}di + \int_0^n \frac{\partial \theta_i}{\partial e_i}di - \left(\frac{\frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial \hat{P}}di}{\int_0^n \frac{\partial \theta_i}{\partial \hat{e}_i}di} + 1\right)\int_0^n \frac{\partial \theta_i}{\partial e_i}di \\ &= \frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial P}di - \frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial \hat{P}}di\frac{\int_0^n \frac{\partial \theta_i}{\partial e_i}di}{\int_0^n \frac{\partial \theta_i}{\partial \hat{e}_i}di}. \end{aligned} \tag{B-5}$$

Using $\frac{\partial P}{\partial Z} < 0$ and the signs established in Lemma A-5, inequality (B-5) is equivalent to

$$\frac{\int_0^n \frac{\partial \theta_i}{\partial P}di}{\int_0^n \frac{\partial \theta_i}{\partial \hat{P}}di} > \frac{\int_0^n \frac{\partial \theta_i}{\partial e_i}di}{\int_0^n \frac{\partial \theta_i}{\partial \hat{e}_i}di},$$

which is in turn equivalent to (15), completing the proof.

Proof of Lemma 3: Differentiation of market clearing (4) with respect to X yields

$$\frac{\partial P}{\partial X}\int_0^n \frac{\partial \theta_i}{\partial P}di + \frac{\partial P}{\partial X}\int_0^n \frac{\partial \theta_i}{\partial \hat{P}}di + \int_0^n \frac{\partial \theta_i}{\partial y_i}di = 0. \tag{B-6}$$

Combined with (A-2), it follows that

$$-\frac{\frac{\partial P}{\partial X}}{\frac{\partial P}{\partial Z}} = -\frac{\int_0^n \frac{\partial \theta_i}{\partial y_i}di}{\int_0^n \frac{\partial \theta_i}{\partial e_i}di + \int_0^n \frac{\partial \theta_i}{\partial \hat{e}_i}di}. \tag{B-7}$$

Substituting in from Lemma A-5,

$$\rho = \frac{\frac{1}{\gamma}\int_0^n \tau_i di}{\int_0^n \left(1 - \frac{\rho}{\gamma}\tau_u\right) di}, \tag{B-8}$$

which rearranges to (16). The comparative static of price efficiency is immediate from the fact that $\frac{1}{n}\int_0^n \tau_i di$ is decreasing in n.

Proof of Proposition 1: The final wealth of agent i, given optimal trading (11), is

$$W_i = e_i P + \frac{\mathbb{E}[X - P|y_i, e_i, P]\,(X - P)}{\gamma \mathrm{var}(X|y_i, e_i, P)}.$$

So by the standard expression for the expected utility of an agent with CARA utility facing normally shocks, combined with simple manipulation, agent i's expected utility at the time of trading is

$$\mathbb{E}\left[u\left(W_i\right)|y_i, e_i, P\right] = -\exp\left(-\gamma\left(e_i P + \frac{1}{2}\frac{\mathbb{E}[X - P|y_i, e_i, P]^2}{\gamma \mathrm{var}(X|y_i, e_i, P)}\right)\right). \tag{B-9}$$

To obtain (18), we proceed in two stages. First, we integrate over realizations of the private signal y_i in (B-9). Second, we integrate over realizations of the price P. Note that the first stage is relatively standard, and similar algebraic arguments can be found in the related literature. Readers familiar with these arguments should proceed directly to the second stage.

For the first stage, define $\xi_i = \mathbb{E}\left[X - P|y_i, e_i, P\right]$ and $A_i = -1/\left(2\mathrm{var}(X|y_i, e_i, P)\right)$. Minor algebraic manipulation of Lemma A-7 implies

$$\mathbb{E}\left[\exp\left(\xi_i^2 A_i\right)|e_i, P\right] = \left(1 - 2A_i \mathrm{var}(\xi_i|e_i, P)\right)^{-\frac{1}{2}} \exp\left(\frac{A_i}{1 - 2A_i \mathrm{var}(\xi_i|e_i, P)}\mathbb{E}[\xi_i|e_i, P]^2\right). \tag{B-10}$$

By the law of total variance,

$$\mathrm{var}(X - P|e_i, P) = \mathrm{var}(\mathbb{E}[X - P|y_i, e_i, P]|e_i, P) + \mathbb{E}[\mathrm{var}(X - P|y_i, e_i, P)|e_i, P]$$

which implies

$$\mathrm{var}(\xi_i|e_i, P) = \mathrm{var}(X|e_i, P) - \mathrm{var}(X|y_i, e_i, P),$$

and so

$$1 - 2A_i \mathrm{var}(\xi_i|e_i, P) = \frac{\mathrm{var}(X|e_i, P)}{\mathrm{var}(X|y_i, e_i, P)} = d_i,$$

where d_i is as defined in (19). Substitution and straightforward manipulation implies that expression (B-10) equals

$$d_i^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{\mathbb{E}[X-P|e_i,P]^2}{\mathrm{var}(X|e_i,P)}\right),$$

and so

$$\mathbb{E}\left[u\left(W_i\right)|e_i,P\right] = -d_i^{-\frac{1}{2}} \exp\left(-\gamma e_i P - \frac{1}{2}\frac{\mathbb{E}[X-P|e_i,P]^2}{\mathrm{var}(X|e_i,P)}\right), \tag{B-11}$$

completing the first stage.

In the second stage, we integrate over realizations of P. Since $P = (P - \mathbb{E}[P|e_i]) + \mathbb{E}[X|e_i] - \mathbb{E}[X-P|e_i]$, the expression in the exponent of (B-11) equals

$$-\frac{1}{2}\frac{\mathbb{E}[X-P|e_i,P]^2}{\mathrm{var}(X|e_i,P)} - \gamma e_i(P-\mathbb{E}[P|e_i]) - \gamma e_i\mathbb{E}[X|e_i] + \gamma e_i\mathbb{E}[X-P|e_i]. \tag{B-12}$$

Denote the expected return $X-P$ given exposure e_i by α_e, and recall that $\mathbb{E}\left[e_i\right] = S$:

$$\alpha_e \equiv \mathbb{E}[X-P|e_i] = \mathbb{E}[X-P] - \frac{\mathrm{cov}(P,e_i)}{\mathrm{var}\left(e_i\right)}\left(e_i - S\right). \tag{B-13}$$

Hence

$$\mathbb{E}[X-P|e_i,P] = \frac{\mathrm{cov}\left(X-P,P|e_i\right)}{\mathrm{var}\left(P|e_i\right)}\left(P-\mathbb{E}[P|e_i]\right) + \alpha_e. \tag{B-14}$$

By substitution and Lemma A-7, the expectation of (B-11) conditional on e_i is given by

$$\begin{aligned}\mathbb{E}\left[u\left(W_i\right)|e_i\right] &= -d_i^{-\frac{1}{2}} D^{-\frac{1}{2}} \exp\left(-\gamma e_i\left(\mathbb{E}\left[X\right] - \alpha_e\right)\right) \\ &\times \exp\left(-\frac{1}{2}\frac{\alpha_e^2}{\mathrm{var}(X|e_i,P)} + \frac{1}{2}\left(\frac{\alpha_e \mathrm{cov}(X-P,P|e_i)}{\mathrm{var}(P|e_i)\mathrm{var}(X|e_i,P)} + \gamma e_i\right)^2 \frac{\mathrm{var}(P|e_i)}{D}\right)\end{aligned} \tag{B-15}$$

where

$$D = 1 + \frac{\mathrm{cov}\left(X-P,P|e_i\right)^2}{\mathrm{var}\left(X|e_i,P\right)\mathrm{var}\left(P|e_i\right)}. \tag{B-16}$$

The law of total variance and (B-14) together yield

$$\mathrm{var}\left(\mathrm{E}\left[X-P|e_i,P\right]|e_i\right) = \mathrm{var}\left(X-P|e_i\right) - \mathrm{var}\left(X|e_i,P\right) = \frac{\mathrm{cov}\left(X-P,P|e_i\right)^2}{\mathrm{var}\left(P|e_i\right)}, \tag{B-17}$$

and substitution into (B-16) delivers (20).

For use below, note also that (B-17) implies that

$$\frac{\mathrm{var}\,(P|e_i)}{D} = \frac{\mathrm{var}\,(P|e_i)\,\mathrm{var}(X|e_i,P)}{D\mathrm{var}(X|e_i,P)} = \frac{\mathrm{var}\,(P|e_i)\,\mathrm{var}\,(X-P|e_i) - \mathrm{cov}\,(X-P,P|e_i)^2}{D\mathrm{var}(X|e_i,P)}$$
$$= \frac{\mathrm{var}\,(X|e_i)\,\mathrm{var}\,(X-P|e_i) - \mathrm{cov}\,(X-P,X|e_i)^2}{D\mathrm{var}(X|e_i,P)}$$
$$= \mathrm{var}\,(X|e_i) - \frac{\mathrm{cov}\,(X-P,X|e_i)^2}{D\mathrm{var}(X|e_i,P)} \qquad \text{(B-18)}$$

where the penultimate equality follows from the fact that for any random variables r_1, r_2,

$$\mathrm{cov}\,(r_1-r_2,r_1)^2 - \mathrm{cov}\,(r_1-r_2,r_2)^2 = \mathrm{var}\,(r_1-r_2)\,(\mathrm{var}\,(r_1) - \mathrm{var}\,(r_2))\,,$$

and the final equality follows from (20).

By a combination of algebraic manipulation and (B-16), (20), (B-18), expected utility (B-15) equals

$$\begin{aligned}
& -(d_i D)^{-\frac{1}{2}} \exp\left(-\gamma e_i \mathbb{E}\,[X] + \frac{1}{2}\frac{\alpha_e^2}{\mathrm{var}(X|P,e_i)}\left(\frac{\mathrm{cov}(X-P,P|e_i)^2}{D\mathrm{var}(P|e_i)\mathrm{var}(X|e_i,P)} - 1\right)\right) \\
\times\ & \exp\left(\gamma e_i \alpha_e \left(\frac{\mathrm{cov}(X-P,P|e_i)}{D\mathrm{var}(X|e_i,P)} + 1\right) + \frac{1}{2}\gamma^2 e_i^2 \frac{\mathrm{var}(P|e_i)}{D}\right) \\
=\ & -(d_i D)^{-\frac{1}{2}} \exp\left(-\gamma e_i \mathbb{E}\,[X] - \frac{1}{2}\frac{\alpha_e^2}{D\mathrm{var}(X|e_i,P)}\right) \\
\times\ & \exp\left(\gamma e_i \alpha_e \frac{\mathrm{cov}(X-P,X|e_i)}{D\mathrm{var}(X|e_i,P)} + \frac{1}{2}\gamma^2 e_i^2\left(\mathrm{var}(X|e_i) - \frac{\mathrm{cov}(X-P,X|e_i)^2}{D\mathrm{var}(X|e_i,P)}\right)\right).
\end{aligned}$$

By further manipulation, and substitution of α_e using (B-13), expected utility (B-15) equals

$$\begin{aligned}
& -(d_i D)^{-\frac{1}{2}} \exp\left(-\gamma e_i \mathbb{E}\,[X] + \frac{\gamma^2 e_i{}^2}{2\tau_X} - \frac{1}{2}\frac{(\alpha_e - \mathrm{cov}(X-P,X)\gamma e_i)^2}{D\mathrm{var}(X|e_i,P)}\right) \\
=\ & -(d_i D)^{-\frac{1}{2}} \exp\left(-\gamma e_i \mathbb{E}\,[X] + \frac{\gamma^2 e_i{}^2}{2\tau_X}\right) \\
\times\ & \exp\left(-\frac{1}{2}\frac{\left(\mathbb{E}[X-P] - \mathrm{cov}(X-P,X)\gamma S - \mathrm{cov}(X-P,X)\gamma\,(e_i - S) - \frac{\mathrm{cov}(P,e_i)}{\mathrm{var}(e_i)}\,(e_i - S)\right)^2}{D\mathrm{var}(X|e_i,P)}\right).
\end{aligned}$$

Substituting Corollary 1's expression for $\mathbb{E}[X-P]$ into this last expression yields (21).

Finally, the fact that each of d_i, D, and Λ can be written as functions of exogenous parameters and price efficiency ρ is established in the proof of Proposition 2.

Proof of Proposition 2: For use below, recall that $\frac{\partial P}{\partial X} = \frac{\text{cov}(X,P)}{\text{var}(X)}$, $\frac{\partial P}{\partial Z} = \frac{\text{cov}(Z,P)}{\text{var}(Z)}$, and hence

$$\text{var}\left(P|e_i\right) = \frac{\text{cov}\left(X,P\right)^2}{\text{var}\left(X\right)^2}\text{var}\left(X\right) + \frac{\text{cov}\left(Z,P\right)^2}{\text{var}\left(Z\right)^2}\text{var}\left(Z|e_i\right), \tag{B-19}$$

$$\text{var}\left(X-P|e_i\right) = \left(\frac{\text{cov}\left(X-P,X\right)}{\text{var}\left(X\right)}\right)^2 \text{var}\left(X\right) + \left(\frac{\text{cov}\left(P,Z\right)}{\text{var}\left(Z\right)}\right)^2 \text{var}\left(Z|e_i\right). \tag{B-20}$$

Note also that Lemmas 1 and A-5 directly imply the following expression for the non-informational effect of prices on aggregate demand:

$$\frac{1}{n}\int_0^n \frac{\partial \theta_i}{\partial P} di = -\frac{1}{\gamma}\frac{1}{\text{cov}\left(X-P,X\right)}. \tag{B-21}$$

We use repeatedly Lemma 3's result that price efficiency decreases in participation n.

To establish the result, we show that each of the three terms d_i, D, and Λ are increasing in participation n. We start with the term d_i, which corresponds to an agent's expected trading gains from her private information. Substitution of (13) delivers

$$d_i = \frac{\tau_X + \rho^2(\tau_Z + \tau_u) + \tau_i}{\tau_X + \rho^2(\tau_Z + \tau_u)}.$$

Because price efficiency ρ is decreasing in participation n, the private gains from information, d_i, are increasing in participation. Economically, when prices convey less information about cash flows X, an agent's private information about X is more valuable.

Next, we consider the term Λ. As a first step, (B-21) implies that

$$\begin{aligned}\frac{\text{cov}(P,e_i)}{\text{var}(e_i)} + \gamma\text{cov}(X-P,X) &= -\gamma\text{cov}\left(X-P,X\right)\frac{\text{var}\left(Z\right)}{\text{var}\left(e_i\right)}\left(-\frac{1}{\gamma\text{cov}\left(X-P,X\right)}\frac{\text{cov}\left(P,e_i\right)}{\text{var}\left(Z\right)} - \frac{\text{var}\left(e_i\right)}{\text{var}\left(Z\right)}\right)\\ &= -\gamma\text{cov}\left(X-P,X\right)\frac{\text{var}\left(Z\right)}{\text{var}\left(e_i\right)}\left(\frac{1}{n}\int_0^n \frac{\partial\theta_i}{\partial P}di\frac{\text{cov}\left(P,Z\right)}{\text{var}\left(Z\right)} - \frac{\text{var}\left(e_i\right)}{\text{var}\left(Z\right)}\right). \quad \text{(B-22)}\end{aligned}$$

Substituting (B-20) and (B-22) into (21), and again using (B-21), gives

$$\Lambda = \frac{\gamma^2 \mathrm{cov}\left(X-P,X\right)^2 \frac{\mathrm{var}(Z)^2}{\mathrm{var}(e_i)^2}\left(\frac{1}{n}\int_0^n \frac{\partial\theta_i}{\partial P}di\frac{\mathrm{cov}(P,Z)}{\mathrm{var}(Z)} - \frac{\mathrm{var}(e_i)}{\mathrm{var}(Z)}\right)^2}{\left(\frac{\mathrm{cov}(X-P,X)}{\mathrm{var}(X)}\right)^2 \mathrm{var}\left(X\right) + \left(\frac{\mathrm{cov}(P,Z)}{\mathrm{var}(Z)}\right)^2 \mathrm{var}\left(Z|e_i\right)}$$

$$= \frac{\frac{\mathrm{var}(Z)^2}{\mathrm{var}(e_i)^2}\left(\frac{1}{n}\int_0^n \frac{\partial\theta_i}{\partial P}di\frac{\mathrm{cov}(P,Z)}{\mathrm{var}(Z)} - \frac{\mathrm{var}(e_i)}{\mathrm{var}(Z)}\right)^2}{\frac{1}{\gamma^2}\frac{1}{\mathrm{var}(X)} + \left(-\frac{\mathrm{cov}(P,Z)}{\mathrm{var}(Z)}\frac{1}{\gamma \mathrm{cov}(X-P,X)}\right)^2 \mathrm{var}\left(Z|e_i\right)}$$

$$= \frac{\frac{\mathrm{var}(Z)^2}{\mathrm{var}(e_i)^2}\left(\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial\theta_i}{\partial P}di - \frac{\mathrm{var}(e_i)}{\mathrm{var}(Z)}\right)^2}{\frac{1}{\gamma^2}\frac{1}{\mathrm{var}(X)} + \left(\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial\theta_i}{\partial P}di\right)^2 \mathrm{var}\left(Z|e_i\right)}. \quad \text{(B-23)}$$

We next show that a key term in the numerator of (B-23) is negative, i.e.,

$$\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial\theta_i}{\partial P}di - \frac{\mathrm{var}\left(e_i\right)}{\mathrm{var}\left(Z\right)} < 0. \quad \text{(B-24)}$$

Inequality (B-24) holds because, by (A-5), it is equivalent

$$\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial\theta_i}{\partial P}di + \frac{\frac{1}{n}\int_0^n \frac{\partial P}{\partial Z}\frac{\partial\theta_i}{\partial \hat{P}}di}{\frac{1}{n}\int_0^n \frac{\partial\theta_i}{\partial \hat{e}_i}di} < 0,$$

and since $\frac{\partial\theta_i}{\partial e_i} = -1$, $\frac{\partial\theta_i}{\partial \hat{P}} > 0$ and $\frac{\partial P}{\partial Z} < 0$, this inequality is in turn equivalent (15), i.e., prices contain more information than exposure shocks.

Note also that, since $\mathrm{cov}(X-P,X) > 0$ (by Lemma 1), equation (B-22) implies that (B-24) is equivalent to

$$\frac{\mathrm{cov}(P,e_i)}{\mathrm{var}(e_i)} + \gamma \mathrm{cov}(X-P,X) > 0, \quad \text{(B-25)}$$

a fact we refer to in the main text.

From (B-24) and (B-23), Λ is decreasing in $\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial\theta_i}{\partial P}di$. Substitution of (A-6) into the market clearing condition (A-2), along with the basic property that the non-informational effect of exposure shocks on demand is $\frac{\partial\theta_i}{\partial e_i} = -1$, yields

$$\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial\theta_i}{\partial P}di = 1 + \frac{\rho}{\gamma}\tau_Z. \quad \text{(B-26)}$$

Hence $\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial\theta_i}{\partial P}di$ is increasing in price efficiency ρ, and hence is decreasing in participation n, implying that Λ is increasing in participation n.

Finally, we consider the term D, which we start by re-expressing. By the law of total variance,

$$\begin{aligned}\operatorname{var}(X|e_i,P)\operatorname{var}(P|e_i) &= \left(\operatorname{var}(X|e_i) - \frac{\operatorname{cov}(X,P|e_i)^2}{\operatorname{var}(P|e_i)}\right)\operatorname{var}(P|e_i)\\ &= \operatorname{var}(X|e_i)\operatorname{var}(P|e_i) - \operatorname{cov}(X,P|e_i)^2.\end{aligned}$$

Substituting in (B-19) gives

$$\operatorname{var}(X|e_i,P)\operatorname{var}(P|e_i) = \frac{\operatorname{cov}(Z,P)^2}{\operatorname{var}(Z)^2}\operatorname{var}(Z|e_i)\operatorname{var}(X). \tag{B-27}$$

Also by (B-19), and making use of (B-21),

$$\begin{aligned}\operatorname{cov}(X-P,P|e_i) &= \operatorname{cov}(X,P|e_i) - \operatorname{var}(P|e_i)\\ &= \frac{\operatorname{cov}(X,P)}{\operatorname{var}(X)}\left(\operatorname{var}(X) - \operatorname{cov}(X,P)\right) - \frac{\operatorname{cov}(Z,P)^2}{\operatorname{var}(Z)^2}\operatorname{var}(Z|e_i)\\ &= \left(\frac{\frac{\operatorname{cov}(X,P)}{\operatorname{var}(X)}}{\frac{\operatorname{cov}(Z,P)}{\operatorname{var}(Z)}} - \frac{\frac{\operatorname{cov}(Z,P)}{\operatorname{var}(Z)}\operatorname{var}(Z|e_i)}{\operatorname{cov}(X-P,X)}\right)\frac{\operatorname{cov}(Z,P)}{\operatorname{var}(Z)}\operatorname{cov}(X-P,X)\\ &= \left(-\rho + \gamma\operatorname{var}(Z|e_i)\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial\theta_i}{\partial P}di\right)\frac{\operatorname{cov}(Z,P)}{\operatorname{var}(Z)}\left(-\frac{1}{\gamma\frac{1}{n}\int_0^n \frac{\partial\theta_i}{\partial P}di}\right)\end{aligned} \tag{B-28}$$

Substitution of (B-27) and (B-28) into (B-16) yields:

$$D = 1 + \frac{\left(-\rho + \gamma\operatorname{var}(Z|e_i)\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial\theta_i}{\partial P}di\right)^2}{\gamma^2\operatorname{var}(Z|e_i)\operatorname{var}(X)\left(\frac{1}{n}\int_0^n \frac{\partial\theta_i}{\partial P}di\right)^2}. \tag{B-29}$$

By (B-26) and the fact that $\operatorname{var}(Z|e_i) = \frac{1}{\tau_Z+\tau_u}$,

$$-\rho + \gamma\operatorname{var}(Z|e_i)\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial\theta_i}{\partial P}di = (\gamma - \rho\tau_u)\operatorname{var}(Z|e_i), \tag{B-30}$$

and so

$$D = 1 + \frac{(\gamma-\rho\tau_u)^2\operatorname{var}(Z|e_i)}{\gamma^2\operatorname{var}(X)\left(\frac{1}{n}\int_0^n \frac{\partial\theta_i}{\partial P}di\right)^2}. \tag{B-31}$$

Note also that (B-30) implies that

$$\gamma - \rho\tau_u > 0, \tag{B-32}$$

since, using $\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial \hat{P}} di = -\frac{\rho}{\gamma \text{var}(Z|e_i)}$ (see Lemma A-5), the LHS of (B-30) equals

$$-\rho \left(1 + \frac{\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial P} di}{\frac{1}{n}\frac{\partial P}{\partial Z}\int_0^n \frac{\partial \theta_i}{\partial \hat{P}} di} \right). \tag{B-33}$$

Expression (B-33) is strictly positive because $\frac{\partial \theta_i}{\partial \hat{P}} > 0$ and demand slopes down (Lemma A-6).

Finally, from Lemma A-5 and (13),

$$\frac{1}{n}\int_0^n \frac{\partial \theta_i}{\partial P} di = -\frac{1}{\gamma}\frac{1}{n}\int_0^n \frac{1}{\text{var}\left(X|y_i, e_i, P\right)} di = -\frac{\tau_X + \rho^2 (\tau_Z + \tau_u) + \frac{1}{n}\int_0^n \tau_i di}{\gamma}. \tag{B-34}$$

So as participation n increases, $\left|\frac{1}{n}\int_0^n \frac{\partial \theta_i}{\partial P} di\right|$ declines, both because price efficiency declines; and because the average signal precision of participating agents, $\frac{1}{n}\int_0^n \tau_i di$, declines.

It then follows from (B-31) and (B-32) that D increases as participation n increases.

Finally, substitution of Lemma 3's expression for ρ into (B-34) yields

$$\left|\frac{1}{n}\int_0^n \frac{\partial \theta_i}{\partial P} di\right| = \frac{\tau_X + \rho^2 \tau_Z + \gamma \rho}{\gamma},$$

thereby establishing that D can be expressed as a function of ρ only, completing the proof.

Proof of Proposition 3: The main complication in the proof is that the equilibrium conditions are different for indexing versus no-indexing equilibria. The basic strategy of the proof is to start by finding total participation in financial markets, n_1, by means of analyzing the fixed points of a function that covers both cases, namely $\max\{g_A, g_I\}$ defined below.

For use throughout the proof, we write $A \succeq_i^{\rho_1, \rho_{-1}} 0$ if agent i prefers full participation to non-participation when price efficiency is (ρ_1, ρ_{-1}), i.e., if $\mathbb{E}\left[\mathcal{U}_i^A(e_i; \rho_1, \rho_{-1})\right] \exp(\gamma\kappa) \geq \mathbb{E}\left[\mathcal{U}_i^0(e_i)\right]$. We define the relations $I \succeq_i^{\rho_1} 0$ and $A \succeq_i^{\rho_{-1}} I$ etc analogously, where "I" corresponds to participation via indexing. Note that the comparison between index-participation and non-participation depends only on ρ_1, and not on ρ_{-1}; while the comparison between full participation and index-participation depends only on ρ_{-1}, and not on ρ_1.

Also for use below, define

$$f(n) \equiv \frac{\gamma}{2\tau_u} - \sqrt{\left(\frac{\gamma}{2\tau_u}\right)^2 - \frac{1}{\tau_u}\frac{1}{n}\int_0^n \tau_i di},$$

i.e., price efficiency associated with participation n (see Lemma 3).

Observe that n_1 is an equilibrium level of participation in the index asset 1 only if it is a fixed point of either the function

$$g_A(n) \equiv \max\left\{i : A \succeq_i^{f(n),f(n)} 0\right\},$$

where both here and below we adopt the adopt the convention that the maximum of an empty set is 0; or of the function

$$g_I(n;\kappa_1) \equiv \max\left\{i : I \succeq_i^{f(n)} 0 \text{ given } \kappa_1\right\}.$$

Moreover, if n_1 is a fixed point of g_A, then $(n_1, n_{-1} = n_1)$ is a no-indexing equilibrium if and only if

$$A \succeq_{n_1}^{f(n_1)} I; \tag{B-35}$$

and if n_1 is a fixed point of g_I, then there is an indexing equilibrium with participation n_1 in the index asset if and only if for some $j \leq n_1$,

$$I \succ_j^{f(j)} A. \tag{B-36}$$

Finally, define

$$g_{AI}(n;\kappa_1) = \max\left\{i : A \succeq_i^{f(n)} I \text{ given } \kappa_1\right\}.$$

By Lemma 3 and Proposition 2, the function g_A (respectively, g_I, g_{AI}) is continuous and weakly increasing in n, and is strictly increasing in the neighborhood of any n for which $g_A(n) \in (0,1)$ (respectively, $g_I(n) \in (0,1)$, $g_{AI}(n) \in (0,1)$).

(A) Let n_1 be a fixed point of $\max\{g_A, g_I\}$ (at least one fixed point exists by Tarski), i.e., $n_1 = \max\{g_A(n_1), g_I(n_1)\}$. To establish the result, we show there is an $n_{-1} \in [0, n_1]$ such that (n_1, n_{-1}) is an equilibrium. First, consider the case in which $g_A(n_1) < n_1$ and $g_I(n_1) = n_1$. So $0 \succ_{n_1}^{f(n_1),f(n_1)} A$ and $I \approx_{n_1}^{f(n_1)} 0$. Hence $I \succ_{n_1}^{f(n_1)} A$, implying $g_{AI}(n_1) < n_1$. So g_{AI} maps $[0, n_1]$ into itself, and hence (by Tarski) has at least one fixed point in $[0, n_1]$, say n_{-1}. By construction, (n_1, n_{-1}) is an equilibrium. Second, consider the case in which $g_A(n_1) = n_1$ and $g_I(n_1) \leq n_1$. So $A \approx_{n_1}^{f(n_1),f(n_1)} 0$ and $0 \succeq_{n_1}^{f(n_1)} I$. Hence $A \succeq_{n_1}^{f(n_1)} I$, establishing that (n_1, n_1) is an equilibrium.

(B) As κ_1 falls, the function g_I strictly increases at any value of n for which $g_I(n) \in (0,1)$. So by Corollary 1 of Milgrom and Roberts (1994), the extremal fixed points of g_I increase. Moreover, condition (B-36) is easier to satisfy. This establishes (i).

For (ii), let $\left(n_1', n_{-1}'\right)$ be the equilibrium with the most indexing at κ_1'. Note that n_{-1}' is the smallest fixed point of $g_{AI}(\cdot;\kappa_1')$. As κ_1 falls, the function g_{AI} decreases. So again by Milgrom and Roberts's Corollary 1, the smallest fixed point of $g_{AI}(\cdot;\kappa_1)$ is weakly smaller than n'_{-1} (and strictly so if $n'_{-1} > 0$). Combined with (i), the result follows.

(C) The function g_A has the same set of fixed points at κ_1 and κ'_1. Condition (B-35) is more demanding to satisfy at κ_1. This completes the proof.

Proof of Lemma 4: Throughout the proof, we use the fact that in equilibrium there exist scalars a and b such that the price of any non-index synthetic asset $k \neq 1$ is $P_k = a + b\rho_{-1}X_k - bZ_k$.

For assets j, k outside the index, the result is almost immediate, since the synthetic and actual assets coincide, and so

$$\tilde{P}_j - \tilde{P}_k = b\rho_{-1}\left(\tilde{X}_j - \tilde{X}_k\right) - b\left(\tilde{Z}_j - \tilde{Z}_k\right), \tag{B-37}$$

and so

$$\text{var}(\tilde{X}_j - \tilde{X}_k | \tilde{P}_j - \tilde{P}_k)^{-1} = \frac{1}{2}\tau_X + \frac{1}{2}\rho_{-1}^2\tau_Z,$$

establishing the result.

For assets j, k covered by the index, synthetic and actual assets differ, and so an additional step is required. The key step is to establish that the relative payoff of actual assets 1 and 2 (and hence, by symmetry, any pair of assets covered by the index) equals the linear combination payoffs of a set of synthetic assets that *does not* include the index asset, specifically,

$$\sum_{i=1}^{\frac{l}{2}} X_{2i} = \frac{\sqrt{l}}{2}\left(\tilde{X}_1 - \tilde{X}_2\right). \tag{B-38}$$

We establish (B-38) below. But taking this equality as given, along with the directly analogous equalities $\sum_{i=1}^{\frac{l}{2}} Z_{2i} = \frac{\sqrt{l}}{2}\left(\tilde{Z}_1 - \tilde{Z}_2\right)$ and $\sum_{i=1}^{\frac{l}{2}} P_{2i} = \frac{\sqrt{l}}{2}\left(\tilde{P}_1 - \tilde{P}_2\right)$, we know

$$\frac{\sqrt{l}}{2}\left(\tilde{P}_1 - \tilde{P}_2\right) = \sum_{i=1}^{\frac{l}{2}} P_{2i} = b\rho_{-1}\sum_{i=1}^{\frac{l}{2}} X_{2i} - b\sum_{i=1}^{\frac{l}{2}} Z_{2i} = b\rho_{-1}\frac{\sqrt{l}}{2}\left(\tilde{X}_1 - \tilde{X}_2\right) - b\frac{\sqrt{l}}{2}\left(\tilde{Z}_1 - \tilde{Z}_2\right),$$

which coincides with (B-37), and hence establishes the result for assets 1 and 2, and hence (by symmetry) for any pair of assets j, k that are covered by the index.

It remains to establish (B-38). To do so, we show that the row vectors A_i of the matrix A satisfy

$$\sum_{i=1}^{\frac{l}{2}} A_{2i} = \frac{\sqrt{l}}{2}\left(1, -1, \mathbf{0}_{m-2}\right), \tag{B-39}$$

where $\mathbf{0}_{m-2}$ denotes a row vector of length $m-2$ in which all entries equal 0. The proof of (B-39) can be rolled into the existing inductive proof of Lemma A-1, where recall that we focus on the case $m = l$, since the extension to $m > l$ is trivial. The base case is $l = 2$, and holds since $A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. For the inductive step, suppose that (B-39) holds for some l; the matrix for the case $2l$ is

$$B = \frac{1}{\sqrt{2}} \begin{pmatrix} A & A \\ A & -A \end{pmatrix}.$$

Hence

$$\sum_{i=1}^{\frac{2l}{2}} B_{2i} = 2\frac{1}{\sqrt{2}}\frac{\sqrt{l}}{2}\left(1, -1, \mathbf{0}_{l-2}, \mathbf{0}_l\right) = \frac{\sqrt{2l}}{2}\left(1, -1, \mathbf{0}_{2l-2}\right),$$

thereby completing the proof.

Proof of Lemma 5: By symmetry of assets covered by the index, it suffices to establish the result for the underlying asset $j = 1$. From Lemma A-1, the underlying asset 1 is related to synthetic assets $1, \ldots, l$ by $\tilde{X}_1 = \frac{1}{\sqrt{l}} \sum_{i=1}^{l} X_i$, with analogous equalities for $\tilde{P}_1$ and $\tilde{Z}_1$. In equilibrium, there are constants a_1, b_1, b_{-1} such that synthetic asset prices are given by $P_1 = a_1 + b_1 \rho_1 X_1 - b_1 Z_1$, and for any $j = 2, \ldots, l$, by $P_j = b_{-1}\rho_{-1} X_j - b_{-1} Z_j$. Moreover, $\rho_1 \leq \rho_{-1}$, where the inequality is strict in any indexing equilibrium. Combining these observations,

$$\begin{aligned}
\tilde{P}_1 &= \frac{1}{\sqrt{l}} \sum_{i=1}^{l} P_i \\
&= \frac{1}{\sqrt{l}} \left(a_1 + b_1\rho_1 X_1 + b_{-1}\rho_{-1} \sum_{i=2}^{l} X_i - b_1 Z_1 - b_{-1} \sum_{i=2}^{l} Z_i \right) \\
&= \frac{1}{\sqrt{l}} \left(\frac{b_1\rho_1 + (l-1)\, b_{-1}\rho_{-1}}{l} \sum_{i=1}^{l} X_i + \frac{l-1}{l} \left(b_1\rho_1 - b_{-1}\rho_{-1}\right) X_1 - \frac{b_1\rho_1 - b_{-1}\rho_{-1}}{l} \sum_{i=2}^{l} X_i \right) \\
&+ \frac{1}{\sqrt{l}} \left(a_1 - \frac{b_1 + (l-1)\, b_{-1}}{l} \sum_{i=1}^{l} Z_i - \frac{l-1}{l} \left(b_1 - b_{-1}\right) Z_1 + \frac{b_1 - b_{-1}}{l} \sum_{i=2}^{l} Z_i \right) \\
&= \frac{1}{\sqrt{l}} \left(\frac{b_1\rho_1 + (l-1)\, b_{-1}\rho_{-1}}{l} \sum_{i=1}^{l} X_i + \frac{b_1\rho_1 - b_{-1}\rho_{-1}}{l} \left((l-1)\, X_1 - \sum_{i=2}^{l} X_i \right) \right) \\
&+ \frac{1}{\sqrt{l}} \left(a_1 - \frac{b_1 + (l-1)\, b_{-1}}{l} \sum_{i=1}^{l} Z_i - \frac{b_1 - b_{-1}}{l} \left((l-1)\, Z_1 - \sum_{i=2}^{l} Z_i \right) \right).
\end{aligned}$$

Note that

$$\begin{aligned}\operatorname{cov}\left(\sum_{i=1}^{l} X_i, (l-1) X_1 - \sum_{i=2}^{l} X_i\right) &= 0 \\ \operatorname{cov}\left(\sum_{i=1}^{l} Z_i, (l-1) Z_1 - \sum_{i=2}^{l} Z_i\right) &= 0.\end{aligned}$$

Hence

$$\begin{aligned}\tilde{P}_1 &= \frac{b_1\rho_1 + (l-1) b_{-1}\rho_{-1}}{l}\tilde{X}_1 + \frac{b_1\rho_1 - b_{-1}\rho_{-1}}{l}\frac{1}{\sqrt{l}}\left((l-1) X_1 - \sum_{i=2}^{l} X_i\right) \\ &+ \frac{a_1}{\sqrt{l}} - \frac{b_1 + (l-1) b_{-1}}{l}\tilde{Z}_1 - \frac{b_1 - b_{-1}}{l}\frac{1}{\sqrt{l}}\left((l-1) Z_1 - \sum_{i=2}^{l} Z_i\right),\end{aligned} \tag{B-40}$$

where $\tilde{X}_1$, $\tilde{Z}_1$, $\left((l-1) X_1 - \sum_{i=2}^{l} X_i\right)$ and $\left((l-1) Z_1 - \sum_{i=2}^{l} Z_i\right)$ are mutually independent. Consequently, since $\rho_1 < \rho_{-1}$,

$$\operatorname{var}\left(\tilde{X}_1|\tilde{P}_1\right)^{-1} < \tau_X + \left(\frac{b_1\rho_1 + (l-1) b_{-1}\rho_{-1}}{b_1 + (l-1) b_{-1}}\right)^2 \tau_Z < \tau_X + \rho_{-1}^2\tau_Z = \operatorname{var}\left(\tilde{X}_k|\tilde{P}_k\right)^{-1},$$

completing the proof.

Proof of Lemma 6: From (26), Lemma A-5, and (B-26),

$$\mathbb{E}\left[X_1 - P_1|P_1\right] = \frac{S_1 + \mathbb{E}\left[Z_1|P_1\right]}{\frac{1}{\gamma}\frac{1}{n_1}\int_0^{n_1}\frac{1}{\operatorname{var}(X_1|Y_{i1},e_{i1},P_1)}di} = -\frac{S_1 + \mathbb{E}\left[Z_1|P_1\right]}{\frac{1}{n_1}\int_0^{n_1}\frac{\partial\theta_{i1}}{\partial P_1}di} = -\frac{S_1 + \mathbb{E}\left[Z_1|P_1\right]}{1 + \frac{\rho_1}{\gamma}\tau_Z}\frac{\partial P_1}{\partial Z_1}.$$

Note that

$$\frac{\partial}{\partial P_1}\mathbb{E}\left[Z_1|P_1\right] = \frac{\operatorname{cov}\left(Z_1, P_1\right)}{\operatorname{var}\left(P_1\right)} = \frac{\operatorname{cov}\left(Z_1, P_1\right)}{\operatorname{var}\left(Z_1\right)}\frac{\operatorname{var}\left(Z_1\right)}{\operatorname{var}\left(P_1\right)} = \frac{\partial P_1}{\partial Z_1}\frac{\operatorname{var}\left(Z_1\right)}{\left(\frac{\partial P_1}{\partial X_1}\right)^2\operatorname{var}\left(X_1\right) + \left(\frac{\partial P_1}{\partial Z_1}\right)^2\operatorname{var}\left(Z_1\right)}.$$

Hence (and using the fact that $\frac{\partial P_1}{\partial Z_1}$ is independent of Z_1)

$$\frac{\partial}{\partial P_1}\mathbb{E}\left[X_1 - P_1|P_1\right] = -\frac{1}{\left(1 + \frac{\rho_1}{\gamma}\tau_Z\right)\left(\rho_1^2\frac{\operatorname{var}(X_1)}{\operatorname{var}(Z_1)} + 1\right)},$$

completing the proof.

C Analysis of public signal economy

For notational transparency, we consider a single-asset version of our economy, and omit all asset subscripts. We note also that the result below does not use the assumption in our model that $\mathbb{E}[e_i]$ is constant across agents i.

Proposition C-1 *Consider the benchmark economy described in subsection 3.1, in which agents do not possess any private information about the asset's cash flow X, but instead all observe a public signal of the form $Y = X + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \tau_\epsilon^{-1})$. In such a setting, each agent's expected utility is decreasing in the precision of the public signal, τ_ϵ.*

Proof: Agent i's terminal wealth is

$$W_i = e_i P + (\theta_i + e_i)(X - P),$$

and he optimally chooses the portfolio

$$\theta_i + e_i = \frac{1}{\gamma} \frac{\mathbb{E}[X|Y] - P}{\operatorname{var}(X|Y)}.$$

So agent i's expected utility at the trading stage is

$$\begin{aligned}
\mathbb{E}[-\exp(-\gamma W_i)|Y, P] &= \mathbb{E}\left[-\exp\left(-\gamma e_i P - \frac{\mathbb{E}[X|Y] - P}{\operatorname{var}(X|Y)}(X - P)\right)|Y, P\right] \\
&= -\exp\left(-\gamma e_i P - \frac{(\mathbb{E}[X|Y] - P)^2}{\operatorname{var}(X|Y)} + \frac{1}{2}\frac{(\mathbb{E}[X|Y] - P)^2}{\operatorname{var}(X|Y)}\right) \\
&= -\exp\left(-\gamma e_i P - \frac{1}{2}\frac{(\mathbb{E}[X|Y] - P)^2}{\operatorname{var}(X|Y)}\right).
\end{aligned}$$

We evaluate

$$\mathbb{E}\left[-\exp\left(-\gamma e_i P - \frac{1}{2}\frac{(\mathbb{E}[X|Y] - P)^2}{\operatorname{var}(X|Y)}\right)|e_i\right]. \tag{C-1}$$

Expanding, this expression equals

$$\mathbb{E}\left[-\exp\left(-\gamma e_i \mathbb{E}[X|Y] + \gamma e_i(\mathbb{E}[X|Y] - P) - \frac{1}{2}\frac{(\mathbb{E}[X|Y] - P)^2}{\operatorname{var}(X|Y)}\right)|e_i\right].$$

By market clearing,

$$\frac{1}{\gamma}\frac{\mathbb{E}[X|Y] - P}{\operatorname{var}(X|Y)} = S + Z,$$

i.e.,

$$\mathbb{E}[X|Y] - P = \gamma \text{var}(X|Y)(S+Z),$$

and so (C-1) equals

$$\mathbb{E}\left[-\exp\left(-\gamma e_i \mathbb{E}[X|Y] + \frac{\gamma^2 \text{var}(X|Y)}{2}\left(2e_i(S+Z) - (S+Z)^2\right)\right)|e_i\right].$$

Moreover,

$$\mathbb{E}[X|Y] = \frac{\tau_X \mathbb{E}[X] + \tau_\epsilon Y}{\tau_X + \tau_\epsilon} = \frac{\tau_\epsilon^{-1}\mathbb{E}[X] + \tau_X^{-1} Y}{\tau_X^{-1} + \tau_\epsilon^{-1}} = \frac{(\text{var}(Y) - \text{var}(X))\mathbb{E}[X] + \text{var}(X)Y}{\text{var}(Y)}.$$

Hence (C-1) equals

$$\mathbb{E}\left[\exp\left(-\gamma e_i \mathbb{E}[X] + \frac{\gamma^2 e_i^2}{2}\frac{\text{var}(X)^2}{\text{var}(Y)} + \frac{\gamma^2 \text{var}(X|Y)}{2}\left(2e_i(S+Z) - (S+Z)^2\right)\right)|e_i\right].$$

By the law of total variance,

$$\text{var}(X) = \text{var}(X|Y) + \text{var}(E[X|Y]) = \text{var}(X|Y) + \frac{\text{var}(X)^2}{\text{var}(Y)}.$$

So (C-1) equals

$$\begin{aligned}
& \mathbb{E}\left[-\exp\left(-\gamma e_i \mathbb{E}[X] + \frac{\gamma^2 e_i^2}{2}\text{var}(X) + \frac{\gamma^2 \text{var}(X|Y)}{2}\left(2e_i(S+Z) - (S+Z)^2 - e_i^2\right)\right)|e_i\right] \\
= \ & \mathbb{E}\left[-\exp\left(-\gamma e_i X - \frac{\gamma^2 \text{var}(X|Y)}{2}(e_i - (S+Z))^2\right)|e_i\right].
\end{aligned}$$

This expression is increasing in $\text{var}(X|Y)$, completing the proof.

D Numerical example of subsection 6.2

Table D-1: Parameters used in the numerical example of subsection 6.2

	Parameter	Value
Reciprocal of var$(\tilde{X}_k)$	τ_X	1
Reciprocal of var$(\tilde{Z}_k)$	τ_Z	9
Reciprocal of var$(\tilde{u}_{ik})$	τ_u	2.25
Mean asset payoff	$\mathbb{E}\left[\tilde{X}_k\right]$	1.5
Per-capita endowment of asset	$\tilde{S}$	0.1
Coefficient of absolute risk aversion	γ	0.6
Assets in economy	m	4
Assets covered by index	l	4
Cost of full participation in financial markets	κ	0.479
Cost of participation via indexing (initial)	κ_1	0.130
Cost of participation via indexing (after reduction)	κ_1	0.1287

Note that condition (8) is satisfied, since

$$\tau_X - 4\gamma^2\left(\tau_Z^{-1} + \tau_u^{-1}\right) = 0.2 > 0.$$

The precision of agents τ_i is displayed in the following figure:

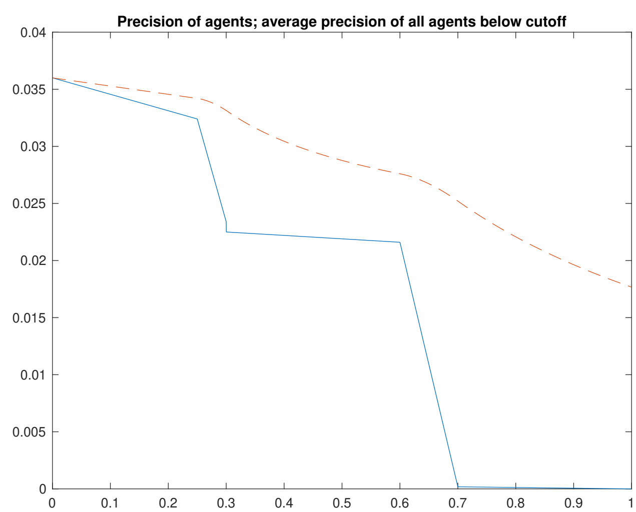

Figure D-1: Precision parameter τ_i. Horizontal axis is identity of agent i. Dashed line shows average precision if all agents with greater precision participate, i.e., $\frac{1}{n}\int_0^n \tau_i di$.

E Online appendix

E.1 Heterogenous risk aversion

Equation (B-7) in the existing proof of Lemma 3 generalizes to

$$\rho = \frac{\int_0^N \frac{\tau_i}{\gamma_i} di}{\int_0^N \left(1 - \frac{\rho}{\gamma_i}\tau_u\right) di},$$

yielding (34).

Lemma 1 generalizes straightforwardly to

$$\int_N \frac{1}{\gamma_i} \frac{1}{\text{var}\left(X|y_i, e_i, P\right)} di = \frac{\int_N \frac{1}{\gamma_i} di}{\text{cov}\left(X - P, X\right)},$$

and so Corollary 1 becomes

$$\mathbb{E}\left[X - P\right] = \frac{S\text{cov}\left(X - P, X\right)}{\frac{1}{|N|}\int_N \frac{1}{\gamma_i} di}.$$

For the remainder of this discussion, we focus on the case in which signal precision is constant across agents. In this case, Lemma 1 becomes simply

$$\frac{1}{\text{var}\left(X|y_i, e_i, P\right)} = \frac{1}{\text{cov}\left(X - P, X\right)},$$

and so

$$\text{cov}\left(X - P, X\right) = \text{var}\left(X|y_i, e_i, P\right) = \frac{1}{\tau_X + \tau_i + \rho^2\left(\tau_Z + \tau_u\right)}.$$

Recall that $\rho = -\frac{\frac{\partial P}{\partial X}}{\frac{\partial P}{\partial Z}}$, so that

$$\begin{aligned}
\frac{\partial P}{\partial Z} = -\frac{1}{\rho}\frac{\partial P}{\partial X} &= -\frac{1}{\rho}\frac{\text{cov}\left(P, X\right)}{\text{var}\left(X\right)} \\
&= -\frac{1}{\rho}\frac{\left(\text{var}\left(X\right) - \text{cov}\left(X - P, X\right)\right)}{\text{var}\left(X\right)} \\
&= -\frac{1}{\rho}\left(1 - \frac{\text{cov}\left(X - P, X\right)}{\text{var}\left(X\right)}\right)
\end{aligned}$$

so that

$$\frac{\text{cov}(P, e_i)}{\text{var}\left(e_i\right)} = \frac{\partial P}{\partial Z}\frac{\text{var}\left(Z\right)}{\text{var}\left(e_i\right)} = -\frac{1}{\rho}\left(1 - \frac{\text{cov}\left(X - P, X\right)}{\text{var}\left(X\right)}\right)\frac{\text{var}\left(Z\right)}{\text{var}\left(e_i\right)}.$$

To evaluate the expected utility of participation, recall from the proof of Proposition 1 that, conditional on e_i, it is given by

$$\begin{aligned} & -\left(d_i D\right)^{-\frac{1}{2}} \exp\left(-\gamma_i e_i \mathbb{E}\left[X\right] + \frac{\gamma_i^2 {e_i}^2}{2\tau_X}\right) \\ \times \quad & \exp\left(-\frac{1}{2}\frac{\left(\mathbb{E}[X-P] - \frac{\text{cov}(P,e_i)}{\text{var}(e_i)}\left(e_i - S\right) - \text{cov}(X-P,X)\gamma_i e_i\right)^2}{D\text{var}(X|e_i,P)}\right). \end{aligned}$$

To take expectations over e_i, we write this expression as $-\left(d_i D\right)^{-\frac{1}{2}}\exp\left(K_0 + K_1 e_i + K_2 e_i^2\right)$, and where K_0, K_1 and K_2 are constants given by

$$\begin{aligned} K_0 &= -\frac{1}{2}\frac{\left(\mathbb{E}[X-P] + \frac{\text{cov}(P,e_i)}{\text{var}(e_i)}S\right)^2}{D\text{var}(X|e_i,P)} \\ K_1 &= -\gamma_i \mathbb{E}\left[X\right] - \frac{\left(\mathbb{E}[X-P] + \frac{\text{cov}(P,e_i)}{\text{var}(e_i)}S\right)\left(-\frac{\text{cov}(P,e_i)}{\text{var}(e_i)} - \text{cov}(X-P,X)\gamma_i\right)}{D\text{var}(X|e_i,P)} \\ K_2 &= \frac{\gamma_i^2}{2\tau_X} - \frac{1}{2}\frac{\left(-\frac{\text{cov}(P,e_i)}{\text{var}(e_i)} - \text{cov}(X-P,X)\gamma_i\right)^2}{D\text{var}(X|e_i,P)}. \end{aligned}$$

E.2 Trading costs instead of fixed participation costs

Here, we consider an alternative specification of costs, with agents paying a transaction cost equal to $\frac{1}{2}\hat{\kappa}\theta_i^2$. So agent i chooses θ_i to maximize expected utility over terminal wealth

$$\left(\theta_i + e_i\right)\left(X - P\right) + e_i P - \frac{1}{2}\hat{\kappa}\theta_i^2.$$

The certainty equivalent is

$$\left(\theta_i + e_i\right)\mathbb{E}\left[X - P|y_i, e_i, P\right] + e_i P - \frac{\gamma}{2}\left(\theta_i + e_i\right)^2 \text{var}\left(X|y_i, e_i, P\right) - \frac{1}{2}\hat{\kappa}\theta_i^2,$$

yielding

$$\mathbb{E}\left[X - P|y_i, e_i, P\right] = \gamma\left(\theta_i + e_i\right)\text{var}\left(X|y_i, e_i, P\right) + \hat{\kappa}\theta_i,$$

and hence

$$\theta_i = \frac{\mathbb{E}\left[X - P|y_i, e_i, P\right] - \gamma e_i \text{var}\left(X|y_i, e_i, P\right)}{\gamma\text{var}\left(X|y_i, e_i, P\right) + \hat{\kappa}}.$$

We characterize the effect of changes in the parameter $\hat{\kappa}$ on price efficiency in the simplest case of heterogeneous signal precision, in which a fraction λ_1 and λ_2 of agents have signal precisions τ_1 and τ_2, where without loss we assume $\tau_1 \geq \tau_2$ In addition, we focus on the effect of small changes in $\hat{\kappa}$ away from 0.

Consider a perturbation to X and Z in which X is changed by $\frac{\partial P}{\partial Z}$ and Z is changed by $-\frac{\partial P}{\partial X}$. By construction, this leaves the price P unchanged. By market clearing,

$$\begin{aligned} 0 &= \lambda_1 \frac{\frac{\tau_1 \frac{\partial P}{\partial Z} - \rho\tau_u \frac{\partial P}{\partial X}}{\tau_X + \tau_1 + \rho^2(\tau_Z + \tau)} + \frac{\gamma \frac{\partial P}{\partial X}}{\tau_X + \tau_1 + \rho^2(\tau_Z + \tau_u)}}{\frac{\gamma}{\tau_X + \tau_1 + \rho^2(\tau_Z + \tau_u)} + \hat{\kappa}} \\ &+ \lambda_2 \frac{\frac{\tau_2 \frac{\partial P}{\partial Z} - \rho\tau_u \frac{\partial P}{\partial X}}{\tau_X + \tau_2 + \rho^2(\tau_Z + \tau)} + \frac{\gamma \frac{\partial P}{\partial X}}{\tau_X + \tau_2 + \rho^2(\tau_Z + \tau_u)}}{\frac{\gamma}{\tau_X + \tau_2 + \rho^2(\tau_Z + \tau_u)} + \hat{\kappa}}. \end{aligned}$$

Hence

$$0 = \lambda_1 \frac{\tau_1 + \rho^2\tau_u - \gamma\rho}{\gamma + \hat{\kappa}\left(\tau_X + \tau_1 + \rho^2\left(\tau_Z + \tau_u\right)\right)} + \lambda_2 \frac{\tau_2 + \rho^2\tau_u - \gamma\rho}{\gamma + \hat{\kappa}\left(\tau_X + \tau_2 + \rho^2\left(\tau_Z + \tau_u\right)\right)}. \tag{E-1}$$

The RHS of (E-1) is strictly positive at $\rho = 0$. Stable equilibria are given by solutions to (E-1) at which the RHS is decreasing in ρ.

The derivative to the RHS with respect to $\hat{\kappa}$ is

$$\begin{aligned} &-\lambda_1 \frac{\left(\tau_1 + \rho^2\tau_u - \gamma\rho\right)\left(\tau_X + \tau_1 + \rho^2\left(\tau_Z + \tau_u\right)\right)}{\left(\gamma + \hat{\kappa}\left(\tau_X + \tau_1 + \rho^2\left(\tau_Z + \tau_u\right)\right)\right)^2} \\ &- \lambda_2 \frac{\left(\tau_2 + \rho^2\tau_u - \gamma\rho\right)\left(\tau_X + \tau_2 + \rho^2\left(\tau_Z + \tau_u\right)\right)}{\left(\gamma + \hat{\kappa}\left(\tau_X + \tau_2 + \rho^2\left(\tau_Z + \tau_u\right)\right)\right)^2}. \end{aligned}$$

At $\hat{\kappa} = 0$, this expression has the same sign as

$$\begin{aligned} &-\lambda_1 \left(\tau_1 + \rho^2\tau_u - \gamma\rho\right)\left(\tau_X + \tau_1 + \rho^2\left(\tau_Z + \tau_u\right)\right) \\ &- \lambda_2 \left(\tau_2 + \rho^2\tau_u - \gamma\rho\right)\left(\tau_X + \tau_2 + \rho^2\left(\tau_Z + \tau_u\right)\right). \end{aligned} \tag{E-2}$$

Also at $\hat{\kappa} = 0$, there is a unique stable equilibrium, with

$$\lambda_1 \left(\tau_1 + \rho^2\tau_u - \gamma\rho\right) = -\lambda_2 \left(\tau_2 + \rho^2\tau_u - \gamma\rho\right),$$

with both sides of this equality positive. So at the equilibrium value of ρ, expression (E-2) has the same sign as

$$-\left(\tau_X + \tau_1 + \rho^2\left(\tau_Z + \tau_u\right)\right) + \left(\tau_X + \tau_2 + \rho^2\left(\tau_Z + \tau_u\right)\right) = \tau_2 - \tau_1 \leq 0.$$

Hence an increase in the cost parameter away from 0 reduces equilibrium price efficiency ρ.