FTG Working Paper Series

The Rise of Factor Investing: "Passive" Security Design and Market Implications

by

Lin William Cong (??)
Shiyang Huang
Douglas Xu

Working Paper No. 00133-00

Finance Theory Group

www.financetheory.com

The Rise of Factor Investing: “Passive” Security Design and Market Implications*

Lin William Cong† Shiyang Huang‡ Douglas Xu§

First draft: April 2016; this draft: June 2023

Abstract

We model financial innovations such as Exchange-Traded Funds, smart beta products, and many index-based vehicles as composite securities (CSs) that facilitate trading the common factors in assets' liquidation values. Through accessing a larger basket of assets in endogenously chosen proportions, CSs reduce investors' duplication of effort in trading multiple securities and attract more factor investors. We characterize analytically how competitive CS designers in equilibrium optimally select liquid underlying assets representative of the factors and find corroborating evidence in ETF data. CS trading entails investors' strategic and active decisions, consequently impounding more systematic information into prices. Their rise creates leads to greater informational efficiency, price variability, and co-movements in the underlying asset markets, as well as potentially heterogeneous effects on liquidity and asset-specific information acquisition/incorporation, depending on the importance of factors for asset value. The predictions explain and reconcile the rich (and often mixed) empirical observations about various types of CSs in the extant literature.

JEL Classification: D40, D82, G11, G14, G23

Keywords: Asset Pricing, ETFs, Indexing, Informational Efficiency, Security Design.

*The authors are grateful to Douglas Diamond, Zhiguo He, Charles Lee, Maureen O'Hara, Wei Xiong, and Mao Ye for many helpful discussions. They also thank Ehsan Azarmsa, David Chapman, Jesse Davis, Taejin Kim, John Loudis, Lubos Pastor, Gideon Saar, Pietro Veronesi, Ivo Welch, Yizhou Xiao, Liyan Yang, Xuewei Yang, Adam Zawadowski, Yao Zeng, and conference and seminar participants at Chicago Booth, University of Exeter, HKUST, INSEAD, China International Conference of Finance (CICF), Financial Intermediation Research Society (FIRS), Econometric Society Meeting (Asia), AsianFA Meeting, NBER Market Microstructure Meeting, and UBC Summer Finance Conference for useful comments and feedback. Cong especially thanks Charles Lee for his work and advice that led to the initial inquiries concerning the topic and the Ewing Marion Kauffman Foundation for generous research funding. This paper subsumes partial results of a 2016 working paper, “Rise of Factor Investing: Asset Prices, Informational Efficiency, and Security Design,” for which Yu Pu and Sean Yin provided excellent research assistance. Send correspondence to will.cong@cornell.edu.

†Cornell University SC Johnson College of Business (Johnson) and NBER

‡The University of Hong Kong School of Economics and Finance

§University of Florida Warrington College of Business

1 Introduction

The last two decades have witnessed a drastic growth in passive investing in both size (from 2% of the U.S. equity market capitalization in 1998 to about 14% in March 2020, Appel, Gormley, and Keim, 2016) and participation (e.g., through the adoption of defined-contribution pension plans, Gomes, Haliassos, and Ramadorai, 2021), accounting for more than a third of mutual fund market and more than half of equity under management. Exchange-traded funds (ETFs) stand out in particular with their fast-growing asset under management (AUM) surpassing $10 trillion U.S. dollars in 2021 and the number of product offerings approaching 10,000 by the end of 2022.[1] Also ever prominent over the past decade is a new generation of ETF products characterized by active and frequent changes in constituent weights, security selections unconstrained by benchmark indices, or special designs catered to investors' attention (Easley, Michayluk, O'Hara, and Putniņš, 2021; Ben-David, Franzoni, Kim, and Moussawi, 2023). Smart beta ETFs constitute the most salient example (Figure 1).[2] The impact of passive investing, ETFs, and smart beta trading on asset prices and the informational efficiency of financial markets remains little understood, and empirical evidence is often mixed. Moreover, research on the security design of these so-called "passive investing" products is missing.

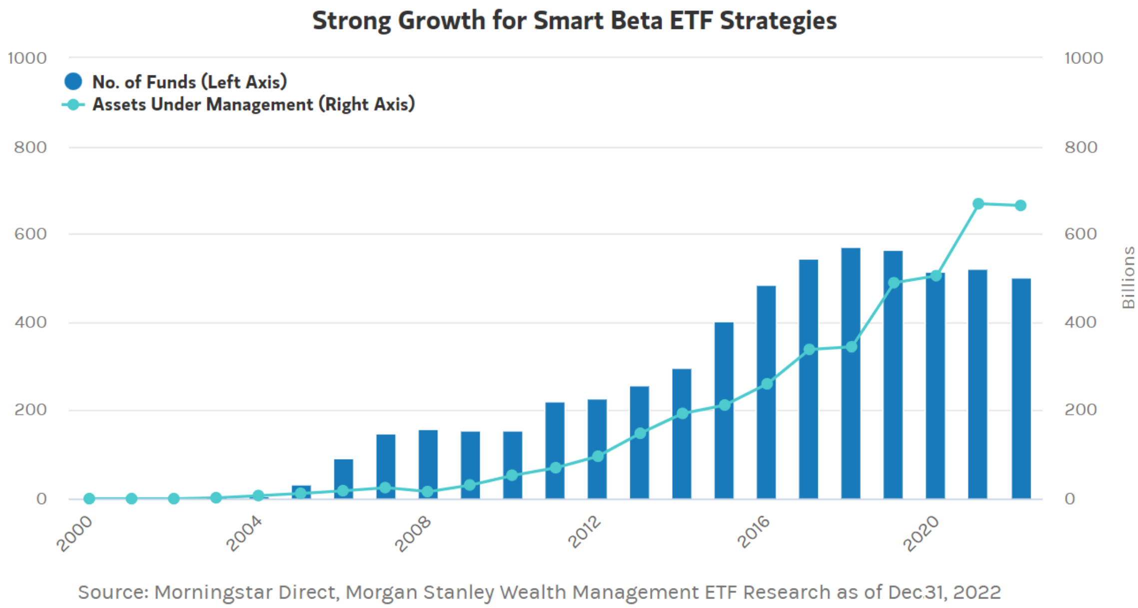
To bridge this knowledge gap, we model passive mutual funds, index ETFs, and smart beta products (which used to be grouped under passive investing until recently), etc., as composite securities (CSs) that provide vehicles for investors to exploit their information on systematic components of assets' quality and liquidation values.[3] Through accessing a larger basket of assets in endogenously-chosen proportions, CSs attract factor investors who want to reap their information rent or hedge against systematic risks, and incorporate greater factor information into asset prices. As the first attempt to analyze the security design aspect of so-called passive investing with endogenous market segmentation and sensible informational frictions, we theoretically derive and empirically verify that the optimal CS design entails underlying asset weights proportional to their factor exposure for representativeness and inversely proportional

[1]According to ETFGI, www.etfgi.com assessed on Dec 28, 2022.

[2]Investors who passively tracked benchmark indices in 2022 likely felt the pain of steep losses. Against a backdrop of potentially elevated inflation and higher interest rates, smart beta strategies become attractive again for boosting portfolio performance without taking undue risk or incurring high management fees.

[3]Our terminology follows Gorton and Pennacchi (1993). CS is a broad concept including a wide variety of financial products, including mortgages and asset-backed securities, real estate investment trusts, etc.

Figure 1: Strong Growth for Smart Beta ETF Strategies



to the equilibrium price impacts (illiquidity).

Our model yields further asset pricing implications that are consistent with recent empirical studies. First, the model predicts that introducing CSs incorporates more factor information and leads to greater informational efficiency, higher price variability, and return co-movements, contrary to the rhetoric that passive investing contributes nothing to price discovery. Second, introducing CSs decreases the price impacts and improves liquidity in underlying asset markets for and only for assets whose systematic element of the liquidation value is more prominent (and thus have more factor speculators trading them). Third, introducing CSs similarly increases endogenous asset-specific information acquisition and pricing efficiency for and only for assets with greater factor exposure and low asset-specific risk.

Conceptually, we argue that much of the so-called passive investing is really factor investing in disguise. While initial passive mutual funds and index ETFs were designed to track exogenously given indices closely and appear distinct from active strategies, the rise of the industry- and characteristic-based ETFs and smart beta strategies fostered the incursion of ETFs into active investing, making them really hybrid form of investment management. Therefore, we offer the first theory, if not the first academic study overall, incorporating and highlighting that passive investing is not so passive after all because (i) investors make active decisions on what and how much factor exposure to have, and (ii) the CS design is endogenous instead of fully index-based or market-weighted. A fast-emerging empirical literature corroborates our view.[4] Some point out that ETFs may deviate from their benchmarks, with traders behaving similarly to active retail investors (Bhattacharya, Loos, Meyer, and Hackethal, 2017). Others demonstrate that index funds and ETFs are active in form and functionality (Easley, Michayluk, O'Hara, and Putniņš, 2021; Akey, Robertson, and Simutin, 2021). Outside the equity market, Koont, Ma, Pastor, and Zeng (2022) document that corporate bond ETFs actively manage their portfolios with cash and only a subset of index assets with large tracking errors. Our factor investing perspective jointly accommodates investors' endogenous participation in "passive investment" and the endogenous design of CS vehicles, offering an explanation for their observed "activeness" and general asset market implications.

Specifically, our model of speculation and liquidity trading features multiple assets, each with a liquidation value that derives from the exposure to a common component, as well as an asset-specific component. Factor speculators who receive informative signals about the common component endogenously participate in multiple asset markets. Asset-speculators, in contrast, receive asset-specific information and opt to trade, if at all, the assets they are informed about in the presence of noise traders. The trading and market-making are modeled à la Kyle (1985). In particular, market makers are specialized and competitive and set prices to break even. We also consider in model extensions endogenous information acquisition, liquidity trading that comes from factor hedging, and the alternative yet practically plausible setting with transparently observed trading volume of CSs.

To focus on general economic insights as opposed to the institutional details of any particular type of CS, we treat CSs as pass-through vehicles that offer weighted bundles of underlying

[4]Rauterberg and Verstein (2013) shows that even for the most "objective" indices, human discretion and value judgment constitute essential ingredients. In legal studies, Robertson (2023) points out that the common interpretation of S&P 500 as a passive index is incorrect because typically more than 500 stocks satisfy the eligibility conditions for inclusion; Sharfman and Deluard (2021) propose a selection risk disclosure to address the discretionary inclusion of stocks in S&P 500; Molk and Robertson (2023) show that funds that track the most prominent index, the S&P 500 do not commit, in a legally enforceable sense, to holding even a representative sample of the underlying index, nor do they commit to replicating the returns of that index, with departures from the indices as a common practice. Li, Liu, and Wei (2023) document that from 1980 to 2018, about 38% of index membership and 97% of the index additions to the S&P 500 index involve discretionary considerations beyond its published rules. More recently, Ben-David, Franzoni, Kim, and Moussawi (2023) shows that providers issuing specialized ETFs track attention-grabbing themes to cater to investor demands.

assets to clients for service fees.[5] The key friction that prevents CSs from being completely redundant is the cost of trading each asset, which could come from monitoring cost, illiquidity, indivisibility of shares, attention cost, etc., which is realistic and endogenously creates market segmentation. A market for CSs naturally arises because of the reduction of duplication of each factor investor's effort trading each asset (henceforth referred to as "trading cost"). CS sponsors are thus financial intermediaries serving a function akin to how banks help lenders avoid the duplication of effort of borrower monitoring (Diamond, 1984). This recognition allows us to understand how CS sponsors compete for customers, which in turn helps rationalize why CSs such as index funds do not simply follow market weights. Moreover, because these factor speculators are not really passive "free-riders," their increased participation in the CS market, which in turn involves trading the underlying assets, affects the way various types of information get impounded in asset prices, with rich asset pricing implications.

We formally define and characterize the unique subgame-perfect *factor investing equilibrium* (FIE) in which CS sponsors first compete for entry by optimally choosing the designs and fees of the CS products offered, followed by a canonical stage of informed trading and market-making in linear strategies. With a competitive CS sponsoring market in which entrant CS sponsors can freely design arbitrary CS products with zero marginal cost (after having incurred the fixed setup cost), any speculating strategy involving simultaneously trading both CSs and underlying assets can be implemented through another properly designed CS. As a result, in equilibrium, all factor speculators trade an optimally designed CS product to exploit their informational advantage about the common component in asset value. Thanks to this property of equilibrium factor speculating, we show that the introduction of CS sponsoring increases the number of factor speculators that effectively trade each underlying asset.

One key prediction of the model concerns the equilibrium design of CS: the weight of an underlying securities is proportional to the common factor exposure and inversely proportional to its equilibrium illiquidity. The factor exposure dictates how effective a CS is for factor investing

[5]We deliberately design CS products as pass-through vehicles for trading a basket of underlying assets. There are broad interpretations of CS products in our model, and the interpretations include passive mutual funds and ETFs. While ETFs are tradable and have a secondary market, passive mutual funds do not have a secondary market. We do not want to limit our applications to ETFs but aim to capture the underlying and common features of passive mutual funds and ETFs–functioning as pass-through vehicles for trading a basket of underlying assets. Even for ETFs, the existence of APs makes ETFs closely track underlying indices, effectively making ETFs as pass-through vehicles.

while the illiquidity consideration allows factor speculators to internalize the collective price impact when they all trade a particular CS. We further consider factor hedgers' and liquidity traders' equilibrium strategies and derive a duality result: the CS design that maximizes factor-speculators' profit is also the one minimizing factor-hedgers' trading costs. We empirically test these novel model predictions focusing on US equity ETFs from January 2000 to December 2008 by analyzing the determinants of stocks' portfolio weights within the ETF.[6] We indeed find that within an ETF basket, there is a positive association between one particular stock's exposure to the ETF index (or the factor it represents) and its portfolio weight, and a negative association between its market illiquidity measure (Amihud, 2002) and portfolio weight.

Our model also delivers rich implications concerning the impact of introducing CSs on informational efficiency and asset pricing in the underlying markets. By increasing the number of factor speculators that effectively trade (via CSs) each underlying asset, we show that introducing CS increases the factor-specific and total efficiency in prices of underlying assets while reducing asset-specific pricing efficiency. Once we endogenize asset speculators' information acquisition, we find that asset-specific pricing efficiency can also improve for assets with high factor exposure and low asset-specific information because their prices contain less asset-specific information before introducing CS to start with. We are thus the first theoretical study to demonstrate that passive investing can increase price efficiency, whereas others typically predict a drop in price efficiency.[7] Moreover, introducing CSs increases the trading price volatility in underlying asset markets because CSs help incorporate more informed trading—an important source of return volatility. Because CSs incorporate more factor information into asset prices, their co-movement across underlying markets increases, too.

Interestingly, we find that the operation of a competitive CS sponsoring market could have mixed effects on the liquidity of the underlying assets, which differs from prior studies but is

[6] We are particularly interested in the excess portfolio weight, which measures how ETFs sponsor deviate from the market value weights, a self-rebalancing benchmark for ETFs not actively adjusting portfolio weights.

[7] For example, Bond and Garcia (2022) studies a competitive REE with exogenously given information endowment and index design and finds that a reduction in indexing cost leads to lower pricing efficiency. Baruch and Zhang (2022) exogenously vary the share of indexers in a conditional CAMP and argue pricing efficiency in terms of measured R^2 decreases. Malikov (2023) obtains similar results with endogenous information acquisition. The only exceptions that derive how passive investing can increase pricing efficiency are two studies subsequent to ours, Lee (2021) and Buss and Sundaresan (2023). Their arguments rely on exogenous investor participation or the interaction of liquidity trading and endogenous active investing, differing from ours on systematic information incorporation.

consistent with recent empirical findings about ETFs. Similar to Subrahmanyam and Titman (1999), the increased number of factor speculators trading via CS the underlying assets influence liquidity through two channels: First, more systematic information incorporated into prices through factor speculators' trading increases adverse selection that market makers face and reduces liquidity (*information inclusion* channel); second, an increased number of factor speculators trading an asset diminishes each participating factor speculator's trading aggressiveness due to the anticipated competition and joint price impact (*competition* channel). When the number of factor speculators is sufficiently high before introducing CSs, the competition effect dominates, improving liquidity. Otherwise, the information inclusion effect dominates, deteriorating market liquidity.[8]

Our paper relates to the fast-growing literature on the economic consequences of indexing and CS trading, especially ETFs. While index investing has existed for almost 50 years, much of the literature on the relationship between index investing, market efficiency, and other market characteristics is relatively recent. Our paper joins the earliest theoretical foundations for the study of CSs: Subrahmanyam (1991) highlights how liquidity traders could be better off trading CSs with mitigated adverse selection under assumptions about the signs of beta, homogeneous securities and equal basket weights. Gorton and Pennacchi (1993) belabor a similar point but focus on risk-averse liquidity traders and do not distinguish systematic versus asset-specific information. Also related are Stambaugh (2014) on the relationship between growth in passive investing and the decline in noise trading, and Yuan (2005) on how asset speculators trade more using CSs as hedging instruments. Specifically concerning ETFs, several studies focus on the liquidity mismatch between ETFs and underlying assets (Pan and Zeng, 2017; Koont, Ma, Pastor, and Zeng, 2022), the limits to the arbitrage of risk-averse, authorized participants (Malamud, 2015), and fragility related to asset tradability (Bhattacharya and O'Hara, 2015).

Unlike these articles, our paper does not rely on risk-aversion, exogenous CS weights, liquidity mismatch, or mispricing due to failure of arbitrage. Furthermore, we endogenize traders' participation in composite securities without exogenously assuming additional noise trading or informed trading when CS is introduced. Our theory not only generalizes to CSs the insight

[8] We also find that introducing CS increases the price impacts for trading low-beta or high-idiosyncratic-risk assets but decreases the price impacts for trading high-beta or low-idiosyncratic-risk assets.

in Gorton and Pennachi (1993) and Subrahmanyam (1991) that futures may provide a preferred venue for uninformed traders by removing the sensitivity to firm-specific informational asymmetries, but it also demonstrates that CSs provide an attractive venue for factor speculators and factor hedgers because of the reduction of duplication of trading cost.[9] Finally, we complement these studies by examining CS design and deriving closed-form solutions.

Our study is broadly related to financial innovation and intermediation.[10] Athanasoulis and Shiller (2000) prove market portfolio access to be welfare-enhancing; Rahi and Zigrand (2009) examine equilibrium security innovation that arbitrageurs use to exploit mispricing across exogoenously segmented markets; Dow (1998) studies introducing an exogenously given security for hedging whereas Shen, Yan, and Zhang (2014) endogenize financial innovations for alleviating collateral/margin constraints; Dai (2018) discusses bundling assets for coordinating investors' information acquisition. We differ in endogenizing the security design without exogenously restricting the securities traded, focusing on CSs, and highlighting the reduction of duplication of trading cost or effort—a major role of financial intermediaries as demonstrated in the banking literature (Diamond, 1984)—as a novel mechanism. Unlike studies such as Koont, Ma, Pastor, and Zeng (2022) that analyze arbitrage in the ETF market and the portfolio composition of creating and redemption baskets, we abstract away from the specific institutional setup of any particular type of ETFs to analyze a similar reduction of duplication applicable to a broad range of CSs (e.g., smart beta products, equity ETFs, etc., and not restricted to those with illiquid underlying assets). Recent articles also on the cost of indexing (Lee, 2021; Bond and Garcia, 2022) exogenously specify the trading costs, whereas we endogenize and microfound the decline in trading cost under endogenous CS sponsor competition. We also allow multiple assets and CS products and analyze the optimal design of CSs for such an intermediation function in a competitive environment, both theoretically and empirically.

All these innovations render our model one of the very few, if not the only, that is consistent with most empirical observations about index funds and ETFs in the literature, many of which appear conflicting. For example, Ben-David, Franzoni, and Moussawi (2018), Madhavan

[9]Earlier studies about the options market also study the endogenous choice of trading venue (Easley, O'hara, and Srinivas, 1998; Chakravarty, Gulen, and Mayhew, 2004).

[10]Our analysis also adds insight to the literature exploring linkages between investor information acquisition in segmented/linked markets (e.g., Cespa and Foucault (2014), Goldstein and Yang (2015), and Dai (2018)), by studying how asset speculators' information acquisition and trading behavior are affected by CS sponsoring.

and Sobczyk (2014), Hamm (2014), Bradley and Litan (2011), and Krause, Ehsani, and Lien (2014) find evidence that ETFs deprive liquidity of the underlying basket with elevated intraday return volatility. However, Ye (2019) finds that corporate bond ETFs improve underlying corporate bond liquidity. Da and Shive (2018) and Leippold, Su, and Ziegler (2015) document increased correlations of underlying assets' returns in the presence of ETFs and index futures. Israeli, Lee, and Sridharan (2017) show that increases in ETF ownership are associated with higher trading costs, greater return synchronicity, reduced firm-specific pricing efficiency, and less information acquisition. Glosten, Nallareddy, and Zou (2019) also find that ETF trading increases co-movement and return synchronicity but argue that ETFs actually increase informational efficiency. (Bai, Philippon, and Savov, 2016; Dávila and Parlatore, 2023) find that the price informativeness has increased, especially for stocks included in indices or more affected by passive investing. Huang, O'Hara, and Zhong (2020); Bhojraj, Mohanram, and Zhang (2020a) also document that industry ETFs can improve information efficiency among stocks with high industry exposure and low idiosyncratic risk.

2 Model Setup and An Illustration

We set up the general model and then discuss a two-asset case with exogenously assumed market segmentation and simplifying assumptions to illustrate key economic intuitions. We solve the general model with endogenous market segmentation in Section 3.

2.1 A Model of Speculative and Liquidity Trading

Assets and liquidation values. There are $K > 1$ underlying assets in the economy. Asset $k \in \{1, 2, \cdots, K\}$ has liquidation value v_k, which derives from its exposure to a common component γ (e.g., a systematic risk factor) and an asset-specific component α_k:

$$v_k = \bar{v}_k + \beta_k \gamma + \alpha_k. \tag{1}$$

β_k is the exposure of Asset k to γ, which represents a shock that affects all assets (e.g., a macroeconomic shock or an industry-wide technology shock). $\gamma \sim \mathcal{N}(0, \sigma_\gamma)$, $\sigma_\gamma > 0$ and $\alpha_k \sim \mathcal{N}(0, \sigma_{\alpha_k})$, $\sigma_{\alpha_k} > 0$ are mutually independent Normal distributions. $\bar{v}_k$ is the expected payoff of Asset k which we normalize to zero without loss of generality. β_k and the ex-ante distributions of γ and α_k, $\forall\ k$ are all agents' common knowledge.

In addition to the underlying assets, composite securities (CSs) can potentially be introduced by CS sponsors (to be described shortly). CSs are bundles of the underlying assets, with weights $\{w_k, k = 1, 2, \cdots, K\}$, subject to $\sum_{k=1}^{K} w_k = 1$. The payoff is simply $\sum_{k=1}^{K} w_k v_k$.

Market participants and information. To focus on the informational aspect of CS trading, we assume that all agents are risk-neutral. The baseline model features three types of investors and potential CS sponsors that interact in the economy. They are:

(i) One representative **asset speculator** for each asset. For Asset k, the asset speculator privately observes α_k and maximizes profit from trading Asset k.[11]

(ii) Numerous profit-maximizing **factor speculators** indexed by i, each endowed with a private signal about γ, i.e., factor speculator i observes $s_i = \gamma + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon)$, $\sigma_\epsilon > 0$.

(iii) An independent group of **liquidity traders** for each asset k with an exogenous aggregate group demand for liquidity n_k, where $n_k \sim \mathcal{N}(0, \sigma_{n_k})$, $\sigma_{n_k} > 0$.

(iv) Competitive **(potential) CS sponsors** designing CSs (determining the weights of underlying assets and management fees on top of the weighted sum of the underlying assets' prices) and deciding on which one(s) to launch to maximize profits, if launching any at all.

(v) One competitive and specialized **market maker** for each underlying asset market. The market maker for Asset k observes the total order flow ω_k and sets an asset price to break even à la Kyle (1985).[12]

Trading an asset or launching a CS is costly, as we discuss shortly. We allow factor speculators and potential CS sponsors to be in abundant supply so that the numbers of factor speculators participating in asset markets and of CS sponsors launching products are endogenously determined under such costly entries.

[11] Considering multiple asset speculators for each asset adds no new insights and is thus not included. We endogenize asset speculators' information acquisition in Section 6.1.

[12] We do not allow the market maker to condition the price on order flows of other assets in the baseline. This is a standard reduced-form way in the literature of capturing any friction precluding the market makers from instantaneously and fully processing and acting on all information in all securities (e.g., Boulatov, Hendershott, and Livdan, 2013). We relax this assumption in Appendix C.

Timeline and trading protocols. The discount factor is set to one, and all agents interact in three stages ($t = 0, 1, 2$). At $t = 0$, each potential CS sponsor decides whether to pay a fixed cost $\hat{C}$ to enter and offers the CS(s) upon entry, all to maximize the anticipated fee revenue subject to at least breaking even (participation constraint). A CS product specifies the portfolio weights $(w_1, \ldots, w_K)$ and the management fee F on top of the underlying asset prices. Factor speculators decide which CS product(s) to purchase. The fee can be contingent on the number of speculator purchases.

At $t = 1$, asset markets open. Fee-paying speculators can trade both (had they chosen so) the underlying assets and shares of CSs from the CS sponsors, who in turn mechanically trade the underlying assets with the corresponding weights according to the CS designs. Other speculators can only trade the underlying assets. Trading of either CS products or underlying assets will each involve a fixed cost $C > 0$. As is standard in market microstructure studies, each market maker observes the total order flows for the asset market in which she specializes. All speculators and CS sponsors submit market orders to the market maker in each underlying market, who then sets the asset price based on the asset's total order flows. At the end of $t = 1$, the CS sponsors deliver the CSs to their clients with the promised compositions.

At $t = 2$, the payoffs are realized for all the assets.

Our deliberate treatment of CSs as pass-through vehicles for trading a basket of underlying assets is worth highlighting. Because the technical barrier to entry is low and CS sponsors are typically not more informed about γ or α_k, they, in practice, are competitive and focused on fulfilling speculators' demands rather than speculating themselves. Moreover, any discrepancies in the prices of CS and its composites are supposedly arbitraged away, transmitting the first-order CS order flows into the weighted order flows of the underlying securities. Our setup, therefore, not only allows us to focus on the first-order implications of CS trading for asset prices and CS design but is also consistent with the real practice: For non-exchange-traded CSs, such as passive mutual funds, a change in demand is directly mapped to demand changes in the underlying assets; in the context of ETFs, the arbitrage by authorized participants largely ensures that the discounts and premiums of ETF share prices over the underlying asset values are sufficiently small (i.e., within the bid-ask spread, see, e.g., Engle and Sarkar, 2006).[13]

[13]Certain CSs, such as ETFs, can deviate in price from their underlying composites, often due to funding or

Key friction and CS trading. The various costs associated with trading underlying and composite securities constitute the key friction in our model. Widely recognized costs in trading include the lack of access to trading opportunities or costly searching for trading counterparties. For example, trading portfolios of illiquid assets, such as corporate bonds, often entails high costs associated with searching and trading over-the-counter (e.g., OHara, Wang, and Zhou, 2018). Retail investors' direct investment in the real estate sector is difficult, with significant entry barriers.[14] Transaction costs for trading small cap stocks or penny stocks easily amount to over 3% (Stoll and Whaley, 1983). Trading costs can also arise from geographical and legal constraints. For example, an average U.S. resident directly trading public companies listed in China has to open local brokerage and bank accounts. The cost of obtaining license for institutions is even higher.

Even for liquid and accessible assets (e.g., public equities), participation cost (e.g., setup cost), information cost (e.g. searching for assets and learning about value-relevant fundamentals), and attention/research cost (e.g., monitoring relevant markets) can be sizable for each trader (Vayanos and Wang, 2013). Market microstructure frictions may also add substantial costs, especially for retail investors and small institutions. For example, due to the indivisibility of shares, especially of the likes of Berkshire and Google, investors face significant frictions to get to the desirable portfolio weights.

CSs emerge because CS sponsors have comparative advantages in mitigating the aforementioned costs. First, as financial intermediaries, their operation reduces the duplication of these costs, similar to how delegated monitors (banks) reduce monitoring costs (Diamond, 1984) or information production costs (Veldkamp, 2006b). In practice, advances in IT enable the economy of scale in marketing/outreach for financial products, information acquisition, and asset trading. Reputable CS sponsors also have favorable access to financial services, not to mention that they may have advantages in recruiting research talent. CS sponsors with large scale of business can also easily divide CS shares so that divisibility of shares of the underlying assets

market illiquidity. This interesting phenomenon, though not our focus, is explored in studies such as Malamud (2015), Bhattacharya and OHara (2017), and Pan and Zeng (2017).

[14] Rahi and Zigrand (2009) estimate that the acquisition costs of property in the UK are around 8%, and in many countries the costs are even higher. In fact, many jurisdictions prohibit foreigners from purchasing property, but some CSs help investors gain exposure to foreign real estate markets. This is also the reason why property total return swaps (TRSs)—a form of CS—have gained popularity.

would no longer be a problem. Consequently, financial institutions such as Vanguard, State Street, and Blackrock offer CSs (e.g., passive mutual funds, ETFs, Smart Beta products) that are often index- or rule-based. They incur the (fixed) cost once for research and then trade the underlying assets while charging clients low management fees.

We model this key friction in reduced form by stipulating a fixed cost $C > 0$ to access and trade any underlying asset or a CS product.[15] While this simple and realistic specification abstracts from some finer details, we show that it yields novel economic insights and rich predictions that are corroborated by empirical evidence. One immediate consequence in the absence of CS is an endogenous market segmentation: Although factor speculators are informed about γ and have incentives to trade all underlying assets (that have loading on γ) directly, in equilibrium they may not do so due to the "trading cost." CSs potentially alleviate market segmentation by allowing each speculator to trade the underlying assets indirectly at low costs.

CS sponsoring and product competition. CS sponsors may be thought of as financial intermediaries and specialists in the packaging and trading of underlying securities. They still incur costs in trading underlying assets, but speculators effectively trade a basket of assets through CS sponsors by paying only the cost of C for trading the CS product and any management fees. Given that the entry cost for CS sponsors includes researching, accessing, monitoring, and trading each of the underlying assets, among others, the entry cost $\hat{C}$ is likely close to $K \cdot C$. When the number of speculators N_{CS} buying from a sponsor is large, $\hat{C}/N_{CS} < C$ naturally, reflecting the comparative advantages of sponsors and their important role in reducing the duplication of effort and attention as intermediaries.

To match real practice, we set the CS sponsor market to be competitive. Although the CS products are multi-dimensional without a strict rank, we introduce two intuitive concepts of dominance (i.e., preferred by speculators) in $t = 0$: (i) For the same product offering, a sponsor charging lower fees dominates another charging higher fees. (ii) For the same fee charged, one sponsor dominates another if her product offering nests the others'. Dominance (i) is intuitive; (ii) helps break speculators' indifference when some CS products are not used in equilibrium. Collectively, they ensure that due to competition, speculators may demand any CS product at

[15] We introduce a trading cost C_A for asset speculators in Section 6.1 to endogenize their participation in each asset market. This is immaterial in the baseline (Sections 3 and 5) where asset speculators' participation is exogenous, which is equivalent to requiring σ_α's to be sufficiently large relative to the cost C_A.

a competitive fee for all practical purposes.[16] This, in turn, allows us to pin down a unique equilibrium whenever a CS product is traded.[17] Finally, we assume that when factor speculators are indifferent among the CS products offered by different CS sponsors, they pick one randomly with equal probabilities.

Equilibrium definition. As is standard in the literature (e.g., Kyle, 1985), we focus on equilibria where speculators and market makers follow linear trading and pricing strategies:

Definition 2.1 (**Generalized Factor Investing Equilibrium (FIE)**). *An FIE is a subgame perfect equilibrium with CS being traded. It consists of* $\left\{\hat{\kappa}_k, \hat{\eta}_k, \hat{\lambda}_k, \hat{N}_k, P_k\right\}_{k=1}^{K}$, $\{w_k^j\}_{k=1}^{K}$, $\{\eta_{CS}^j\}$, $\{N_{CS}^j\}$, *and* $\{F^j\}$, *where* $j \in \mathbb{J}$ *indexes the countable set of CSs offered, such that:*

1. *Entrant CS sponsors offer CS product* $j \in \mathbb{J}$ *at* $t = 0$ *by specifying the weights* $\left(w_1^j, w_2^j, \ldots, w_K^j\right)$ *and fee* F^j *to maximize her anticipated fee revenue at* $t = 1$ *when the product is launcned. A sponsor enters only if she expects to at least break even.*

2. *The asset speculator for Asset* k *submits order* $x_k = \hat{\kappa}_k \cdot \alpha_k$ *at* $t = 1$ *given signal* α_k *to maximize her expected trading profit.*

3. $\hat{N}_k$ *factor speculators directly trade Asset* k *at* $t = 1$ *by each submitting an order* $\hat{y}_k = \hat{\eta}_k \cdot s$ *given a signal* s*, to break even net of trading costs* (C)*;*

4. N_{CS}^j *factor speculators choose to trade via the* j*th CS product, with an order* $y_{CS,j} = \eta_{CS,j} \cdot s$ *given signal* s*, to break even after CS fees and trading costs* ($C + F^j$)*;*

5. *Upon receiving a total order* ω_k*, the market maker for Asset* k *breaks even by setting* $P_k(\omega_k) = \lambda_k^{CS}\omega_k$.

[16] While large, passive index products continue to flourish, specialized ETFs provide exposure to everything from AI (e.g., AIEQ, BIKR, and ROBO), to religion (e.g., BIBL), to social sentiments (e.g., BUZZ). Motif, a successful fintech startup later acquired by Charles Schwab, allows users to suggest any investment themes to trade using Motif's ETFs.

[17] This setup essentially allows the CS sponsors to offer a full menu of designs. In Appendix B, we relax (ii) and show how the unique equilibrium in this baseline setting still constitutes an equilibrium with desirable properties (e.g., Pareto-undominated and participation-maximizing) in general, even when the CS sponsors do not offer a full menu from which the speculators subsequently choose.

The equilibrium F^j and N_{CS}^j are related because at $t = 0$, a potential CS sponsor solves:

$$\max_{\left\{F^j, \left(w_1^j, \ldots, w_K^j\right)\right\}} \sum_j F^j \cdot N_{CS}^j - \hat{C}, \tag{2}$$

where F^j is the fee she charges and N_{CS}^j is the anticipated equilibrium number of factor speculators choosing the jth CS product, which would be influenced by F^j. Because CS sponsors are competitive, they launch products and charge fees to just break even, a common pricing outcome in the literature (e.g., Veldkamp, 2006a,b; Van Nieuwerburgh and Veldkamp, 2010), and we later show in equilibrium only one CS sponsor enters.

2.2 CS Design and Informational Efficiency: An Illustration

Before we fully characterize the equilibrium, we illustrate the key economic insights and intuition concerning CS designs and the informational efficiency of asset prices. Trading a small loss of generality for transparency and clarity, let us specialize in this illustration to: (1) only two underlying assets, i.e., $K = 2$, (2) perfect signals, i.e., $\sigma_\epsilon^2 = 0$, (3) no asset-specific information asymmetry, i.e., $\sigma_{\alpha_1} = \sigma_{\alpha_2} = 0$ and that the underlying assets are symmetric other than the loading on the common component, i.e., (4) $\sigma_{n_1}^2 = \sigma_{n_2}^2 = \sigma_n^2$ and (5) $\beta_1 > \beta_2 > 0$.

Without CS, denote by N_k the number of factor speculators trading Asset k ($k = 1, 2$). The market maker for Asset 1 (MM1) receives total order flows $\omega_1 = N_1 \eta_1 \gamma + n_1$, and sets price $P_1 = \mathbb{E}[\beta_1 \gamma | \omega_1] = \lambda_1 \omega_1$. Market maker for Asset 2 (MM2) acts similarly and sets $P_2 = \lambda_2 \omega_2$, where $\omega_2 = N_2 \eta_2 \gamma + n_2$. Because $\sigma_\alpha^2 = 0$, asset speculators do not trade. The optimization problem for a factor speculator who trades Asset k becomes:

$$\Pi_k^F \equiv \max_{y_k} \mathbb{E}\left[y_k \left(\beta_k \gamma - P_k(\omega_k)\right) | \gamma\right]. \tag{3}$$

The solution follows from standard Kyle (1985)-style models:

$$Y_k(\gamma) = \frac{\beta_k \gamma}{(N_k + 1)\lambda_k} \quad \text{and} \quad \lambda_k = \frac{N_k \beta_k \eta_k \sigma_\gamma^2}{N_k^2 \eta_k^2 \sigma_\gamma^2 + \sigma_n^2}.$$

The above equation system yields:

$$\lambda_k = \frac{\beta_k \sigma_\gamma}{\sigma_n} \frac{\sqrt{N_k}}{N_k + 1} \quad \text{and} \quad \Pi_k^F = \frac{\beta_k \sigma_\gamma \sigma_n}{(N_k + 1)\sqrt{N_k}}.$$

First, the expected trading profit of factor speculators is increasing in β_k. Second, the expected trading profit of factor speculators is decreasing in N_k. These results suggest that high-β assets would have more factor speculators trading them in equilibrium. Absent CS, with N (≥ 2) potential factor speculators and $\frac{\beta_2 \sigma_\gamma \sigma_n}{2} < C < \frac{\beta_1 \sigma_\gamma \sigma_n}{(N+1)\sqrt{N}}$, all factor speculators trade Asset 1 only.

The optimal CS choice. Consider introducing CS with portfolio weight w_k on Asset k, where $k \in \{1, 2\}$ and $w_1 + w_2 = 1$. With more than two speculators trading this CS product, the management fee, $\frac{\hat{C}}{N_{CS}}$, is smaller than C (since $\hat{C} < 2C$). In other words, using CS to access Asset 2 is less costly than trading Asset 2 directly. For illustration, in what follows, we assume that all factor speculators who participate in trading would only trade via CS (proven in Section 3.3 for a perfectly competitive CS sponsoring market). Furthermore, because asset speculators face adverse selection in markets that they are not informed about, in general, they abstain from trading CS involving multiple assets.

We denote the choice of one specific CS product chosen by the jth factor speculator in the CS market as $\{w_{k,j}\}_{k \in \{1,2\}}$, where $\sum_{k=1}^{2} w_{k,j} = 1$. The jth factor speculator then chooses the CS product(s) to trade and the amount to trade. Mathematically, she solves:

$$\max_{y_{CS,j}, \{w_{k,j}\}_{k \in \{1,2\}}} E\left[\sum_{k=1}^{2} y_{CS,j} w_{k,j} \left(\beta_k \gamma - \lambda_k^{CS} \left(\sum_{i \in J \text{ and } i \neq j} \eta_{CS,i} w_{k,i} \gamma + n_k + y_{CS,j} w_{k,j} \right) \right) \middle| \gamma \right]$$

subject to $\sum_{k=1}^{2} w_{k,j} = 1$. Here $\eta_{CS,j} * w_{k,j}$ is the effective trading aggressiveness of CS traders in asset market k. Let $y_{CS,j} w_{k,j} = y_{CS,k}$ and $\eta_{CS,i} w_k = \eta_{CS,i,k}$, the above the optimization problem is equivalent to:

$$\max_{\eta_{CS,j,k}} E\left[\sum_{k=1}^{2} y_{CS,k} \left(\beta_k \gamma - \lambda_k^{CS} \left(\sum_{i \in J \text{ and } i \neq j} \eta_{CS,i,k} \gamma + n_k + y_{CS,k} \right) \right) \middle| \gamma \right],$$

which implies that the optimal trading strategy (with the symmetry among factor speculator that trade CS products) is:

$$\widehat{\eta}_{CS,k} = \frac{\beta_k - \lambda_k^{CS}(N_{CS}-1)\widehat{\eta}_{CS,k}}{2\lambda_k^{CS}} \quad \Leftrightarrow \quad \widehat{\eta}_{CS,k} = \frac{\beta_k}{(N_{CS}+1)\lambda_k^{CS}}.$$

Solving this, we get the choice of asset weights satisfying $w_1^S : w_2^S = \left(\beta_1/\lambda_1^{CS}\right) : \left(\beta_2/\lambda_2^{CS}\right)$, where superscript "S" indicates symmetry and hence identical asset weights choice among factor speculators, and the price impact $\lambda_k^{CS} = \frac{\beta_k \sigma_\gamma}{\sigma_n}\frac{\sqrt{N_{CS}}}{N_{CS}+1}$. The commonly desired CS product is one that weights assets based on their exposure to the factor scaled by the assets' illiquidity in the underlying markets. This is intuitive: After all, CS is a vehicle for factor investing, and the factor exposure should matter when designing its weight; but because the CS coordinates factor speculators to trade in its designed proportion, the price impacts have to be taken into consideration so that the trading costs (which get passed onto the investors) are minimized. We later show this insight to be general and robust.

The expected trading profit for factor speculators that trade CS products is:

$$\Pi_{CS}^F = \sum_{k=1}^{2} \Pi_{CS,k}^F,$$

in which $\Pi_{CS,k}^F = \frac{\beta_k \sigma_\gamma \sigma_{n_k}}{(N_{CS}+1)\sqrt{N_{CS}}}$ is the expected trading profit that a factor speculator earns from asset market k ($k = 1, 2$) through trading CS. With CS, the equilibrium net profit for a factor speculator becomes $\hat{\Pi}^F = \Pi_{CS,1}^F + \Pi_{CS,2}^F - C - F$. When $\frac{\hat{C}}{N_{CS}} < \Pi_{CS,2}^F < C$, all factor speculators trade both underlying assets indirectly via CS products. Relative to the net trading profit without CS, which is $\Pi_1^F - C$, the incremental benefit of trading CS is twofold. First, the"factor access" is profitable because after introducing CS, factor speculators can trade Asset 2 indirectly via CSs and generate additional trading profit, $\Pi_{CS,2}^F$, leveraging her private information regarding the systematic factor γ which is also relevant for Asset 2. Second, as more factor speculators trade CS products, the management fees and trading cost F are lowered via a "duplication reduction."Overall, introducing CS products allows factor speculators to trade assets with lower costs and trade some otherwise unattractive assets (e.g., Asset 2 with low β), which in turn leads to more factor speculators entering the market, i.e., $N_{CS} > N_1$.

Informational efficiency. Comparing informational efficiency in the economies without CS and with CS, it is clear that CS increases the participation of factor speculators in both asset markets, which has important implications on price impacts, information efficiency or return variability, particularly when factor speculators receive perfectly correlated private information about the systematic component γ. Specifically, with more factor speculators participating in trading on their private information ($N_{CS} > N_k$ for $k = 1, 2$), it follows that in this special case, the introduction of CS products always lowers the price impact in both asset markets.

The increase in the number of factor speculators that effectively trade in each asset market k also has an impact on the pricing efficiency in both asset markets. In particular, the factor-specific pricing efficiency, captured by $Var(\gamma|P_k)$, is determined as

$$Var(\gamma|P_k) = Var(\gamma) - \frac{Cov(\gamma, P_k)^2}{Var(P_k)} = Var(\gamma) - \frac{N_k}{N_k + 1}\sigma_\gamma^4$$

in this special case with $\sigma_{\alpha_k}^2 = \sigma_\epsilon^2 = 0$, where N_k is the number of factor speculators effectively trading in asset market k (in the case with CS trading, $N_k = N_{CS}$). The introduction of CS trading, which increases the number of speculators that effectively trade in both asset markets, thus improves the factor-specific pricing efficiency in both markets. The return variability

$$Var(P_k) = \frac{N_k \beta_k^2 \sigma_\gamma^2}{N_k + 1}$$

in each asset market also increases after the introduction of CS, effectively raising the number of factor speculators trading on both assets. Similarly, with perfectly correlated private signals on the systematic factor γ across factor speculators, the price comovement of the two assets is:

$$COV(P_1, P_2) = \frac{N_1 N_2 \beta_1 \beta_2}{(N_1 + 1)(N_2 + 1)}\sigma_\gamma^2$$

before the introduction of CS production, and becomes:

$$C\hat{O}V(P_1, P_2) = \frac{N_{CS}^2 \beta_1 \beta_2}{(N_{CS} + 1)^2}\sigma_\gamma^2$$

after the CS sponsor starts to operate. With the equilibrium number of factor speculators trading CS satisfying $N_{CS} > N_k$ for both $k = 1, 2$, it follows that the price co-movement across asset markets also increases after CS trading is introduced.

3 Equilibrium Characterization

We now characterize the factor investing equilibrium (FIE) under the general setting. To understand the impact of CS trading, we compare an economy without CS sponsors (and thus without CS products) with an economy in the presence of CS trading.

3.1 Equilibrium without CS

The benchmark economy without CS trading could be viewed as a special case in which N_{CS} is set exogenously to zero instead of being determined endogenously. To determine the number of factor speculators endogenously trading in each asset market, we first postulate that in equilibrium, there are N_k factor speculators trading in asset market k ($k = 1, \ldots, K$). We then derive the optimal trading strategies at $t = 1$ for all speculators as well as the pricing rules of market makers, based on which we can determine equilibrium trading profit and N_k.

Specifically, we denote the linear trading strategies by asset speculator and factor speculators in market k by $\kappa_k \alpha_k$ and $\eta_k s_i$, respectively. The total market order of Asset k is:

$$\omega_k = \kappa_k \alpha_k + \sum_{i=1}^{N_k} \eta_k s_i + n_k. \tag{4}$$

The market maker of Asset k sets the price $P_k(\omega_k) = E(v_k|\omega_k) = \lambda_k \omega_k$, where

$$\lambda_k = \frac{\kappa_k \sigma_{\alpha_k}^2 + N_k \beta_k \eta_k \sigma_\gamma^2}{\kappa_k^2 \sigma_{\alpha_k}^2 + N_k^2 \eta_k^2 \sigma_\gamma^2 + N_k \eta_k^2 \sigma_\epsilon^2 + \sigma_{n_k}^2}. \tag{5}$$

Given this pricing rule by market makers, the asset speculator for Asset k then solves:

$$\max_{x_k} E\left[x_k \left(\alpha_k + \beta_k \gamma - \lambda_k \left(x_k + \sum_{i=1}^{N_k} \eta_k s_i + n_k \right) \right) \middle| \alpha_k \right], \tag{6}$$

which implies the optimal order and trading aggressiveness as:

$$x_k = \frac{1}{2\lambda_k}\alpha_k \quad \text{and} \quad \kappa_k = \frac{1}{2\lambda_k}. \tag{7}$$

Similarly, factor speculator i submits orders on Asset k to solve:

$$\max_{y_{k,i}} E\left[y_{k,i}\left(\alpha_k + \beta_k\gamma - \lambda_k\left(\kappa_k\alpha_k + \sum_{i'=1,i'\neq i}^{N_k} \eta_k s_{i'} + n_k + y_{k,i}\right)\right)\middle| s_i\right]. \tag{8}$$

in which she takes as given other factor speculators' equilibrium trading. Exploiting the symmetry in trading aggressiveness η_k of those factor speculators trading, we get:

$$y_{k,i} = \frac{\beta_k - \lambda_k(N_k-1)\eta_k}{2\lambda_k}\frac{\sigma_\gamma^2}{\sigma_\gamma^2+\sigma_\epsilon^2}s_i \quad \text{and} \quad \eta_k = \frac{\beta_k}{\lambda_k\left(N_k+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}\right)}. \tag{9}$$

Solving (7) and (9), we can derive:

Lemma 3.1. *The equilibrium price impact of trading Asset k is* $\lambda_k = \frac{1}{\sigma_{n_k}}\sqrt{\frac{\sigma_{\alpha_k}^2}{4} + \beta_k^2\frac{N_k(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}$ *and the expected profit of a factor speculator from trading Asset k is*

$$\Pi_k^F = \frac{(\sigma_\gamma^2+\sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2\sqrt{\frac{\sigma_{\alpha_k}^2}{4} + \beta_k^2\frac{N_k(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}}. \tag{10}$$

The price impact λ_k is increasing in σ_γ^2 and $\sigma_{\alpha_k}^2$, but is decreasing in $\sigma_{n_k}^2$. Meanwhile, the expected profit from trading Π_k^F is increasing with β_k, $\sigma_{n_k}^2$ and σ_γ^2 but is decreasing with $\sigma_{\alpha_k}^2$ and N_k. The solution is intuitive and reminiscent of those in Subrahmanyam and Titman (1999) and Lee (2013).

With factor speculators' endogenous entry, the equilibrium N_k should exactly make them break even trading Asset k, i.e., $\Pi_k^F = C$.[18] Since Π_k is decreasing with N_k, we have:

Proposition 3.1. *Without CS, a unique equilibrium ensues, in which the equilibrium* N_k *is increasing in* β_k, $\sigma_{n_k}^2$, *and* σ_γ^2, *but is decreasing in* $\sigma_{\alpha_k}^2$, *for each Asset k.*

[18] We follow the standard practice in prior literature (e.g., Subrahmanyam and Titman, 1999; Lee, 2013) to focus on continuous solutions of N_k to simplify the analysis. All results would go through when considering only integer N_k, but the equilibrium N_k in market k is pinned down by the (N_k+1)th factor speculator's entry decision, which adds no further insights despite the complication.

When σ_γ^2 is higher, factor speculators have greater information advantage and thus stronger incentives to enter the financial market. Meanwhile, when $\beta_k(> 0)$ is higher, Asset k's payoff has more systematic exposure, and factor speculators can thus better exploit their private information about γ. Also intuitively, when $\sigma_{n_k}^2$ is lower or $\sigma_{\alpha_k}^2$ is larger, market makers face a higher degree of adverse selection, and thus the price impact is higher, which in turn discourage factor speculators from trading this underlying asset.

3.2 Subgame Equilibrium with CS Trading

In Section 3.1, factor speculators can only trade the underlying securities to exploit their private information. Due to trading costs, they may not trade all underlying assets (i.e., there exists Asset k with $N_k = 0$ in equilibrium), particularly those low-β or high-σ_α assets as suggested by Proposition 3.1. We now introduce CS sponsors, who can provide the service of packaging and bundling underlying securities to help factor speculators better utilize their information advantages regarding the systematic factor γ. Conjecture that only one CS product with weight $\{w_k\}_{k=1}^K$ is traded by factor speculators in equilibrium (which we verify later), and take the number of factors speculators trading the CS, N_{CS}, and the number of factors speculators trading Asset k directly, $\hat{N}_k$, as given, we characterize the equilibrium outcomes at $t = 1$ first.

We denote the asset speculator's trading strategy in Asset k as $\hat{x}_k = \hat{\kappa}_k \alpha_k$, the ith factor speculator's strategy as $\hat{y}_{k,i} = \hat{\eta}_k s_i$ for those directly trade Asset k, and the jth factor speculator's strategy for trading CS products as $y_{CS,j} = \eta_{CS} s_j$. Then the market maker of Asset k receives total order flows:

$$\omega_k = \hat{\kappa}_k \alpha_k + \sum_{i \in I_k} \hat{\eta}_k s_i + \sum_{j \in I_{CS}} \eta_{CS} w_k s_j + n_k, \tag{11}$$

where I_k is the set of factor speculators that submit orders directly in the Asset k and I_{CS} is the set of factor speculators who trade via CS. Note, unlike in the illustrative example characterized in Section 2.2, here we have not assumed any exclusiveness between I_k and I_{CS}—namely, it is possible that $I_k \cap I_{CS} \neq \emptyset$ in equilibrium. Denote the number of factor speculators in I_k by $\hat{N}_k$ and the number of factor speculators that trade via CS by N_{CS}. Since the market makers' pricing rules depend on the information structure in the order flows, it is important to know constitutes of I_k and I_{CS}. Intuitively, for each factor speculator in I_{CS}, she chooses from the full menu of CS products to trade the underlying assets (as a result of the competitive CS sponsoring market) and thus she never has incentives to pay additional costs to trade them directly. Formally, there is thus no overlap between I_k and I_{CS}:

Lemma 3.2. *In equilibrium, $I_k \cap I_{CS} = \emptyset$, for all $k = 1, \dots, K$.*

Given this aforementioned order structure, the market maker of Asset k sets the price $P_k(\omega_k) = E(v_k|\omega_k) = \lambda_k^{CS}\omega_k$, where

$$\lambda_k^{CS} = \frac{\hat{\kappa}_k\sigma_{\alpha_k}^2 + \left(N_{CS}w_k\eta_{CS} + \hat{N}_k\hat{\eta}_k\right)\beta_k\sigma_\gamma^2}{\hat{\kappa}_k^2\sigma_{\alpha_k}^2 + \left(N_{CS}w_k\eta_{CS} + \hat{N}_k\hat{\eta}_k\right)^2\sigma_\gamma^2 + \hat{N}_k\hat{\eta}_k^2\sigma_\epsilon^2 + N_{CS}(\eta_{CS}w_k)^2\sigma_\epsilon^2 + \sigma_{n_k}^2}. \tag{12}$$

Rationally anticipating the above pricing rule adopted by market makers and the equilibrium trading strategies adopted by other traders, the asset speculator in market k solves

$$\max_{\hat{x}_k} E\left[\hat{x}_k\left(\alpha_k + \beta_k\gamma - \lambda_k^{CS}\left(\sum_{i\in I_k}\hat{\eta}_k s_i + \sum_{j\in I_{CS}}\eta_{CS}w_k s_j + n_k + \hat{x}_k\right)\right)\middle|\, \alpha_k\right],$$

which implies that the optimal trading strategy is:

$$\hat{x}_k^* = \frac{1}{2\lambda_k^{CS}}\alpha_k \quad \text{and} \quad \hat{\kappa}_k = \frac{1}{2\lambda_k^{CS}}. \tag{13}$$

For factor speculator i who directly trades in Asset k, she solves

$$\max_{\hat{y}_{k,i}} E\left[\hat{y}_{k,i}\left(\alpha_k + \beta_k\gamma - \lambda_k^{CS}\left(\hat{\kappa}_k\alpha_k + \sum_{i'\in I_k \text{ and } i'\neq i}\hat{\eta}_k s_{i'} + \sum_{j\in I_{CS}}\eta_{CS}w_k s_j + n_k + \hat{y}_{k,i}\right)\right)\middle|\, s_i\right],$$

which gives the optimal trading strategy (with the symmetry among factor speculator trading asset k directly):

$$\hat{\eta}_k = \frac{\beta_k - \lambda_k^{CS}\left((\hat{N}_k - 1)\hat{\eta}_k + N_{CS}\eta_{CS}w_k\right)}{2\lambda_k^{CS}}\frac{\sigma_\gamma^2}{\sigma_\gamma^2 + \sigma_\varepsilon^2}. \tag{14}$$

Finally, we solve the optimal trading strategy of the jth factor speculator that submits orders in the CS market. As we mentioned earlier, due to competition CS sponsors in equilibrium will effectively provide a full list of CS products that factor speculators can choose from. We denote the choice of one specific CS product chosen by the jth factor speculator in the CS market as $\{w_{k,j}\}_{k=1}^{k=K}$, where $\sum_{k=1}^{K} w_{k,j} = 1$. In this sense, the jth factor speculator needs to choose which CS product to trade and how many shares of this CS product to trade. Mathematically, she solves:

$$\max_{y_{CS,j},\{w_{k,j}\}_k} E\left[\sum_{k=1}^{K} y_{CS,j}w_{k,j}\left(\beta_k\gamma - \lambda_k^{CS}\left(\begin{array}{c}\hat{\kappa}_k\alpha_k + \sum_{i\in I_k}\hat{\eta}_k s_i + n_k + \\ \sum_{j'\in I_{CS} \text{ and } j'\neq j}\eta_{CS,j'}w_{k,j'}s_{j'} + y_{CS,j}w_{k,j}\end{array}\right)\right)\middle| s_j\right],$$

subject to $\sum_{k=1}^{K} w_{k,j} = 1$.

It is worth noting that $\eta_{CS,j} * w_{k,j}$ is the effective trading aggressiveness of CS trader j in asset market k. Let $y_{CS,j,k} \equiv y_{CS,j} \cdot w_{k,j}$ and $\eta_{CS,j}w_k \equiv \eta_{CS,k}$, the above the optimization problem is equivalent to:

$$\max_{\{y_{CS,j,k}\}_k} E\left[\sum_{k=1}^{K} y_{CS,j,k}\left(\beta_k\gamma - \lambda_k^{CS}\left(\begin{array}{c}\hat{\kappa}_k\alpha_k + \sum_{i\in I_k}\hat{\eta}_k s_i + n_k + \\ \sum_{j'\in I_{CS} \text{ and } j'\neq j}\eta_{CS,j'}w_{k,j'}s_{j'} + y_{CS,j,k}\end{array}\right)\right)\middle| s_j\right],$$

which implies that the optimal trading strategy (with the symmetry among factor speculator that trade CS products, i.e., $\eta_{CS,j,k} = \eta_{CS,j',k} \equiv \eta_{CS,k}$) is:

$$\widehat{\eta}_{CS,k} = \frac{\beta_k - \lambda_k^{CS}(\hat{N}_k\hat{\eta}_k + (N_{CS}-1)\widehat{\eta}_{CS,k})}{2\lambda_k^{CS}}\frac{\sigma_\gamma^2}{\sigma_\gamma^2 + \sigma_\varepsilon^2}. \tag{15}$$

With each factor speculator that trades via CSs adopting the effective trading aggressiveness in Asset k as above, it is verified that all factor speculators that trade via the CS sponsor will indeed choose CS products of the same weight design. Combining (12), (13), (14) and (15), we derive the equilibrium price impact (λ_k), factor speculators' trading strategies ($\hat{\eta}_{CS,k}$ and $\hat{\eta}_k$) and expected trading profits (Π_k^F and Π_k^{CS}):

Proposition 3.2. *The trading aggressiveness of CS traders in Asset k is the same as that of* *factor speculators who directly trade in Asset k. Specifically, in equilibrium, we have:*

$$\hat{\eta}_{CS,k} = \hat{\eta}_k = \frac{\beta_k}{\lambda_k^{CS}\left(\hat{N}_k + N_{CS} + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}\right)}; \tag{16}$$

The price impact in the asset market k is:

$$\lambda_k^{CS} = \frac{1}{\sigma_{n_k}}\sqrt{\frac{\sigma_{\alpha_k}^2}{4} + \beta_k^2\frac{(\hat{N}_k + N_{CS})(\sigma_\gamma^2 + \sigma_\epsilon^2)\sigma_\gamma^4}{[(\hat{N}_k + N_{CS} + 1)\sigma_\gamma^2 + 2\sigma_\epsilon^2]^2}}; \tag{17}$$

The expected trading profit for factor speculators that trade Asset k directly is:

$$\hat{\Pi}_k^F = \frac{(\sigma_\gamma^2 + \sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{\left[(\hat{N}_k + N_{CS} + 1)\sigma_\gamma^2 + 2\sigma_\epsilon^2\right]^2\sqrt{\frac{\sigma_{\alpha_k}^2}{4} + \beta_k^2\frac{(\hat{N}_k + N_{CS})(\sigma_\gamma^2 + \sigma_\epsilon^2)\sigma_\gamma^4}{[(\hat{N}_k + N_{CS} + 1)\sigma_\gamma^2 + 2\sigma_\epsilon^2]^2}}}; \tag{18}$$

The expected trading profit for factor speculators that trade CS products is $\Pi_{CS}^F = \sum_{k=1}^K \hat{\Pi}_k^F$.

Several implications of Proposition 3.2 are worth highlighting. First, when factor speculators trade underlying assets indirectly via CSs, they can achieve the same effective trading aggressiveness as those factor speculators that trade underlying assets directly. That is, $\hat{\eta}_{CS,k} = \hat{\eta}_k$. Second, the total trading profit for factor speculators trading CS products (Π_{CS}^F) is the sum of the profit from trading each single underlying asset ($\hat{\Pi}_k^F$). Third, the price impact in each Asset k depends on the effective total number of factor speculators trading this asset, $\hat{N}_k + N_{CS}$.

3.3 Factor Investing Equilibrium

Having characterized the subgame equilibrium at $t = 1$ taking as given $\hat{N}_k$ and N_{CS}, as well as the CS sponsors' entry decisions and product offerings at $t = 0$, we now solve for $\hat{N}_k$ and N_{CS} and characterize the equilibrium at $t = 0$. In particular, we endogenize sponsors' CS design and the equilibrium number of factor speculators trading different assets.

In any equilibrium, $\hat{N}_k$ should make the expected profit from directly trading Asset k equal the trading cost. Similarly, in an FIE, N_{CS} makes the expected profit in trading CS products equal the sum of trading costs and management fees. In other words, $\hat{\Pi}_k^F - C = 0$ and $\Pi_{CS}^F - C - F = 0$, where $F = \frac{\hat{C}}{N_{CS}}$ as discussed in Section 2.1. Because in equilibrium

$\Pi^F_{CS} > \hat{\Pi}^F_k$, trading CS allows factor speculators to better utilize their private information than trading a single underlying. So factor speculators need to trade off the additional fee cost F with C for trading each additional underlying asset. The CS sponsor competition and the avoidance of duplication of trading costs can render F sufficiently low, such that factor speculators would prefer CS over underlying assets. We formalize the result next.

Lemma 3.3. *In an FIE, factor speculators all trade CS products and* $\hat{N}_k = 0 \ \forall \ k = 1, \ldots, K$.

In other words, in an economy with a competitive CS sponsoring market, an underlying asset cannot exist such that there is still a positive number of factor speculators trading in it. The intuition for this important result can be understood as follows. Consider an equilibrium in which a CS product with design $\mathbf{w}$ and fee F is offered by the sponsor on $t = 0$ and traded by factor speculators on $t = 1$. Factor speculators' equilibrium trading profit from this CS product is $C + F$ (which is greater than C, since $F > 0$ to allow the sponsor break even). Suppose in equilibrium, a factor speculator is still directly trading certain underlying assets (say, Asset 1) in the subgame equilibrium at $t = 1$, which thus implies the factor's equilibrium trading profit from directly trading Asset 1 is C.

In this equilibrium, it must be the case that CS product constructed as any arbitrary linear combination of only Asset 1 and the CS product $\mathbf{w}$ with zero fee is included in the list of the CS products offered by the operating CS sponsor in equilibrium. Otherwise, a potential CS sponsor could then enter on date $t = 0$ and offer the list of CS products (including $\mathbf{w}$) currently offered by the incumbent CS sponsor plus a full spectrum of CS products constructed as linear combinations of Asset 1 and CS $\mathbf{w}$ with zero fee. Under the dominance assumption in Section 2.1, all speculators will then switch to this newly entrant CS sponsor, who offers a larger feasibility set for security trading choices on date $t = 1$. But these would imply a profitable deviation in the subgame equilibrium at $t = 1$ for the factor speculator who previously trades only Asset 1 in equilibrium.

With this property of the equilibrium outcome, we can further characterize the changes in the participation of factor speculators before and after the introduction of CS trading.

Lemma 3.4. *For any Asset k, we have* $N_{CS} \geq N_k$*, where* N_{CS} *is the equilibrium number of factor speculators on CS and* N_k *is the equilibrium number of factor speculators on Asset k in the economy without CS.*

The above lemmas allow us to characterize the generalized equilibrium with CS sponsoring and trading (formally defined as the FIE in Section 2.1), as summarized next.

Proposition 3.3. *There is a unique FIE in which* $\hat{N}_k = 0$ *and* $N_{CS} \geq N_k$ *for* $k = 1, ..., K$*. In this equilibrium, one CS sponsor enters the market and offers a CS product with fee* $F = \frac{\hat{C}}{N_{CS}}$ *and the weights of the underlying securities satisfying:*

$$w_1 : w_k = \frac{\beta_1}{\lambda_1^{CS}} : \frac{\beta_k}{\lambda_k^{CS}}, \tag{19}$$

where $k = 2, ..., K$*.*

In an economy with perfectly competitive CS sponsoring, all factor speculators are attracted to trading via CS to exploit their private information on γ. In addition, the introduction of CS sponsoring market can weakly increase the number of factor speculators exploiting their private information in each asset market.

In order to attract the maximum number of factor speculators, a (potential) CS sponsor would need to choose the weights of each constituent asset optimally in her CS design. Specifically, the optimal weight of each asset in the CS product helps factor speculators achieve their desired effective trading aggressiveness in each underlying asset, as characterized in Eq. (15). The optimal weight assigned to each underlying asset is positively related to its factor loading β while negatively related to the price impact λ in the asset market.

4 Information and Asset Pricing Implications

We now discuss how introducing CS affects asset prices and the informational efficiency of financial markets. As shown in Section 3.3, introducing CS sponsors can weakly increase the number of factor speculators exploiting private information on the systematic component of asset value in each underlying asset market. When more factor speculators participate in asset markets, one would expect factor-specific information to be more impounded into asset prices, which affects asset-specific information in prices and other market outcomes such as liquidity, volatility, and correlations in asset valuations.

4.1 Informational Efficiency

Over the past decades, almost all publicly traded companies saw a sharp rise in passive ownership, which includes ETFs and index funds. The resulting changes in pricing efficiency affect resource allocation and investors' wealth dynamics. Managerial compensation also relies heavily on stock prices and may not be effective if firm prices are not informative. More generally, whether and how financial markets incorporate relevant information in the economy is a central question in economics and finance. To this end, our model helps us understand how the rise of CS affects the pricing efficiency of the underlying assets.

Specifically, since the asset payoffs have both systematic and asset-specific components, we are interested in three types of efficiencies: asset-specific efficiency, measured by $1/Var(\alpha_k|P_k)$, factor-specific efficiency, measured by $1/Var(\gamma|P_k)$, and total efficiency, measured by $1/Var(v_k|P_k)$. The naive view that passive ownership reduces the informational content of asset prices because passive investors lack the incentive to acquire asset-specific information neglects (2) and (3). We demonstrate that considering the nuances in the type of efficiency can help rationalize puzzling empirical patterns observed in the data.

In our setting, both market makers' pricing rules and factor speculators' trading strategies affect informational efficiency.[19] Proposition 3.2 reveals that when no factor speculators directly trade in underlying asset markets, market makers set $P_k = \lambda_k \omega_k$, where $\omega_k = \kappa_k \alpha_k + \sum_{i=1}^{N_{CS}} \eta_{CS,k} s_i + n_k$. Meanwhile, the trading aggressiveness of asset speculators and factor speculators are $\kappa_k = \frac{1}{2\lambda_k}$ and $\eta_{CS,k} = \frac{\beta_k}{\lambda_k \left(N_{CS}+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}\right)}$, respectively. Asset k then has price:

$$P_k = \frac{\alpha_k}{2} + \frac{N_{CS}\beta_k}{N_{CS}+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}}\gamma + \sum_{i=1}^{N_{CS}} \frac{\beta_k}{N_{CS}+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}}\epsilon_i + \lambda_k^{CS} n_k. \tag{20}$$

In this pricing function, the second term represents the factor-specific component that becomes more dominant as N_{CS} increases. We have the following proposition regarding the impact of introducing CS on pricing efficiency in the underlying asset markets:

Proposition 4.1. *Introducing CS increases factor-specific efficiency and total efficiency but decreases asset-specific efficiency in asset prices.*

[19] In Section 6.1, we further endogenize information acquisition and asset speculators' participation.

This result is intuitive because CS trading allows factor speculators to better exploit their private information regarding the systematic factor and thus encourages the participation of factor speculators, which naturally brings in more factor-specific information. Meanwhile, when the asset prices contain more factor-specific information, they are less sensitive to asset-specific information—which decreases the pricing efficiency of asset-specific components. Overall, the total price efficiency in each underlying asset is also improved when more factor speculators enter and trade using their factor information via CS, which more than offsets the reduced asset-specific information.

These predictions are consistent with a large literature of empirical studies on ETFs. For example, Glosten, Nallareddy, and Zou (2021) find that ETF trading increases information efficiency for small firms and firms with imperfect access to capital markets by incorporating aggregate information into stock prices, but find no such effect for big stocks. Consistent with their findings, our model reveals that the relatively illiquid asset (e.g., due to high variance of asset-specific components, or low β, or low variance of noise trading) experiences a larger increase in systematic informational efficiency. The decreased asset-specific information efficiency associated with CS introduction is also consistent with Israeli, Lee, and Sridharan (2017), which documents that firms experiencing a 1% increase in ETF ownership experience a 21% reduction in the magnitude of their future earnings response coefficients, a measure of the association between current firm-specific returns and future firm-specific earnings.[20] Bhojraj, Mohanram, and Zhang (2020a) show that sector ETFs have improved informational efficiency by facilitating the transmission of information. Studies such as Bai, Philippon, and Savov (2016) and Farboodi, Matray, Veldkamp, and Venkateswaran (2022) also document that firms with high shares of institutional investors, who tend to be passive, have more informative prices.

Paul Samuelson has long made the interesting conjecture that there is more informational inefficiency for macro information than for micro information, which has come to be known as the Samuelson's Dictum.[21] A number of studies demonstrate that the hypothesized macro

[20]Note that Subrahmanyam (1991) predicts that the introduction of a basket tends to increase the number of security analysts for the most heavily weighted securities in the basket, and prices of such securities will become more informative in the security-specific component, which contradicts the empirical findings. In Section 6.1, we endogenize the participation decision of asset speculators in each underlying asset market and find that the introduction of CS could have a surprising mixed impact on asset-specific efficiency.

[21]See Samuelson et al. (1998) and his letter to John Campbell and Robert Shiller (Shiller, 2000, p. 243).

inefficiency and micro efficiency in Samuelson's dictum can and do arise naturally in equilibrium (e.g. Jung and Shiller, 2005; Glasserman and Mamaysky, 2023). Interpreting how effectively factor information is reflected in prices as macro efficiency, our findings demonstrate that the rise of factor investing and passive investing moderate such a situation, in contrast to the findings in Gârleanu and Pedersen (2022). That said, the general welfare implications of pricing efficiency are complicated and are beyond our paper. For example, if the capital providers are using the market information to decide how much capital they provide to the firm for real production, then the fact prices reflect more systematic information potentially lowers the allocative efficiency to specific firms (despite improving allocative efficiency to certain sectors or styles), as discussed in (Veldkamp and Wolfers, 2007; Goldstein and Yang, 2014).

4.2 Price Impact and Liquidity

Equally important is the effect of introducing CS trading on the price impact (a measure of market liquidity) of trading each specific underlying asset, which turns out to depend on the relationship between the number of CS-trading factor speculators and the parameters governing the uncertainties concerning asset payoffs.

Proposition 4.2. *(i) If $N_{CS} \leq \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$, introducing CS increases the price impacts for trading each underlying asset. That is, $\lambda_k^{CS} > \lambda_k$ for all k, where λ_k^{CS} and λ_k are the price impacts in the market for trading underlying Asset k after and before introducing CS, respectively.*
(ii) If $N_{CS} > \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$, introducing CS increases the price impact (i.e., $\lambda_k^{CS} > \lambda_k$) of trading assets with $N_k < \frac{\left(\frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}\right)^2}{N_{CS}}$, but decreases the price impact (i.e., $\lambda_k^{CS} < \lambda_k$) of trading assets with $N_k > \frac{\left(\frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}\right)^2}{N_{CS}}$, where N_k is the number of factor speculators trading on asset k before introducing CS.

In the first case, where the number of factor speculators trading CS in equilibrium is relatively small, introducing CS trading unambiguously increases the price impact in all underlying asset markets. In the second case, where the equilibrium number of factor speculators trading CS is sufficiently high, introducing CS trading could instead have a mixed effect on the price impact in the underlying asset markets. The effect in the first case still holds for asset markets that have relatively small N_k, which is the number of factor speculators trading in it before the introduction of CS. In contrast, for asset markets with relatively high N_k, introducing CS actually reduces the price impact and improves liquidity.

The economic mechanisms for the above effects are similar to that in Subrahmanyam and Titman (1999). When more factor speculators with dispersed information participate in asset markets, there are two opposite effects on market liquidity (price impact). First, when the information is diverse (e.g., σ_ϵ relatively large compared to σ_γ), increasing the number of informed factor speculators increases adverse selection faced by market makers. Consequently, price impact increases through an *information inclusion* effect. On the other hand, an increased number of factor speculators trading in equilibrium also encompasses a *competition* effect, which reduces the trading aggressiveness of each factor speculator and hence leads to a lowered price impact. Intuitively, when the number of factor speculators is sufficiently high, the *competition* effect dominates the information effect, leading to decreased price impact and improved liquidity. In contrast, when the number of factor speculators is low, the *information inclusion* effect dominates the information effect, leading to increased price impact and deteriorated liquidity.

As discussed in Section 3.1, the number of factor speculators N_k trading in market k depends on the factor loading β_k and α_k. Two cross-sectional predictions directly follow:

Corollary 4.1. *When all assets have the same σ_α^2 and σ_n^2, and $N_{CS} > \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$, introducing CS increases the price impacts of trading low-beta assets but decreases the price impacts of trading high-beta assets.*

Corollary 4.2. *When all assets have the same β and σ_n^2, and $N_{CS} > \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$, introducing CS increases the price impacts of trading high σ_α^2 assets but decreases them for low σ_α^2 assets.*

Intuitively, in asset markets associated with relatively low factor loading β or relatively high asset-specific component volatility σ_α^2, the number of factor speculators before CS introduction is likely to be small. As such, based on Proposition 4.2, with an increased number of factor speculators exploiting their private information after CS trading becomes available, the information effect dominates, and price impact in these markets increases. In contrast, for assets associated with high β or low σ_α^2, the number of factor speculators trading in these markets (and the informativeness of market orders) are likely to be already high even before the introduction of CS trading, such that the information effect of further increasing the participation by factor speculators is likely to be limited. In such markets, the competition effect dominates, and introducing CS trading would decrease price impact.

Our results highlight that the impact can be heterogeneous and depends on characteristics of the underlying assets such as the factor exposure and idiosyncratic noise level.[22]

4.3 Return Variability and Co-Movements

Our model has clear implications for the impact of CS trading on the return variability and return co-movements in the underlying asset markets. We define the asset return variability of Asset k as $Var(P_k)$ and define the return co-movement between Assets i and j as $corr(P_i, P_j)$. When factor speculators are trading, the asset price incorporates more systematic information and becomes more sensitive to fundamental innovations in γ, increasing return variability. Moreover, the increased number of factor speculators trading CS increases common-factor-related information in all underlying assets, which should increase return co-movement.

Proposition 4.3. *Introducing CS increases the return variability and co-movement in the underlying asset markets.*

While earlier studies in the context of index products arrive at similar conclusions, we emphasize the role of systematic information. Subrahmanyam (1991) predicts that the introduction of a basket will have no effect on the variability of price changes of individual securities, but our model predicts that CSs increase the volatility of the underlying securities. This view is consistent with Ben-David, Franzoni, and Moussawi (2018) who find that stocks included in ETSs (CSs) exhibit significantly higher intraday and daily volatility.[23] The authors argue that ETFs attract a new layer of demand shocks to the stock market due to their high liquidity. Our model demonstrates an alternative channel: it is possible that stock price variance increases

[22] These results have implications on welfare on liquidity traders. In our model, factor speculators earn zero trading profits in the equilibrium as a result of free entry. But this is not the case for liquidity traders. As liquidity traders' welfare depends on price impact, introducing CS can either increase or decrease liquidity traders' welfare, depending on assets' factor exposure and asset-specific components.

[23] In a more recent study, Jiang, Vayanos, and Zheng (2022) find empirically that flows into passive funds raise the largest firms' return volatility the most, which our model predicts given that larger firms tend to have larger β loadings on systematic factors (Chan and Chen, 1991).

because innovations in the systematic component gets impounded into asset prices more when investors with more accurate signals on γ migrate to ETF trading.

Moreover, our model predictions are corroborated by empirical findings in Crawford, Roulstone, and So (2012); Da and Shive (2018) and Glosten, Nallareddy, and Zou (2021), which document that ETF trading increases return co-movement among underlying stocks. Consistent with our model, Glosten, Nallareddy, and Zou (2021) find that ETF trading increases co-movement and synchronicity which is partly attributable to the timely incorporation of information about systemic components in earnings.

5 Composite Security Design and Emperical Evidence

5.1 Implications for CS Design

There has been little discussion on the optimal CS design, which Proposition 3.3 directly speaks to. Our theory helps reconcile ETFs' passive indexing with their active role that the median turnover rate of U.S. index ETFs is as big as 16% per year, and 37% of ETFs use self-designed indices to save execution costs (e.g., Li, 2021).[24] Even index-based CSs do not use market weights and incur high turnovers, as the sponsors' optimal security design implies.

Instead, as predicted by Proposition 3.3, rational CS sponsors should offer products with weights of the underlying assets proportional to the asset exposure to a factor (β) and inversely proportional to some illiquidity measure (λ). Although a large amount of empirical literature provides support for our model's asset pricing implications, we have to investigate whether our model predictions regarding the equilibrium CS design bear out in the data.[25] When we specialize in U.S. equity ETFs, we formally hypothesize that there is a positive association between one particular stock's exposure to the ETF index (or the factor it represents) and its portfolio weight and a negative association between its market illiquidity and its weight.

[24]For example, the Vanguard S&P Small-Cap 600 Growth ETF (VIOG) has a 35% turnover rate.

[25]More recently, Brogaard, Heath, and Huang (2023) document that exchange-traded funds (ETFs) "sample" their indexes, systematically underweighting or omitting illiquid index stocks, as our model predicts.

5.2 Data on Equity ETFs in the United states

Our empirical exercise mainly uses two data sets. The first data set contains information on U.S. equity ETFs, including their categories (industry ETFs or smart beta ETFs). The second data set contains the holding information of underlying stocks for each ETF. Our sample period is from January 2000 to December 2018. We construct our data in the following steps.

We first obtain a list of U.S. equity ETFs from the CRSP Survivor-Biased-Free Mutual Fund database. We identify a fund as an ETF if the "et_flag" of the fund is "F." Additionally, we require these funds to have a CRSP share-code of either "44" or "73." To focus on non-synthetic U.S. equity ETFs, we drop ETFs whose names contain "bond," "bear," or "hedged."

Next, we obtain ETF holding data from Thomson Thomson-Reuters Mutual Fund Holding database (S12) and CRSP Mutual Fund database. For each ETF, we first merge with the holding data from Thomson-Reuters Mutual Fund Holding database using the MFLINKS tables. Our final sample consists of 361 ETFs with valid holding data from 2000 to 2018.

We construct the key variables as follows. The main variable of interest is the excess portfolio weight of a given holding stock in its parent ETF on a holding reporting date. To calculate the excess portfolio weight, we first define a benchmark portfolio weight as the "value-weighted portfolio weight" calculated based on the holding stock's market capitalization on the reporting date. Then, we define excess portfolio weight as the actual portfolio weight minus the benchmark portfolio weight. We multiply the excess portfolio weight by 100 throughout our empirical analysis. The motivation for such an excess portfolio weight is that if ETF sponsors do not actively adjust portfolio weights of stocks with ETFs, they only passively follow the simple rule as many passive indices (e.g., S&P500, Russell 2000) and should construct the portfolios based on stocks' market capitalization. The excess portfolio weight can measure how ETF sponsors deviate from their "passive" choices.

To test the prediction on the optimal composite security design in Proposition 3.3, we focus on two important determinants of ETF portfolio weights: the holding stock's price impact and loading on the parent ETF. Here, we define price impact as the average daily Amihud (2002) illiquidity measure within the three-month window ending on the reporting date. We multiply illiquidity by 10^8 throughout our empirical analysis. The ETF loading is estimated in a rolling window. Specifically, for one specific stock's loading on its parent ETF on a holding reporting date, we regress daily stock returns on daily returns of its parent ETF (excluding this stock) in the twelve-month window ending at the reporting date.

We also consider a set of control variables. Firm size (Ln(Mktcap)) is measured by the natural logarithm of market capitalization on the reporting date. Book-to-market ratio (BM) is the ratio of book equity to market value of equity, measured at the latest fiscal year end prior to the reporting date. Institutional ownership (IO) is the ratio of shares held by 13-F institutions to the total shares outstanding in CRSP, measured at the latest quarter-end as of the reporting date. Missing IO is replaced by a value of zero. Past twelve-month return (MOM) is the cumulative returns in the twelve-month window ending on the reporting date. Analyst coverage (#Analyst) is defined as the number of distinct analysts who make fiscal year one earnings forecast for the stock in the calendar year prior to the reporting date. The analyst earnings forecast data are from I/B/E/S Unadjusted Detail file. Idiosyncratic volatility (IVOL) is defined as the standard deviation of daily return residuals relative to Fama-French three factors in the month of the reporting date.

Table 1 reports the basic statistics. It is clear that the average excess portfolio weight is zero. Interestingly, there are large variations in the excess portfolio with a standard deviation of 0.9905, suggesting that some economic forces are at work.

5.3 Empirical Findings

We now empirically test how stock characteristics affect ETF sponsors' choices of portfolio weights within ETFs. To test Proposition 3.3, we run the following panel regression:

$$w_{ijt} = \alpha_0 + \alpha_1 \cdot \beta_{ijt-1} + \alpha_2 \cdot \lambda_{it-1} + \alpha_3 \cdot X + \epsilon_{ijt}, \tag{21}$$

where w_{ijt} is the excess portfolio weight on stock i in ETF j at quarter t, β_{ijt-1} is the stock i's loading on factor j prior to quarter t, λ_{it-1} is Amihud (2002)'s illiquidity measure of stock i prior to quarter t. X represents the set of control variables: Firm size (Ln(Mktcap)), Book-to-market ratio (BM), Institutional ownership (IO), Past twelve-month return (MOM), Analyst coverage (#Analyst), Idiosyncratic volatility (IVOL). Across all specifications, we include ETF and time fixed effects and calculate standard errors double clustered by ETF and time.

Table 1: **Summary Statistics.** This table reports summary statistics on the excess portfolio weight and other characteristics of ETF holding stocks. The unit of observation is ETF-stock pair on each ETF holding reporting date during 2010-2018. Panel A reports statistics on excess portfolio weight of holding stocks for all ETFs. The excess portfolio weight is defined as the actual portfolio weight minus the benchmark portfolio weight under value-weighting scheme. Panel B reports characteristics of ETF holding stocks. Illiquidity is the average daily Amihud (2002) illiquidity measure of the holding stock in the three-month window ending at the reporting date (multiplied by 10^8). Beta is the stock's loading on its parent ETF estimated in the twelve-month window ending on the holding reporting date. Other stock characteristics reported in Panel B are market capitalization (Ln(Mktcap)), book-to-market ratio (BM), past twelve-month return (MOM), institutional ownership (IO), number of analyst coverage (#Analyst), and idiosyncratic volatility (IVOL) of the stock as of the reporting date.

Panel A: Excess Portfolio Weight (%)					
Sample	Mean	SD	25th	50th	75th
All ETFs	0.0000	0.9905	−0.0046	0.0013	0.0165
Panel B: Holding Stock Characteristics					
Variable	Mean	SD	25th	50th	75th
Illiquidity	1.1080	3.4909	0.0177	0.0831	0.4184
Beta	1.0248	0.3967	0.7610	0.9907	1.2555
Ln(Mktcap)	21.5565	1.7511	20.3276	21.4557	22.7071
BM	0.5862	0.6515	0.2733	0.4707	0.7537
MOM	0.1535	0.5173	−0.0804	0.1093	0.3113
IO	0.7416	0.2424	0.6287	0.7926	0.9062
#Analyst	13.7339	10.5068	6.0000	11.0000	20.0000
IVOL	0.0165	0.0132	0.0086	0.0129	0.0200

Table 2 reports the results. We find evidence consistent with Proposition 3.3. Specifically, as shown in Table 2, within one ETF, an underlying stock's excess portfolio weight is significantly and negatively associated with Amihud (2002) illiquidity measure but is significantly and positively associated with the stock's loading on ETF returns. It terms of economic magnitudes, an increase in Amihud (2002) illiquidity/stock's loading from their 25th to 75th percentile value is associated with -0.01%/0.02% increase in excess portfolio weight. In comparison, the difference in 25th and 75th percentile value of Excess portfolio weight is 0.02%. Overall, the results in Table 2 provide strong supporting evidence for Proposition 3.3.

6 Discussion and Extensions

6.1 Endogenous Information Acquisition and Asset Speculation

As shown in Section 4.1, introducing CS always harms asset-specific informational efficiency. However, some recent empirical studies (e.g., Huang, O'Hara, and Zhong, 2020; Bhojraj, Mohanram, and Zhang, 2020b) document that introducing ETF may also help the market incorporate firm-specific information. To reconcile these empirical facts and generalize our framework, we extend our analysis to endogenously study the information acquisition and trading participation decisions by asset speculators.

We assume that the potential asset speculator in each underlying Asset k faces a discrete choice of whether to incur a fixed cost C_A (e.g., attention cost or information acquisition cost) to become informed about Asset k and thus trade in the asset market k. Paying for this information acquisition cost gives her a perfect signal on the asset-specific component α_k in the liquidation value. Otherwise, she remains uninformed about α_k and chooses not to trade.

Based on our analysis in Section 3.2, she participates in the market if and only if :

$$\frac{\sigma^2_{\alpha_k}\sigma_{n_k}}{4\sqrt{\frac{\sigma^2_{\alpha_k}}{4}+\beta_k^2\frac{N_k(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} > C_A, \tag{22}$$

where N_k is the number of factor speculators effectively trading in underlying asset k (with or without CS trading). Absent factor speculators, it is easy to see Eq. (22) reduces to

Table 2: **Panel regressions of excess portfolio weight.** This table reports the results on the panel regression of excess portfolio weight on stock characteristics. The dependent variable is excess portfolio weight of a given ETF holding stock on a reporting date. The key independent variables are Amihud (2002) illiquidity measure of the stock (Illiquidity) and the stock's loading on its parent ETF (Beta). The control variables include market capitalization (Ln(Mktcap)), book-to-market ratio (BM), past twelve-month return (MOM), institutional ownership (IO), number of analyst coverage (#Analyst), and idiosyncratic volatility (IVOL) of the stock as of the reporting date. The variable definitions are in Table 1. ETF and time (year-quarter of the reporting date) fixed effects are included. t-statistics are computed based on standard errors clustered by ETF and time. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
Illiquidity	-0.0127***	-0.0190***	-0.0125***	-0.0193***
	(-4.54)	(-4.62)	(-4.50)	(-4.59)
Beta	0.0436***	0.0296***	0.0459***	0.0283***
	(2.99)	(2.86)	(3.17)	(2.77)
Ln(Mktcap)	-0.0895***	-0.1627***	-0.0913***	-0.1664***
	(-5.49)	(-4.88)	(-5.48)	(-4.82)
BM	0.0018	-0.0080*	0.0026	-0.0063
	(0.53)	(-1.69)	(0.75)	(-1.41)
Mom	0.0093*	0.0290***	0.0093*	0.0296***
	(1.95)	(3.11)	(1.75)	(3.15)
IO	0.2682***	0.2898***	0.2711***	0.2864***
	(5.93)	(5.80)	(5.95)	(5.86)
#Analysts	-0.0012	-0.0029**	-0.0011	-0.0026**
	(-1.13)	(-2.24)	(-0.95)	(-2.06)
IVOL	-1.2642***	-1.3320***	-1.2860***	-1.2336***
	(-4.43)	(-4.47)	(-4.34)	(-4.36)
Controls	Yes	Yes	Yes	Yes
ETF FE	No	Yes	No	Yes
Time FE	No	No	Yes	Yes
No. Obs.	3,547,316	3,547,316	3,547,316	3,547,316
Adj. R^2	0.02	0.04	0.02	0.04

$\sigma_{\alpha_k}\sigma_{n_k}/2 > C_A$. To avoid the trivial case of asset speculators not participating, which renders them irrelevant, we maintain $\sigma_{\alpha_k}\sigma_{n_k}/2 > C_A$ throughout the remainder of the section.

Endogenous participation benchmark absent CS trading. We introduce Asset Speculators' endogenous participation and speculation in an economy without CS trading. The characterization for this equilibrium is more complicated than that studied in Section 3, where all asset speculators are assumed to be endowed with perfect information about α_k and thus will always trade in the asset market. In Appendix A1, we provide a detailed analysis in which Proposition A1.1 fully characterizes all possible equilibria under this setting of endogenous asset speculation. From Proposition A1.1, we derive how asset speculators' participation depends on asset characteristics (β or σ_α^2):

Proposition 6.1. *Assume all assets have the same $\sigma_{n_k}^2$, in an economy without CS sponsoring, asset speculators tend to participate in asset markets with low β and high $\sigma_{\alpha_k}^2$. Specifically,*

1. *When all assets have the same σ_α^2, there exist $\beta_H > \beta_L > 0$ such that asset speculators always trade assets with $\beta_k < \beta_L$ and asset speculators never trade assets with $\beta_k > \beta_H$, where β_H and β_L are given in the Appendix.*

2. *When all assets have the same β_k, there exist $\sigma_H^2 > \sigma_L^2 > 0$ such that asset speculators always trade assets with $\sigma_{\alpha_k}^2 > \sigma_H^2$ and asset speculators never trade assets with $\sigma_{\alpha_k}^2 < \sigma_L^2$, where σ_H^2 and σ_L^2 are given in the Appendix.*

Intuitively, when all assets have the same σ_α^2 and $\sigma_{n_k}^2$, for assets with high factor loading β, factor speculators would adopt a relatively more aggressive trading strategy to exploit their private information of factor γ. This results in a larger price impact for asset speculators to trade these assets, which in turn deters their participation.

Similarly, when all assets have the same β_k and $\sigma_{n_k}^2$, for assets with high σ_α^2, asset speculators' private information is more valuable. Therefore, they are more likely to participate in trading in these asset markets to exploit their private information. Mathematically, a higher σ_α^2 means that the price impact would be lower for asset speculators to submit orders, holding other variables fixed.

Impact of CS trading. We next examine how introducing CS trading affects the participation decisions of asset speculators, which naturally hinges on how liquidity changes in the underlying asset markets. For example, if price impact decreases after CS introduction, asset speculators can better exploit their private information with lower price impact and thus are more willing to participate in asset markets after the introduction of a CS sponsoring market.

As shown in the baseline setting (Section 4.2) with exogenous asset speculator participation, introducing CS has mixed effects on price impact in underlying asset markets, which depends on β and σ_α^2. In what follows, we continue focusing on these asset characteristics, β and σ_α^2.

To understand the role of β, we assume for simplicity that all underlying assets have same σ_α^2 and $\sigma_{n_k}^2$. We first show that the heterogeneous effect of CS introduction on price impact in the economy with endogenous participation of asset speculators is similar to that in Section 4.2, and then we characterize how introducing CS affects the participation of asset speculators. Specifically, in Appendix A2 (Lemma A1.1), we show that in this equilibrium with endogenous participation of asset speculators, the effect of CS trading on liquidity (price impact) in the underlying assets market is similar to that characterized in Corollary 4.1. Lemma A1.1 leads to the following proposition:

Proposition 6.2. *When $N_{CS} > \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$, there exists a cut-off value β^* such that introducing CS weakly decreases (increases) the participation of asset speculators and asset-specific information efficiency for assets with $\beta_k < \beta^*$ ($\beta_k > \beta^*$).*

For assets with relatively low factor loading, the introduction of CS trading may force asset speculators to quit the market in which they used to be active and acquire private information. Asset-specific information efficiency likely decreases in such markets then because the order received by market makers in these markets contains weakly less asset-specific information. Conversely, introduction of CS sponsoring may encourage asset speculators to enter those asset markets with high factor loading, where they used to refrain from trading.

The intuitions for Proposition 6.2 are reminiscent of that for Proposition 4.2. As shown in Proposition 4.2, introducing CS will effectively increase the number of factor speculators in all underlying assets, which has two effects (information effect vs. competition effect) on price impacts. The net effect of CS trading on price impacts depends on the number of factor speculators trading on assets before introducing CS. When the number of factor speculators is already high before CS introduction, the competition effect dominates, and thus, introducing CS decreases price impact. When the number of factor speculators is low before CS introduction, the information effect dominates, and thus, introducing CS increases price impact.

For assets with high factor loading β, the number of factor speculators trading on them is likely to be already high before the CS introduction. Thus, the increased number of factor speculators due to CS introduction turns out to decrease the price impact of submitting orders in these asset markets due to the competition effect. For assets with low β, the number of factor speculators trading on them is likely to be low before CS introduction. Thus for these assets, the increased number of factor speculators due to CS introduction increases price impact, as more information is brought into the orders received by market makers.

These effects of CS trading on price impacts naturally determine the participation decisions of asset speculators. In the scenario where CS sponsoring cost is sufficiently low such that $N_{CS} > \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$, as suggested by item (ii) of Lemma A1.1, introducing CS sponsoring increases the price impact in underlying asset markets for those with β below the threshold $\beta^\star$ while decreases it for those below the threshold.[26] As a result of this mixed effect on price impact, we thus obtain this two-sided effect of CS sponsoring on asset speculators' participation as summarized in Proposition 6.2.

Regarding how the effect of CS varies with σ_α^2, we similarly assume that all underlying assets have the same β and $\sigma_{n_k}^2$. In Appendix A2 (Lemma A1.2), we show that in this equilibrium with endogenous participation of asset speculators, the effect of CS trading on price impact in the underlying assets market is similar to that characterized in Corollary 4.2. Based on Lemma A1.2, we have the following proposition.

Proposition 6.3. *When $N_{CS} > \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$, there exists a cut-off value σ^* such that introducing CS weakly decreases (increases) the participation of asset speculators and asset-specific information efficiency for assets with $\sigma_k^2 > \sigma^{*2}$ ($\sigma_k^2 < \sigma^{*2}$).*

The economic intuitions for Lemma A1.2 are also from Proposition 4.2. For assets with high $\sigma_{\alpha_k}^2$, the number of factor speculators trading on them is low before CS introduction. Thus, the increased number of factor speculators due to CS introduction increases price impact thanks

[26]In the case of $N_{CS} < \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$, introducing CS trading always increase price impact, which makes the endogenous participate of asset speculator less interesting.

to the dominating information effect. For assets with low $\sigma^2_{\alpha_k}$, the number of factor speculators trading on them is already high before CS introduction. Thus, the increased number of factor speculators due to CS introduction decreases price impact as the competition effect dominates.

Overall, we find that when asset speculators' participation is endogenous, introducing CS sponsoring can potentially encourage asset speculators to acquire asset-specific information about and trade underlying assets with high β or low σ^2_α. These assets are a priori not attractive to asset speculators, and thus, their prices should reflect low asset-specific information before CS becomes available. However, with CS, asset speculators start to trade the underlying assets, and consequently, asset-specific information efficiency improves.[27] These predictions are corroborated by recent empirical findings by Huang, O'Hara, and Zhong (2020), who use a difference-in-difference analysis around the inception events of industry ETFs and find that high-industry-risk-exposure stocks (defined as a high ratio of $\beta^2\sigma^2_\gamma$ to σ^2_α) experience more improvement in information efficiency after the inception of industry ETFs.

Lee (2021) similarly shows that decreased passive fees make it cheaper to participate in the market, improve liquidity (and thus the value of active investing), and encourage active investing. Our asset speculators exactly correspond to active investors, and we show that the prediction is more nuanced—it only holds for assets with high exposure to systematic information and with low asset-specific innovations. In our setting, these assets mostly attract factor speculators instead of liquidity traders or asset speculators. Introducing CSs improves liquidity as we show in Section 4.2, and allows more asset-speculator to join.

6.2 Factor Hedgers, Liquidity Trading, and Duality

To demonstrate the robustness of our results about CS security design, we extend the model by introducing a separate group of noise traders who participate in the financial markets for liquidity motives—the factor liquidity trader. Specifically, similar to asset liquidity traders in each underlying asset market, we model this group of liquidity traders as one representative who has an exogenous need τ for exposure to factor γ, where τ follows a normal distribution $\tau \sim \mathcal{N}(0, \sigma^2_\tau)$. The interpretation is that hedgers might be endowed with assets with certain

[27] Goldstein and Yang (2015) show that due to the complementarities in different pieces of information about asset value, greater information diversity in the economy improves price informativeness. We do not rely on risk aversion, and we clarify when factor information and asset-specific information acquisition are complementary.

risk exposure, and their goal is to unload that exposure.[28]

Like asset liquidity traders, the factor liquidity trader is also uninformed of the realization of factor γ and trade to minimize expected cost in achieving the required factor exposure. With the presence of this factor liquidity trader, the order follow received by the market maker in Asset k thus becomes:

$$\omega_k = \hat{\kappa}_k \alpha_k + \sum_{i \in I_k} \hat{\eta}_k s_i + \sum_{j \in I_{CS}} \eta_{CS} w_k s_j + \frac{\tau_k}{\beta_k} + n_k, \tag{23}$$

where τ_k is the factor exposure gained from trading via Asset k such that $\sum_{k=1}^{K} \tau_k = \tau$.

Similar to our analysis of the equilibrium with CS sponsor in Section 3, it can be shown that the factor liquidity trader must also be trading CS product rather than underlying assets in equilibrium. In the subgame equilibrium at $t = 1$, the factor liquidity trader thus minimizes her trading cost by optimally choosing the CS product $\{w_k^H\}_{k=1}^{K}$ and her trading volume y_H:

$$\min_{y_H, \left\{w_k^H\right\}_k} E\left[\sum_{k=1}^{K} y_H w_k^H \lambda_k^{CS} \left(\hat{\kappa}_k \alpha_k + \textstyle\sum_{i \in I_k} \hat{\eta}_k s_i + \sum_{j \in I_{CS}} \eta_{CS,j} w_{k,j}^S s_j + n_k + y_H w_k^H \right)\right] \tag{24}$$

$$\textit{subject to:} \qquad y_H \sum_{k=1}^{K} w_k^H \beta_k = \tau,$$

where $\left\{w_{k,j}^S\right\}_{k=1}^{K}$ denotes the CS product chosen by the j'th factor speculator and the factor liquidity trader takes as given the equilibrium trading strategies of speculators. Denote by $y_H w_k^H \equiv y_k^H$ the factor hedger's effective trading volume in Asset k, then from Eq. (24) it can be shown that the optimal choice of $\left\{y_k^H\right\}_{k=1}^{K}$ satisfies

$$y_1^H : y_k^H = \frac{\beta_1}{\lambda_1^{CS}} : \frac{\beta_k}{\lambda_k^{CS}},$$

[28] If they own the properly designed ETFs, they trade ETFs; if they own underlying securities, they can sell them too. At the daily horizon, such noise trading on ETFs would be corrected by the Authorized Participants. So, our discretionary liquidity traders do not have hedging needs based on the baskets as assumed in Subrahmanyam (1991). Mutual funds and index funds satisfying the liquidity needs of their clients may do so, but it would be hard to imagine passive mutual funds trading ETFs. Moreover, if they do not already own ETFs, they cannot sell ETFs to unload their assets because they are not authorized participants. This assumption thus is reasonable for pure passive funds but would not apply to sector index funds or other smart beta tilts.

for $k = 2, \ldots, K$. It is worthy highlighting that the above optimal trading volume allocation across assets for the factor liquidity trader resembles the weights of the optimal CS product chosen by factor speculators as in Eq. (19). More formally, we have the following duality:

Proposition 6.4. *In the unique FIE, factor speculators and the factor liquidity trader choose the same CS product, with weights* $\{w_k\}_{k=1}^K$ *satisfying:*

$$w_1 : w_k = \frac{\beta_1}{\lambda_1^{CS}} : \frac{\beta_k}{\lambda_k^{CS}}, \quad \forall\ k = 2, \ldots, K. \tag{25}$$

In other words, this CS with weights given in Eq. (25) serves a dual role—it maximizes the expected trading profits for factor speculators while it also minimizes the expected loss from trading for the factor liquidity trader.[29] Our conclusion about CS design remains robust.

7 Conclusion

Despite the drastic growth of passive investing (e.g., passive mutual funds, ETFs, smart beta products) in the past two decades, how to design such composite securities (CSs) remains little understood. Moreover, their impact on asset prices and the informational efficiency of financial markets are also unclear, with mixed empirical evidence. We model CSs as pass-through vehicles for investors to trade underlying assets subject to a realistic effort or trading cost. By reducing the duplication of each investor's effort cost of trading each security, CSs attract factor investors to exploit their information on the systematic component of asset value or hedge against exposure to systematic factors.

Our model features endogenous CS offering, market participation, informed trading, and price setting. It conceptually underscores how the so-called passive investing is actually active factor investing: the CS design and trading are active choices by CS sponsors and investors.

[29]This extension with factor liquidity traders is also helpful for understanding the economic consequences of trading transparency (i.e., underlying asset market makers observe and set prices contingent on CS volume) on financial markets without violating the market structure in Boulatov, Hendershott, and Livdan (2013). We discuss this alternative specification in Appendix C. Naturally, observing asset-related order flows (from asset speculators and liquidity traders), and factor-related order flows (from factor speculators and liquidity traders separately can help make makers learn more information about asset payoffs but negatively affect speculators' trading profits. Consequently, the optimal portfolio weights do not depend on the conventional price impacts but depend on the price impacts associated with factor-related order flows.

Concerning optimal CS design, we show in closed-form that CS sponsors optimally select liquid and representative assets, i.e., the portfolio weights of underlying assets are proportional to the assets' factor exposure and negatively proportional to their illiquidity (measured by Kyle (1985)'s price impact). We verify this implication in the U.S. equity ETFs.

Our model also yields novel implications for asset prices and informational efficiency. First, introducing CSs incorporates more factor information and leads to higher pricing efficiencies, price variability, and co-movements in the underlying asset markets. Second, introducing CSs can decrease (increase) the price impacts of underlying assets when the underlying assets have a high (low) number of factor-informed traders initially. Third, in the extension with endogenous asset-specific information acquisition, we find that introducing CSs can either increase or decrease asset-specific information acquisition and incorporation, depending on the asset's exposure to common shocks and idiosyncratic risk. These predictions are consistent with most of the recent empirical findings and reconciles mixed evidence in the literature.

References

Akey, Pat, Adriana Robertson, and Mikhail Simutin, 2021, Closet active management of passive funds, Rotman School of Management Working Paper.

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, Journal of Financial Markets 5, 31–56.

Appel, Ian R, Todd A Gormley, and Donald B Keim, 2016, Passive investors, not passive owners, Journal of Financial Economics 121, 111–141.

Athanasoulis, Stefano G, and Robert J Shiller, 2000, The significance of the market portfolio, The Review of Financial Studies 13, 301–329.

Bai, Jennie, Thomas Philippon, and Alexi Savov, 2016, Have financial markets become more informative?, Journal of Financial Economics 122, 625–654.

Baruch, Shmuel, and Xiaodi Zhang, 2022, The distortion in prices due to passive investing, Management Science 68, 6219–6234.

Ben-David, Itzhak, Francesco Franzoni, Byungwook Kim, and Rabih Moussawi, 2023, Competition for attention in the etf space, The Review of Financial Studies 36, 987–1042.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2018, Do etfs increase volatility?, The Journal of Finance 73, 2471–2535.

Bhattacharya, Ayan, and Maureen O'Hara, 2015, Can etfs increase market fragility? effect of information linkages in etf markets, Effect Of Information Linkages in ETF Markets (October 10, 2015).

Bhattacharya, Ayan, and Maureen OHara, 2017, Can ETFs increase market fragility? effect of information linkages in ETF markets, SSRN Electronic Journal.

Bhattacharya, Utpal, Benjamin Loos, Steffen Meyer, and Andreas Hackethal, 2017, Abusing etfs, Review of Finance 21, 1217–1250.

Bhojraj, Sanjeev, Partha Mohanram, and Suning Zhang, 2020a, Etfs and information transfer across firms, Journal of Accounting and Economics 70, 101336.

Bhojraj, Sanjeev, Partha S. Mohanram, and Suning Zhang, 2020b, Etfs and information transfer across firms, Rotman School of Management Working Paper.

Bond, Philip, and Diego Garcia, 2022, The equilibrium consequences of indexing, The Review of Financial Studies 35, 3175–3230.

Boulatov, Alex, Terrence Hendershott, and Dmitry Livdan, 2013, Informed trading and portfolio returns, Review of Economic Studies 80, 35–72.

Bradley, Harold, and Robert E Litan, 2011, Etfs and the present danger to capital formation, Ewing Marion Kauffman Foundation Research Paper.

Brogaard, Jonathan, Davidson Heath, and Da Huang, 2023, Etf sampling and index arbitrage, Available at SSRN 3510359.

Buss, Adrian, and Savitar Sundaresan, 2023, More risk, more information: How passive ownership can improve informational efficiency, The Review of Financial Studies 36, 4713–4758.

Cespa, Giovanni, and Thierry Foucault, 2014, Illiquidity contagion and liquidity crashes, The Review of Financial Studies 27, 1615–1660.

Chakravarty, Sugato, Huseyin Gulen, and Stewart Mayhew, 2004, Informed trading in stock and option markets, The Journal of Finance 59, 1235–1257.

Chan, Kam C, and Nai-Fu Chen, 1991, Structural and return characteristics of small and large firms, The journal of finance 46, 1467–1484.

Crawford, Steven S, Darren T Roulstone, and Eric C So, 2012, Analyst initiations of coverage and stock return synchronicity, The Accounting Review 87, 1527–1553.

Da, Zhi, and Sophie Shive, 2018, When the bellwether dances to noise: Evidence from exchange-traded funds, European Financial Management 24, 136–168.

Dai, Liang, 2018, Asset bundling and information acquisition of investors with different expertise, Journal of Economic Theory 175, 447–490.

Dávila, Eduardo, and Cecilia Parlatore, 2023, Identifying price informativeness, .

Diamond, Douglas W, 1984, Financial intermediation and delegated monitoring, The review of economic studies 51, 393–414.

Dow, James, 1998, Arbitrage, hedging, and financial innovation, The Review of Financial Studies 11, 739–755.

Easley, David, David Michayluk, Maureen O'Hara, and Tālis J Putniņš, 2021, The active world of passive investing, Review of Finance 25, 1433–1471.

Easley, David, Maureen O'hara, and Pulle Subrahmanya Srinivas, 1998, Option volume and stock prices: Evidence on where informed traders trade, The Journal of Finance 53, 431–465.

Engle, R.F., and D. Sarkar, 2006, Premium-discounts and exchange traded funds, Journal of Derivatives pp. 27–45.

Farboodi, Maryam, Adrien Matray, Laura Veldkamp, and Venky Venkateswaran, 2022, Where has all the data gone?, The Review of Financial Studies 35, 3101–3138.

Gârleanu, Nicolae, and Lasse Heje Pedersen, 2022, Active and passive investing: Understanding samuelsons dictum, The Review of Asset Pricing Studies 12, 389–446.

Glasserman, Paul, and Harry Mamaysky, 2023, Investor information choice with macro and micro information, The Review of Asset Pricing Studies 13, 1–52.

Glosten, Lawrence, Suresh Nallareddy, and Yuan Zou, 2019, Etf trading and informational efficiency of underlying securities, Management Science (forthcoming).

——— , 2021, Etf activity and informational efficiency of underlying securities, Management Science 67, 22–47.

Goldstein, Itay, and Liyan Yang, 2014, Market efficiency and real efficiency: The connect and disconnect via feedback effects, .

——— , 2015, Information diversity and complementarities in trading and information acquisition, The Journal of Finance 70, 1723–1765.

Gomes, Francisco, Michael Haliassos, and Tarun Ramadorai, 2021, Household finance, Journal of Economic Literature 59, 919–1000.

Gorton, Gary B, and George G Pennacchi, 1993, Security baskets and index-linked securities, Journal of Business pp. 1–27.

Hamm, Sophia JW, 2014, The effect of etfs on stock liquidity, Available at SSRN 1687914.

Huang, Shiyang, Maureen O'Hara, and Zhuo Zhong, 2020, Innovation and Informed Trading: Evidence from Industry ETFs, The Review of Financial Studies 34, 1280–1316.

Israeli, Doron, Charles Lee, and Suhas A Sridharan, 2017, Is there a dark side to exchange traded funds (etfs)? an information perspective, Review Of Accounting Studies 22, 1048–1083.

Jiang, Hao, Dimitri Vayanos, and Lu Zheng, 2022, Passive investing and the rise of mega-firms, NBER Working Paper.

Jung, Jeeman, and Robert J Shiller, 2005, Samuelson's dictum and the stock market, Economic Inquiry 43, 221–228.

Koont, Naz, Yiming Ma, Lubos Pastor, and Yao Zeng, 2022, Steering a ship in illiquid waters: Active management of passive funds, Discussion paper, National Bureau of Economic Research.

Krause, Timothy, Sina Ehsani, and Donald Lien, 2014, Exchange-traded funds, liquidity and volatility, Applied Financial Economics 24, 1617–1630.

Kyle, Albert S, 1985, Continuous auctions and insider trading, Econometrica: Journal of the Econometric Society pp. 1315–1335.

Lee, Jeongmin, 2021, Passive investing and price efficiency, Working Paper.

Lee, Samuel, 2013, Active investment, liquidity externalities, and markets for information, Working Paper.

Leippold, Markus, Lujing Su, and Alexandre Ziegler, 2015, Do index futures and etfs affect stock return correlations?, Available at SSRN 2620955.

Li, Kun, Xin Liu, and Shang-Jin Wei, 2023, The role of credit rating purchases in sp 500 membership decisions, Working paper.

Li, Sida, 2021, Should passive investors actively manage their trades?, Available at SSRN 3967799.

Madhavan, Ananth, and Aleksander Sobczyk, 2014, Price dynamics and liquidity of exchange-traded funds, Available at SSRN 2429509.

Malamud, Semyon, 2015, A dynamic equilibrium model of etfs, Swiss Finance Institute Research Paper.

Malikov, George, 2023, Information, participation, and passive investing, Participation, and Passive Investing.

Molk, Peter, and Adriana Robertson, 2023, Discretionary investing by passives&p 500 funds, Yale Journal on Regulation, Forthcoming.

OHara, Maureen, Yihui Wang, and Xing Zhou, 2018, The execution quality of corporate bonds, Journal of Financial Economics 130, 308–326.

Pan, Kevin, and Yao Zeng, 2017, ETF arbitrage under liquidity mismatch, ESRB Working Paper Series 59 European Systemic Risk Board.

Rahi, Rohit, and Jean-Pierre Zigrand, 2009, Strategic financial innovation in segmented markets, The Review of Financial Studies 22, 2941–2971.

Rauterberg, Gabriel, and Andrew Verstein, 2013, Index theory: The law, promise and failure of financial indices, Yale J. on Reg. 30, 1.

Robertson, Adriana Z, 2023, The (mis) uses of the s&p 500, The University of Chicago Business Law Review 2, 3.

Samuelson, Paul A, et al., 1998, Summing up on business cycles: opening address, in Conference Series-Federal Reserve Bank of Boston vol. 42 pp. 33–36. Federal Reserve Bank of Boston.

Sharfman, Bernard S, and Vincent Deluard, 2021, How discretionary decision-making impacts the financial performance and legal disclosures of s&p 500 funds, Brook. L. Rev. 87, 915.

Shen, Ji, Hongjun Yan, and Jinfan Zhang, 2014, Collateral-motivated financial innovation, The Review of Financial Studies 27, 2961–2997.

Shiller, Robert C, 2000, Irrational exuberance, Philosophy and Public Policy Quarterly 20, 18–23.

Stambaugh, Robert F, 2014, Presidential address: Investment noise and trends, The Journal of Finance 69, 1415–1453.

Stoll, Hans R, and Robert E Whaley, 1983, Transaction costs and the small firm effect, Journal of Financial Economics 12, 57–79.

Subrahmanyam, Avanidhar, 1991, A theory of trading in stock index futures, Review of Financial Studies 4, 17–51.

——— , and Sheridan Titman, 1999, The going-public decision and the development of financial markets, The Journal of Finance 54, 1045–1082.

Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2010, Information acquisition and under-diversification, The Review of Economic Studies 77, 779–805.

Vayanos, Dimitri, and Jiang Wang, 2013, Chapter 19 - market liquiditytheory and empirical evidence *, vol. 2 of Handbook of the Economics of Finance . pp. 1289–1361 (Elsevier).

Veldkamp, Laura, 2006a, Information markets and the comovement of asset prices, Review of Economic Studies 73, 823–845.

——— , 2006b, Media frenzies in markets for financial information, American Economic Review 96, 577–601.

——— , and Justin Wolfers, 2007, Aggregate shocks or aggregate information? costly information and business cycle comovement, Journal of Monetary Economics 54, 37–55.

Ye, Shuai, 2019, How do etfs affect the liquidity of the underlying corporate bonds?, Working Paper.

Yuan, Kathy, 2005, The liquidity service of benchmark securities, Journal of the European Economic Association 3, 1156–1180.

Appendix: Proofs of Lemmas and Propositions

Proof of Lemma of 3.1

We first solve the price impact and then solve the expected profit of factor speculators. Inserting the expressions of κ_k and η_k in the expression of λ_k yields:

$$\lambda_k = \frac{\frac{1}{2\lambda_k}\sigma^2_{\alpha_k} + N_k\beta_k\sigma^2_\gamma \frac{\beta_k}{\lambda_k(N_k+1+2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma})}}{\frac{1}{4\lambda_k^2}\sigma^2_{\alpha_k} + N_k^2\sigma^2_\gamma \frac{\beta_k^2}{(N_k+1+2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma})^2}\frac{1}{\lambda_k^2} + N_k\sigma^2_\epsilon \frac{\beta_k^2}{(N_k+1+2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma})^2}\frac{1}{\lambda_k^2} + \sigma^2_{n_k}}. \tag{A-1}$$

After simplifying the above equation, we solve λ_k as:

$$\lambda_k = \frac{1}{\sigma_{n_k}}\sqrt{\frac{\sigma^2_{\alpha_k}}{4} + \beta_k^2 \frac{N_k(\sigma^2_\gamma + \sigma^2_\epsilon)\sigma^4_\gamma}{[(N_k+1)\sigma^2_\gamma + 2\sigma^2_\epsilon]^2}}. \tag{A-2}$$

For the expected profit,

$$\begin{aligned}
\Pi_k^F &= \frac{(1+\frac{\sigma^2_\epsilon}{\sigma^2_\gamma})^2\beta_k^2}{(N_k+1+2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma})^2\lambda_k}\frac{\sigma^4_\gamma}{\sigma^2_\gamma+\sigma^2_\epsilon} \\
&= \frac{(\sigma^2_\gamma+\sigma^2_\epsilon)^2\beta_k^2}{[(N_k+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2\frac{1}{\sigma_{n_k}}\sqrt{\frac{\sigma^2_{\alpha_k}}{4} + \beta_k^2\frac{N_k(\sigma^2_\gamma+\sigma^2_\epsilon)\sigma^4_\gamma}{[(N_k+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2}}}\frac{\sigma^4_\gamma}{\sigma^2_\gamma+\sigma^2_\epsilon} \\
&= \frac{(\sigma^2_\gamma+\sigma^2_\epsilon)\beta_k^2\sigma^4_\gamma\sigma_{n_k}}{[(N_k+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2\sqrt{\frac{\sigma^2_{\alpha_k}}{4} + \beta_k^2\frac{N_k(\sigma^2_\gamma+\sigma^2_\epsilon)\sigma^4_\gamma}{[(N_k+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2}}}
\end{aligned}$$

Proof of Proposition of 3.1

The equilibrium N_k is the solution to the equation:

$$\Pi_k^F = C.$$

where Π_k^F is as given in Lemma 3.1. It is clear that Π_k^F is decreasing with N_k. Thus, there is one unique solution to the above equation. Meanwhile, since Π_k^F is increasing with β_k, σ^2_γ, and $\sigma^2_{n_k}$ but is decreasing with $\sigma^2_{\alpha_k}$, when β_k or σ^2_γ, or $\sigma^2_{n_k}$ increases, N_k^* should increase to make $\Pi_k^F = C$ hold in the equilibrium. When $\sigma^2_{\alpha_k}$ increases, N_k^* should decrease to make $\Pi_k^F = C$ hold in the equilibrium.

Proofs of Lemma of 3.2

Suppose by way of contradiction that in equilibrium there exists an Asset $k^\star$ such that $I_{k^\star} \cap I_{CS} \neq \emptyset$ and suppose factor speculator $i^\star \in I_{k^\star} \cap I_{CS}$. Denote by $y_{k^\star,i^\star}$ the equilibrium trading volume of this factor speculator $i^\star$ in Asset market $k^\star$ and by $y_{CS,i^\star}$ the equilibrium trading volume in CS. Furthermore, denote by $\{w_{k,i^\star}\}_{k=1}^K$ the weights of the CS product chosen by this factor speculator $i^\star$ in equilibrium, with the fee charged being $F^\star$.

Now consider the following synthetic CS with weights

$$\{\hat{w}_k\}_{k=1}^K = \frac{y_{k,i^\star}}{y_{k,i^\star} + y_{CS,i^\star}} \cdot \left(0, \cdots, \underbrace{1}_{k^\star th}, \cdots, 0\right) + \frac{y_{CS,i^\star}}{y_{k,i^\star} + y_{CS,i^\star}} \cdot \{w_{k,i^\star}\}_{k=1}^K \tag{A-3}$$

Then it must be the case that this synthetic CS product constructed as above is offered by the operating CS sponsor in the equilibrium. Otherwise, a new CS sponsor can enter by offering the same list of the CS products offered by the incumbent CS sponsor plus this synthetic CS product with design $\{\hat{w}_k\}_{k=1}^K$ and fee $\hat{F} = F^\star + \epsilon$ where $\epsilon > 0$ is arbitrarily small, which has a marginal cost of zero to the CS sponsor. Under our dominance assumption, this deviation is a profitable one for this new entrant CS sponsor given the anticipation that after she offers this more inclusive menu of CS products, all factor speculators who trade CS products in equilibrium will be attracted. This include the factor speculator $i^\star$, who previous incurs a total trading cost of $2C + F^\star$ but can replicate her asset trading via this synthetic CS product with a total cost of $C + F^\star + \epsilon$.

However, if such a synthetic CS product with weight given in Eq. (A-3) is offered by CS sponsors in their date 0 game play, then a deviation in the speculators' subgame play on date 1 exists—the factor speculator $i^s tar$ will deviate to trade this synthetic CS product, which is *feasible* in this subgame equilibrium. Therefore, it is impossible that in equilibrium there exists a factor speculator simultaneously trades in underlying asset markets and CS product.

Proof of Proposition 3.2

The proof of this proposition is similar to that of Lemma of 3.1 and is omitted for brevity.

Proof of Lemma 3.3

Following the same argument, we can further show that in equilibrium there cannot be any factor speculators that directly trade in underlying assets. To see this,

Similar to our proof of Lemma 3.2, we prove this result following way of contradiction. Consider an equilibrium in which a CS product of design $\mathbf{w} = (w_1, \cdots, w_K)$ with positive fee charged ($F > 0$) is offered by the sponsor on $t = 0$ and traded by factor speculators on $t = 1$. There must exist such an CS product since otherwise CS sponsor will not be able to break even. Suppose in equilibrium there exists factor speculator that still directly trades certain underlying assets (say, Asset 1) in the subgame equilibrium at $t = 1$. From Lemma 3.2, it follows that this factor speculator is not trading the CS product. In this equilibrium, for the factor speculator who directly trades the underlying Asset 1, she earns an equilibrium trading profit of $\Pi = C$ by competitive entry. Factor speculators' equilibrium trading profit from the CS product $\mathbf{w}$ is $\Pi_{CS}^F = C + F > C$, since a positive fee $F > 0$ needs to be charged so that the CS sponsor can break even. As such, there must exist an Asset k^* ($k^* \neq 1$) such that factor speculators are making a strictly positive trading profit $\hat{\Pi}_k^*$ from Asset k^* through trading the CS product $\mathbf{w}$.

In this equilibrium, under our dominance assumption in 2.1, a synthetic CS product consists of just Asset 1 and Asset k^* with zero fee must be offered by the incumbent CS sponsor. Otherwise, a potential CS sponsor could then enter on date $t = 0$ and offer the list of CS products (including $\mathbf{w}$) offered by the incumbent CS sponsor plus a full spectrum of CS products constructed as a linear combination of Asset 1 and Asset k^* with zero fee charged (since the marginal cost of offering such synthetic securities is zero).

However, with such a feasible set of CS product choices in the subgame played on $t = 1$, there must exist a profitable deviation for a factor speculator who previously trade directly in Asset 1. This is because this factor speculator can replicate her previous trading positive in Asset 1 (which guarantees a profit C) while also gain access to profits $\hat{\Pi}_k^* > 0$ (at least partially) from Asset k^* via trading a CS product constructed as a properly chosen combination of Asset 1 and Asset k^*.

Proofs of Lemma of 3.4

Following the proof of Theorem 3.3, in the equilibrium we must have $\hat{N}_k = 0$ for all $k = 1, \ldots, N$.

We use the method of proof by contradiction to show that $N_{CS} \geq N_1$. Assuming $N_{CS} < N_1$, the expected trading profit for factor speculators that trade asset k directly is:

$$\hat{\Pi}_1^F = \frac{(\sigma_\gamma^2 + \sigma_\epsilon^2)\beta_1^2\sigma_\gamma^4\sigma_{n_k}}{[(N_{CS}+1)\sigma_\gamma^2 + 2\sigma_\epsilon^2]^2\sqrt{\frac{\sigma_{\alpha_k}^2}{4} + \beta_k^2\frac{(N_{CS})(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}}. \tag{A-4}$$

When $N_{CS} < N_1$, we have

$$\hat{\Pi}_1^F - C > \frac{(\sigma_\gamma^2 + \sigma_\epsilon^2)\beta_1^2\sigma_\gamma^4\sigma_{n_k}}{[(N_1+1)\sigma_\gamma^2 + 2\sigma_\epsilon^2]^2\sqrt{\frac{\sigma_{\alpha_k}^2}{4} + \beta_k^2\frac{(N_1)(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_1+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} - C = 0. \tag{A-5}$$

This suggests that potential factor speculators that have not trade asset 1 will deviate from the equilibrium and participate asset 1, which contradicts the equilibrium that $\hat{N}_1 = 0$. Thus, the only equilibrium in this case should have $N_{CS} \geq N_1$.

Proof of Proposition 3.3

The proof for this proposition is directly from the expressions of $\hat{\eta}_{CS,k}$ (see Equation 16 in Proposition 3.2).

Proof of Proposition 4.1

There are three parts in this proof. In the first part, we show that CS trading decrease asset-specific efficiency. In the second part, we show that CS trading increases factor-specific efficiency. In the third part, we show that CS trading increase the total efficiency.

Part 1: According to the projection theorem, for $Var(\alpha_k|P_k)$, we have:

$$Var(\alpha_k|P_k) = Var(\alpha_k) - \frac{Cov(\alpha_k, P_k)^2}{Var(P_k)}.$$

Here: $Cov(\alpha_k, P_k) = \sigma^2_{\alpha_k}$ and

$$Var(P_k) = \frac{\sigma^2_{\alpha_k}}{2} + \beta_k^2 \sigma_\gamma^2 \frac{N_k}{N_k + 1 + 2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma}}. \tag{A-6}$$

It is clear that $Var(P_k)$ is increasing with N_k. Thus, $Var(\alpha_k|P_k)$ is increasing with N_k. This implies that introducing CS trading increases the number of factor speculators and decreases the asset-specific information efficiency.

Part 2: According to the projection theorem, for $Var(\gamma|P_k)$,we have:

$$Var(\gamma|P_k) = Var(\gamma) - \frac{Cov(\gamma, P_k)^2}{Var(P_k)}$$

Here: $Cov(\gamma, P_k) = \frac{N_k\beta_k}{N_k+1+2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma}}\sigma^2_\gamma$. Therefore, we have:

$$\frac{Cov(\gamma, P_k)^2}{Var(P_k)} = \sigma^2_\gamma \left[\frac{N_k\beta_k}{N_k + 1 + 2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma}}\sigma^2_\gamma\right]^2 / \left[\frac{\sigma^2_{\alpha_k}}{2} + \frac{N_k\beta_k^2\sigma^2_\gamma}{N_k + 1 + 2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma}}\right].$$

It is not difficult to get that $Cov(\gamma, P_k)^2/Var(P_k)$ is increasing with N_k. This implies that introducing CS trading increases the number of factor speculators and increases the factor-specific information efficiency.

Part 3: According to the projection theorem, for $Var(\alpha_k + \beta_k\gamma|P_k)$, we have:

$$Var(\alpha_k + \beta_k\gamma|P_k) = Var(\alpha_k + \beta_k\gamma) - \frac{Cov(\alpha_k + \beta_k\gamma, P_k)^2}{Var(P_k)}$$

Since $Cov(\alpha_k + \beta_k\gamma, P_k) = \frac{\sigma^2_{\alpha_k}}{2} + \frac{N_k\beta_k^2}{N_k+1+2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma}}\sigma^2_\gamma$, we can get that:

$$\frac{Cov(\alpha_k + \beta_k\gamma, P_k)^2}{Var(P_k)} = \frac{\sigma^2_{\alpha_k}}{2} + \frac{N_k\beta_k^2}{N_k + 1 + 2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma}}\sigma^2_\gamma.$$

This implies that introducing CS trading increases the number of factor speculators and increases the total information efficiency.

Proof of Proposition 4.2

Introducing CS effectively increases the number of factor speculators in each asset. Thus, we could do comparative statics with N_k in asset k. Examining λ_k shows that λ_k increases with N_k when $N_k \leq \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$ but decreases with N_k when $N_k > \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$. In this sense, when $N_{CS}^* \leq \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$, it suggests that $N_k \leq \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$ and CS increases N_k, which in turn increases λ_k^{CS}.

When $N_{CS}^* > (\sigma_\gamma^2 + 2\sigma_\epsilon^2)/\sigma_\gamma^2$, we have two cases. First, if $N_k N_{CS}^* < \left[(\sigma_\gamma^2 + 2\sigma_\epsilon^2)/\sigma_\gamma^2\right]^2$, $\frac{N_k(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2} > \frac{N_{CS}^*(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}^*+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}$. From the expressions of price impacts, we can obtain that the price impact in the economy with CS is higher than that without CS.

Second, if $\left[(\sigma_\gamma^2 + 2\sigma_\epsilon^2)/\sigma_\gamma^2\right]^2 < N_k N_{CS}^*$, $\frac{N_k(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2} < \frac{N_{CS}^*(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}^*+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}$. From the expressions of price impacts, we can obtain that the price impact in the economy with CS is lower than that without CS.

Proof of Proposition 4.3

This proof has two parts. In the first part, we show that introducing CS increase return variance. In the second part, we show that introducing CS increase return co-movement.

Part 1: From the proof of Proposition 4.1, we have

$$Var(P_k) = \frac{\sigma_{\alpha_k}^2}{2} + \beta_k^2 \sigma_\gamma^2 \frac{N_k}{N_k + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}}.$$

This suggests that introducing CS increases the return variance.

Part 2: We first calculate the return correlation coefficient between any two underlying assets before CS trading. For any two underlying assets (i,j), the return covariance is maximized when they have same factor speculators. Assuming $N_i \leq N_j$, we have

$$COV(P_i, P_j) \leq \frac{N_i\beta_i}{N_i + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}} \frac{N_j\beta_j}{N_j + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}} \sigma_\gamma^2 \quad \text{(A-7)}$$

$$+ \frac{\beta_i\sqrt{N_i}}{N_i + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}} \frac{\beta_j\sqrt{N_j}}{N_j + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}} \sigma_\epsilon^2. \quad \text{(A-8)}$$

The correlation coefficient can be written as:

$$corr(P_i, P_j) \leq c_1 \cdot c_2 \cdot \sigma_\gamma^2 + d_1 \cdot d_2 \cdot \sigma_\epsilon^2 \quad \text{(A-9)}$$

where

$$c_1 = \frac{1}{\sqrt{\frac{\sigma_{\alpha_i}^2}{2} + \beta_i^2\sigma_\gamma^2 \frac{N_i}{N_i+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}}}} \frac{N_i\beta_i}{N_i + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}},$$

$$c_2 = \frac{1}{\sqrt{\frac{\sigma_{\alpha_j}^2}{2} + \beta_j^2\sigma_\gamma^2 \frac{N_j}{N_j+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}}}} \frac{N_j\beta_j}{N_j + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}},$$

$$d_1 = \frac{1}{\sqrt{\frac{\sigma^2_{\alpha_i}}{2} + \beta_i^2\sigma_\gamma^2 \frac{N_i}{N_i+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}}}} \frac{\beta_i\sqrt{N_i}}{N_i+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}},$$

$$d_2 = \frac{1}{\sqrt{\frac{\sigma^2_{\alpha_j}}{2} + \beta_j^2\sigma_\gamma^2 \frac{N_j}{N_j+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}}}} \frac{\beta_j\sqrt{N_j}}{N_j+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}}.$$

It is clear that c_1 and d_1 are increasing with N_i, and c_2 and d_2 are increasing with N_j.

Now we can consider the effect of CS trading on the correlation coefficient between asset i and j. Introducing CS trading has two effects. First, introducing CS trading increases the effective numbe of factor speculators trading on underlying assets. Second, when factor specualtors trade underlying assets indireclty via CSss, the effective number of factor speculators trading on any two assets overlap more. These two effects work together and increase c_1, c_2, d_1 and d_2, which increases the return correlation between asset i and j.

Proofs of Proposition 6.1

We start with first part of the proposition. In what follows, we characterizes the sufficient conditions to ensure the participation or non-participation of asset speculators.

(1) To characterize the sufficient condition to ensure the participation of asset speculators, we focus on the first case in Proposition A1.1 ($N_k^{C1} < N_k^{C2} < n_{k1} < n_{k2}$). Since n_{k1} is decreasing with β_k and N_k^{C2} is increasing with β_k, there exists a cutoff β_L such that $N_k^{C2} < n_{k1}$ for all $\beta_k < \beta_L$. In this sense, when $\beta_k < \beta_L$, asset speculators participate asset markets.

(2) To characterize the sufficient condition to ensure non-participation of asset speculators, we focus on the third case in Proposition A1.1 ($n_{k1} < N_k^{C1} < N_k^{C2} < n_{k2}$). Since n_{k1} is decreasing with β_k and both N_k^{C1} and N_k^{C2} are increasing with β_k, there exists a cutoff β_{H1} such that $n_{k1} < N_k^{C1}$ for all $\beta_k > \beta_{H1}$. Meanwhile, we need to find a sufficient condition to ensure $N_k^{C2} < n_{k2}$. Given that N_k^{C2} is the solution to $\frac{\sqrt{(\sigma_\gamma^2+\sigma_\epsilon^2)}\beta_k\sigma_\gamma^2\sigma_{n_k}}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]\sqrt{N_k}} = C$, the sufficient condition for $N_k^{C2} < \frac{b+\sqrt{b^2-4(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2}}{2}$ is equivalent to:

$$\frac{\sqrt{(\sigma_\gamma^2+\sigma_\epsilon^2)}\beta_k\sigma_\gamma^2\sigma_{n_k}}{\left[(\frac{b+\sqrt{b^2-4(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2}}{2}+1)\sigma_\gamma^2+2\sigma_\epsilon^2\right]\frac{b+\sqrt{b^2-4(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2}}{2}} < C. \tag{A-10}$$

Since LHS of Equation (A-10) goes to zero when β_k goes to ∞, there exists a sufficient cutoff β_{H2} such that the above equation holds for all $\beta_k > \beta_{H2}$. This suggests that when β_k is higher than $\max(\beta_{H1}, \beta_{H2})$ (denoted as β_H), asset speculators do not trade underlying assets.

Similarly, we follow the same procedure to prove the second part of this proposition. Again, we characterizes the sufficient conditions to ensure the participation or non-participation of asset speculators.

(1) To characterize the sufficient condition to ensure the participation of asset speculators, we focus on the first case in Proposition A1.1 ($N_k^{C1} < N_k^{C2} < n_{k1} < n_{k2}$). Since n_{k1} is increasing with $\sigma^2_{\alpha_k}$ and N_k^{C2} is not a function of $\sigma^2_{\alpha_k}$, there exists a cutoff σ^2_H such that $N_k^{C2} < n_{k1}$ for all $\sigma^2_{\alpha_k} > \sigma^2_H$. In this sense, when $\sigma^2_{\alpha_k} > \sigma^2_H$, asset speculators trade underlying assets directly.

(2) To characterize the sufficient condition to ensure the non-participation of asset speculators, we focus on the third case in Proposition A1.1 ($n_{k1} < N_k^{C1} < N_k^{C2} < n_{k2}$). Since n_{k1} is increasing with $\sigma^2_{\alpha_k}$ and N_k^{C1} is increasing with increasing with $\sigma^2_{\alpha_k}$, there exists a cutoff σ^2_{L1} such that $n_{k1} < N_k^{C1}$ for $\sigma^2_{\alpha_k} < \sigma^2_{L1}$. Meanwhile, we need to find a sufficient condition to ensure $N_k^{C2} < n_{k2}$. Since n_{k2} is decreasing with $\sigma^2_{\alpha_k}$ and N_k^{C2} is not a function of $\sigma^2_{\alpha_k}$, there exists a cutoff σ^2_{L2} such that $N_k^{C2} < n_{k2}$ for all $\sigma^2_{\alpha_k} < \sigma^2_{L2}$. Thus, when $\sigma^2_{\alpha_k} < \min(\sigma^2_{L1}, \sigma^2_{L2})$ (denoted as σ^2_L) , asset speculators do not trade underlying assets.

Proofs of Proposition 6.2

Depending on β_L (asset speculators trade assets with β_k lower than β_L), there are two potential cases: $\beta_L < \beta^*$ and $\beta_L > \beta^*$. We prove this proposition case by case as follows.

Case 1: $\beta_L < \beta^*$

From the definition of β_L (see Proposition 6.1), we know that asset speculators trade assets with β_k lower than β_L. For assets with β_k lower than β_L, the equilibrium number of factor speculators before introducing CS is determined by the following condition:

$$\frac{(\sigma^2_\gamma + \sigma^2_\epsilon)\beta_k^2 \sigma^4_\gamma \sigma_{n_k}}{[(N_k+1)\sigma^2_\gamma + 2\sigma^2_\epsilon]^2 \sqrt{\frac{\sigma^2_{\alpha_k}}{4} + \beta_k^2 \frac{N_k(\sigma^2_\gamma+\sigma^2_\epsilon)\sigma^4_\gamma}{[(N_k+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2}}} = C, \tag{A-11}$$

which suggests that N_k is a function of β_k and is increasing with β_k.

In this case, because $\beta_L < \beta^*$, we can easily infer that $N_L \equiv N_k(\beta_L)$ is lower than $(\frac{\sigma^2_\gamma+2\sigma^2_\epsilon}{\sigma^2_\gamma})^2/N^*_{CS}$. Thus, introducing CS will increase the price impact (from the proof of Lemma A1.1) and lower expected profit of asset speculators due to increased price impact. Consequently, asset speculators' incentives to trade these assets weakly decreases.

Now we pin down which assets will lose asset speculators. For assets with $\beta_k = \beta_L$, asset speculator trades this asset before CS trading and her expected trading profit is:

$$\Pi^A_L = \frac{\sigma^2_\alpha \sigma_n}{4\sqrt{\frac{\sigma^2_\alpha}{4} + \beta_L^2 \frac{N_L(\sigma^2_\gamma+\sigma^2_\epsilon)\sigma^4_\gamma}{[(N_L+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2}}} = C_A. \tag{A-12}$$

Examining asset k with $\beta_k > \beta_L \sqrt{\frac{N_L(\sigma^2_\gamma+\sigma^2_\epsilon)\sigma^4_\gamma}{[(N_L+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2} / \frac{N^*_{CS}(\sigma^2_\gamma+\sigma^2_\epsilon)\sigma^4_\gamma}{[(N^*_{CS}+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2}}$ but $\beta_k \le \beta_L$, after CS introduction, the expected profit of asset speculators (if trading asset k) is lower than $\Pi^A_L(C_A)$ and thus asset speculator will not trade these assets. This analysis shows that introducing CS trading indeed deters the participation of asset speculators on some assets with β_k lower than β_L. Intuitively, when asset speculators exit asset market k, there is no asset-specific information in the asset price, which suggests that asset-specific information efficiency deceases after CS introduction. .

Case 2: $\beta_L > \beta^*$

Given the association between N_k and β_L, we know that $N_k(\beta_L)$ is higher than $(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2/N_{CS}^*$. Thus, introducing CS will decrease the price impact (from the proof of Lemma A1.1) and increase expected profit of asset speculators due to decreased price impact. Consequently, asset speculators' incentives to trade these assets weakly increases.

Now we pin down which assets will attract asset speculators after CS introduction. For assets with $\beta_k = \beta_L$, asset speculator trades this asset before CS trading and her expected trading profit is:

$$\Pi_L^A = \frac{\sigma_\alpha^2\sigma_n}{4\sqrt{\frac{\sigma_\alpha^2}{4} + \beta_L^2\frac{N_L(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_L+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} = C_A. \tag{A-13}$$

Now for asset k with $\beta_k > \beta_L$ but $\beta_k \leq \beta_L\sqrt{\frac{N_L(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_L+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}/\frac{N_{CS}^*(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}^*+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}$, after CS introduction, the expected profit of asset speculators (if trading asset k) is higher than Π_L^A (C_A) and thus asset speculators start to trade these assets. Intuitively, for assets with $\beta_k > \beta_L$, asset speculator does not trade on it and there is no asset-specific information in the asset price. After CS introduction, asset speculators start to trade asset k and there is asset-specific information in the asset price, which suggests that asset-specific information efficiency increases after CS introduction.

Finally, we examine whether asset speculators exit the markets on assets with β_k lower than β^*. For assets with with β_k lower than β^*, introducing CS will increase the price impact on these assets and decrease expected profit of asset speculators. This implies that potentially asset speculators will exit on these assets. However, the asset speculators in these assets are still making positive profits and will not leave the market. We prove this argument as follows.

After introducing CS, for assets with β_k lower than β^*, their profit is

$$\frac{\sigma_\alpha^2\sigma_n}{4\sqrt{\frac{\sigma_\alpha^2}{4} + \beta^2\frac{N_{CS}^*(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}^*+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}}, \tag{A-14}$$

where β_k lower than β^* (or β_L) and $\frac{N_{CS}^*(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}^*+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}$ is lower $\frac{N_L(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_L+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}$. This suggests that profit is still higher than C_A and thus asset speculators will not exist the markets.

Proofs of Proposition 6.3

Depending on σ_H^2 (asset speculators trade assets with σ_k^2 higher than σ_H^2), there are two potential cases: $\sigma_H^2 > \sigma^{2*}$ and $\sigma_H^2 < \sigma^{2*}$. We prove this proposition case by case as follows.

Case 1: $\sigma_H^2 > \sigma^{2*}$

From the definition of σ_H^2 (see Proposition 6.1), we know that asset speculators participate asset k with $\sigma_{\alpha_k}^2$ higher than σ_H^2. From the association between N_k and $\sigma_{\alpha_k}^2$, when $\sigma_H^2 > \sigma^{2*}$, we have that $N_k(\sigma_H^2)$ is lower than $(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2/N_{CS}^*$. Thus, for asset k with $\sigma_{\alpha_k}^2$ higher than σ_H^2, introducing CS will increase the price impact (from the proof of Lemmas A1.1 and A1.2) and lower expected profit of asset speculators. Consequently, the incentive of asset speculators to trade asset with $\sigma_{\alpha_k}^2$ higher than σ_H^2 weakly decreases.

We turn to pin down which assets will lose asset speculators. As the definition of σ_H^2, we have

$$\Pi_H^A = \frac{(\sigma_H^2)^2\sigma_n}{4\sqrt{\frac{\sigma_H^2}{4} + \beta^2\frac{N_H(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_H+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} = C_A. \tag{A-15}$$

We now examine asset k with $\sigma_{\alpha_k}^2 \geq \sigma_H^2$ but $\sigma_{\alpha_k}^2 < \sigma_C^2$, where σ_C^2 satisfies

$$\frac{(\sigma_C^2)^2\sigma_n}{4\sqrt{\frac{\sigma_C^2}{4} + \beta^2\frac{N_{CS}^*(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}^*+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} = \frac{(\sigma_H^2)^2\sigma_n}{4\sqrt{\frac{\sigma_H^2}{4} + \beta^2\frac{N_H(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_H+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}}. \tag{A-16}$$

For this asset, the expected profit of asset speculators (if trading asset k) is lower than C_A and thus will not participate. This analysis shows that introducing CS trading indeed deters the participation of asset speculators on some assets with $\sigma_{\alpha_k}^2$ higher than σ_H^2. Intuitively, when asset speculators exit asset market k, there is no asset-specific information in the asset price, which suggests that asset-specific information efficiency deceases after CS introduction.

Case 2: $\sigma_L^2 < \sigma^{2*}$

Given the association between between N_k and $\sigma_{\alpha_k}^2$, when $\sigma_L^2 < \sigma^{2*}$, we have that $N_k(\sigma_L^2)$ is higher than $(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2/N_{CS}^*$. Thus, for asset k with $\sigma_{\alpha_k}^2$ lower than σ_H^2, introducing CS will lower the price impact and increase expected profit of asset speculators. Consequently, the incentive of asset speculators to trade asset with $\sigma_{\alpha_k}^2$ higher than σ_H^2 weakly increases.

We turn to pin down which assets will lose asset speculators. As the definition of σ_H^2, we have

$$\Pi_H^A = \frac{\sigma_H^2\sigma_n}{4\sqrt{\frac{\sigma_H^2}{4} + \beta_H^2\frac{N_H(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_H+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} = C_A. \tag{A-17}$$

We examine asset k with $\sigma_{\alpha_k}^2 > \sigma_{CC}^2$ but $\sigma_{\alpha_k}^2 \leq \sigma_L^2$, where σ_C^2 satisfies

$$\frac{(\sigma_{CC}^2)^2\sigma_n}{4\sqrt{\frac{\sigma_{CC}^2}{4} + \beta^2\frac{N_{CS}^*(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}^*+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} = \frac{(\sigma_L^2)^2\sigma_n}{4\sqrt{\frac{\sigma_L^2}{4} + \beta^2\frac{N_L(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_L+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}}. \tag{A-18}$$

After introducing CS, if asset speculator trades asset k, her profit is higher than Π_H^A and thus will start to trade asset k.

Finally, we examine whether asset speculators exit the market of asset k with $\sigma_{\alpha_k}^2$ higher than σ^{2*}. For assets with with $\sigma^2_{\alpha_k}$ higher than σ^{2*}, introducing CS will increase the price impact and decrease expected profit of asset speculators. This implies that potentially asset speculators will exit on these assets. However, asset speculators in these assets are still making positive profits and will not exit the market. We prove this argument as follows:

After introducing CS, for assets with $\sigma^2_{\alpha_k}$ higher than σ^{2*}, the trading profit of asset speculator is

$$\frac{\sigma^2_{\alpha_k}\sigma_n}{4\sqrt{\frac{\sigma^2_{\alpha_k}}{4}+\beta^2\frac{N^*_{CS}(\sigma^2_\gamma+\sigma^2_\epsilon)\sigma^4_\gamma}{[(N^*_{CS}+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2}}}, \tag{A-19}$$

where $\sigma^2_{\alpha_k}$ is higher than σ^{2*} (or σ^2_L) and $\frac{N^*_{CS}(\sigma^2_\gamma+\sigma^2_\epsilon)\sigma^4_\gamma}{[(N^*_{CS}+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2}$ is lower $\frac{N_L(\sigma^2_\gamma+\sigma^2_\epsilon)\sigma^4_\gamma}{[(N_L+1)\sigma^2_\gamma+2\sigma^2_\epsilon]^2}$. This suggests that the trading profit of asset speculators is still higher than C_A and thus asset speculators will not exit the market.

Proofs of Proposition 6.4

First, we argue that in the equilibrium with CS sponsoring under this extended model, the factor liquidity trader must be also only trading CS product in equilibrium. This is because whenever the factor liquidity trader is trading certain underlying asset, competitive CS sponsoring at $t = 0$ (together with the dominance assumption) implies that a CS product consists of only this underlying asset with zero fee must be offered in equilibrium, which hence implies a feasible deviation for the factor liquidity trader.

As such, the trading problem solved by the factor liquidity trader can be formulated as the optimal choice of CS product design $\{w^H_k\}_k$ and trading volume y_H to minimize the expected trading loss in Eq. (24) subject to the factor exposure coverage constraint $y_H\sum_{k=1}^K w^H_k\beta_k = \tau$. By defining $y_H w^H_k \equiv y^H_k$, this optimization problem can be transformed into the following portfolio choice problem:

$$\min_{\{y^H_k\}_k} E\left[\sum_{k=1}^K y^H_k\lambda^{CS}_k\left(\ \hat{\kappa}_k\alpha_k+\textstyle\sum_{i\in I_k}\ \hat{\eta}_k s_i+\sum_{j\in I_{CS}}\eta_{CS,j}w^S_{k,j}s_j+n_k+y^H_k\ \right)\right],$$

subject to constraint

$$\sum_{k=1}^K\beta_k y^H_k = \tau$$

Since factor liquidity trader is uninformed of both γ and all α_k's, the objective of the factor liquidity trader can be reduced to $\min_{\{y^H_k\}_k}\sum_{k=1}^K\lambda^{CS}_k\left(y^H_k\right)^2$. This implies that the optimal portfolio $\{y^H_k\}_k$ chosen by the factor liquidity trader satisfy $y^H_1 : y^H_k = \frac{\beta_1}{\lambda^{CS}_1} : \frac{\beta_k}{\lambda^{CS}_k}$, for all $k = 2,\ldots,K$. Hence in equilibrium, the CS product traded by the factor liquidity trader has a weight design $\{w^H_k\}^K_{k=1}$ satisfying $w^H_1 : w^H_k = \frac{\beta_1}{\lambda^{CS}_1} : \frac{\beta_k}{\lambda^{CS}_k}$, which coincides with that chosen by factor speculators as suggested by Proposition 3.3.

Internet Appendices: Additional Analyses and Discussion

Appendix A. Characterization of Equilibrium with Endogenous Participation by Asset Speculators

In this section, we provide detailed analysis for our investigation of the equilibrium in which asset speculators make endogenous decisions regarding information acquisition and participation in asset market trading.

A1. Characterizing Equilibrium Without CS Sponsoring

We first characterize this equilibrium before the introduction of CS sponsoring. Compared to our analysis in Section 3, this analysis involves more possible cases for consideration as both factor speculators and asset speculators are making endogenous participation decisions, which in turn affects each other's equilibrium payoff. The following proposition fully characterizes all possible scenarios of equilibrium that could arise in such an economy.

Proposition A1.1. *Denote by N_k^* the number of factor speculators that trade asset k in equilibrium without CS trading. There are five possible cases in asset market k:*

(1) When $N_k^{C1} < N_k^{C2} < n_{k1} < n_{k2}$, there is one unique equilibrium. In the equilibrium, $N_k^* = N_k^{C1}$ and asset speculator participates asset market k.

(2) When $N_k^{C1} < n_{k1} < N_k^{C2} < n_{k2}$, two equilibria will exists. In the first equilibrium, $N_k^* = N_k^{C1}$ and asset speculator participates. In the second equilibrium, $N_k^* = N_k^{C2}$ and asset speculator does not participate asset market k.

(3) $n_{k1} < N_k^{C1} < N_k^{C2} < n_{k2}$, there is one unique equilibrium. In this equilibrium, $N_k^* = N_k^{C2}$ and asset speculator does not participate asset market k.

(4) $n_{k1} < N_k^{C1} < n_{k2} < N_k^{C2}$, there is no equilibrium.

(5) $n_{k1} < n_{k2} < N_k^{C1} < N_k^{C2}$, there is one unique equilibrium. In the equilibrium, $N_k^* = N_k^{C1}$ and asset speculator participates asset market k.

Here, N_k^{C1}, N_k^{C2}, n_{k1} and n_{k2} are functions of (β_k, σ_γ^2, σ_ϵ^2, $\sigma_{\alpha_k}^2$, $\sigma_{n_k}^2$). N_k^{C1} is always smaller than N_k^{C2}, and n_{k1} is always smaller than n_{k2}

Briefly, N_k^{C1} is the equilibrium number of factor speculators trading asset k when asset speculator participates asset market k, and N_k^{C2} is the equilibrium number of factor speculators trading asset k when asset speculator does participates asset market k. As we show in Proposition A1.1, N_k^{C1} is always smaller than N_k^{C2}. Intuitively, when asset speculator participates asset market k, adverse selection is more severe and price impact is higher, which decreases the incentive of factor speculators to trade asset k.

Proof of Proposition A1.1

The equilibrium is determined by whether asset speculators earn positive trading profits. In this proof, we take two steps. For each underlying assets, we first characterize the condition which ensures positive profits of asset and factor speculators. We then characterize the equilibrium.

Step 1: The condition which ensures positive profits of asset speculators is as follows:

$$\Pi_k^A = \frac{\sigma_{\alpha_k}^2 \sigma_{n_k}}{4\sqrt{\frac{\sigma_{\alpha_k}^2}{4} + \beta_k^2 \frac{N_k(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} > C_A. \tag{A-20}$$

To ensure the above inequality, N_k should satisfy $N_k < n_{k1}$ or $N_k > n_{k2}$, where

$$n_{k1} = \frac{b - \sqrt{b^2 - 4(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2}}{2} \text{ and } n_{k2} = \frac{b + \sqrt{b^2 - 4(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2}}{2}, \tag{A-21}$$

$$b = \frac{\beta_k^2(\sigma_\gamma^2 + \sigma_\epsilon^2)}{(\frac{\sigma_{\alpha_k}^2 \sigma_{n_k}}{4C_A})^2 - \frac{\sigma_{\alpha_k}^2}{4}} - 2\frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}. \tag{A-22}$$

We now characterize the condition which ensures the breaking-even of factor speculators. Denote by Π_k^F the trading profit of a factor speculator if asset speculator trades asset k. The condition is as follows:

$$\Pi_k^F = \frac{(\sigma_\gamma^2 + \sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{[(N_k+1)\sigma_\gamma^2 + 2\sigma_\epsilon^2]^2\sqrt{\frac{\sigma_{\alpha_k}^2}{4} + \beta_k^2 \frac{N_k(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} = C \tag{A-23}$$

Since Π_k^F is decreasing with N_k, there is one unique solution to the above equation and we denote the solution as N_k^{C1}.

If asset speculator does not trade asset k, the expected profit is denoted as $\Pi_{k,N}^F$. The condition is as follows:

$$\Pi_{k,N}^F = \frac{(\sigma_\gamma^2 + \sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{[(N_k+1)\sigma_\gamma^2 + 2\sigma_\epsilon^2]^2\sqrt{\beta_k^2 \frac{N_k(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} = C \tag{A-24}$$

Since Π_k^F is decreasing with N_k, there is one unique solution to the above equation and we denote the solution as N_k^{C2}. It is clear that $N_k^{C1} < N_k^{C2}$.

Step 2: We characterize the equilibrium, which are determined by n_{k1}, n_{k2}, N_k^{C1} and N_k^{C2}.

(1) When $N_k^{C1} < N_k^{C2} < n_{k1} < n_{k2}$, there ins one unique equilibrium where $N_k^* = N_k^{C1}$ and asset speculator trades asset k. We prove it by contradiction. Assuming that asset speculator does not trade asset k, the number of factor speculator should be N_k^{C2}. Since $N_k^{C2} < n_{k1}$, the asset speculator has incentive to deviate and participate the asset market k, which contradicts the assumption that asset speculator does not trade asset k. Thus, asset speculator will participate asset market k and $N_k^* = N_k^{C1}$.

(2) When $N_k^{C1} < n_{k1} < N_k^{C2} < n_{k2}$, two equilibria will exists. First, $N_k^* = N_k^{C1}$ and asset speculator trades asset k. Second, $N_k^* = N_k^{C2}$ and no asset speculator. It is straightforward that asset speculator will not deviate their participation decision in each equilibrium.

(3) When $n_{k1} < N_k^{C1} < N_k^{C2} < n_{k2}$, the only equilibrium is that asset speculator does not trade asset k. We prove it by contradiction. Assuming that asset speculator trades asset k, her profit is positive, which suggests that the number of factor speculator, N_k should be lower than n_{k1} or be higher than n_{k2}. Meanwhile, if N_k is lower than n_{k1}, we have $N_k < N_k^{C1}$ given $n_{k1} < N_k^{C1}$ and $\Pi_{k,N}^F(N_k)$ is positive. In this sense, factor speculators outside the markets will deviate from the equilibrium and participate asset market k, which contradicts the equilibrium.

If N_k is higher than n_{k2}, we have $N_k > N_k^{C2}$ given $N_k^{C2} < n_{k2}$ and $\Pi_{k,N}^F$ is negative. In this sense, factor speculators in the market will deviate from the equilibrium and does not trade asset k.

(4) When $n_{k1} < N_k^{C1} < n_{k2} < N_k^{C2}$, there are no equilibria. If the equilibrium is that asset speculator trades asset k, N_k^* should be N_k^{C1}. But because $N_k^{C1} < n_{k2}$, asset speculator has negative profit and has incentive to exit the market.

If the equilibrium is that asset speculator does not trade asset k, N_k^* should be N_k^{C2}. But because $N_k^{C2} > n_{k2}$, asset speculator has positive profit and has incentive to trade asset k.

(5) When $n_{k1} < n_{k2} < N_k^{C1} < N_k^{C2}$, the equilibrium is $N_k^* = N_k^{C1}$ and asset speculator trades asset k. It is straightforward that asset speculator will not deviate their participation decision in the equilibrium.

A2. Impact of CS Trading

This section provides detailed analysis for the effect of introducing CS sponsoring on the price impact in underlying asset markets, paralleling our analysis in Section 4.2 where asset speculators are exogenously assumed to always participate in trading.

Lemma A1.1. *When all underlying assets have the same σ_α^2 and $\sigma_{n_k}^2$, the effect of CS introduction on price impact is as follows:*
(i) If $N_{CS} < \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$, introducing CS increases price impact on all underlying assets.
(ii) If $N_{CS} > \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$, introducing CS has heterogeneous effects on price impact on underlying assets. That is, there exists a cut-off value β^*, such that it increases price impacts of low-beta assets (e.g., assets with β_k below β^*) and decreases price impacts of high-beta assets (e.g., assets with β_k above β^*).

Proof of Lemma A1.1

We use two steps to prove this lemma. In the first step, we prove the results for $N_{CS} < \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$. In the second step, we prove the results for $N_{CS} > \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$. In the proof, we mainly use a proof by contradiction and focus on Asset 1 for illustration.

Step 1: $N_{CS} < \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$

There are two potential cases: (1) asset speculator trades asset 1 before CS trading; (2) asset speculator does not trade asset 1 before CS trading.

In Case (1), assuming that introducing CS decreases price impact, asset speculator still trades asset 1 because she can get better trading profit due to decreased price impact. However, since $N_{CS} < \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$, when the number of factor speculators increases from N_1 to N_{CS}, the price impact increases. This contradicts the assumption about decreased price impact.

In Case (2), assuming that introducing CS decreases price impact, the asset speculator may or may not trade asset 1. Since $N_{CS} < \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$, when the number of factor speculators increases from N_1 to N_{CS}, the price impact increases even if asset speculator does not trade asset 1. This contradicts the assumption about decreased price impact. We use similar steps to prove that introducing CS increases price impact in this case.

Step 2: $N_{CS} > \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$

We first prove that introducing CS will increase price impact on assets with $N_k \cdot N_{CS}^* < (\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2$ but will decrease price impact on assets with $N_k \cdot N_{CS}^* > (\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2$.

For assets with $N_k \cdot N_{CS}^* < (\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2$, we prove it by contradiction. Assuming that the price impact decreases after CS introduction, the number of asset speculators weakly increases since asset speculators can earn higher trading profits due to lower price impact. But when $N_k \cdot N_{CS}^* < (\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2$, in the expression of price impact, we always have

$$\frac{N_k(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2} < \frac{N_{CS}^*(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}^*+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}. \tag{A-25}$$

Combining the weakly increasing participation of asset speculators, price impact always increases after CS introduction. This contradicts the assumption about decreased price impact. Thus, we can conclude that the price impact increases after CS trading for assets with $N_k \cdot N_{CS}^* < (\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2$.

For assets with $N_k \cdot N_{CS}^* > (\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2$, we also prove it by contradiction. Assuming that price impact increases after CS introduction, the number of asset speculators weakly decreases since asset speculators can earn lower trading profits due to higher price impact. But when $N_k \cdot N_{CS}^* > (\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2$, in the expression of price impact, we always have

$$\frac{N_k(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2} > \frac{N_{CS}^*(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}^*+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}. \tag{A-26}$$

Combining the weakly decreasing participation of asset speculators, price impact always decreases after CS introduction. This contradicts the assumption about increased price impact. Thus, we can conclude that the price impact decreases after CS trading for assets with $N_k \cdot N_{CS}^* > (\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})^2$.

Finally, based on Proposition 3.1 that N_K is increasing with β_k, there exists a cutoff β^*

such that introducing CS will increase price impact on assets with β_k lower than β^* but will decrease price impact on assets with β_k higher than β^*. ■

Lemma A1.2. *When all underlying assets have the same β and $\sigma^2_{n_k}$, the effect of CS introduction on price impact is as follows:*
(i) If $N_{CS} < \frac{\sigma^2_\gamma + 2\sigma^2_\epsilon}{\sigma^2_\gamma}$, introducing CS increases price impact for all underlying asset market.
(ii) If $N_{CS} > \frac{\sigma^2_\gamma + 2\sigma^2_\epsilon}{\sigma^2_\gamma}$, introducing CS has heterogeneous effects on the price impact in the underlying asset market. That is, there exists a cut-off value σ^*, such that it decreases price impacts of low-σ^2_α assets (e.g., assets with $\sigma^2_{\alpha_k}$ below σ^{*2}) and increases price impacts of high-σ^2_α (e.g., assets with $\sigma^2_{\alpha_k}$ above σ^{*2}).

Proof of Lemma A1.2

This proof of this lemma is similar to the proof of Lemma A1.1 but only uses the results in Proposition 3.1 that N_K is decreasing with $\sigma^2_{\alpha_k}$. ■

Appendix B. Alternative Specifications of the CS Offering and Trading Game

We relax the dominance concept (ii) in Section 3 in one aspect: CS sponsors provide only one CS product to maximize the expected trading profits of factor speculators. As we will show in the following analysis, the equilibrium in this economy is equivalent to the economy in Section 3.

We take the following two steps. In the first step, we show that the CS speculators' effective trading aggressive, asset prices, and trading profits of factor speculators in this economy are the same as those in Proposition 3.2. In the second step, we characterize the condition that all factor speculators trade CS products and obtain the results in Proposition 3.3.

In the first step, we show that taking the number of factor speculators $\hat{N}_k$ $(k = 1, ..., K)$ and $\hat{\eta}_k$ as given in Proposition 3.2, the CS sponsors design only one CS product to maximize the expected trading profits of CS speculators.

Taking the number of factor speculators $\hat{N}_k$ $(k = 1, ..., K)$ and $\hat{\eta}_k$ as given in Proposition 3.2, we solve the trading strategy of CS traders. Specifically, the jth factor speculators that trade CS choose $y_{CS,j}(= \beta_{CS,j} s_{CS,j})$ to maximize her expected trading profits. The optimization is as follows:

$$\max_{y_{CS,j}} E\left[\sum_{k=1}^{K} y_{CS,j} w_k \left(\beta_k \gamma - \lambda_k^{CS} \left(\begin{array}{c} \hat{\kappa}_k \alpha_k + \sum_{i \in I_k} \hat{\eta}_k s_{k,i} + n_k \\ \sum_{i \in J \text{ and } i \notin j} y_{CS,i} w_k s_{CS,i} + y_{CS,j} w_k \end{array} \right)\right) \middle| s_{CS,j} \right]. \quad \text{(A-27)}$$

FOC with $y_{CS,j}$ yields:

$$y_{CS,j} = \frac{\sum_{k=1}^{K} w_k \left[\beta_k - \lambda_k^{CS}(\sum_{i\in I_k} \ \hat{\eta}_k + \sum_{i\in J \text{ and } i\notin j} y_{CS,i}w_k)\right]}{2(\sum_{k=1}^{K} \lambda_k^{CS} w_k^2)} \frac{\sigma_\gamma^2}{\sigma_\gamma^2+\sigma_\epsilon^2} s_{CS,j}. \tag{A-28}$$

Give the symmetry among CS traders ($\beta_{CS,j} = \beta_{CS,i}$ for $i \neq j$), we can get:

$$\beta_{CS,j} = \frac{\sum_{k=1}^{K} w_k \left[\beta_k - \lambda_k^{CS} \sum_{i\in I_k} \ \hat{\eta}_k\right]}{(N_{CS}+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})(\sum_{k=1}^{K} \lambda_k^{CS} w_k^2)}. \tag{A-29}$$

Inserting the expression of $\beta_{CS,j}$ in the expected trading profit yields:

$$\Pi_{CS}^F = \frac{(1+\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})^2}{(N_{CS}+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})^2} \frac{\left(\sum_{k=1}^{K} w_k \left[\beta_k - \lambda_k^{CS}(\sum_{i\in I_k} \ \hat{\eta}_k)\right]\right)^2}{(\sum_{k=1}^{K} \lambda_k^{CS} w_k^2)} \frac{\sigma_\gamma^4}{\sigma_\gamma^2+\sigma_\epsilon^2}. \tag{A-30}$$

Since the CS sponsor market is competive, CS sponsors choose $(w_1, ..., w_K)$ to maximze Π_{CS}^F. Otherwise, there always exists another CS sponsor who will enter the CS sponsoring market and provide better CS products to factor speculators. The CS sponsor's optimization problem is summarized as:

$$\max_{\{w_k\}} \frac{(1+\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})^2}{(N_{CS}+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})^2} \frac{\left(\sum_{k=1}^{K} w_k \left[\beta_k - \lambda_k^{CS}(\sum_{i\in I_k} \ \hat{\eta}_k)\right]\right)^2}{(\sum_{k=1}^{K} \lambda_k^{CS} w_k^2)} \frac{\sigma_\gamma^4}{\sigma_\gamma^2+\sigma_\epsilon^2}$$

$$\text{subject to : } \sum_{k=1}^{K} w_k = 1.$$

Using the Lagrange approach and assuming L is the Lagrange multiplier, FOC with w_k yields:

$$\frac{d\Pi_{CS}^F}{dw_k} - L = 0,$$

where

$$\begin{aligned} \frac{d\Pi_{CS}^F}{dw_k} \ = \ & \frac{(1+\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})^2}{(N_{CS}+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})^2} \frac{2\left(\sum_{k=1}^{K} w_{k,j} \left[\beta_k - \lambda_k^{CS}(\sum_{i\in I_k} \ \hat{\eta}_k)\right]\right)\left[\beta_k - \lambda_k^{CS}(\sum_{i\in I_k} \ \hat{\eta}_k)\right]}{(\sum_{k=1}^{K} \lambda_k^{CS} w_{k,j}^2)} \\ & - \frac{(1+\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})^2}{(N_{CS}+1+2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})^2} \frac{2\left(\sum_{k=1}^{K} w_{k,j} \left[\beta_k - \lambda_k^{CS}(\sum_{i\in I_k} \ \hat{\eta}_k)\right]\right)^2 \lambda_k^{CS} w_{k,j}}{(\sum_{k=1}^{K} \lambda_k^{CS} w_{k,j}^2)^2}. \end{aligned}$$

This suggests that in the equilibrium, $\frac{d\Pi^F_{CS}}{dw_k}$ should be the same across weights over different assets. In the following proposition, we show that the equilibrium CS product design (w_k) and the effective trading aggressive of CS traders ($\hat{\eta}_{CS,k}$) are the same as those in Section 3 (Proposition 3.2).

Proposition A1.2. *Taking $\hat{N}_k$,$\hat{\eta}_k$, N_{CS} and λ^{CS}_k as given, the optimal design of CS product is as follows:*

$$w_k : w_l = \hat{\eta}_{CS,k} : \hat{\eta}_{CS,l}, \tag{A-31}$$

where $\hat{\eta}_{CS,k} = \frac{\beta_k}{\lambda^{CS}_k\left(\hat{N}_k+N_{CS}+1+2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma}\right)}$. Meanwhile, the effective trading aggressive of jth CS speculator in asset k ($y_{CS,j}w_k$) is:

$$y_{CS,j}w_k = \frac{\beta_k}{\lambda^{CS}_k(\hat{N}_k + N_{CS} + 1 + 2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma})}. \tag{A-32}$$

Proof: Inserting $\hat{\eta}_k = \frac{\beta_k}{\lambda^{CS}_k\left(\hat{N}_k+N_{CS}+1+2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma}\right)}$ (from Proposition 3.2) into the expression of $\frac{d\Pi^F_{CS}}{dw_k}$, we have :

$$\beta_k - \lambda^{CS}_k(\sum_{i\in I_k}\hat{\eta}_k) = \frac{(N_{CS}+1+2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma})\beta_k}{\hat{N}_k + N_{CS} + 1 + 2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma}},$$

$$\begin{aligned}&\left(\sum_{k=1}^{K} w_{k,j}\left[\beta_k - \lambda^{CS}_k(\sum_{i\in I_k}\hat{\eta}_k)\right]\right)\lambda^{CS}_k w_{k,j}\\ = \;& \frac{1}{(\sum_{k=1}^{K}\hat{\eta}_{CS,k})^2}\frac{(N_{CS}+1+2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma})\beta_k}{\hat{N}_k + N_{CS} + 1 + 2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma}}\left(\sum_{k=1}^{K}\frac{\beta^2_k}{\lambda^{CS}_k(\hat{N}_k + N_{CS} + 1 + 2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma})^2}\right).\end{aligned}$$

$$\sum_{k=1}^{K}\lambda^{CS}_k w^2_{k,j} = \frac{1}{(\sum_{k=1}^{K}\hat{\eta}_{CS,k})^2}\sum_{k=1}^{K}\frac{\beta^2_k}{\lambda^{CS}_k(\hat{N}_k + N_{CS} + 1 + 2\frac{\sigma^2_\epsilon}{\sigma^2_\gamma})^2}.$$

Inserting the above equations into $\frac{d\Pi^F_{CS}}{dw_k}$, we can get $\frac{d\Pi^F_{CS}}{dw_k} = 0$. This suggests that the portfolio weights in this proposition satisfy the first-order condition of optimal CS product designs.

Given the expressions of portfolio weight, the effective trading aggressive of jth CS speculator in asset k ($y_{CS,j} w_k$) is:

$$\begin{aligned} y_{CS,j} w_k &= w_k \frac{\sum_{k=1}^{K} w_k \left[\beta_k - \lambda_k^{CS} \sum_{i \in I_k} \hat{\eta}_k\right]}{(N_{CS} + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})(\sum_{k=1}^{K} \lambda_k^{CS} w_k^2)} \\ &= \frac{1}{\lambda_k^{CS}(N_{CS} + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})} \frac{(N_{CS} + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})\beta_k}{\hat{N}_k + N_{CS} + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2}} \\ &= \frac{\beta_k}{\lambda_k^{CS}(\hat{N}_k + N_{CS} + 1 + 2\frac{\sigma_\epsilon^2}{\sigma_\gamma^2})}. \end{aligned}$$

Proposition A1.3. *When* $\min(N_1, N_2, ...N_k) > \frac{\sigma_\gamma^2 + 2\sigma_\epsilon^2}{\sigma_\gamma^2}$ *and* $\widehat{C}$ *is sufficiently small, factor speculators only trade CS.*

Proof: We take three steps to prove this proposition.

Step 1: we prove that, at most, only one asset has factor speculators directly trading on it. To simplify the analysis, we first denote the profits for factor speculators trading on the asset and CS is as follows:

$$\Pi_k^F = \frac{\beta_k^2 \sigma_\gamma^2}{(\hat{N}_k + N_{CS} + 1)^2 \lambda_k^{CS}}$$

and

$$\Pi_{CS}^F = \sum_{k=1}^{K} \frac{\beta_k^2 \sigma_\gamma^2}{(\hat{N}_k + N_{CS} + 1)^2 \lambda_k^{CS}} = \sum_{k=1}^{K} \Pi_k^F$$

In the equilibrium, the profit of factor speculators trading on the asset and CS is a function of $\hat{N}_k$ and N_{CS}. That is, we have $\Pi_k^F(N_k, N_{CS})$ and $\Pi_{CS}^F(N_1, ..., N_K, N_{CS})$. In the equilibrium, we should have

$$\Pi_{CS}^F(\hat{N}_1, ..., \hat{N}_k, N_{CS}) = C + F > 0, \text{ where } F = \frac{K\widehat{C}}{N_{CS}}$$

$$\Pi_{CS}^F(\hat{N}_1, ..., \hat{N}_k, , N_{CS} + 1) - C - \frac{K\widehat{C}}{N_{CS} + 1} < 0$$

$$\Pi_k^F(\hat{N}_k, N_{CS}) - C = 0$$

$$\Pi_k^F(\hat{N}_k + 1, N_{CS}) - C < 0.$$

We show that, at most, only one asset has factor speculators directly trading on it by contradiction. Assuming there exist M assets having a positive number of factor speculators. Sorting these assets by the number of factor speculators on them. Without loss of generality, we assume that $N_1 > N_2 > N_3.... > N_M$. If this is one equilibrium, we always have

$$\Pi_i^F(N_i, N_{CS}) - C = 0 \text{ for } i = 1, ..., M.$$

Now consider another equilibrium, let $N_{CS}^{'} = N_{CS} + N_2$,$N_1^{'} = N_1 - N_2$ and $N_i^{'} = 0$ for $i = 2, ..., M$ we still have the following:

$\Pi_1^F(N_1 - N_2, N_{CS} + N_2)$ does not change, and it still equals C.

But

$$\begin{aligned}
\Pi_{CS}^F &= \sum_{k=1}^{K} \Pi_k^F(N_1^{'}, ..., N_K^{'}, N_{CS}^{'}) - C - \frac{K\widehat{C}}{N_{CS} + N_2} \\
&> \Pi_1^F(N_1, N_{CS}) + \Pi_2^F(N_i, N_{CS}) - C - \frac{K\widehat{C}}{N_{CS}} \\
&> C - \frac{K\widehat{C}}{N_{CS}} \\
&> 0
\end{aligned}$$

This suggests that the trading profits of CS speculators can further increases and thus cannot be an equilibrium.

Step 2: We characterize the condition under which all factor speculators trade CS in this step. We first show that the equilibrium N_{CS} is unique and decreases with $\widehat{C}$. We discuss two cases: $N_{CS} \leq N_1$ and $N_{CS} > N_1$ (noted: N_1 is the equilibrium number of factor speculators in Asset 1 in Proposition 3.1 without CS.)

Case 1: $N_{CS} \leq N_1$

If the equilibrium is $N_{CS} \leq N_1$, for the factor speculators that only trade asset 1, her net profit is zero $\frac{\beta_k^2\sigma_\gamma^2}{(\hat{N}_1+N_{CS}+1)^2\lambda_k} = C$. For CS traders, her net profit is

$$\begin{aligned}
\Pi_{CS}^F &= \sum_{k=1}^{K} \Pi_k^F - C - \frac{K\widehat{C}}{N_{CS}} \\
&= \sum_{k=2}^{K} \frac{(\sigma_\gamma^2 + \sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{[(N_{CS}+1)\sigma_\gamma^2 + 2\sigma_\epsilon^2]^2\sqrt{\frac{\sigma_{\alpha_k}^2}{4} + \beta_k^2\frac{(N_k+N_{CS})(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_k+N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}} - \frac{K\widehat{C}}{N_{CS}} \\
&= \frac{1}{N_{CS}}(f - K\widehat{C})
\end{aligned}$$

where: $f(N_{CS}) = \sum_{k=2}^{K} \frac{N_{CS}(\sigma_\gamma^2+\sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2\sqrt{\frac{\sigma_{\alpha_k}^2}{4}+\beta_k^2\frac{(N_{CS})(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}}$

We prove that $f - K\widehat{C} = 0$ has one unique solution when $N_{CS} > N_2$.

Rearranging f yields:

$$f=\sum_{k=2}^{K}\frac{(\sigma_\gamma^2+\sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{\sqrt{\frac{\sigma_{\alpha_k}^2}{4}(\frac{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}{N_{CS}})^2+\beta_k^2(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4\frac{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}{N_{CS}}}}$$

Let $f_i=\frac{1}{\sqrt{\frac{\sigma_{\alpha_k}^2}{4}(\frac{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}{N_{CS}})^2+\beta_k^2(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4\frac{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}{N_{CS}}}}$ and $l=\frac{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}{N_{CS}}$

We have $\frac{df_i}{dN_{CS}}=-\frac{1}{2}m_i^{-3/2}(\frac{\sigma_{\alpha_k}^2}{2}l+\beta_k^2(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4)(\sigma_\gamma^4-\frac{(\sigma_\gamma^2+2\sigma_\epsilon^2)^2}{N_{CS}^2})$, which suggests that $\frac{df}{dN_{CS}}=-\frac{1}{2}\{\sum_{k=2}^{K}\frac{(\sigma_\gamma^2+\sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{2m_i^{3/2}}(\frac{\sigma_{\alpha_k}^2}{2}l+\beta_k^2(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4)\}(\sigma_\gamma^4-\frac{(\sigma_\gamma^2+2\sigma_\epsilon^2)^2}{N_{CS}^2})$, where $m_i=\frac{\sigma_{\alpha_k}^2}{4}(\frac{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}{N_{CS}})^2+\beta_k^2(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4\frac{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}{N_{CS}}$.

This means that $\frac{df}{dN_{CS}}<0$ only if $N_{CS}>\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$. In fact, under the assumption that $\min(N_1, N_2, ...N_k)>\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$, we have $N_{CS}>\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$. This implies that f is decreasing with N_{CS}. In this sense, to make the equality $f(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})-K\widehat{C}=0$ holds, N_{CS} is decreasing with $\widehat{C}$. Meanwhile, as $N_2>\min(N_1, N_2, ...N_k)>\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$, we can get that $f(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2})-K\widehat{C}>0$. Since $f(N_{CS})-K\widehat{C}$ becomes negatives when N_{CS} goes to infinity, according to the intermediate value theorem, there exists one unique solution N_{CS}^*.

Case 2: $N_{CS}>N_1$

If the equilibrium is $N_{CS}>N_1$, there are no factor speculators who only trade Asset 1 as their net profit becomes lower than zero when more factor speculators trade on Asset 1 now. For CS traders, their net profit is:

$$\begin{aligned}\Pi_{CS}^F &= \sum_{k=1}^{K}\Pi_k^F-C-\frac{K\widehat{C}}{N_{CS}}\\ &= \sum_{k=1}^{K}\frac{(\sigma_\gamma^2+\sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2\sqrt{\frac{\sigma_{\alpha_k}^2}{4}+\beta_k^2\frac{(N_{CS})(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4}{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}}}-C-\frac{K\widehat{C}}{N_{CS}}\\ &= \frac{1}{N_{CS}}(g-K\widehat{C})-C\end{aligned}$$

Now we let

$$g=\sum_{k=1}^{K}\frac{(\sigma_\gamma^2+\sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{\sqrt{\frac{\sigma_{\alpha_k}^2}{4}(\frac{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}{N_{CS}})^2+\beta_k^2(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4\frac{[(N_{CS}+1)\sigma_\gamma^2+2\sigma_\epsilon^2]^2}{N_{CS}}}}$$

Now $\frac{dg}{dN_{CS}}=-\frac{1}{2}\{\sum_{k=1}^{K}\frac{(\sigma_\gamma^2+\sigma_\epsilon^2)\beta_k^2\sigma_\gamma^4\sigma_{n_k}}{2m_i^{3/2}}(\frac{\sigma_{\alpha_k}^2}{2}l+\beta_k^2(\sigma_\gamma^2+\sigma_\epsilon^2)\sigma_\gamma^4)\}(\sigma_\gamma^4-\frac{(\sigma_\gamma^2+2\sigma_\epsilon^2)^2}{N_{CS}^2})$.

Again, as we can see, $\frac{dg}{dN_{CS}}<0$ only if $N_{CS}>\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$. In fact, under the assumption that

$\min(N_1, N_2, ...N_k) > \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$, we have $N_{CS} > \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$. This implies that g decreases with N_{CS}. In this sense, to make the equality $g(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}) - K\widehat{C} = 0$ holds, N_{CS} is decreasing with $\widehat{C}$. Meanwhile, as $N_2 > \min(N_1, N_2, ...N_k) > \frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}$, we can get that $g(\frac{\sigma_\gamma^2+2\sigma_\epsilon^2}{\sigma_\gamma^2}) - K\widehat{C} > 0$. Since $g(N_{CS}) - K\widehat{C}$ becomes negatives when N_{CS} goes to infinity, according to the intermediate value theorem, there exists one unique solution N_{CS}^*.

Step 3: we show that when $\widehat{C}$ is sufficiently small, Case 1 in Step 2 does not exist, suggesting that factor speculators choose to trade only CS. In fact, since N_{CS} is decreasing with $\widehat{C}$ and N_{CS} goes to infinity when $\widehat{C}$ becomes to zero, there is a cutoff $\widehat{C^F}$ satisfying $N_{CS} > N_1$ when $\widehat{C} < \widehat{C^F}$. Meanwhile, as we show in Step 2, even if when $\widehat{C} < \widehat{C^F}$, there is one unique solution of N_{CS} in the equilibirum.

Appendix C. CS Trading Transparency

CS products have potential heterogeneity in whether market makers can observe order flows from CSs and can distinguish order flows from CSs and order flows from assets. While passive mutual funds only reveal the portfolio and order flow at monthly frequency at best, the shares outstanding and the weights of ETFs are all available at daily frequency, if not higher. Moreover, the ETF arbitrage process also makes authorized participants and fund sponsors visible to the market makers. It is interesting to study the impact of such trading transparency. We model this distinguishing feature of trading transparency by allowing market makers to observe perfectly the order flow from CS sponsors.

Again, we focus on linear equilibria. With the asset speculator adopting strategy $\hat{\kappa}_k \alpha_k$ and factor speculator j adopting effective trading strategy $\eta_{CS,k} s_j$ in Asset k, the market maker in Asset k observes the total order flow:

$$\omega_k = \hat{\kappa}_k \alpha_k + \sum_{j \in I_{CS}} \eta_{CS,k} s_j + n_k + \tau_k$$

and the total order flow via CSs

$$m_k = \sum_{j \in I_{CS}} \eta_{CS,k} s_j + \tau_k,$$

where τ_k is the effective trading volume of the factor liquidity trader in Asset k.

With this information set, the market maker sets the price

$$P_k = p_k(\omega_k, m_k) \equiv \bar{v} + \lambda_{1,k}\omega_k + \lambda_{2,k} m_k. \tag{A-33}$$

It can be shown that

$$\lambda_{1,k} \equiv \frac{\hat{\kappa}_k \sigma_{\alpha_k}^2}{\hat{\kappa}_k^2 \sigma_{\alpha_k}^2 + \sigma_{n_k}^2}, \quad \lambda_{2,k} \equiv \frac{N_{CS}\eta_{CS,k}\beta_k\sigma_\gamma^2}{\eta_{CS,k}^2\left(N_{CS}^2\sigma_\gamma^2 + N_{CS}\sigma_\epsilon^2\right) + \sigma_{\tau_k}^2} - \frac{\hat{\kappa}_k \sigma_{\alpha_k}^2}{\hat{\kappa}_k^2 \sigma_{\alpha_k}^2 + \sigma_{n_k}^2}. \tag{A-34}$$

Given this pricing rule of market makers, the asset speculator in Asset k solves

$$\max_{x_k} E\left[x_k\left(\alpha_k + \beta_k\gamma - \lambda_{1,k}(x_k + m_k + n_k) - \lambda_{2,k}m_k\right)\middle|\, \alpha_k\right],$$

which implies:

$$x_k = \frac{\alpha_k}{2\lambda_{1,k}} \quad \Rightarrow \quad \hat{\kappa}_k = \frac{1}{2\lambda_{1,k}}.$$

While asset speculators' profit maximization stays almost the same, it becomes structurally different for factor speculators who trade using CSs in equilibrium. After observing signal $s_j = \gamma + \epsilon_j$, a factor speculator optimally chooses the CS product $\{w_{k,j}\}_{k=1}^K$ to trade and the trading volume y_j. Following our analysis in Section 3.2, by defining $y_{k,j} \equiv w_{k,j} \cdot y_j$, the factor speculator's optimization problem can be formulated as:

$$\max_{\{y_{j,k}\}_k} E\left[\sum_{k=1}^K y_{j,k}\left(\beta_k\gamma - \lambda_{1,k}\left(\begin{array}{c}\hat{\kappa}_k\alpha_k + n_k + y_{j,k} + \tau_k \\ \sum_{j'\in I_{CS} \text{ and } j'\neq j}\eta_{j',k}s_{j'}\end{array}\right) - \lambda_{2,k}\left(\sum_{j'\in I_{CS} \text{ and } j'\neq j}\eta_{j',k}s_{j'} + y_{j,k} + \tau_k\right)\right)\middle|\, s_j\right],$$

which implies

$$\eta_{CS,k} = \frac{\beta_k}{(N_{CS}+1)(\lambda_{1,k}+\lambda_{2,k})}\frac{\sigma_\gamma^2}{\sigma_\gamma^2+\sigma_\epsilon^2}.$$

The following proposition characterizes the equilibrium with transparent CS trading.

Proposition A1.4. *In the equilibrium with transparent CS trading, the (effective) trading aggressiveness of the asset speculator in Asset k and factor speculators in CS trading are*

$$\hat{\kappa}_k = \frac{\sigma_{n_k}}{\sigma_{\alpha_k}}, \quad \eta_{CS,k} = \sqrt{\frac{\sigma_{\tau_k}^2}{N_{CS}\sigma_\gamma^2 + N_{CS}^2\sigma_\epsilon^2}},$$

respectively, and the design of CS product is irrelevant. The equilibrium trading profit of asset speculator in Asset k and factor speculators in CS trading are:

$$\Pi_k = \frac{\sigma_{\alpha_k}\sigma_{n_k}}{2}, \quad \Pi_{CS}^F = \sum_{k=1}^K \beta_k \frac{\sigma_{\tau_k}\sigma_\gamma^2}{\sqrt{N_{CS}\sigma_\gamma^2 + N_{CS}^2\sigma_\epsilon^2}}\left(\frac{\sigma_\gamma^2}{(N_{CS}+1)(\sigma_\gamma^2+\sigma_\epsilon^2)} + \frac{\sigma_\epsilon^2}{\sigma_\gamma^2+\sigma_\epsilon^2}\right).$$

Notably, with transparent CS trading in some ETF markets, the equilibrium features an endogenous segmentation between asset-specific speculation and factor speculation. In particular, the equilibrium trading aggressiveness and expected profit of factor speculators are unaffected by or related to those of the asset speculator in each underlying asset. This contrasts sharply with our baseline analysis where CS trading volume is unobservable to market makers of underlying assets. There, as suggested by our analysis in Section 3 (Lemma 3.1 and Proposition 3.2), the equilibrium trading strategy and profit making of asset speculators and factor speculators are intertwined with each other. Relatedly, unlike in our analysis in Section 6.1 where introducing CS sponsoring could affect the information acquisition decisions by asset-specific speculators, such impact will be absent when CS trading volume can be transparently observed to market participants.

While we do not formally study the optimal CS product designs, it is intuitive that the optimal portfolio weights depend on $\lambda_{2,k}$ as factor speculators only care about this specific price impact and their trading strategy is related only to this price impact. Since the analysis of this potential extension largely deviates from our focus, we do not push further but leave it for further study.

Proof of Proposition A1.4

We start with characterizing the equilibrium pricing rule of market makers. For the market maker in Asset k who observes the total order flow:

$$\omega_k = \hat{\kappa}_k \alpha_k + \sum_{j \in I_{CS}} \eta_{CS,k} s_j + n_k + \tau_k$$

and the total order flow via CSs

$$m_k = \sum_{j \in I_{CS}} \eta_{CS,k} s_j + \tau_k,$$

she sets the price of Asset k according to

$$\begin{aligned}
P_k &= E\left[\bar{v}_k + \beta_k \gamma + \alpha_k | \omega_k = \hat{\kappa}_k \alpha_k + \sum_{j \in I_{CS}} \eta_{CS,k} s_j + n_k + \tau_k, m_k = \sum_{j \in I_{CS}} \eta_{CS,k} s_j + \tau_k\right] \\
&= \bar{v}_k + \beta_k E\left[\gamma | m_k = \sum_{j \in I_{CS}} \eta_{CS,k} s_j + \tau_k\right] + E\left[\alpha_k | \omega_k - m_k = \hat{\kappa}_k \alpha_k + n_k\right] \\
&= \bar{v}_k + \frac{N_{CS} \eta_{CS,k} \beta_k \sigma_\gamma^2}{\eta_{CS,k}^2 \left(N_{CS}^2 \sigma_\gamma^2 + N_{CS} \sigma_\epsilon^2\right) + \sigma_{\tau_k}^2} \cdot m_k + \frac{\hat{\kappa}_k \sigma_{\alpha_k}^2}{\hat{\kappa}_k^2 \sigma_{\alpha_k}^2 + \sigma_{n_k}^2} \cdot (\omega_k - m_k) \\
&\equiv \bar{v} + \lambda_{1,k} \omega_k + \lambda_{2,k} m_k,
\end{aligned}$$

where $\lambda_{1,k}$ and $\lambda_{2,k}$ are as given in Eq. (A-34).

With market maker adopting the pricing rule $P_k = p_k(\omega_k, m_k) \equiv \bar{v} + \lambda_{1,k} \omega_k + \lambda_{2,k} m_k$, the asset speculator in Asset k who observes α_k solves

$$\max_{x_k} E\left[x_k \left(\alpha_k + \beta_k \gamma - \lambda_{1,k}(x_k + m_k + n_k) - \lambda_{2,k} m_k\right) | \, \alpha_k\right],$$

which implies

$$x_k = \frac{\alpha_k}{2\lambda_{1,k}} \quad \Rightarrow \quad \hat{\kappa}_k = \frac{1}{2\lambda_{1,k}}.$$

After observing signal $s_j = \gamma + \epsilon_j$, a factor speculator optimally chooses the CS product $\{w_{k,j}\}_{k=1}^{K}$ to trade and the trading volume y_j. Following our analysis in Section 3.2, by defining $y_{k,j} \equiv w_{k,j} \cdot y_j$, the factor speculator's optimization problem can be formulated as

$$\max_{\{y_{j,k}\}_k} E\left[\sum_{k=1}^{K} y_{j,k}\left(\beta_k\gamma - \lambda_{1,k}\left(\begin{array}{c}\hat{\kappa}_k\alpha_k + n_k + y_{j,k} + \tau_k \\ \sum_{j'\in I_{CS} \text{ and } j'\neq j}\eta_{j',k}s_{j'}\end{array}\right) - \lambda_{2,k}\left(\sum_{j'\in I_{CS} \text{ and } j'\neq j}\eta_{j',k}s_{j'} + y_{j,k} + \tau_k\right)\right)\middle| s_j\right],$$

which implies

$$\eta_k = \frac{\beta_k}{(N_{CS}+1)(\lambda_{1,k}+\lambda_{2,k})}\frac{\sigma_\gamma^2}{\sigma_\gamma^2+\sigma_\epsilon^2}.$$

Plug in the expressions of $\lambda_{1,k}$ and $\lambda_{2,k}$ are as in Eq. (A-34), we get

$$\eta_{CS,k} = \sqrt{\frac{\sigma_{\tau_k}^2}{N_{CS}\sigma_\gamma^2 + N_{CS}^2\sigma_\epsilon^2}}$$

As such, the equilibrium trading profit for the asset speculator of Asset k is

$$\begin{aligned}\Pi_k &= E\left[\hat{\kappa}_k\alpha_k\left(\alpha_k - \frac{1}{2\hat{\kappa}_k}\hat{\kappa}_k\alpha_k\right)\right] \\ &= E\left[\frac{1}{2}\hat{\kappa}_k\alpha_k^2\right] \\ &= \frac{\sigma_{n_k}\sigma_{\alpha_k}}{2},\end{aligned}$$

and the equilibrium trading profit for the factor speculator in CS trading is

$$\begin{aligned}\Pi_{CS}^F &= E\left[\sum_{k=1}^{K}\eta_{CS,k}\gamma\left(\beta_k\gamma - (\lambda_{1,k}+\lambda_{2,k})N_{CS}\eta_{CS,k}\gamma\right)\right] \\ &= E\left[\sum_{k=1}^{K}\eta_{CS,k}\gamma\left(\beta_k\gamma - \frac{\beta_k N_{CS}}{N_{CS}+1}\frac{\sigma_\gamma^2}{\sigma_\gamma^2+\sigma_\epsilon^2}\gamma\right)\right] \\ &= \sum_{k=1}^{K}\beta_k\frac{\sigma_{\tau_k}\sigma_\gamma^2}{\sqrt{N_{CS}\sigma_\gamma^2+N_{CS}^2\sigma_\epsilon^2}}\left(\frac{\sigma_\gamma^2}{(N_{CS}+1)(\sigma_\gamma^2+\sigma_\epsilon^2)} + \frac{\sigma_\epsilon^2}{\sigma_\gamma^2+\sigma_\epsilon^2}\right)\end{aligned}$$