Three Essays on Information Economics and Delegated Portfolio Management

by

George Fedorovich Malikov

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in The University of Michigan
2021

Doctoral Committee:

Professor Paolo Pasquariello, Chair
Associate Professor David A. Miller
Professor Gregory Smith Miller
Professor Amiyatosh Kumar Purnanandam
Professor Uday Rajan

George Malikov

gmalikov@umich.edu

ORCID iD: 0000-0001-5310-5541

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor and mentor Paolo Pasquariello for his unwavering support, encouragement, and guidance throughout the years. From sparking my curiosity in the subject matter that ultimately laid the foundation for this dissertation, to guiding me through the ups-and-downs of research and the academic career, Paolo was always there. His warmth, enthusiasm, and pragmatic advice were invaluable to my progress through the PhD program. I will always fondly remember our lively research discussions in his office, the vibrant atmosphere of his PhD and MBA courses, and the healthy competitiveness that he brought as a goalkeeper to the local recreational soccer league.

I am grateful to Uday Rajan for being incredibly generous with his time in talking through my job market paper and other research ideas. From pointing out theoretical subtleties, to asking big picture questions, his comments and suggestions were always relevant and valued. I am thankful to Amiyatosh Purnanandam for always taking the time to meet with me and asking difficult and pertinent questions that ultimately pushed my research further. I am truly appreciative of the participation by David Miller and Greg Miller in my Committee, as I benefited tremendously from their feedback and advice.

The open-door policy and the academic rigor of the Finance Department at Ross played a crucial role in my progress through the PhD program. The countless opportunities to present, discuss, and solicit feedback were of essence. I am also deeply grateful to my fellow PhD students, including Thomas Flanagan, Jay Kahn, Emil Lakkis, Giacomo Meille, Taeuk Seo, and Zhengyang Xu, for always willing to help.

Most importantly, this dissertation would not have been possible without the unconditional love and support of my wife Katya, my daughter Veronica, my parents Fedor and Katya, and my brother Andrei. It is to them that I dedicate my PhD.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF APPENDICES

Appendix

ABSTRACT

This dissertation studies the implications of technological innovation in the financial markets to asset prices, market quality, and the delegated portfolio management sector. Three such developments are considered: index funds (Chapter 1), quantitative mutual funds (Chapter 2), and exchange traded products (Chapter 3). For index funds, I study the drivers of the growth in passive investing and its effect on market efficiency. For quantitative mutual funds and exchange traded products, I explore their impact on various measures of market quality such as price discovery, market efficiency, and liquidity.

Chapter 1 investigates the impact of fundamental information acquisition costs on price informativeness and passive investing. Within a noisy rational expectations equilibrium (REE) model of multiple risky assets and a redundant market index, I define passive investing as the optimal decision to: 1) free-ride on the information acquisition efforts of active traders in the index asset, and 2) forgo all stock picking strategies. Falling information costs have the dual effect of lowering the cost of market timing, decreasing passive share, and lowering the cost of stock picking, increasing passive share. If the stock picking effect dominates the market timing effect, passive share increases in tandem with greater price informativeness. I exploit the Security and Exchange Commission's eXtensible Business Reporting Language mandate as a negative shock to information costs to provide suggestive evidence that falling information costs may be contributing to the rise in passive investing.

Chapter 2, joint with Paolo Pasquariello, studies the effects of quantitative equity investing, an increasingly popular investment style, on financial market quality. Within a noisy REE model of strategic speculation with two informed market participants, we define discretionary investing as fully strategic trading and quantitative investing as partially or fully myopic via its reliance on a backtested trading strategy. Growth in quantitative investing is modeled as both the introduction of and the greater backtest adherence by an informed speculator. The introduction of an additional speculator generally benefits financial market quality. The effect of greater backtest adherence depends on whether it leads to more or less aggressive trading than discretion, the former improving, while the latter worsening market quality. If it is more aggressive, market quality broadly benefits with greater quantitative investing; if it is less aggressive, market quality deteriorates.

Chapter 3 explores the implications of the growth in exchange traded funds (ETFs) and the associated arbitrage trading on price discovery and market liquidity. The introduction of arbitrage trading to segmented markets with otherwise diverging prices averages noise trading across markets, as the arbitrageur buys (sells) in the market with excess noise supply (demand). This smoothing results in less informed trading due to lower camouflage for the speculators, and lower liquidity due to greater adverse selection concerns for the market makers. The introduction of an ETF that attracts a threshold level of incremental noise trading leads to unambiguous improvements in the market quality of the underlying security, as the arbitrageur connects the synthetic and underlying markets by averaging noise trading across markets. I highlight the differential effects on market quality of stand-alone arbitrageurs and market makers jointly serving as arbitrageurs, with the former leading to greater informed trading intensity for the speculators and greater adverse selection for the market makers, and the latter having the opposite effect.

CHAPTER I

Information, Participation, and Passive Investing

1.1 Introduction

The dramatic growth in the share of passive versus active equity investment management has sparked a vibrant debate about both its causes and its effects on price discovery and market efficiency.[1,2]

A natural and popular explanation for the rise in passive investing, one consistent with standard models of price formation in the literature, is the relative decline in the cost of passive versus active asset management. For instance, within Grossman and Stiglitz (1980, "GS80"), greater relative costs of active investing would lead to more passive investing in equilibrium. However, in such models this growth would be accompanied by less informative prices, which is at odds with empirically documented trends of increasing price informativeness (Bai et al. (2016)) especially for the larger firms in the economy most affected by indexing strategies (Farboodi et al. (2018)).[3]

[1] According to Investment Company Institute (2020), as of 2019, US index and exchange traded funds (ETFs) accounted for 49% of assets under management by the US open-end equity fund industry, up from 6% in 1996.

[2] See, for example, O'Hara (2003) and Boehmer and Kelley (2009) for some evidence on the importance of asset management in price discovery and Bond et al. (2012) for a review of the literature on the real effects of financial markets.

[3] By price informativeness I refer to the extent to which prices reflect fundamentals. In my setting this will be the inverse of payoff variances conditional on prices. An example of an empirical coun-

The observed joint growth in passive share and price informativeness poses a challenge for simple cost-based explanations.

My first contribution is theoretical: I highlight conditions under which a *decline* in active versus passive investment costs leads to *both* a growing passive share and more informative prices. The result is driven by two realistic extensions to the traditional GS80 framework. First, I incorporate multiple risky assets to separately consider market timing and stock picking investment strategies for active investing. Second, I differentiate between participation and information costs for stock picking strategies. Within this setup, falling fundamental information costs have the dual effect of increasing the fraction of investors pursuing market timing strategies – thereby decreasing passive share – and decreasing the fraction of investors participating in stock picking strategies – thereby increasing passive share. To the extent that the stock picking effect on passive share dominates the market timing effect, falling fundamental information costs lead to both an endogenously rising passive share and greater price informativeness.

My second contribution is empirical: I provide supportive evidence that the falling information cost channel is at least partially responsible for the rise in passive investing. I utilize the eXtensible Business Reporting Language (XBRL) mandate implemented by the Securities and Exchange Commission (SEC) as a regulatory shock to information costs. In the literature (e.g. Dong et al. (2016), Jaskowski and Rettl (2018)), this event is interpreted as a negative exogenous shock to the cost of acquiring fundamental information. Exploiting the advantageous phase-in design of the regulation, I find that lower information costs *cause* a deterioration in the stock-picking performance of active strategies, consistent with the main testable implication of my model. This decline in performance coincides with flows out of active and into passive strategies with comparable investment objectives. To my knowledge, my study is the first to

terpart would be the slope from a regression of future realized fundamentals on current market prices (e.g. Bai et al. (2016)).

attempt to empirically link the rise in passive investing with falling information costs.

To describe the theoretical effect of falling information costs on passive investing, I first identify what it means to be optimally passive. In this paper, I focus on the purest and most prominent form of passive investing: the sole allocation to a broad market aggregate without information acquisition. In this setting, a passive investor participates in the stock market via an index fund and avoids all active management. Since active management is a combination of market timing and stock picking strategies, a passive investor: 1) allocates to a broad market index, 2) forgoes market timing, and 3) forgoes stock picking. Taken together, the first and second conditions imply that passive investors optimally free-ride on the information acquisition efforts of active traders in a market aggregate. The third condition states that passive investors optimally do not participate in stock selection strategies. These determinants of passive investing inform the structure of my stylized model for the asset management sector in which passive investing arises endogenously as an optimal portfolio allocation.

My framework builds on GS80 to allow for both passive investing and the different types of active management. To distinguish between market timing and stock picking strategies, I introduce N risky securities and a redundant index asset, whose expected payoff mimics the aggregate expected payoff to the market. As in GS80, each of the $N+1$ securities has a fundamental payoff, which can be learned for a corresponding information acquisition cost. I employ an asset rotation inspired by Bond and Garcia (2020) to decompose the investment set into the index asset and $N-1$ orthogonal long-short portfolios, labeled as stock picking strategies. The stock picking strategies are costly to participate in, which is meant to resemble the costly nature of active management over-and-above information acquisition costs (e.g. trading and market impact costs, excessive fees, overhead, moral hazard, etc.).[4] The rotation allows me

[4]My model generalizes to costly participation for the index asset as well without affecting its main takeaways. In this case, stock market non-participation will also emerge, which is not the focus of

to directly map the investors' equilibrium portfolio allocations in each of the synthetic securities to either the passive strategy or various types of active investment strategies. Within this setup, market timing is defined as investing in the index security with information acquisition. Stock picking is reflected as participation in *at least* one of the stock picking strategies, either with or without information acquisition.[5]

In equilibrium, the fractions of investors pursuing the passive and various active strategies are determined by two key forces: 1) *strategic substitutability in information acquisition* and 2) *strategic substitutability in stock picking participation*. The former is the key outcome of GS80 and directly applies to the index asset: as more investors acquire information, the value of information acquisition declines, uniquely identifying the equilibrium fraction of market timers in the economy. The latter, to my knowledge, is novel and is essential to exploring passive investing. If participating stock pickers optimally select their information acquisition strategies, more stock picking participation results in lower gains to participation for all. Taken together, the two forces ensure the uniqueness and existence of a linear equilibrium for participation and information acquisition in stock picking. I then explore the effects of changing information costs on passive share through their effect on the optimal investment in the index and stock picking strategies.

The effect of lower information costs for the index asset is standard (i.e. as in GS80): lower costs lead to a greater fraction of investors acquiring information and more informative prices. This has the immediate implication of a lower passive share as more investors pursue a market timing strategy. Thus, as noted earlier, this channel cannot explain the observed improvement in price informativeness accompanying a

this paper. Therefore, for exposition, I assume that all investors participate in the index asset.

[5] Uninformed stock picking allows investors to deviate from the passive index without information acquisition. As originally pointed out by Admati (1985) and more recently emphasized by Biais et al. (2010) and Garleanu and Pedersen (2020), within models of noisy rational expectations the unconditional expected market portfolio need not be the optimal portfolio for an uninformed investor due to the uncertainty associated with the supply of securities.

greater passive share in the US. The effect of lower information costs on stock picking is twofold. On the one hand, lower costs incentivize information acquisition, which motivates a greater fraction of participating investors to become informed. On the other hand, because a greater fraction of investors are informed, prices become more reflective of fundamentals, which, in turn, reduces the gains to participation. As expected profits for stock pickers turn negative, some choose to forgo stock picking altogether, resulting in a greater passive share. In this scenario, passive share and price informativeness increase in tandem. The question of which of these forces dominates is ultimately an empirical one and motivates me to take the prediction to the data.

As a basis for my tests, I propose the following institutional mechanism for the effect of lower information costs on passive investing. As information costs fall, the first and most immediate effect is that more active investors acquire information, leading to more informative prices, and ultimately resulting in deteriorating performance for active funds. Consequently, investors allocating to active funds internalize the performance decline and take money out leading to capital outflows from active investing. Finally, some of the investors pulling money out of active strategies, reallocate to passive strategies resulting in capital inflows to passive funds. I attempt to identify each of the three effects (i.e. lower active performance, active outflow, and passive inflow) in the data.

Empirically identifying the effect of lower information costs on passive investing is nontrivial since fundamental information costs are endogenously determined and typically not easily observable. I address these concerns by exploiting SEC's "Interactive Data to Improve Financial Reporting" rule as an exogenous and negative shock to information costs. This regulation, originally introduced in 2009, required firms to provide their financial statements in an interactive format utilizing XBRL, which made financial information easier to export into spreadsheets, to machine read with off-the-shelf software, and to compare it across firms (via a standard list of tags developed by XBRL US and reviewed by the Financial Accounting Standards Board [FASB]). To the extent that investors were previously paying third-party data providers for data manipulation or were spending their own time on it, they can now reduce these costs by automating the data collection process.[6, 7]

To establish the effect of information costs on the performance of active management, I rely on the staggered phase-in design of the regulation. The first phase applied to companies with a public equity float of over \$5bn: these companies were required to adopt XBRL reporting for financial statements for fiscal periods ending on or after June 15, 2009. The second phase applied to companies with public equity float between \$700mm and \$5bn: these firms had to utilize XBRL for filings for fiscal periods ending on or after June 15, 2010. Finally, all remaining filers had to comply for fiscal periods ending on or after June 15, 2011.

I exploit cross-sectional variation in active mutual funds' exposure to the information cost shock as measured by the fraction of their assets invested in securities exposed to each phase of the regulation. I argue that greater exposure to the shock will lead to weaker future fund performance as measured by a fund's Carhart alpha (Carhart (1997)). To establish causality, I group funds based on their exposures to either phase 1 or their joint exposure to phases 2 and 3. I take a subset of funds that are either exposed predominantly to phase 1 (group 1, $\geq 75\%$ of total net assets) or predominantly to phases 2 and 3 (group 2, $\geq 75\%$ of total net assets) and track their relative performance through time utilizing an event study approach.

[6] See e.g. Kim et al. (2012), Blankespoor et al. (2014), Dong et al. (2016), Jaskowski and Rettl (2018), Bhattacharya et al. (2018) for evidence that the XBRL mandate results in greater information acquisition by market participants.

[7] Implicit to my analysis is also the assumption that lower information costs leads to greater price informativeness in the affected securities. Several theories have been proposed demonstrating that greater transparency or lower information costs need not imply more informative prices (Banerjee et al. (2018) and Dugast and Foucault (2018)).

I find that the performance of the two groups of funds is comparable both during the pre-treatment period and during phases 1 and 2 of the treatment. However, as phase 3 is implemented, group 1 funds begin to significantly underperform group 2 funds. The underperformance steadily declines for approximately two years, after which the gap starts to shrink, ultimately converging in another two years. The trend suggests that the information cost shock affects fund performance with a lag of approximately two years, which is consistent with the notion that changes to information costs need time to be internalized by investors and ultimately be reflected in more informative prices and declining fund performance.

Both group 1 and group 2 funds exhibit fund outflows, which are especially pronounced in the periods after the respective implementation of phase 1 and phases 2 and 3 of the XBRL mandate. While I'm unable to tease out a causal interpretation for the active fund outflows, a closer look at the flows to passive funds suggests that perhaps investors are rotating out of active and into passive funds. More specifically, I apply a similar grouping based on exposure for passive funds and focus on those passive funds that are most comparable to the affected active funds as dictated by their investment objective. I find that group 1 passive funds see significant inflows right around the time of deteriorating performance for group 1 active funds.

Overall, my paper highlights a novel theoretical link between falling information costs and the rise in passive investing and provides supportive empirical evidence for this relationship in US equity markets. Thus, my contribution to the literature is twofold. On the theoretical side, I contribute to the literature on endogenous information acquisition, the aggregation of information and its incorporation into prices, and the effects of both on portfolio allocation and investor behavior.[8] On the empirical side, I add to the growing literature that explores the effects of developments in information

[8]E.g. GS80,Verrecchia (1982),Kacperczyk et al. (2016),Garleanu and Pedersen (2017)

technology and financial innovation on asset prices and market efficiency.[9]

I proceed as follows. Section 1.2 discusses the relevant literature. Section 1.3 introduces the model. Section 1.4 explains the solution. Section 1.5 discusses the comparative statics. Section 1.6 presents the empirical analysis. Section 1.7 concludes.

1.2 Literature Review

Reconciling falling information costs with growth in passive investing is challenging within canonical noisy rational expectations models since lower costs imply more information acquisition and therefore more active investing (GS80, Verrecchia (1982)). Hence, the more recent theories, which explain the rise in passive investing by relying on the lower relative costs to indexing strategies as the primary driver, yield lower equilibrium price informativeness (Peress (2005), Bond and Garcia (2020), Garleanu and Pedersen (2020)).[10] Multiple asset models with information choice yield similar implications (Nieuwerburgh and Veldkamp (2009), Nieuwerburgh and Veldkamp (2010), Veldkamp (2011), Kacperczyk et al. (2016), Abis (2020), Kacperczyk et al. (2018)). In these studies, lower information costs can be viewed as greater information processing capacity; however, this would have a similar effect of more information acquisition implying more active management. More broadly my work contributes to the class of models exploring the aggregation of information in markets with asymmetric information (Hellwig (1980), Diamond and Verrecchia (1981)) and investing across multiple securities (Admati (1985)).

Several papers – including Subrahmanyam (1991), Bhattacharya and O'Hara (2018) Gorton and Pennacchi (1993), and Cong and Xu (2016) – consider passive investing

[9]E.g. Dong et al. (2016), Jaskowski and Rettl (2018), Zhu (2019), Katona et al. (2019)

[10]Bond and Garcia (2020) do find that while price efficiency declines for the index asset, the relative price efficiency of individual stocks rises. Within my framework, since I am specifically altering information acquisition costs with endogenous information acquisition, lower information costs lead to greater aggregate and relative price efficiency.

within a Kyle (1985) framework in which sophisticated market participants are strategic and internalize the impact of their trades. These papers generally point to greater factor information in prices due to the introduction of an index and the potential mitigation of adverse selection by uninformed traders. The latter two also emphasize the decreasing price efficiency in individual security prices as a result of composite security trading. Unlike these studies, I focus specifically on information costs, and how they may impact passive share.

Over the last few decades, fundamental data have become more abundant, more easily accessible, and less expensive, especially thanks to regulatory overhauls. For example, the SEC has implemented such mandates as EDGAR and XBRL to make data more easily accessible and machine readable for investors. These reforms have spurred much research on the effects of declining information costs on investment management with some pointing towards greater price informativeness for the affected securities (Kim et al. (2012), Dong et al. (2016), Jaskowski and Rettl (2018)) while others find greater information asymmetry (Blankespoor et al. (2014)). Advancements in information technology have also yielded alternative fundamental data ranging from retail parking lot satellite imagery (Zhu (2019)) to cell phone ping data (Froot et al. (2017)). Recent studies argue that such data may increase the information advantage of the more sophisticated traders who have access to it (Katona et al. (2019)) leading to greater income inequality (Mihet (2020)). However, other studies have documented broad improvements to price informativeness (Bai et al. (2016), Brogaard et al. (2019), Davila and Parlatore (2019), Martineau (2020)), especially to the larger firms in the economy most affected by passive investing (Farboodi et al. (2018)). To the extent that passive investing is at least partially driven by falling information costs, the presented theoretical framework would imply more informative prices for the affected securities, consistent with many of the studies mentioned here.

A growing literature has focused on the direct effects of passive investing on price informativeness with mixed results. On the one hand, Glosten et al. (2017) find that greater ETF ownership leads to more accurate incorporation of accounting information suggesting that price efficiency is improving with greater passive share. On the other hand, when investigating Russell index reconstitutions, Coles et al. (2018) find that weak-form price efficiency deteriorates due to passive investing. Similarly, Israeli et al. (2017) report that ETF ownership results in decreasing price efficiency, while Cremers et al. (2016) find that more indexing usually leads to more active investing with greater alpha. Sammon (2020) further contributes to the debate by demonstrating that prices are becoming less informative around earnings announcements due to greater passive ownership. While I do not measure price informativeness explicitly, my model suggests that the portion of the rise in passive investing that is driven by falling information costs should be accompanied by greater market efficiency.

More broadly, my paper contributes to the literature on flows and performance for active mutual funds (e.g. Jensen (1968), Carhart (1997), Sirri and Tufano (1998), French (2008), Fama and French (2010), Berk and van Binsbergen (2015)). Furthermore, various studies have attempted to identify the sources of skill and alpha for active mutual funds (e.g. Coval and Moskowitz (2001), Kacperczyk et al. (2005), Baker et al. (2010)). While the literature has provided a variety of explanations for differences in performance across funds, I focus on a very specific driver of performance, namely information costs, and provide suggestive evidence to demonstrate that this force is in effect.

Lastly, my theory sidesteps some realistic aspects of the mutual fund industry and the effects of changing information costs. In particular, I do not model the asset management sector in detail and capture all relevant frictions in reduced form (i.e. through participation costs). Thus, my results on the growth of passive investing due to falling information costs likely form a lower-bound relative to a model with a more realistic asset management sector. In the presence of moral hazard (Brown and Davies (2017)) or search frictions (Garleanu and Pedersen (2017)), the index asset would likely become even more attractive leading to even greater passive growth. Similarly, introducing economies of scale and differential skill for active management (Berk and Green (2004)) may also lead to greater passive share.

1.3 The Model

The goal of the model is to explore the effect of changes to the relative costs of active and passive investing and the ensuing implications to the prevalence of indexing strategies and price informativeness. To this end, the model will feature utility-maximizing traders who allocate between a risk-free asset and a set of risky securities, among which is an aggregate market index. Two types of costs to active investing will be introduced: 1) the cost of acquiring fundamental information (i.e. receiving a more precise signal about the future payoff) and 2) any other costs associated with implementing the active strategy (i.e. a cost to participate in the strategy). The main purpose of the exercise is to investigate how changes to information and participation costs affect the optimal portfolio allocations of traders in the economy, and, in particular, the fraction of traders pursuing a pure indexing strategy.

1.3.1 Model Setup

The setting is a two-period endowment economy with multiple risky assets and a continuum of traders indexed by the interval $j \in [0, 1]$. Endowments are assumed to be W_0 of cash for all traders and traders have CARA preferences with risk aversion of $1/\tau$. Assets are priced and exchanged in the first period based on the demands of the traders and the total supply of securities. Liquidation values are realized and consumed in the second period. Markets are assumed to be competitive in that an individual agent's demand cannot impact the price. The agents have rational expectations in that the structure of the model is common knowledge.

1.3.2 Assets

The market contains a risk-free asset with price and gross return normalized to 1 and $i = 1, \dots, N$ risky assets with prices $P = [P_1, \dots, P_N]'$ and liquidation values $v = [v_1, \dots, v_N]'$. Liquidation values are governed by a fundamental payoff θ and noise payoff u:

$$v = \theta + u$$

where $\theta = [\theta_1, \dots, \theta_N]'$, and $u = [u_1, \dots, u_N]'$. Payoffs are independent with fundamentals governed by $\theta \sim N(\mu_\theta, \Sigma_\theta)$ with $\Sigma_\theta = \text{diag}[\sigma^2_{\theta_1}, \dots, \sigma^2_{\theta_N}]$, noise governed by $u \sim N(0, \Sigma_u)$ with $\Sigma_u = \text{diag}[\sigma^2_{u_1}, \dots, \sigma^2_{u_N}]$, and $\sigma_{\theta_i u_j} = 0$ for any assets i, j. *Expected* supply for each asset is given by the vector $X = [X_1, X_2, \dots, X_N]'$ and noise in the supply will be introduced for synthetic assets defined below. I assume independent fundamental and noise payoffs for exposition. Lemma A.2, highlighted in Appendix A.9, demonstrates that in the case of correlated fundamentals, the payoff space can be re-spanned with portfolios of underlying securities whose noise and fundamental payoffs are orthogonal.

In addition to the individual assets, a redundant, value-weighted index, is offered in the market with payoff governed by:

$$X'v = X_1\theta_1 + X_2\theta_2 + \cdots + X_N\theta_N + X_1u_1 + X_2u_2 + \cdots + X_Nu_N$$

The security-specific weights of the index asset correspond to the expected supply of each security, which ensures that this is a value-weighted index.[11,12]

1.3.3 Passive Investing and Information Acquisition

Traders are ex-ante identical and optimally choose to become passive. Passive traders are defined as those who avoid all active management and restrict their investment only to the index asset, which is free to trade, without acquiring information on its fundamental payoff. All other traders are referred to as "active." Active traders, by definition, participate in at least one active strategy, which will be introduced below. For any risky asset i, the fundamental θ_i can be learned for an *information acquisition cost* c_i.[13] Without information acquisition, traders attempt to learn the fundamental payoffs from prices. For tractability, for all assets I assume constant *information quality*, $\frac{\sigma^2_{\theta_i}}{\sigma^2_{u_i}} = n$, and a constant inverse relationship between fundamental uncertainty and expected supply, $X_i^2 \sigma^2_{\theta_i} = q$.[14]

1.3.4 Introducing Synthetic Assets

The index asset is redundant. Since I am after portfolio choice in its presence, it is useful to construct non-redundant synthetic assets that span the original payoff space

[11] A value-weighted index will hold a fixed *percentage* of shares outstanding in all of the securities within the index. The prices will then ensure that percent of capital allocations are based on market capitalization. Thus, absent rebalancing considerations, relative share allocations are determined by shares outstanding for each firm in the index.

[12] The index is meant to resemble a broad market aggregate resembling the entire economy. The portfolio weights of the index asset have to be proportional to the expected supply of each security for the proofs that follow. See Garleanu and Pedersen (2020) for a further discussion on the optimality of an index based on expected supply, as well as Pedersen (2018) for a discussion on the importance of the uncertainty in asset supply.

[13] The structure of the model is slightly different from models based on Admati (1985) in which traders receive heterogenous signals of the form $s_{ji} = v_i + \epsilon_{ji}$ where the signal noise $\sigma_{\epsilon_{ji}}$ varies by trader j. In such models individual signals aggregate and are reflected through the realization of the fundamental v_i entering the price function. In my setting, agents choose whether to acquire a fixed signal of the form $\theta_i = v_i - u_i$. See Veldkamp (2011) for more discussion of the difference between the various types of information acquisition models.

[14] Taken together, these two assumptions imply that assets with lower payoff uncertainty have greater expected supply, which is consistent with security issuers catering to the risk aversion of the traders.

such that the index asset is included. Note that there are infinitely many ways of constructing such assets; I choose a specific approach that allows me to both include the index asset and orthogonalize the noise and fundamental payoffs of the synthetic assets. This greatly simplifies the theoretical results to follow, as I am able to analyze the assets on an asset-by-asset basis, since their payoffs and prices will not be correlated by construction. In particular, I follow an approach similar to that of Bond and Garcia (2020) and introduce synthetic assets with payoffs $\tilde{v}_1, \ldots, \tilde{v}_N$ as portfolios of the original assets as defined by the following mapping:

$$\begin{bmatrix} \tilde{v}_1 \\ \tilde{v}_2 \\ \tilde{v}_3 \\ \tilde{v}_4 \\ \vdots \\ \tilde{v}_N \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & \ldots & X_{N-1} & X_N \\ \frac{N-1}{X_1\sigma^2_{\theta_1}} & -\frac{1}{X_2\sigma^2_{\theta_2}} & -\frac{1}{X_3\sigma^2_{\theta_3}} & -\frac{1}{X_4\sigma^2_{\theta_4}} & \ldots & -\frac{1}{X_{N-1}\sigma^2_{\theta_{N-1}}} & -\frac{1}{X_N\sigma^2_{\theta_N}} \\ 0 & \frac{N-2}{X_2\sigma^2_{\theta_2}} & -\frac{1}{X_3\sigma^2_{\theta_3}} & -\frac{1}{X_4\sigma^2_{\theta_4}} & \ldots & -\frac{1}{X_{N-1}\sigma^2_{\theta_{N-1}}} & -\frac{1}{X_N\sigma^2_{\theta_N}} \\ 0 & 0 & \frac{N-3}{X_3\sigma^2_{\theta_3}} & -\frac{1}{X_4\sigma^2_{\theta_4}} & \ldots & -\frac{1}{X_{N-1}\sigma^2_{\theta_{N-1}}} & -\frac{1}{X_N\sigma^2_{\theta_N}} \\ \vdots & & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \ldots & \frac{1}{X_{N-1}\sigma^2_{\theta_{N-1}}} & -\frac{1}{X_N\sigma^2_{\theta_N}} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ \vdots \\ v_N \end{bmatrix}$$

Synthetic asset $i = 1$ is the index asset while synthetic assets $i = 2, \ldots, N$ are long-short portfolios.[15] I refer to the long-short portfolios as stock picking strategies, since their payoffs reveal the relative values of securities. Synthetic assets have prices $\tilde{P} = [\tilde{P}_1, \ldots, \tilde{P}_N]'$, fundamentals $\tilde{\theta} = [\tilde{\theta}_1, \ldots, \tilde{\theta}_N]$, and noise $\tilde{u} = [\tilde{u}_1, \ldots, \tilde{u}_N]$. For any synthetic asset i, the fundamentals refer to the linear combination of θ_i's in the payoff and the noise refers to the linear combination of u_i's in the payoff. Note that fundamental and noise payoffs of the synthetic assets are uncorrelated by construction due to the assumptions of constant information quality and constant fundamental uncertainty to supply ratios. I label the distributional assumptions of the synthetic

[15] I will continue to index the assets by i not $\tilde{i}$ in the subscripts of variable names, however a tilde above a variable implies that it is synthetic.

payoffs by $\tilde{\theta} \sim N(\mu_{\tilde{\theta}}, \Sigma_{\tilde{\theta}})$ and $\tilde{u} \sim N(\mu_{\tilde{u}}, \Sigma_{\tilde{u}})$. To be able to analyze the synthetic assets independently, noise supply is assumed to be independent across the synthetic assets and is governed by $\tilde{x} \sim N(0, \Sigma_{\tilde{x}})$ with $\Sigma_{\tilde{x}} = \text{diag}[\sigma^2_{\tilde{x}_1}, \ldots, \sigma^2_{\tilde{x}_N}]$. The synthetic asset *expected* supply vector is determined by: $\tilde{X} = Z^{-1}X$ where each column of the matrix Z corresponds to the portfolio weights of each synthetic asset (e.g. the first column is the vector $[X_1, X_2, \ldots, X_N]'$). The solution for the synthetic asset expected supply vector is $\tilde{X} = [1, 0, 0, \ldots, 0]'$ (i.e. the index asset has expected supply of 1 while all other assets have expected supply of 0).

All traders invest in the index asset for free.[16] Participation in stock picking strategies is costly over-and-above information acquisition costs, reflected by participation costs $\tilde{k}_i$ for synthetic assets $i = 1, \ldots, N$, aggregated into vector $\tilde{k}$ (note that the first element $\tilde{k}_1 = 0$ since the index asset is free to trade).[17]. Costly participation refers to the incremental costs incurred by active funds such as trading commissions, market impact costs, overhead expenditures, tax filings, etc. The fundamental component of each synthetic asset i can be learned for cost $\tilde{c}_i$ for $i = 1, \ldots, N$ aggregated into vector $\tilde{c}$. For consistency, it must be the case that $\sum_{i=1}^{N} c_i = \sum_{i=1}^{N} \tilde{c}_i$. I do not take a stand on the exact mapping from the original to the synthetic information costs. For my results to go through, I need falling information costs to the original assets to translate into falling information costs for the index asset and at least one stock picking strategy.

Passive traders only invest in the index asset without information acquisition, i.e. they optimally free-ride on the information acquisition efforts of active traders in the index asset. Active traders pursue at least one active strategy, i.e. they either acquire information about the index asset (market timing), participate in at least one stock

[16]This assumption is without loss of generality for the main theoretical conclusions. With costly participation, aggregate stock market non-participation would emerge, which is not the focus of this paper. See Peress (2005) for more details on the relationship between information and participation costs in a market index.

[17]Subscripts refer to the elements of the specified vector. E.g., I_{j_i} refers to the i^{th} element of vector I_j

picking strategy (stock picking) or do both.

1.4 Model Solution

The model is solved in four steps. First, I hold fixed the fraction of traders participating and informed by asset (exogenous "participation" and "information") to identify the equilibrium price function at which all traders optimize given their type. Next, I endogenize the information acquisition of active traders (i.e. the "information" decision), to determine what fraction of traders optimally acquires information. Third, I endogenize the participation decision (i.e. the "participation" decision), to determine what fraction of traders optimally participates in the stock picking strategies. Finally, I map the fractions of traders acquiring information across all assets and participating in the stock picking strategies to my definition of passive share.

1.4.1 Optimal Portfolio Choice

Trader j forms a portfolio D_j to maximize the expected utility of period two wealth (W_{2j}) conditional on their fundamental information set F_j and participation set R_j:

$$D_j = \arg\max_{\tilde{D}_j} \mathbb{E}[-\exp(-\frac{W_{2j}}{\tau})|F_j, R_j] \tag{1.1}$$

The participation set indicates the prices of the synthetic assets in which trader j is investing. For example, if trader j is active in assets 1,2, and $N-1$, the participation vector would be $R_j = [\tilde{P}_1, \tilde{P}_2, \tilde{P}_{N-1}]$. I denote the indices of the securities in which trader j is investing by r_j (in the example, $r_j = [1, 2, N-1]$). Note that for all traders $\tilde{P}_1 \in R_j$ since it is free to participate in the trading of the index asset. Even though the trader observes all prices, due to the orthogonality of assets, the prices of assets in which the trader does not participate do not inform the investment decision once the participation set has been determined. Therefore the prices of the other assets can be excluded from the conditioning variables.

The fundamental information set is composed of a subset of the fundamental vector $\tilde{\theta}$ for which information is acquired. The vector F_j indicates the realizations of the fundamentals the trader will learn after incurring the associated information costs. For example, if trader j pays $\tilde{c}_2$ and $\tilde{c}_N$ to learn the realizations of $\tilde{\theta}_2$ and $\tilde{\theta}_N$ respectively, the information set vector F_j would be $[\tilde{\theta}_2, \tilde{\theta}_N]$. Note that for any passive trader, $F_j = \varnothing$, since passive traders, by definition, do not acquire information. I aggregate the total information of trader j in the information-participation set I_j, where each element I_{j_i} represents the information relevant for security i: for an informed trader $I_{j_i} = \{\tilde{P}_i, \tilde{\theta}_i\}$, for an uninformed trader $I_{j_i} = \{\tilde{P}_i\}$, and for a non-participating trader $I_{j_i} = \{\varnothing\}$.

To map information and participation decisions to absolute costs, I define $\mathbb{1}_{R_j} = [\mathbb{1}_{\{\tilde{P}_1 \in R_j\}}, \ldots, \mathbb{1}_{\{\tilde{P}_N \in R_j\}}]'$ and $\mathbb{1}_{F_j} = [\mathbb{1}_{\{\tilde{\theta}_1 \in F_j\}}, \ldots, \mathbb{1}_{\{\tilde{\theta}_N \in F_j\}}]'$. Conditional on participation and information decisions, period two wealth for trader j is given by:[18]

$$W_{2j} = W_0 - \mathbb{1}'_{F_j} \tilde{c} - \mathbb{1}'_{R_j} \tilde{k} + D'_j \big((\tilde{v} - \tilde{p}) \odot \mathbb{1}_{R_j} \big) \tag{1.2}$$

The synthetic assets are orthogonal across all dimensions relevant to prices and payoffs: noise supply, fundamental payoffs, and noise payoffs. Therefore, due to fixed endowments and CARA preferences, optimal portfolio choice can be considered on an asset-by-asset basis.

For any asset i in trader j's participation set (i.e. $\{\tilde{P}_i\} \in I_{j_i}$), optimal demands

[18] The symbol $\odot$ signifies element-by-element vector multiplication. In this context it is necessary to limit investments to assets in the participation set R_j.

conditional on the information set I_{j_i} are given by the standard demand function:

$$D_{j_i} = \frac{\tau(\mathbb{E}[\tilde{v}_i|I_{j_i}] - \tilde{P}_i)}{\sigma^2_{\tilde{v}_i|I_{j_i}}} \tag{1.3}$$

where in the case of a trader j uninformed in asset i ($\{\tilde{\theta}_i\} \notin I_{j_i}$)

$$D_{j_i} = \frac{\tau(\mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i)}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \tag{1.4}$$

and in the case of a trader j informed in asset i ($\{\tilde{\theta}_i\} \in I_{j_i}$)

$$D_{j_i} = \frac{\tau(\tilde{\theta}_i - \tilde{P}_i)}{\sigma^2_{\tilde{u}_i}} \tag{1.5}$$

For any asset i not in trader j's participation set (i.e. $\{\tilde{P}_i\} \notin I_{j_i}$), demands of trader j are zero:

$$D_{j_i} = 0 \tag{1.6}$$

All else equal, demands are greater when the conditional expected payoffs are greater or the conditional payoff variance is lower.

1.4.2 Exogenous Information and Participation

Given the optimal demand functions specified above, I can impose market clearing for each asset. For each synthetic asset i, I label the fraction of non-participating traders as $\tilde{\gamma}_i$ (aggregated in vector γ) and the fraction of informed traders among the participating traders as $\tilde{\lambda}_i$ (aggregated in vector λ). Note that since all traders participate in the trading of the index asset, $\tilde{\gamma}_1 = 0$. Given expected supply vectors for the synthetic assets, noise supplies, and optimal demand functions provided by Eq.

(1.4), (1.5), and (1.6), market clearing implies:

$$(1-\tilde{\gamma}_i)\tilde{\lambda}_i \frac{\tau(\tilde{\theta}_i - \tilde{P}_i)}{\sigma^2_{\tilde{u}_i}} + (1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\frac{\tau(\mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i)}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} = \tilde{X}_i + \tilde{x}_i \tag{1.7}$$

It is immediate that an equilibrium does not exist in the full non-participation case ($\tilde{\gamma}_i = 1$), since there is no demand to clear the supply of the risky asset i. Therefore, I assume that $\tilde{\gamma}_i < 1$ for all i going forward. I can rewrite the market clearing condition by moving all of the variables known to the uninformed traders to the right-hand side to arrive at:

$$\begin{aligned}\tilde{\theta}_i - \frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\tilde{x}_i &= \tilde{P}_i + \frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\tilde{X}_i - \left(\frac{1}{\tilde{\lambda}_i} - 1\right)\frac{(\mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i)}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \\ &\equiv S_i(\tilde{P}_i) \text{ if } \tilde{\gamma}_i < 1 \text{ and } \tilde{\lambda}_i > 0\end{aligned} \tag{1.8}$$

and define

$$S_i(\tilde{P}_i) = \tilde{x}_i \text{ if } \tilde{\lambda}_i = 0 \tag{1.9}$$

$S_i(\tilde{P}_i)$ is a sufficient statistic for prices and greatly simplifies the derivation of the equilibrium price function because $\mathbb{E}[\tilde{v}_i|\tilde{P}_i] = \mathbb{E}[\tilde{v}_i|S_i(\tilde{P}_i)]$ and $\mathrm{Var}[\tilde{v}_i|\tilde{P}_i] = \mathrm{Var}[\tilde{v}_i|S_i(\tilde{P}_i)]$. In Appendix A.1, I derive the following:

Lemma I.1 (Equilibrium: Exogenous Information and Participation Strategy)**.** *For any synthetic asset i, fixed $\tilde{\lambda}_i$ and fixed $\tilde{\gamma}_i < 1$ there exists a price function of the form $\tilde{P}_i(\tilde{\gamma}_i, \tilde{\lambda}_i, \tilde{\theta}_i, \tilde{x}_i) = q_1(\tilde{\gamma}_i, \tilde{\lambda}_i) + q_2(\tilde{\gamma}_i, \tilde{\lambda}_i)\tilde{\theta}_i + q_3(\tilde{\gamma}_i, \tilde{\lambda}_i)\tilde{x}_i$, where q_1, q_2 and q_3 are constants, such that for all realizations of $\tilde{\theta}_i$ and $\tilde{x}_i$, Eq.* (1.7) *is satisfied.*

$S_i(\tilde{P}_i)$ demonstrates that prices provide a noisy signal for the fundamental $\tilde{\theta}_i$. As in GS80, it is useful to think of $S_i(\tilde{P}_i)$ as a mean preserving spread of $\tilde{\theta}_i$:

$$\mathbb{E}[S_i(\tilde{P}_i)|\tilde{\theta}_i] = \tilde{\theta}_i \tag{1.10}$$

and

$$\text{Var}[S_i(\tilde{P}_i)|\tilde{\theta}_i] = \frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2} \text{ if } \tilde{\lambda}_i > 0 \tag{1.11}$$

Greater volatility of the noise payoff, lower fraction of informed traders, higher risk-aversion, higher volatility of noise trader supply, all lead to a less precise price system. In addition, lower participation (i.e., greater $\tilde{\gamma}_i$) also leads to lower precision under the assumption of exogenous information and participation.

I have demonstrated that an equilibrium exists in which traders are exogenously assigned information and participation decisions. Next, I allow traders to optimally choose whether to become informed about asset fundamentals.

1.4.3 Endogenous Information, Exogenous Participation

I now consider endogenous information acquisition, i.e. traders optimally make information acquisition decisions, with preassigned participation across multiple assets. When making the information acquisition decision, traders do not know the specific realization of the fundamental $\tilde{\theta}_i$ or the specific realization of price $\tilde{P}_i$. Therefore, in making the decision of whether to acquire information, traders need to evaluate the unconditional expected utility of acquiring information (i.e. for all potential realizations of $\tilde{\theta}_i$ and $\tilde{P}_i$) relative to the unconditional expected utility of inferring the information from prices (i.e. for all potential realizations of $\tilde{P}_i$).

For the results that follow, I will rely on an identity derived in Admati and Pfleiderer (1987), which allows me to evaluate the unconditional expected utility of a trader with a specific information set. In particular, the unconditional expected utility (EU) of trader j with information-participation set I_j is:

$$EU_{I_j} = -\frac{|\Sigma_{\tilde{v}|I_j}|^{1/2}}{|\Sigma_{\tilde{v}-\tilde{P},R_j}|^{1/2}} \exp\left(-\frac{W_0}{\tau} + \frac{\mathbb{1}'_{F_j}\tilde{c}}{\tau} + \frac{\mathbb{1}'_{R_j}\tilde{k}}{\tau} - \frac{1}{2}\mu'_{\tilde{v}-\tilde{P},R_j}\Sigma^{-1}_{\tilde{v}-\tilde{P},R_j}\mu_{\tilde{v}-\tilde{P},R_j}\right) \tag{1.12}$$

where $\mu_{\tilde{v}-\tilde{P},R_j}$ is the unconditional expected return of the assets in trader j's participation set, $\Sigma_{\tilde{v}-\tilde{P},R_j}$ is the unconditional variance-covariance matrix of returns of the assets in trader j's participation set, and $\Sigma_{\tilde{v}-\tilde{P}|I_j} = \Sigma_{\tilde{v}|I_j}$ is the conditional variance-covariance matrix of returns of the assets in trader j's participation set. The unconditional expected utility is composed of the expected utility of investing without conditioning on prices or fundamentals (the term in the exponent) and the uncertainty reduction associated with conditioning on any information acquired by the trader (the ratio of determinants). Recall that utilities are negative; therefore a reduction in variance implies a greater expected utility.

Due to orthogonality of the synthetic assets, I can rewrite the unconditional expected utility as:

$$EU_{I_j} = -\left(\prod_{i\in r_j}\sqrt{\frac{\sigma^2_{\tilde{v}_i|I_{j_i}}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}}\right) \exp\left(-\frac{W_0}{\tau} + \frac{\mathbb{1}'_{F_j}\tilde{c}}{\tau} + \frac{\mathbb{1}'_{R_j}\tilde{k}}{\tau} - \sum_{i\in r_j}\frac{\mathbb{E}[\tilde{v}_i - \tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right) \tag{1.13}$$

An immediate consequence of Eq. (1.13) is that for a given participation set R_j, the ratio of the unconditional expected utility of becoming informed in asset i (information-participation set I^*_j) to being uninformed in asset i (information-participation set I_j) is

$$\frac{EU_{I^*_j}}{EU_{I_j}} = \sqrt{\frac{\sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}} \exp\left(\frac{\tilde{c}_i}{\tau}\right) \tag{1.14}$$

The lower uncertainty associated with the acquisition of fundamental information ($\sigma_{\tilde{u}_i}/\sigma_{\tilde{v}_i|\tilde{P}_i}$) is offset by the utility lost due to the cost spent on information ($\exp(\tilde{c}_i/\tau)$). Assuming optimal demands by traders given information and participation sets and market clearing, I can derive a closed form solution for the ratio of expected utilities provided by Eq. (1.14). This result is summarized in the following Lemma.

Lemma I.2 (Utility Change due to Information Acquisition)**.** *Given the equilibrium price function and conditions specified in Lemma I.1, for trader j participating in synthetic asset i ($i \in r_j$), the ratio of the expected utility of being informed about i ($I^*_{j_i} = \{\tilde{P}_i, \tilde{\theta}_i\}$) to being uninformed ($I_{j_i} = \{\tilde{P}_i\}$) is monotone and increasing in $\tilde{\lambda}_i$ for a fixed $\tilde{\gamma}_i$.*

The ratio referred to in Lemma I.2 is given by the following formula:

$$\begin{aligned}
\frac{EU_{I^*_j}}{EU_{I_j}} &= \exp\left(\frac{\tilde{c}_i}{\tau}\right)\sqrt{\frac{\sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}} \\
&= \exp\left(\frac{\tilde{c}_i}{\tau}\right)\sqrt{\frac{\tau^2(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2\sigma^2_{\tilde{\theta}_i} + \sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i} + \sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}} \\
&\equiv f_i(\tilde{\gamma}_i, \tilde{\lambda}_i)
\end{aligned} \tag{1.15}$$

The result immediately follows from properties of conditional normal distributions (see Eq. (A.2)) and is consistent with GS80.[19] The orthogonality of assets allows me to consider the utility ratio on an asset-by-asset basis, which implies that Eq. (1.15) holds for all synthetic assets i.

Lemma I.2 highlights the strategic substitutability in information acquisition. The more traders acquire information in a particular asset, the lower are the relative gains to information acquisition. I can therefore classify the equilibrium in information ac-

[19] Although, as noted by Veldkamp (2011), GS80 do not explicitly derive the closed form solution for the ratio of utilities.

quisition. If there are no informed traders, when gains to information are greatest, and it is still not beneficial to acquire information, no one becomes informed. If everyone is informed, and it is still worthwhile to acquire information, everyone remains informed. Finally, if information acquisition is attractive but not for all participating traders, then the fraction of informed traders grows exactly to the point at which the expected utility to becoming informed is equal to the expected utility of remaining uninformed. I call an equilibrium with endogenous information acquisition an information equilibrium. The information equilibrium will be characterized by the fraction of traders optimally acquiring information in every asset, the exogenously determined participants in every asset, and a price function that clears the market at traders' optimal demands.

Proposition I.1 (Equilibrium: Endogenous Information, Exogenous Strategy)**.** *For a fixed level of participation given by vector γ, there exists an information equilibrium in synthetic assets in which traders optimally make information acquisition decisions across all assets in which they participate and all conditions of Lemma I.1 are satisfied. The equilibrium is unique due to the monotonicity of $f_i(\tilde{\gamma}_i, \tilde{\lambda}_i)$ in $\tilde{\lambda}_i$.*

The equilibrium highlighted in Proposition I.1 will be characterized by the following conditions. For each synthetic asset i, the following pairs $(\tilde{\gamma}_i, \tilde{\lambda}_i)$ and price functions $\tilde{P}_i$ define the information equilibrium:

$$\begin{cases} \text{If } f_i(\tilde{\gamma}_i, 0) \geq 1 \text{ then } \tilde{P}_i(\tilde{\gamma}_i, 0, \tilde{\theta}_i, \tilde{x}_i) \text{ and } (\tilde{\gamma}_i, 0) \text{ is an information equilibrium} \\ \text{If } f_i(\tilde{\gamma}_i, \tilde{\lambda}_i^*) = 1 \text{ then } \tilde{P}_i(\tilde{\gamma}_i, \tilde{\lambda}_i^*, \tilde{\theta}_i, \tilde{x}_i) \text{ and } (\tilde{\gamma}_i, \tilde{\lambda}_i^*) \text{ is an information equilibrium} \\ \text{If } f_i(\tilde{\gamma}_i, 1) \leq 1 \text{ then } \tilde{P}_i(\tilde{\gamma}_i, 1, \tilde{\theta}_i, \tilde{x}_i) \text{ and } (\tilde{\gamma}_i, 1) \text{ is an information equilibrium} \end{cases}$$

I refer to the first equilibrium as "fully-uninformed", the second as the "interior," and the third as the "fully-informed" information equilibrium. Proposition I.1 directly follows from Lemma I.2. To understand the forces present in the information equilibrium, it is helpful to proceed as in GS80 by defining the following measures for each synthetic asset i:

$$\tilde{m}_i = \left(\frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\right)^2 \frac{\sigma^2_{\tilde{x}_i}}{\sigma^2_{\tilde{\theta}_i}} \tag{1.16}$$

$$\tilde{n}_i = \frac{\sigma^2_{\tilde{\theta}_i}}{\sigma^2_{\tilde{u}_i}} \tag{1.17}$$

It is simple to demonstrate that the squared correlation between the function of price $S_i(\tilde{P}_i)$ that contains all information about $\tilde{\theta}_i$ and $\tilde{\theta}_i$ is

$$\rho^2_{S_i(\tilde{P}_i),\tilde{\theta}_i} = \frac{1}{1+\tilde{m}_i} \tag{1.18}$$

Therefore, price informativeness is inversely related to $\tilde{m}_i$. Specifically, holding all model parameters fixed, the total number of informed traders in asset i given by $(1-\tilde{\gamma}_i)\tilde{\lambda}_i$ determines the informativeness of the price system. Furthermore, $\tilde{n}_i$ determines the quality of the informed trader's information because:

$$\rho^2_{\tilde{\theta}_i,\tilde{v}_i} = \frac{\tilde{n}_i}{1+\tilde{n}_i} \tag{1.19}$$

As in GS80, utilizing Lemma I.2 for an interior equilibrium, I arrive at

$$\tilde{m}_i = \frac{\exp\left(\frac{2\tilde{c}_i}{\tau}\right) - 1}{1+\tilde{n}_i - \exp\left(\frac{2\tilde{c}_i}{\tau}\right)} \tag{1.20}$$

Applying Eq. (1.18)

$$1-\rho^2_{S_i(\tilde{P}_i),\tilde{\theta}_i} = \frac{\exp\left(\frac{2\tilde{c}_i}{\tau}\right) - 1}{\tilde{n}_i} \tag{1.21}$$

Eq. (1.20) and (1.21) highlight the properties of the interior information equilibrium.

Falling information acquisition costs, higher information quality, and lower risk aversion (higher τ) all lead to a more informative price system. Any change in parameters while keeping information acquisition cost, risk aversion, and information quality constant does not change the price informativeness of the price system. In an information equilibrium, for a fixed level of participation $1-\tilde{\gamma}_i$, changes in these parameters would lead to a change in fraction informed $\tilde{\lambda}_i$ to exactly offset their effect on price informativeness.

Proposition I.1 also allows me to determine the relationship between levels of participation and information acquisition by the participating traders. For each synthetic asset i, given an exogenous participation level $1-\tilde{\gamma}_i$, the following conditions specify the amount of information acquisition that takes place among the participating traders:

$$\begin{cases} \text{If } \tilde{c}_i \geq \frac{\tau}{2}\log(1+\tilde{n}_i) & \text{then } \tilde{\lambda}_i^* = 0 \\ \text{If } \frac{\tau}{2}\log(1+\frac{\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{n}_i+\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}\tilde{n}_i) < \tilde{c}_i < \frac{\tau}{2}\log(1+\tilde{n}_i) & \text{then } \tilde{\lambda}_i^* = \frac{\sigma_{\tilde{u}_i}\sigma_{\tilde{x}_i}}{(1-\tilde{\gamma}_i)\tau}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1}-\frac{1}{\tilde{n}_i}} \\ \text{If } \tilde{c}_i \leq \frac{\tau}{2}\log(1+\frac{\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{n}_i+\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}\tilde{n}_i) & \text{then } \tilde{\lambda}_i^* = 1 \end{cases} \tag{1.22}$$

For sufficiently high information acquisition costs (relative to levels of risk aversion and information quality) no trader is incentivized to acquire information regardless of participation levels, resulting in a fully-uninformed equilibrium. In this equilibrium price informativeness does not change in response to changing participation levels. For a sufficiently low level of participation $1-\tilde{\gamma}_i$, even with all participating traders acquiring information, it is still worthwhile to do so, resulting in a fully-informed equilibrium. In this equilibrium, by Eq. (1.18), greater levels of participation result in more informative prices (lower m_i). Finally, with a sufficiently high level of participation and sufficiently low information costs, the expected utility of being informed is exactly equal to the expected utility of being uninformed, resulting in an interior equilibrium. In this equilibrium, greater participation levels lead to less information acquisition by participating traders, leading to a fixed number of informed traders and a constant price informativeness.

1.4.4 Endogenous Information and Participation

By Proposition I.1, for every level of participation there is an information equilibrium. We now have to solve for the optimal participation decision, which in tandem with an information equilibrium will be referred to as the "overall equilibrium." A trader's decision to participate in a particular asset is driven by the tradeoff between the expected utility gains to adding another risky asset i to their portfolio and the utility losses associated with the participation cost. Since the asset is in an information equilibrium, a trader will also take into consideration optimal information acquisition: to the extent that it is beneficial to acquire information (i.e. in a fully-informed equilibrium), the trader will do so. I utilize Eq. (1.13) to evaluate the ratio of the unconditional expected utility of a trader participating in asset i (information-participation set I_j^*, fundamental information set F_j^*) to the expected utility of a trader not participating in asset i (information-participation set I_j):

$$\frac{EU_{I_j^*}}{EU_{I_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|I^*_{j_i}}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}} \exp\left(\frac{\mathbb{1}_{F^*_{j_i}}\tilde{c}_i}{\tau} + \frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right) \tag{1.23}$$

Since all traders participate in the index asset for free, the participation decision only affects the stock picking strategies ($i > 1$). Optimal information acquisition is provided by conditions specified in Eq. (1.22). Furthermore, because the stock picking strategies are in zero expected supply, the unconditional risk-premium for each stock picking strategy is zero (see Appendix A.2). Taken together, I can rewrite the ratio of expected utilities of participating and not-participating in stock picking strategy $i > 1$ as:

$$\begin{cases} \text{If } \tilde{c}_i \geq \frac{\tau}{2}\log(1+\tilde{n}_i) & \text{then } \frac{EU_{I_j^*}}{EU_{I_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}}\exp\left(\frac{\tilde{k}_i}{\tau}\right) \\ \text{If } \frac{\tau}{2}\log(1+\frac{\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{n}_i+\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}\tilde{n}_i) < \tilde{c}_i < \frac{\tau}{2}\log(1+\tilde{n}_i) & \text{then } \frac{EU_{I_j^*}}{EU_{I_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}}\exp\left(\frac{\tilde{k}_i}{\tau}\right) \\ \text{If } \tilde{c}_i \leq \frac{\tau}{2}\log(1+\frac{\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{n}_i+\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}\tilde{n}_i) & \text{then } \frac{EU_{I_j^*}}{EU_{I_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}}\exp\left(\frac{\tilde{k}_i}{\tau}+\frac{\tilde{c}_i}{\tau}\right) \end{cases} \tag{1.24}$$

I summarize Eq. (1.24) in the following Lemma.

Lemma I.3 (Utility Change Due to Participation)**.** *Suppose participation $1-\tilde{\gamma}_i$ is fixed for each synthetic asset i and consider the resulting information equilibrium as specified by Proposition I.1. In this equilibrium the ratio of the unconditional expected utility of a trader j participating in asset i (information-participation set I_j^*) to not participating in asset i (information-participation set I_j) is as follows:*

1. *In a fully-uninformed or interior information equilibrium:*

$$\frac{EU_{I_j^*}}{EU_{I_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}}\exp\left(\frac{\tilde{k}_i}{\tau}\right) \equiv g_i(\tilde{\gamma}_i,\tilde{\lambda}_i)$$

2. *In a fully-informed equilibrium*

$$\frac{EU_{I_j^*}}{EU_{I_j}} = \sqrt{\frac{\sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}}\exp\left(\frac{\tilde{c}_i}{\tau}+\frac{\tilde{k}_i}{\tau}\right) \equiv g_i(\tilde{\gamma}_i,\tilde{\lambda}_i)$$

The utility lost due to participation and, potentially, information costs (exponential term) is offset by the diversification benefits of adding another asset to the portfolio (ratio of conditional and unconditional standard deviations). I can now formulate the equilibrium in information and participation.

Proposition I.2 (Equilibrium: Endogenous Information and Participation)**.** *There exists an overall equilibrium in synthetic assets in which for every asset i traders optimally make information and participation decisions resulting in $1-\tilde{\gamma}_i^*$ traders optimally participating and, conditional on levels of participation, $\tilde{\lambda}_i^*$ of the participating traders optimally acquiring information as specified by the conditions of Eq.* (1.22)*. The ratio of the utility change due to participation is monotone and decreasing in $\tilde{\gamma}_i$ ensuring the uniqueness of an overall equilibrium within the class of linear price functions.*

Proposition I.2 highlights the *strategic substitutability in participation.* In an information equilibrium, the value of participating in stock picking decreases with participation. The more traders choose to participate in the stock picking strategies, the less valuable it is to participate. Conversely, the fewer traders invest in the stock picking strategies, the more valuable it is to be active in them. If even with all traders participating in synthetic asset i the gains to participation outweigh the costs, all traders will participate in equilibrium: $\tilde{\gamma}_i^* = 0$. These results stems from the intuition that optimally informed and uninformed traders benefit at the expense of noise traders. The amount of trading by the noise traders is exogenous and results in a fixed profit pool for the participating active traders. To the extent that there are more informed and uninformed traders in a particular asset, the profit pool on a per-trader basis shrinks resulting in the diminishing attractiveness of participating in the asset to begin with.

An immediate consequence of the proof of Proposition I.2, are closed-form solutions for equilibrium participation levels as outlined in Corollary I.3.

Corollary I.3 (Closed-Form Solutions for Equilibrium Participation Levels)**.** *The closed-form solution for the equilibrium participation level in any synthetic asset $i > 1$,* *in which all conditions of Proposition I.2 are satisfied, is as follows:*

$$\begin{cases} \text{If } \frac{1}{\exp(2\tilde{c}_i/\tau)-1} \leq \frac{1}{\tilde{n}_i} \text{ then } \tilde{\gamma}_i = \max\left\{0, 1 - \frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{\tilde{n}_i+1}{\exp(2\tilde{k}_i/\tau)-1}}\right\} \\ \text{If } \frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\exp(2(\tilde{k}_i+\tilde{c}_i)/\tau)-1} < \frac{1}{\tilde{n}_i} < \frac{1}{\exp(2\tilde{c}_i/\tau)-1} \text{ then } \tilde{\gamma}_i = \\ \max\left\{0, 1 - \frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}}\left(\frac{\exp(\tilde{c}_i/\tau)}{\sqrt{\exp(2\tilde{k}_i/\tau)-1}} - \sqrt{\tilde{n}_i+1-\exp(2\tilde{c}_i/\tau)}\right)\right\} \\ \text{If } \frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\exp(2(\tilde{k}_i+\tilde{c}_i)/\tau)-1} \geq \frac{1}{\tilde{n}_i} \text{ then } \tilde{\gamma}_i = \max\left\{0, 1 - \frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{1}{\exp(2(\tilde{k}_i+\tilde{c}_i)/\tau)-1}}\right\} \end{cases} \tag{1.25}$$

For a fixed level of risk aversion and information quality, participation levels are determined by the relative magnitude of information and participation costs for the stock picking strategies. All else equal, falling information acquisition costs move the information equilibrium from fully-uninformed (first case), to an interior (second case), to a fully-informed (third case) equilibrium. Similarly, a rising participation cost moves the equilibrium from interior to a fully-informed equilibrium. When the equilibrium participation value is in the interior, I can utilize Eq. (1.21) and (A.18) to rewrite $\tilde{\gamma}_i^*$ as

$$\tilde{\gamma}_i^* = 1 - \frac{\sigma_{\tilde{x}_i}\sqrt{1-\rho^2_{S_i(\tilde{P}_i),\tilde{\theta}_i}}}{\tau\sqrt{\exp(2\tilde{k}_i/\tau)-1}}\left(\sigma_{\tilde{v}_i|\tilde{P}_i} - \sqrt{\sigma^2_{\tilde{v}_i} - \sigma^2_{\tilde{v}_i|\tilde{P}_i}}\sqrt{\exp(2\tilde{k}_i/\tau)-1}\right) \tag{1.26}$$

Since, $\rho_{S_i(\tilde{P}_i),\tilde{\theta}_i}$ and $\sigma_{\tilde{v}_i|\tilde{P}_i}$ both determine price informativeness, in an interior equilibrium participation is jointly determined with price informativeness, risk aversion, participation costs, and noise trading variability. All else equal, in an interior equilibrium, higher price informativeness correlates with lower participation and vice-versa. Note that in the case of no informed traders, $\rho^2_{S_i(\tilde{P}_i),\tilde{\theta}_i} = 0$, $\sigma^2_{\tilde{v}_i} = \sigma^2_{\tilde{v}_i|\tilde{P}_i}$, and Eq. (1.26) simplifies to the solution for the fully-uninformed equilibrium.

The emergence of uninformed stock picking in the uninformed and interior information equilibria with endogenous participation occurs due to the mean-variance inefficiency of the index asset for an uninformed investor. In models of noisy rational expectations, the expected market portfolio may not be the optimal portfolio for an uninformed investor due to the noise in security-level supply (e.g. Admati (1985), Biais et al. (2010), Garleanu and Pedersen (2020)). There is value in deviating from the passive index as such deviation gives investors access to the noise trading demands in the underlying securities, even if the deviation is based only on market prices.[20] Uninformed stock pickers are able to more accurately infer the supply of individual securities and benefit from providing liquidity to the security-level noise traders.

1.4.5 Participation Levels and Passive Share

In defining passive share, I take the view that passive investors:

1. *Optimally free-ride on the information acquisition efforts of active investors in the market index, i.e. invest in the index asset without information acquisition.*

2. *Forgo stock picking, i.e. do not participate in the stock picking strategies.*

The conditions above map naturally to the framework at hand. Within my model there are up to $2 \times 3^{N-1}$ possible participation and information strategies corresponding to information acquisition in the index asset (2 possibilities total) and participation and information acquisition in the stock picking strategies (3 possibilities per asset across $N - 1$ assets). Portfolio allocations are defined probabilistically and map to the fractions (equivalently, numbers) of traders allocating to each portfolio. Since the allocations across synthetic assets are independent by construction, the fraction of traders pursuing both conditions 1 and 2 is the fraction of traders pursuing condition 1

[20] Pedersen (2018) provides various examples for why a passive fund may be at a disadvantage relative to an active fund when maintaining a value-weighted index. Examples include share buybacks, secondary offerings, index additions and deletions during which passive investors will trade at less favorable prices than active investors.

times the fraction of traders pursuing condition 2. Passive share Γ within the structure of the model equals to the number of traders uninformed in the market index and not participating in any stock picking strategy:

$$\Gamma(\tilde{c},\tilde{k}) = \underbrace{\big(1-\tilde{\lambda}_1(\tilde{c}_1)\big)}_{\text{Fraction uninformed in the index asset}} \times \underbrace{\prod_{i=2}^{N}\tilde{\gamma}_i(\tilde{k}_i,\tilde{c}_i)}_{\text{Fraction not participating in stock picking strategies}} \tag{1.27}$$

The former term specifies the fraction (i.e. number) of uninformed traders in the market index (in which all traders participate) and is defined by Eq. (1.22), which can be rewritten as:

$$\begin{cases} \text{If } \tilde{c}_1 \geq \frac{\tau}{2}\log(1+n_1) & \text{then } \tilde{\lambda}_1 = 0 \\ \text{If } f_1(0,\tilde{\lambda}_1) = 1 & \text{then } \tilde{\lambda}_1 = \frac{\sigma_{\tilde{u}_1}\sigma_{\tilde{x}_1}}{\tau}\sqrt{\frac{1}{\exp(2\tilde{c}_1/\tau)-1} - \frac{1}{n_1}} \\ \text{If } \tilde{c}_1 \leq \frac{\tau}{2}\log(1+\frac{\sigma^2_{\tilde{u}_1}\sigma^2_{\tilde{x}_1}}{\tau^2 n_1+\sigma^2_{\tilde{u}_1}\sigma^2_{\tilde{x}_1}}n_1) & \text{then } \tilde{\lambda}_1 = 1 \end{cases} \tag{1.28}$$

The latter term specifies the fraction (i.e. number) of traders not participating in the stock picking strategies as provided by Eq. (1.25):

$$\begin{cases} \text{If } \frac{1}{\exp(2\tilde{c}_i/\tau)-1} \leq \frac{1}{\tilde{n}_i} \text{ then } \tilde{\gamma}_i = \max\left\{0, 1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{\tilde{n}_i+1}{\exp(2\tilde{k}_i/\tau)-1}}\right\} \\ \text{If } \frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\exp(2(\tilde{k}_i+\tilde{c}_i)/\tau)-1} < \frac{1}{\tilde{n}_i} < \frac{1}{\exp(2\tilde{c}_i/\tau)-1} \text{ then } \tilde{\gamma}_i = \\ \max\left\{0, 1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}}\left(\frac{\exp(\tilde{c}_i/\tau)}{\sqrt{\exp(2\tilde{k}_i/\tau)-1}} - \sqrt{\tilde{n}_i+1-\exp(2\tilde{c}_i/\tau)}\right)\right\} \\ \text{If } \frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\exp(2(\tilde{k}_i+\tilde{c}_i)/\tau)-1} \geq \frac{1}{\tilde{n}_i} \text{ then } \tilde{\gamma}_i = \max\left\{0, 1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{1}{\exp(2(\tilde{k}_i+\tilde{c}_i)/\tau)-1}}\right\} \end{cases} \tag{1.29}$$

I can now identify the effects of information and participation costs on passive share.

1.5 Comparative Statics

I investigate the effects of information acquisition costs and participation costs on the equilibrium level of passive investing as given by Eq. (1.27), (1.28), and (1.29) and price informativeness. Price informativeness is defined as the inverse payoff variance conditional on prices. As demonstrated in Appendix A.3, conditional variances are provided by:

$$\begin{cases} \text{Fully-uninformed equilibrium:} & \sigma^2_{\tilde{v}_i|\tilde{P}_i} = \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} \\ \text{Interior equilibrium:} & \sigma^2_{\tilde{v}_i|\tilde{P}_i} = \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} \\ \text{Fully-informed equilibrium:} & \sigma^2_{\tilde{v}_i|\tilde{P}_i} = \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \frac{\sigma^2_{\tilde{\theta}_i}}{1+\frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2\sigma^2_{\tilde{\theta}_i}(1-\tilde{\gamma}_i)^2}} \end{cases} \tag{1.30}$$

1.5.1 Information Acquisition Costs and Passive Share

The primary goal of this paper is to identify the relationship between the costs of acquiring fundamental information and passive investing. According to my definition, a passive trader must: 1) free-ride on the information acquisition of others in the index asset and 2) not participate in stock picking. The traditional setting of GS80, sheds light on the first requirement. To the extent that passive investors are information free-riders in the index, greater information acquisition costs lead to more uninformed traders, which leads to a greater passive share and lower price informativeness.

However, GS80 are silent on the effects of information costs on participation in the trading of a risky asset. What differentiates my work from GS80 is the incorporation of a participation cost into the costly information framework.[21] In GS80 participation is free, therefore all traders invest in the asset. However, in reality, participating in

[21]Peress (2005) also does this, however, his model contains a single risky asset and is designed to explore aggregate stock market participation. As in GS80, within his framework, greater information costs would also lead to more passive investing due to the single-asset nature of the model.

stock picking is a costly endeavor relative to indexing. Stock-picking funds usually have higher turnover, incur greater trading commissions, and participate in stocks which may have a greater market impact relative to an index asset. Therefore, it seems reasonable to assume that in addition to potentially incurring a cost to acquire fundamental information, traders must incur non-zero participation costs to invest in the stock picking strategies.

I apply this reasoning to the stock picking strategies and assume that stock picking requires some trading cost over-and-above a simple indexing strategy. In this setting, falling information costs for stock picking affect the attractiveness of participating in stock picking unconditionally due to two forces. On one hand, as information becomes cheaper, more participating traders become informed and prices become more informative. On the other hand, precisely due to the fact that more participating traders become informed and prices are more informative, participating in stock picking becomes less attractive for all participating traders. Lemma I.4, demonstrated in Appendix A.5, highlights the relationship between gains to participation and information costs under the assumption of fixed participation.

Lemma I.4 (Effect of Information Costs on Unconditional Expected Utility)**.** *For any stock picking strategy i ($i > 1$) and a fixed level of non-participation $\tilde{\gamma}_i < 1$, changes to information cost $\tilde{c}_i$ have the following effect on the unconditional expected utility of a participating trader j (information-participation set I_j^*) across the three information equilibria:*

1. *Fully-uninformed:* $\frac{EU_{I_j^*}}{\partial \tilde{c}_i} = 0$
2. *Interior:* $\frac{EU_{I_j^*}}{\partial \tilde{c}_i} > 0$
3. *Fully-informed:* $\frac{EU_{I_j^*}}{\partial \tilde{c}_i} < 0$

As information costs fall, some previously uninformed active traders become informed (as in GS80), while others optimally choose to not participate in the trading of the asset. The net effect of these two forces is to equate the expected utility gains to participation in stock picking and the utility losses to participation costs. Price informativeness continues to improve since the total number of informed traders continues to grow with lower information costs.

Below a certain level of information acquisition costs for the stock picking strategies, all participating traders become informed as specified by the final condition of Eq. (1.25). Beyond this point, a further decline in information costs accrues directly to the participating traders without affecting price informativeness, since the number of informed traders remains the same. At this point, participating in stock picking becomes attractive yet again, and some of the traders on the sidelines rotate back into the stock picking strategy by paying both the participation and the information acquisition cost. The number of investors switching back is exactly such that the expected utility gain to participation is offset by the utility losses due to the costs spent. Price informativeness continues to improve with lower costs until full participation in the stock picking strategy is achieved.

The discussion above assumes an interior or fully-informed equilibrium in which information costs affect trading decisions. In the corner solutions, with full participation, falling information costs have no impact on price informativeness because the total number of informed traders does not change in these regions. Proposition I.4 summarizes the effects of information acquisition costs on passive share and the resulting price informativeness.

Proposition I.4 (Effect of Information Costs on Passive Share and Price Informativeness)**.** *Equilibrium passive share* $\Gamma(\tilde{c}, \tilde{k})$ *defined by Eq.* (1.27) *has the following relationship with information costs across the three information equilibria for the index* *and stock picking strategies ($i > 1$) respectively:*

1. *Fully-uninformed:* $\frac{\partial \Gamma(\tilde{c},\tilde{k})}{\partial \tilde{c}_1} = 0, \frac{\partial \Gamma(\tilde{c},\tilde{k})}{\partial \tilde{c}_i} = 0$
2. *Interior:* $\frac{\partial \Gamma(\tilde{c},\tilde{k})}{\partial \tilde{c}_1} > 0, \frac{\partial \Gamma(\tilde{c},\tilde{k})}{\partial \tilde{c}_i} \leq 0$
3. *Fully-informed:* $\frac{\partial \Gamma(\tilde{c},\tilde{k})}{\partial \tilde{c}_1} = 0, \frac{\partial \Gamma(\tilde{c},\tilde{k})}{\partial \tilde{c}_i} \geq 0$

Concurrently, price informativeness changes as follows:

1. *Fully-uninformed:* $\frac{\partial \sigma^{-2}_{\tilde{v}_1|\tilde{P}_1}}{\partial \tilde{c}_1} = 0, \frac{\sigma^{-2}_{\tilde{v}_i|\tilde{P}_i}}{\partial \tilde{c}_i} = 0$
2. *Interior:* $\frac{\partial \sigma^{-2}_{\tilde{v}_1|\tilde{P}_1}}{\partial \tilde{c}_1} < 0, \frac{\partial \sigma^{-2}_{\tilde{v}_i|\tilde{P}_i}}{\partial \tilde{c}_i} \leq 0$
3. *Fully-informed:* $\frac{\partial \sigma^{-2}_{\tilde{v}_1|\tilde{P}_1}}{\partial \tilde{c}_1} = 0, \frac{\partial \sigma^{-2}_{\tilde{v}_i|\tilde{P}_i}}{\partial \tilde{c}_i} \leq 0$

The results of Proposition I.4 with regards to passive share follow directly from the definition of passive share (Eq. (1.27)), the closed form solution for the uninformed fraction in the index asset (Eq. (1.28)), and the closed form solution for the stock picking participation levels (Eq. (1.29)). The results relating to price informativeness follow from the equilibrium passive share and Eq. (1.30).

1.5.2 Participation Costs and Passive Share

Participation costs for stock picking have no effect on the market index, which is free to trade. Therefore, their effect on passive share only comes through participation levels in the stock picking strategies. The higher the cost to participate in stock picking (equivalently the lower the cost to index), the fewer traders participate in the stock picking strategies. This, in turn, increases passive share as more traders optimally forgo participation in the stock picking strategies. Price informativeness is only impacted for sufficiently high participation costs, as dictated by the third condition of Eq. (1.27), for which the stock picking strategy is in a fully-informed equilibrium.

Beyond this cost level, greater participation costs lead to lower participation and fewer informed traders, resulting in deteriorating price informativeness by Eq. (1.30). I summarize the relationship between participation costs and passive investing in the following Proposition.

Proposition I.5 (Effect of Participation Costs on Passive Share and Price Informativeness)**.** *Equilibrium passive share* $\Gamma(\tilde{c}, \tilde{k})$ *defined by Eq.* (1.27) *has the following relationship with participation costs across the three information equilibria for the stock picking strategies* ($i > 1$) *respectively:*

1. *Fully-uninformed:* $\frac{\partial\Gamma(\tilde{c},\tilde{k})}{\partial\tilde{k}_i} \geq 0$

2. *Interior:* $\frac{\partial\Gamma(\tilde{c},\tilde{k})}{\partial\tilde{k}_i} \geq 0$

3. *Fully-informed:* $\frac{\partial\Gamma(\tilde{c},\tilde{k})}{\partial\tilde{k}_i} \geq 0$

Concurrently, price informativeness changes as follows:

1. *Fully-uninformed:* $\frac{\partial\sigma^{-2}_{\tilde{v}_1|\tilde{P}_1}}{\partial\tilde{k}_i} = 0$

2. *Interior:* $\frac{\partial\sigma^{-2}_{\tilde{v}_1|\tilde{P}_1}}{\partial\tilde{k}_i} = 0$

3. *Fully-informed:* $\frac{\partial\sigma^{-2}_{\tilde{v}_1|\tilde{P}_1}}{\partial\tilde{k}_i} \leq 0$

Proposition I.5 follows directly from Eq. (1.27), (1.29), and (1.30).

1.5.3 Asset Pricing Implications of Information and Participation Costs

So far I have discussed the effect that information and participation costs have on equilibrium portfolio allocations and price informativeness. I now explore the effects of changes to these costs on observable return characteristics: market risk and risk premium and individual security expected returns, variances, and covariances.

1.5.3.1 Asset Pricing Implications to the Synthetic Assets

As demonstrated in Appendix A.2, the unconditional expected returns for the synthetic assets are given by

$$\mathbb{E}[\tilde{v}_i - \tilde{P}_i] = \tilde{X}_i \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i \tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \right)^{-1} \tag{1.31}$$

Expected returns are determined by the aggregate expected supply of the risky asset ($\tilde{X}_i$) and the inverse of the weighted average precision of the informed and uninformed participating traders. An immediate consequence is that all stock picking strategies have a risk premium of zero since they are in net zero expected supply. For the market index, index information acquisition costs will affect the risk premium as follows:

Lemma I.5 (Effect of Information Costs on Expected Returns)**.** *Changes to information costs across the three information equilibria for the index and stock picking strategies ($i > 1$) respectively have the following effect on expected returns:*

1. *Fully-uninformed:* $\frac{\partial \mathbb{E}[\tilde{v}_1 - \tilde{P}_1]}{\partial \tilde{c}_1} = 0$, $\frac{\partial \mathbb{E}[\tilde{v}_i - \tilde{P}_i]}{\partial \tilde{c}_i} = 0$
2. *Interior:* $\frac{\partial \mathbb{E}[\tilde{v}_1 - \tilde{P}_1]}{\partial \tilde{c}_1} > 0$, $\frac{\partial \mathbb{E}[\tilde{v}_i - \tilde{P}_i]}{\partial \tilde{c}_i} = 0$
3. *Fully-informed:* $\frac{\partial \mathbb{E}[\tilde{v}_1 - \tilde{P}_1]}{\partial \tilde{c}_1} = 0$, $\frac{\partial \mathbb{E}[\tilde{v}_i - \tilde{P}_i]}{\partial \tilde{c}_i} = 0$

Lemma I.5 is demonstrated in Appendix A.6. In an interior equilibrium, greater information costs lead to a noisier price system resulting in higher expected returns. Lemma I.6, demonstrates the effect of information and participation costs on unconditional return variances:

Lemma I.6 (Effect of Information and Participation Costs on Return Variances)**.** *Changes to information costs and participation costs across the three information equilibria for the index and stock picking strategies ($i > 1$) respectively have the following effect on return volatility:*

1. *Fully-uninformed:* $\frac{\partial \text{Var}(\tilde{v}_1 - \tilde{P}_1)}{\partial \tilde{c}_1} = 0$, $\frac{\partial \text{Var}(\tilde{v}_i - \tilde{P}_i)}{\partial \tilde{c}_i} = 0$, $\frac{\partial \text{Var}(\tilde{v}_i - \tilde{P}_i)}{\partial \tilde{k}_i} \geq 0$

2. *Interior:* $\frac{\partial \text{Var}(\tilde{v}_1 - \tilde{P}_1)}{\partial \tilde{c}_1} > 0$, $\frac{\partial \text{Var}(\tilde{v}_i - \tilde{P}_i)}{\partial \tilde{c}_i} > 0$, $\frac{\partial \text{Var}(\tilde{v}_i - \tilde{P}_i)}{\partial \tilde{k}_i} > 0$

3. *Fully-informed:* $\frac{\partial \text{Var}(\tilde{v}_1 - \tilde{P}_1)}{\partial \tilde{c}_1} = 0$, $\frac{\partial \text{Var}(\tilde{v}_i - \tilde{P}_i)}{\partial \tilde{c}_i} > 0$, $\frac{\partial \text{Var}(\tilde{v}_i - \tilde{P}_i)}{\partial \tilde{k}_i} \geq 0$

Lemma I.6 is demonstrated in Appendix A.6. Greater information and participation costs generally lead to a noisier price system, resulting in greater unconditional variances of returns. Armed with the properties of the expected returns and variances of the synthetic assets, I can analyze the asset pricing implications to the broad market and the underlying securities.

1.5.3.2 Asset Pricing Implications to the Original Assets

It is important to explore the implications of changing information and participation costs on the return properties of the original assets. I reconstruct the original assets using portfolios of the synthetic assets and derive the comparative statics in Appendix A.8.

1.5.4 An Illustration

I now illustrate the equilibrium properties of my model through a simple numerical example. There are two risky assets in the market, resulting in two synthetic assets: the market index and a stock picking strategy. There are up to two possible investment strategies in the index (participating informed or participating uninformed) and three possible investment strategies in the stock picking strategy (not participating, participating informed, participating uninformed), resulting in a maximum of $2 \times 3^1 = 6$ potential information and participation strategies, one of which is purely passive and the others are active. The five active strategies can be categorized based on their participation and information acquisition in stock picking. Figure 2.1 demonstrates the equilibrium fraction of traders pursuing the various investment strategies and the implications to price informativeness for changing information and participation costs, while Figure 1.2 highlights the implications to cross-sectional and aggregate returns. Passive share is in blue and is defined as the fraction of traders optimally allocating to the index portfolio without acquiring information and optimally not participating in the stock picking strategy.

Figure 2.1a demonstrates the effect of a falling stock picking information acquisition cost on equilibrium strategy participation. For sufficiently high information costs, no information acquisition takes place. However, because the benefit of participation in stock picking outweighs its cost, all traders participate in stock picking resulting in no passive share and active portfolios that include uninformed stock picking, highlighted by the orange. As mentioned earlier, this is due to the benefits to deviating from the passive index, even if the deviation is based solely on market prices: there are gains to providing liquidity to the noise traders in the stock-picking assets. As information costs continue to fall they reach a certain threshold at which information acquisition begins to take place for some stock picking participants: at first we see the emergence of active portfolios that contain informed stock picking in yellow. In the initial stage the participation constraint is slack; therefore all traders continue to participate in stock picking, some informed, some not. At a certain level of information costs (just below 2.5 on the graph), prices have gotten so informative that it no longer makes sense for all traders to participate, leading some to exit stock picking just to the point where the participation constraint binds for the rest. This leads to the emergence of both a pure market timing portfolio (in purple) and a passive portfolio (in blue). As information costs fall further, the share of both continues to grow. At a certain point, information costs have fallen so low, that the entire stock picking sector is informed. At this point further decreases to information costs directly accrue to the participating traders since they no longer alter the informativeness of the price system. Here, some traders are once again incentivized to switch back from both pure market timing and passive to stock picking leading to the inverted U-shaped curve for passive share.

Figure 2.1b demonstrates the price informativeness of the stock picking strategy with falling information costs. Initially, price informativeness stays constant since there is no information acquisition in the asset. Once information acquisition begins, price informativeness begins to grow. As the peak passive share is reached and all stock pickers are informed, price informativeness grows at a lower pace because the non-participating traders are the ones switching to participation and acquiring information. Since they have to overcome both the participation cost hurdle and the information acquisition cost hurdle, they do not switch to information acquisition as quickly as originally participating uninformed traders. Price informativeness for the original assets follows a similar path.

As shown in Figure 1.2a, the index risk and return dynamics are not affected by stock picking information costs. However, greater stock picking information costs lead to a noisier price system in the relative values of securities, which results in increasing variances for the returns of the underlying assets and more negative covariances, as demonstrated in Figure 1.2b.

Figures 2.1c and 2.1d show the equilibrium strategy participation and price informativeness in the presence of rising information costs for the index asset. As information costs increase, more and more traders switch from acquiring information to free-riding on the acquisition efforts of others resulting in deteriorating price informativeness and a growing passive share. For sufficiently high levels of index information costs, no one is informed about the market aggregate and changes to costs have no further impact on passive share and price informativeness.

Lower information acquisition in the market index leads to a noisier price system, which directly impacts the market risk premium and variance, shown in Figure 1.2c. This, in turn, directly affects the underlying assets, which are all implicitly long the overall market, resulting in both greater expected returns, greater volatility, and greater covariance.

Figures 2.1e and 2.1f show the equilibrium strategy participation in the presence of a rising stock picking participation cost. For sufficiently low participation costs, the relative benefit of participation outweighs its cost resulting in no passive share since everyone participates in stock picking. In this region price informativeness for neither the stock picking strategy nor the index asset is affected. Beyond a certain participation cost threshold, some traders begin to optimally forgo participation in stock picking, leading to fewer participating traders, resulting in greater benefits to information acquisition for those who remain. As some uninformed participants acquire information, price informativeness remains constant despite the growing share of passive capital. Beyond a certain point, all participants become informed, and further increases to participation costs lead to deteriorating price informativeness as the aggregate share of informed capital starts to shrink.

Participation constraints only affect the stock picking strategy, and thereby have no impact on market risk and return dynamics as demonstrated in Figure 1.2e. Decreasing participation increases the noise in the price system for the stock picking strategy, thereby increasing its return variance. This, in turn, has the effect of (in this example, very slightly) increasing the return variance of both underlying securities and lowering their covariance, shown in Figure 1.2f.

1.6 Empirical Analysis

The primary insight from the theory presented is that falling fundamental information costs have a dual effect on passive share. On one hand, they increase the prevalence of market timing strategies, which reduces passive share. On the other hand, they decrease stock picking participation, which increases passive share. If the stock picking force dominates the market timing force, passive share increases and price informativeness rises due to falling information costs. I provide suggestive evidence for the existence of this channel in the data.

Although the model is static in nature and does not explicitly incorporate the active vs. passive investment management sectors, its underlying logic would suggest the following institutional mechanism. As information costs fall for individual firms, more active investors acquire information about their fundamentals. This greater information acquisition results in prices that are more reflective of future fundamentals, which in turn leads to lower gains to being an active investor. As investors internalize the lower expected performance, some choose to take money out of active funds, resulting in capital outflows from active investment strategies. Finally, some investors will choose to reallocate their funds to passive investment strategies.

Guided by this logic, I identify a quasi-exogenous negative information cost shock and conduct the following tests. First, I explore the performance of active funds around the time of the shock, and exploit its staggered design for causal inference of the effect of information costs on active performance. Second, I analyze the flows out of active funds around the time of the regulation. Third, I examine the flows into passive funds around the same time frame.

1.6.1 The Setting and Test Design

One of the primary goals of the SEC is to "promote efficient and transparent capital markets" while embracing technological advancement (SEC (2009)). Historically, the Commission has pursued this through various regulations, most notably through the implementation of EDGAR in 1993, which required public firms to submit their public filings electronically. The focus of this paper is a more recent regulation aimed at making financial information more useful to investors.

In January of 2009 the SEC announced the adoption of Rule S7-11-08: "Interactive Data to Improve Financial Reporting." The mandate requires firms to supplement their regular financial filings made through EDGAR with financial statements provided in an interactive format utilizing the eXtensible Business Reporting Language. The benefits of this regulation are twofold. First, investors can now directly download financial information into spreadsheets or analyze it with a variety of commercial off-the-shelf software. Second, XBRL relies on a standard taxonomy (i.e. a set of tags for different financial items) developed in accordance with U.S. generally accepted accounting principles (US GAAP) and reviewed by the Financial Accounting Standards Board (FASB) and the Commission. This facilitates the interpretation of financial information, makes it more comparable across firms, and simplifies the automation of financial analysis. To the extent that investors were previously paying third-party data providers for the services of information extraction and aggregation or were devoting meaningful amounts of time to performing these tasks themselves, the regulation at least partially alleviates these burdens. Therefore, I interpret the XBRL mandate as a negative shock to information costs.

To identify the causal effect of lower information costs on passive investing, I rely on the phase-in design of the XBRL mandate. Companies had to comply with the regulation according to the following timeline:

- Phase 1: Filers with public equity float > \$5bn: for 10-Q/K with financial statements for fiscal periods ending on or after June 15, 2009
- Phase 2: Filers with public equity float between \$700mm and \$5bn: for 10-Q/K with financial statements for fiscal periods ending on or after June 15, 2010

- Phase 3: All remaining filers: for 10-Q/K with financial statements for fiscal periods ending on or after June 15, 2011

In my primary tests, I exploit cross-sectional heterogeneity in passive and active fund-level exposure to the XBRL mandate. Specifically, I take the universe of active and passive domestic equity funds and estimate the fraction of their total net assets allocated to companies affected by phase 1, 2, and 3 of the regulation as of the second calendar quarter of 2009. I limit my analysis to funds that were either predominantly invested in phase 1 securities (group 1) or those that were primarily invested in phase 2 and 3 securities (group 2). To test the active fund performance effect of lower information costs, I conduct an event study in which I track the relative performance of group 1 and group 2 active funds. Similarly, I perform event studies on the flows out of group 1 and group 2 active funds, and flows into group 1 and group 2 passive funds.

1.6.2 Data

I rely on EDGAR index files to come up with a sample of XBRL adopters through time. In particular, I utilize the "full index" data, which contain the list of all companies submitting XBRL filings by quarter, including company CIK, filing type, and filing date.[22] This allows me to track companies' adoption of XBRL through time and I classify adopting companies into phase 1, phase 2, or phase 3 based on the dates of their initial XBRL filings and the number of XBRL filings made subsequently. I exclude companies that participated in the voluntary program prior to the ruling. My final sample contains 448 phase 1 firms, 996 phase 2 firms, and 3,353 phase 3 firms with the public equity data required for the analysis presented here.[23,24]

[22] See https://www.sec.gov/edgar/searchedgar/accessing-edgar-data.htm for more information.

[23] The SEC estimated approximately 500 filers to fall into the phase 1 group, however, this number also includes foreign issuers and voluntary adopters.

[24] I thank Ekaterina Volkova (University of Melbourne) for sharing her CIK-CUSIP mapping based on 13D and 13G filings. The mapping served as a helpful complement to the CIK-CUSIP mapping

I utilize the Center for Research in Security Prices Mutual Fund Database (CRSP MFDB) to construct my active and passive mutual fund samples.[25] I limit the analysis to broad market domestic equity funds and include both index and exchange traded funds as passive funds in the main specifications.[26] I only include those active and passive funds whose data fully overlaps with the three year phase-in of the XBRL mandate. I include summary statistics for both the fund-level and fund share-class level samples for active funds (Tables 1.1-1.4 and 1.5-1.7 respectively) and the fund share-class level sample for passive funds (Tables 1.8-1.9).

I utilize the CRSP MFDB fund-level holdings data to estimate XBRL exposure.[27] Specifically, I use a fund's holdings data closest to the beginning of phase 1 in June of 2009, which is typically the holdings disclosure for calendar 2Q2009. I then estimate the exposure of fund i to phase j of XBRL as

$$exposure_{i,j} = \frac{\sum_{k=1}^{N} \mathbb{1}\{\text{security } k \in \text{ phase j}\} \times \$holdings_{i,\approx 6/2009,k}}{TNA_{i,\approx 6/2009}} \tag{1.32}$$

where $j = 1, 2, 3$ specifies the phase of XBRL to which a security k may potentially belong, security $k = 1, \ldots, N$ indexes the securities in which fund i is invested and $TNA_{i,\approx 6/2009}$ are the total net assets for fund i in the month closest to June of 2009. I classify funds into groups (specified by the variable "TRT") based on their estimated exposures to the various phases of XBRL:

provided by Compustat through WRDS.

[25] I gratefully acknowledge Doshi et al. (2015) for making their code publicly available, which was extremely helpful in getting up to speed on using the CRSP MFDB and TR S12 databases.

[26] I include active funds with the following CRSP objective codes: EDCL, EDCM, EDCS, EDCI, EDYG, EDYB, EDYI. Passive funds are those with the same objective codes that were also identified as index funds ($INDEX\_FUND\_FLAG = D$) or exchange traded funds ($ET\_FLAG = F$). I go through the passive funds by hand and take out misclassified active funds.

[27] I also run the analysis using the Thomson Reuters (TR) s12 database with similar results, however I rely on CRSP as my main source since, as pointed out by Zhu (2020), between 2008-2015 58% of new US equity mutual fund classes from CRSP cannot be matched to the TR s12 database and index funds are more likely to be missing in TR.

- Group 1 ($TRT_i = 1$): $exposure_{i,1} \geq 0.75$

- Group 2 ($TRT_i = 0$): $exposure_{i,2} + exposure_{i,3} \geq 0.75$

- Group 3 ($TRT_i = 2$): all other funds with $exposure_{i,1}+exposure_{i,2}+exposure_{i,3} \geq 0.75$

Group 1 (or phase 1) funds are those with oversized exposure to phase 1 of the XBRL mandate. Group 2 (or phase 2-3) funds have oversized exposure to phases 2 and 3 of XBRL. Finally, group 3 (or phase 1-2-3) funds have exposure spread out over the three phases. By construction, I limit the analysis to funds with at least 75% of TNA allocated to XBRL-exposed securities. Tables 1.3,1.7, and 1.10 demonstrate the exposure of active funds on a share-level and fund-level, and passive funds on a share-level to the three phases of the XBRL mandate by group. As can be seen, group 1 and group 2 funds have minimal overlap by phase exposure, with approximately 5-8% TNA exposure overlap between the two groups. By relying on a categorical grouping as opposed to variation in phase exposure through time, I avoid well-documented issues associated with security-level holdings data.

I further rely on CRSP MFDB to extract data on fund total net assets, returns, and expenses. Fund flow for fund i in month t is estimated as common in the literature (e.g. Sirri and Tufano (1998), Lou (2012)):

$$flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}(1 + r_{i,t}) - MRG_{i,t}}{TNA_{i,t-1}} \tag{1.33}$$

where $TNA_{i,t}$ are the total net assets of fund i in month t, $r_{i,t}$ are the returns of fund i in month t, and $MRG_{i,t}$ are the total net assets of any funds acquired by fund i in month t. Following Lou (2012), I use a smoothing procedure to identify merger dates. Consistent with prior studies, fund-month level observations are required to be a minimum of \$15mm TNA, both current month and past month data must exist for flows to be imputable, and flows are winsorized at the 1% level from both tails. Flows analysis is conducted on a fund share-class level basis since flows to and from different share classes (e.g. institutional vs. retail) may vary.

Fund performance is evaluated on a 36-month rolling basis utilizing fund-level returns against the Carhart (1997) four-factor model:

$$r_{i,t}^{e} = \alpha_{i,t} + \beta_{i,t}^{m} r_t^{m} + \beta_{i,t}^{HML} r_t^{HML} + \beta_{i,t}^{SMB} r_t^{SMB} + \beta_{i,t}^{UMD} r_t^{UMD} + \varepsilon_{i,t} \tag{1.34}$$

where $r_{i,t}^{e}$ is the net return of fund i in month t in excess of the one-month Treasury bill rate, r_t^{m}, r_t^{HML}, r_t^{SMB}, r_t^{UMD} are the market, value, size, and momentum factors respectively.[28] I require 36 months of historical data to evaluate the alphas for any given month. Performance analysis is conducted on a fund-level basis to avoid duplicated performance measures for different share classes.

1.6.3 Results

Guided by the proposed mechanism, I ask three questions through the lens of the mutual fund data. First, do lower information costs lead to lower alphas for active funds? Two, if so, are funds flowing out of the affected active strategies? Three, are passive strategies seeing inflows as a result? As such, I conduct tests along three dimensions: 1) active mutual fund performance, 2) active mutual fund flows, and 3) passive fund flows. As a starting point, I examine the viability of the following event study specification to track the difference in these measures through time by month, before, during, and after the XBRL mandate:

$$y_{i,t} = \gamma_0 + \sum_{t=2}^{T} \beta_t (\mathbb{1}_t \times TRT_i) + \gamma_1 x_{i,t-1} + \eta_t + \nu_i \tag{1.35}$$

[28] The data are obtained from the Fama-French Monthly Research Factors dataset available through WRDS and sourced from Ken French's website.

where $y_{i,t} \in \{\alpha_{i,t}, flow_{i,t}\}$ is defined in Eq. (1.33) and (1.34), $x_{i,t-1} \equiv \log(TNA_{i,t-1})$ is a measure of fund size, $\mathbb{1}_t$ is an indicator function for each month t, and η_t and ν_i are month and fund fixed effects. The performance specification would be conducted just for active funds, while the flow specification would be conducted for active and passive funds separately. In all specifications I would track the difference, measured by β_t, in the outcome variable between funds primarily affected by the first phase of the XBRL mandate ($TRT_i = 1$) to those primarily exposed to the second and third phases of the regulation ($TRT_i = 0$). First, I highlight the econometric complications that would arise within the specification above. Second, I propose a solution to these complications.

Fund performance and fund flows may be explained by various factors such as portfolio manager skill, fund size, fund objective, and market conditions. To isolate the effect of information costs on performance and flows, one needs to control for these various observable and unobservable fund-level and time characteristics. The latter is resolved via time fixed effects η_t. One may be inclined to resolve the former via fund fixed effects ν_i and lagged fund size $x_{i,t-1}$. However, as pointed out by Pástor et al. (2015), this results in a finite sample bias due to a contemporaneous correlation between innovations in $\alpha_{i,t}$ or $flow_{i,t}$ and innovations in assets under management. Intuitively, greater performance or greater flows also imply growth in total net assets, which in turn leads to a negative bias in the coefficient capturing the effect of fund size on performance or flows.

To address the finite sample bias I apply a recursive demeaning procedure based on Moon and Phillips (2000), applied to the fund performance vs. size relationship by Pástor et al. (2015), and further updated by Zhu (2018). Borrowing notation from Zhu (2018), I first construct forward-demeaned variables for performance, flows, and fund size:

$$\bar{y}_{i,t} = y_{i,t} - \frac{1}{T_i - t + 1} \sum_{s=t}^{T_i} y_{i,s}$$
$$\bar{x}_{i,t-1} = x_{i,t-1} - \frac{1}{T_i - t + 1} \sum_{s=t}^{T_i} x_{i,s-1}$$

where t spans 1 to $T_i - 1$ with T_i being the month of the final observation for fund i. The forward demeaning procedure absorbs fund fixed effects, however introduces a correlation between the lagged forward-demeaned size and the demeaned innovation in flows or performance. To resolve the latter, I conduct a two-stage least squares analysis whereby in the first stage I regress the forward-demeaned measure of fund size on its current realization

$$\bar{x}_{i,t-1} = \psi + \rho x_{i,t-1} \tag{1.36}$$

and in the second stage I conduct the event study initially introduced in Eq. (1.35) using the fitted values from above

$$\bar{y}_{i,t} = \sum_{t=1}^{T} \beta_t (\mathbb{1}_t \times TRT_i) + \gamma \bar{x}^*_{i,t-1} + \eta_t \tag{1.37}$$

where $\bar{x}^*_{i,t-1}$ is the fitted value from the first stage regression. As pointed out by Pástor et al. (2015), the relevance and exclusion conditions of the instrument are likely met as $x_{i,t-1}$ is embedded in the dependent variable in the first stage and $x_{i,t-1}$ is unlikely to be correlated with innovations to the forward-demeaned performance and flows. The month fixed effects will span $t = 2, \ldots, T$ thereby imposing the zero-intercept condition of the specification above, which is appropriate given the forward-demeaned nature of the variables. The interaction coefficients β_t will capture the monthly differential performance and flows to group 1 versus group 2 funds after accounting for observable and unobservable fund-specific characteristics and aggregate time trends.

The results are highlighted in Figures 1.3-1.5. Figures 1.3a and 1.3b document the time trends and event study results for the performance of active mutual funds. In the pre-treatment period and during phases 1 and 2, the two groups of active funds exhibit broadly similar performance as can be seen both through the time series plot as well as the event study analysis. However, starting in the second half of 2011, the performance of the two groups diverges dramatically with the funds initially affected by XBRL seeing deteriorating performance relative to funds affected in the later stages of the regulation. The result would suggest that lower information costs for phase 1 firms have propagated through the market, leading to greater information acquisition, greater price informativeness, and, ultimately, declining performance by group 1 funds. The process took about two years after the initial implementation of the regulation. In another two years, group 1 performance begins to rebound relative to group 2 performance, suggesting that the lower information costs are now affecting group 2 firms. Finally, as the lower information costs work through the system for group 2 funds, the relative performance of groups 1 and 2 converges.

To ensure that the identified trend in performance is not an artifact of omitted variables or fund-level aggregation issues, I focus on a closely matched sample of group 1 and group 2 funds, paired both by total net assets (group 1 funds' TNA has to be within 10% of group 2 funds' TNA) and CRSP investment objective code. This set of funds is much smaller, as highlighted in table 1.4, however the sample is tightly matched with roughly the same fund characteristics for group 1 and group 2 funds. The inclusion of time fixed effects and the recursive demeaning procedure introduced above further addresses concerns regarding the causal inference. Figure 1.6 demonstrates the results of the analysis. The trends in performance are consistent with the patterns for the full sample and are even more convincing. Taken together, the two sets of results provide suggestive evidence for the negative effects of lower information costs on active fund performance.

The next step in the mechanism are flows out of active funds due to their performance decline. Figures 1.4a and 1.4b display both the cumulative flows out of active funds and the results of the event study as applied to monthly flows. Although the event study does not highlight a clear pattern in differential flows between group 1 and group 2 funds, a visual inspection of the cumulative flows suggests that there is a lead-lag relationship comparable to the patterns in the relative performance of the funds. Outflows are initially greater for group 1 funds, suggestive of investors allocating out of these funds due to their weaker performance. This is followed by greater outflows from group 2 funds, once again in-line with their performance decline.

The final step in the mechanism would suggest that investors allocating out of active funds would shift their freely-available capital to passive funds. Without individual-level transaction data this step is difficult to test since capital flows cannot be traced to individual investors. However, operating under the assumption that investors allocating out of active and into passive would prefer a passive fund with similar characteristics to the active fund (e.g. similar investment objective), the analysis demonstrated in figures 1.5a and 1.5b provides some supportive evidence. Specifically, passive funds in group 1 see greater inflows around the time of under-performance by group 1 active funds. Furthermore, this trend is reversed as group 2 active funds under-perform group 1 active funds. Although this evidence is at most suggestive, it aligns well with the mechanism described.

In summary, the empirical results presented provide causal evidence on the effect of information costs on active performance. When coupled with an investigation of flows out of active and into passive funds a picture emerges that is consistent with the proposed effect of falling information costs on passive investing.

1.7 Conclusion

Many critics of passive investing view capital flows to index products as uninformed capital and worry about the implications to price efficiency and the functionality of financial markets. To assuage these concerns, I document a novel theoretical link between two strong trends in the data, not easily reconcilable within standard models: 1) increasing passive share and 2) decreasing information costs and greater price informativeness.

In my model, traders optimize across two dimensions: information acquisition and stock picking participation. The dual forces of strategic substitutability in information acquisition and participation dictate the equilibrium levels of information acquisition and passive investing. An equilibrium outcome of my framework is that for interior information equilibria, falling fundamental information costs have the joint effect of decreasing passive share via a greater prevalence of market timing strategies and increasing passive share via greater non-participation in stock picking strategies. If the stock picking channel dominates the market timing channel, passive share rises endogenously in tandem with more informative prices. I provide suggestive empirical evidence on the existence of this channel, utilizing the XBRL mandate as a negative shock to information costs.

Alternative drivers for a rising passive share include falling costs to indexing and greater costs to market timing strategies, both of which lead to decreasing price informativeness. Given the differential outcomes to market efficiency, further empirically detangling the drivers of passive investing presents an exciting avenue for future research.

1.8 Tables and Figures

Table 1.1: Active, Fund Level: TNA by Phase Exposure ($mm)

Phase Exposure	count	mean	p5	p25	p50	p75	p95
Phase 1	782	1,814	20	78	287	1,091	6,622
Phase 1-2-3	574	1,058	18	57	172	607	4,060
Phase 2-3	470	564	18	56	184	555	2,056
Total	1826	1,255	18	63	215	789	4,586

Analysis performed for active funds on the fund-level basis. Funds are grouped based on XBRL Phase exposure estimated using available holdings data closest to June 30, 2009.

Table 1.2: Active, Fund Level: Objective by Phase (count)

CRSP Objective Codes	Ph. 1	Ph. 2-3	Ph. 1-2-3	Total
Domestic_Equity_Growth	468	25	211	704
Domestic_Equity_Growth_Income	236	11	72	319
Domestic_Equity_Income	69	2	25	96
Domestic_Equity_Micro_Cap	0	17	3	20
Domestic_Equity_Mid_Cap	8	34	239	281
Domestic_Equity_Small_Cap	1	381	24	406
Total	782	470	574	1826

Analysis performed for active funds on the fund-level basis. Funds are grouped based on XBRL Phase exposure estimated using available holdings data closest to June 30, 2009.

Table 1.3: Active, Fund Level: Exposure by Phase (%)

Phase Exposure	Ph. 1	Ph. 2	Ph. 3	Ph. 2-3	Ph. 1-2-3
Phase 1	85.5	7.3	0.3	7.6	93.1
Phase 1-2-3	51.5	36.8	4.3	41.1	92.6
Phase 2-3	4.8	60.3	27.3	87.5	92.3
Total	54.2	30.1	8.5	38.6	92.8

Analysis performed for active funds on the fund-level basis for the paired sample. Funds are grouped based on XBRL Phase exposure estimated using available holdings data closest to June 30, 2009.

Table 1.4: Active, Fund Level, Paired: TNA by Phase Exposure ($mm)

Phase Exposure	count	mean	p5	p25	p50	p75	p95
Phase 1	21	713	57	106	240	865	2,943
Phase 2-3	21	696	58	106	244	902	2,744
Total	42	704	57	106	242	902	2,943

Analysis performed for active funds on the fund-level basis for the paired sample. Funds are grouped based on XBRL Phase exposure estimated using available holdings data closest to June 30, 2009.

Table 1.5: Active, Class Level: TNA by Phase Exposure ($mm)

Phase Exposure	count	mean	p5	p25	p50	p75	p95
Phase 1	1616	912	17	40	137	549	3,157
Phase 1-2-3	1099	578	17	30	88	302	2,167
Phase 2-3	843	313	16	28	94	303	1,105
Total	3558	667	17	34	109	381	2,329

Analysis performed for active funds on the share class-level basis. Funds are grouped based on XBRL Phase exposure estimated using available holdings data closest to June 30, 2009.

Table 1.6: Active, Class Level: Objective by Phase (count)

CRSP Objective Codes	Ph. 1	Ph. 2-3	Ph. 1-2-3	Total
Domestic_Equity_Growth	905	40	397	1342
Domestic_Equity_Growth_Income	530	20	126	676
Domestic_Equity_Income	164	3	49	216
Domestic_Equity_Micro_Cap	0	25	3	28
Domestic_Equity_Mid_Cap	16	65	477	558
Domestic_Equity_Small_Cap	1	690	47	738
Total	1616	843	1099	3558

Analysis performed for active funds on the share class-level basis. Funds are grouped based on XBRL Phase exposure estimated using available holdings data closest to June 30, 2009.

Table 1.7: Active, Class Level: Exposure by Phase (%)

Phase Exposure	Ph. 1	Ph. 2	Ph. 3	Ph. 2-3	Ph. 1-2-3
Phase 1	85.3	7.2	0.3	7.5	92.8
Phase 1-2-3	52.3	36.4	3.9	40.3	92.6
Phase 2-3	5.0	61.4	25.6	87.0	92.0
Total	56.1	29.0	7.4	36.4	92.5

Analysis performed for active funds on the share class-level basis. Funds are grouped based on XBRL Phase exposure estimated using available holdings data closest to June 30, 2009.

Table 1.8: Passive, Class Level: TNA by Phase Exposure ($mm)

Phase Exposure	count	mean	p5	p25	p50	p75	p95
Phase 1	282	1,739	18	47	176	655	5,277
Phase 1-2-3	46	618	16	29	67	498	2,304
Phase 2-3	154	571	16	26	98	330	3,018
Total	482	1,258	17	37	127	587	4,154

Analysis performed for passive funds on the share class-level basis. Funds are grouped based on XBRL Phase exposure estimated using available holdings data closest to June 30, 2009.

Table 1.9: Passive, Class Level: Objective by Phase (count)

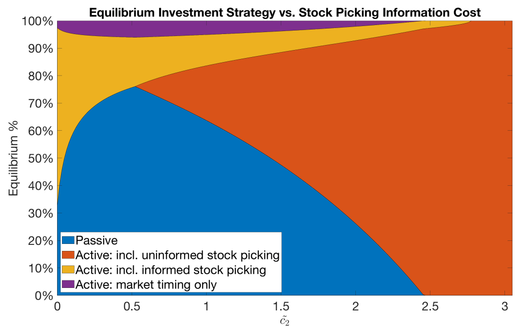
CRSP Objective Codes	Ph. 1	Ph. 2-3	Ph. 1-2-3	Total
Domestic_Equity_Growth	101	4	14	119
Domestic_Equity_Growth_Income	18	0	1	19
Domestic_Equity_Income	10	0	4	14
Domestic_Equity_Large_Cap	153	0	0	153
Domestic_Equity_Micro_Cap	0	3	0	3
Domestic_Equity_Mid_Cap	0	59	27	86
Domestic_Equity_Small_Cap	0	88	0	88
Total	282	154	46	482

Analysis performed for passive funds on the share class-level basis. Funds are grouped based on XBRL Phase exposure estimated using available holdings data closest to June 30, 2009.

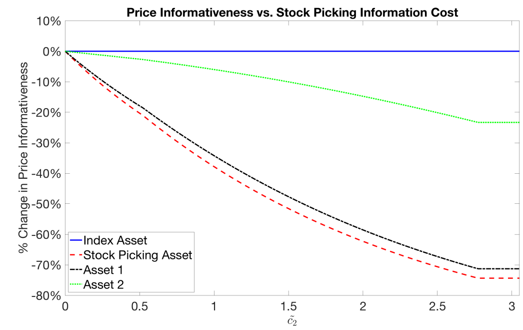
Table 1.10: Passive, Class Level: Exposure by Phase (%)

Phase Exposure	Ph. 1	Ph. 2	Ph. 3	Ph. 2-3	Ph. 1-2-3
Phase 1	88.8	7.2	0.4	7.6	96.5
Phase 1-2-3	47.9	43.5	5.0	48.5	96.4
Phase 2-3	5.3	63.6	25.9	89.5	94.8
Total	58.2	28.7	9.0	37.7	95.9

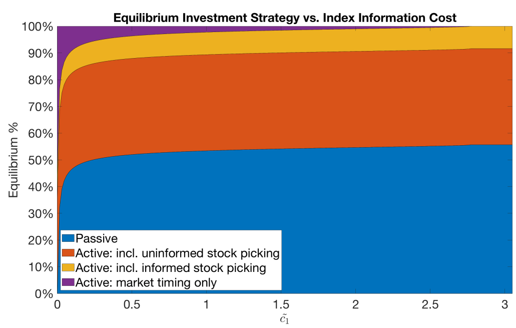
Analysis performed for passive funds on the share class-level basis. Funds are grouped based on XBRL Phase exposure estimated using available holdings data closest to June 30, 2009.



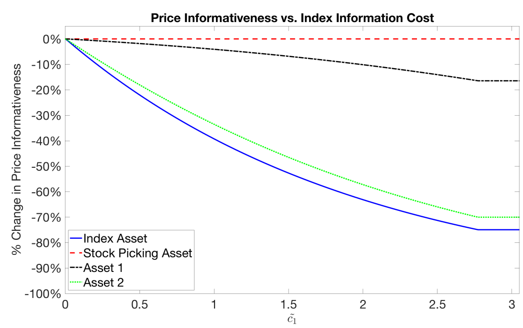
(a) Stock Picking Info. Cost: Participation by Strategy



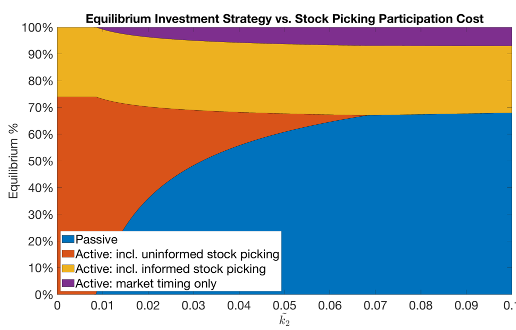
(b) *Rising* Stock Picking Info. Cost: Price Inf.



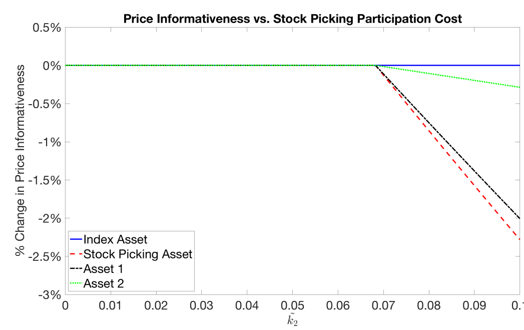
(c) Index Information Cost: Participation by Strategy



(d) *Rising* Index Info. Cost: Price Inf.

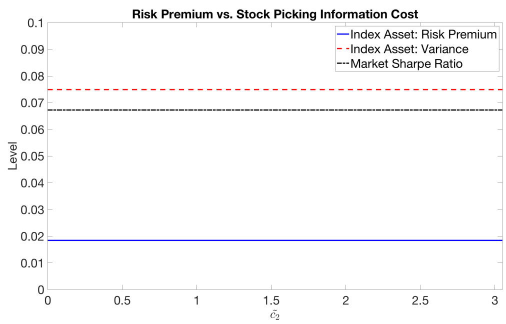


(e) Participation Cost: Participation by Strategy

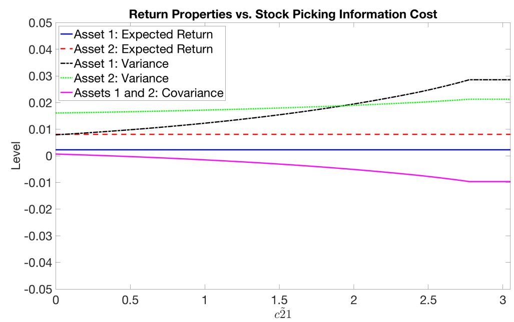


(f) *Rising* Part. Cost: Price Inf.

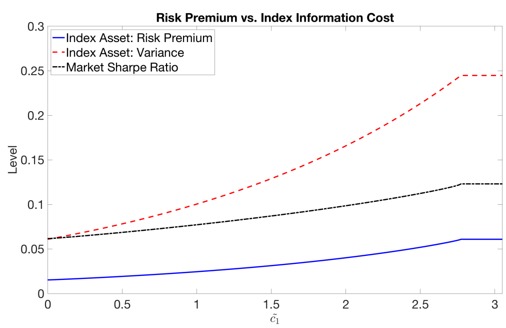
Figure 1.1: Effect of Information and Participation Costs on Equilibrium Strategy Participation and Price Informativeness.



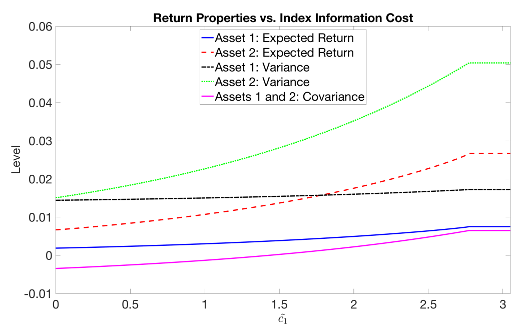
(a) Stock Picking Info. Cost: Market Return Effects



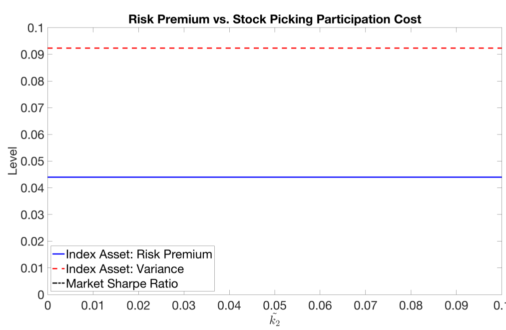
(b) *Rising* Stock Picking Info. Cost: Returns



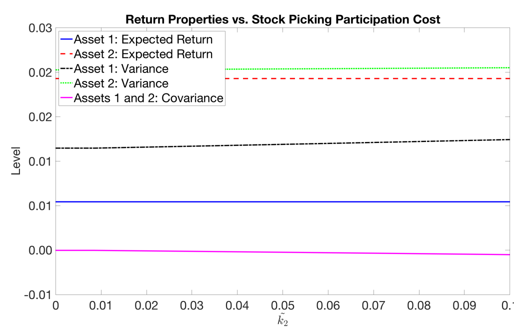
(c) Index Info. Cost: Market Return Effects



(d) *Rising* Index Info. Cost: Returns

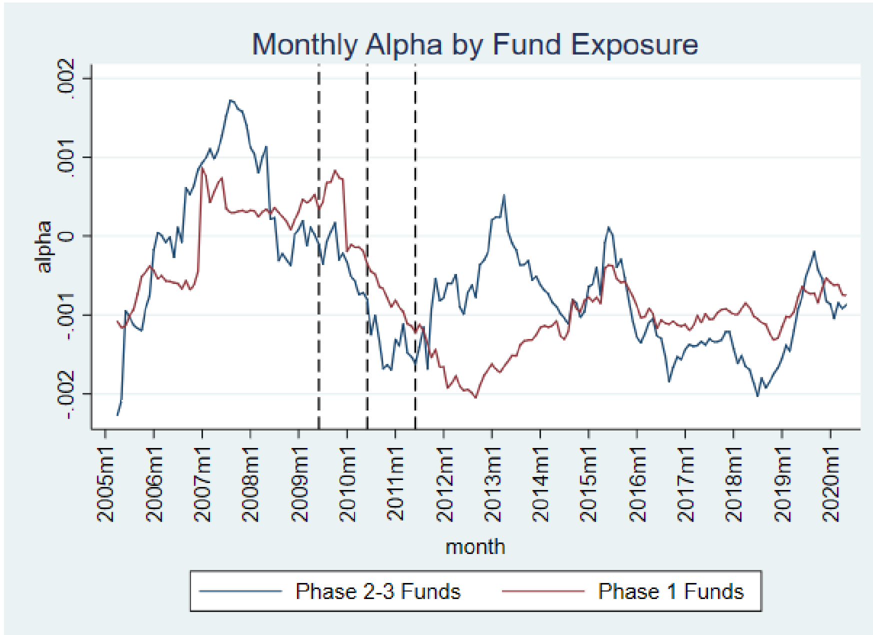


(e) Participation Cost: Market Return Effects

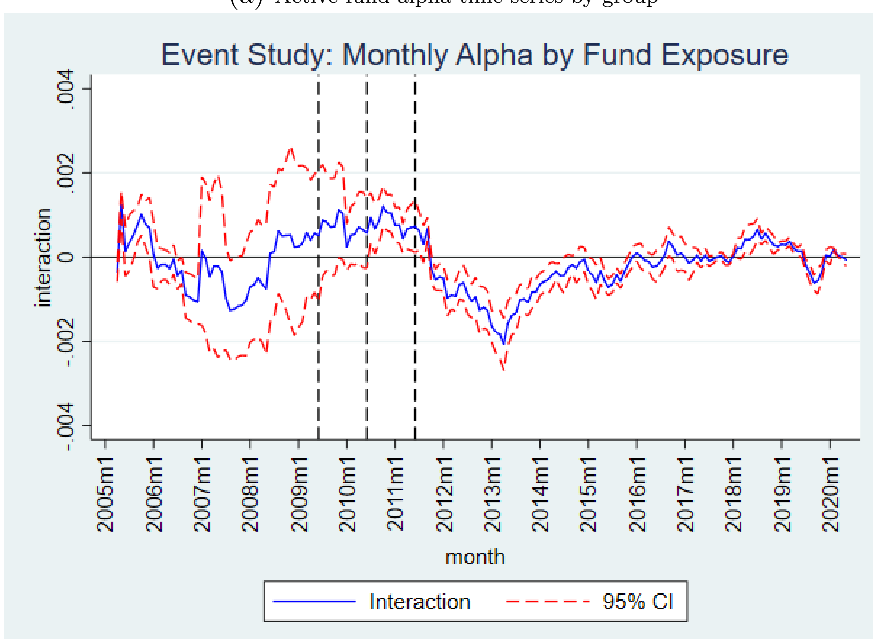


(f) *Rising* Part. Cost: Returns

Figure 1.2: Effect of Information and Participation Costs on the Market Risk Premium and the Cross-Section of Asset Returns.



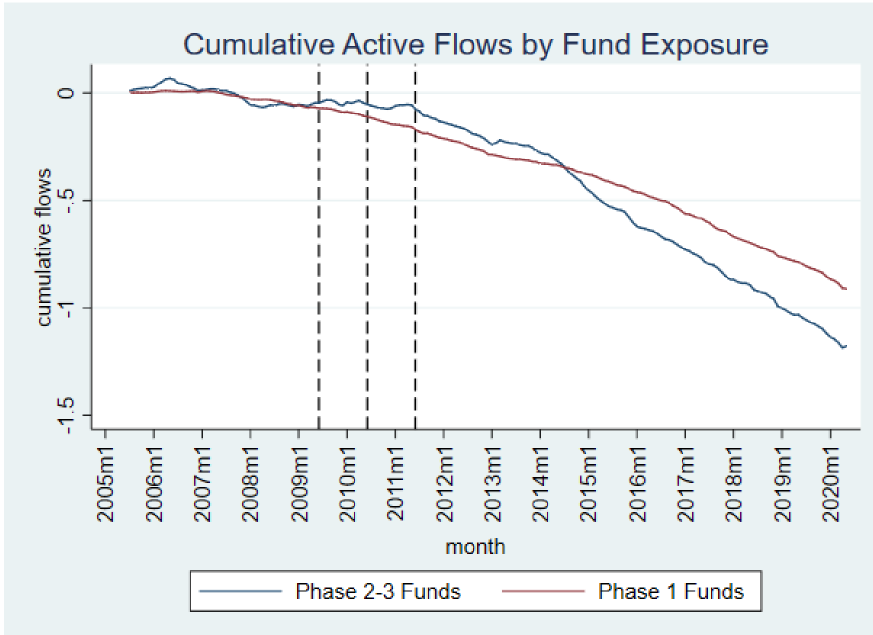
(a) Active fund alpha time series by group



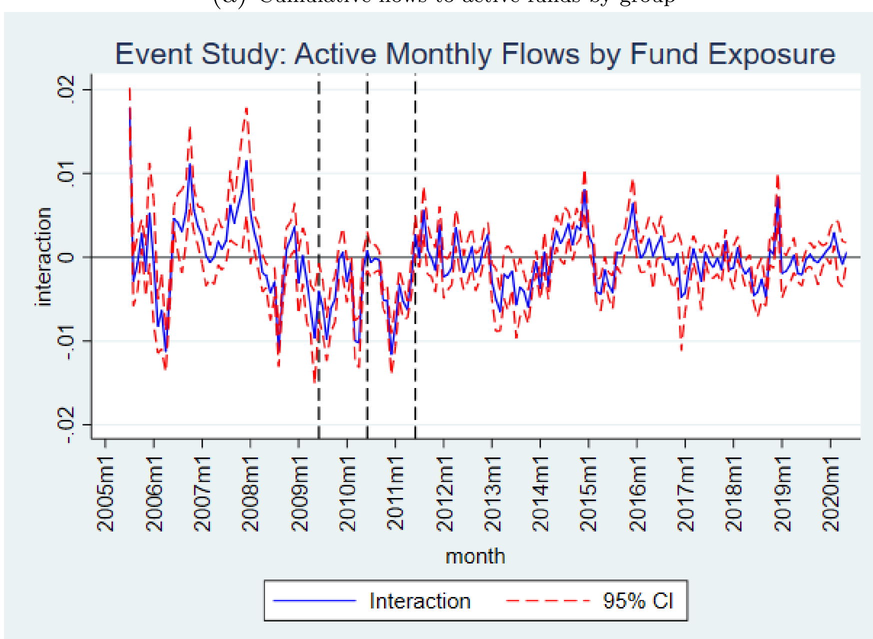
(b) Relative performance of group 1 and 2 funds

Figure 1.3: Effect of the XBRL Mandate on Active Fund Performance.

The dotted lines signify the dates of the initial implementation of each phase of XBRL in sequential order from left to right. Heteroskedasticity-robust standard errors are clustered by month × investment objective.



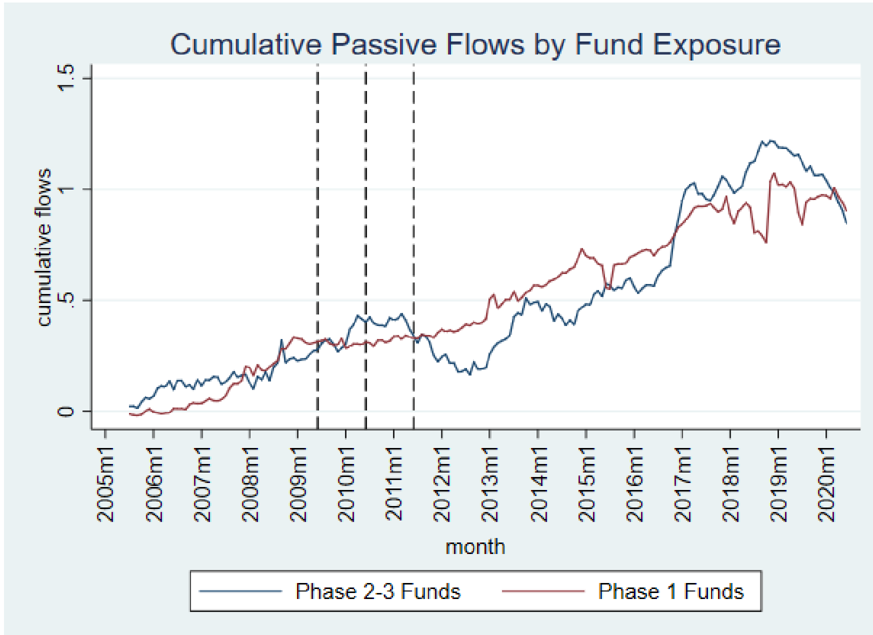
(a) Cumulative flows to active funds by group



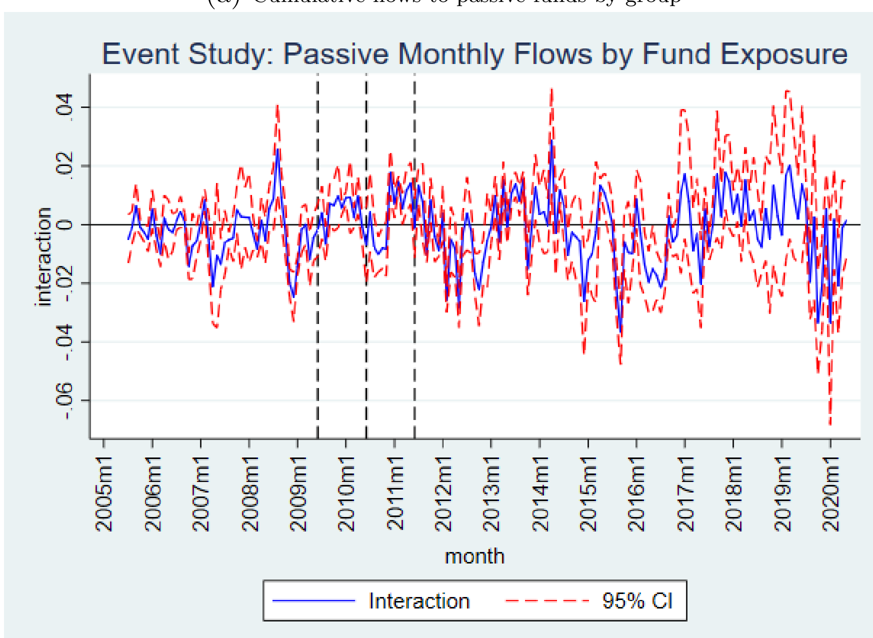
(b) Monthly difference between group 1 and 2 fund flows

Figure 1.4: Effect of the XBRL Mandate on Active Fund Flows.

The dotted lines signify the dates of the initial implementation of each phase of XBRL in sequential order from left to right. Heteroskedasticity-robust standard errors are clustered by month $\times$ investment objective.



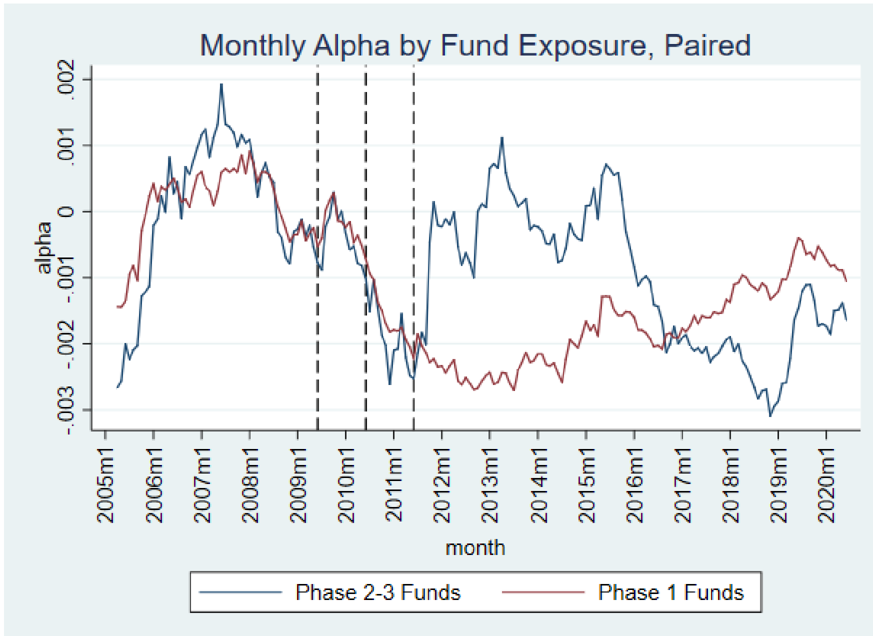
(a) Cumulative flows to passive funds by group



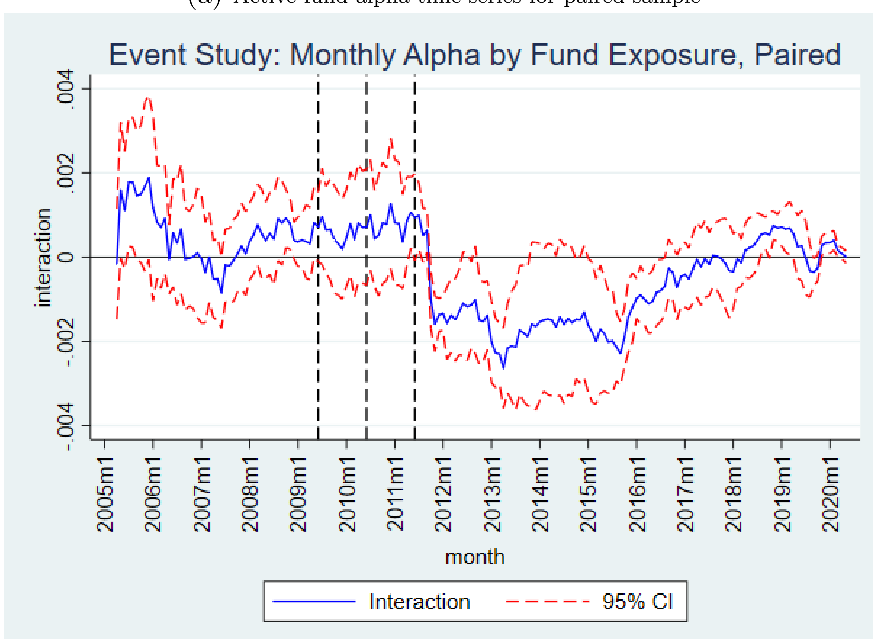
(b) Monthly difference between group 1 and 2 fund flows

Figure 1.5: Effect of the XBRL Mandate on Passive Fund Flows.

The dotted lines signify the dates of the initial implementation of each phase of XBRL in sequential order from left to right. Heteroskedasticity-robust standard errors are clustered by month $\times$ investment objective.



(a) Active fund alpha time series for paired sample



(b) Relative perf. of group 1 and 2 funds, paired sample

Figure 1.6: Active Fund Performance for a Closely-Matched Set of Funds.

The dotted lines signify the dates of the initial implementation of each phase of XBRL in sequential order from left to right. Heteroskedasticity-robust standard errors are clustered by month $\times$ investment objective.

CHAPTER II

Quants, Strategic Speculation, and Financial Market Quality

2.1 Introduction

Over the last few decades, quantitative investing has become increasingly popular in the financial markets. More generally, quantitative investing takes many forms, from high frequency algorithmic market-making to fundamentally-oriented quantitative strategies. Of particular note is the rapid growth in the fundamental equity quantitative investing trading sector. While such strategies were previously available primarily through hedge funds to select clients, they are now gaining a wider prominence and are being offered through traditional mutual funds and ETFs to ordinary investors (see Abis (2020) and Beggs et al. (2019) for further evidence). Quantitative strategies are generally based on quantitatively-disciplined trading rules, which results in both benefits, via greater trading discipline, and pitfalls, via strategy-crowding and myopia.[1] The goal of this study is to develop a theoretical framework for exploring the strategic interaction between humans and machines in the financial markets and the implications of greater fund automation for their quality.

[1] For an example of myopia, see Jason Zweig, "The Stock Got Crushed. Then the ETFs Had to Sell." The Wall Street Journal, 31 Jan. 2020.

Many fundamentally-oriented quantitative investment strategies are based on backtests. A defining characteristic of backtested strategies is their reliance on historical data to identify a signal which is predictive of future risk-adjusted returns. The investor then relies on the estimation to translate a real-time signal into a trading rule. An attractive aspect of this approach is its inherent discipline: the trading rule is not as susceptible to human emotion due to its automated nature. However, this greater discipline comes at a cost. Precisely because the trading strategy relies on historical data, it cannot take into account the effect that the implementation of the strategy will have on other market participants. For example, if other fundamental investors know that the machines will be receiving an inaccurate signal, they may trade against them to take advantage of the expected mispricing. It is this inherent myopia of backtested strategies that we focus on as the primary differentiator between quantitative and discretionary investing.

In principle, the effects of quantitative investing on financial market quality are twofold. On the one hand, incremental informed speculators are entering the market. On the other hand, those speculators adhere to quantitative strategies, which may be myopic in nature. We build on the canonical strategic speculator framework of Kyle (1985) to incorporate the two dimensions. First, we use Kyle's economy with a single imperfectly informed speculator, as the benchmark, i.e. before quantitative investing (Economy 1). In this baseline setting, we label the single speculator as the discretionary investor ("DI"), as his fully-strategic trading strategy is meant to resemble the trading strategies of the more traditional fundamental equity investment funds. Economy 1 also serves as the backtesting environment for the quantitative investor ("QI"). She uses the data generating process from Economy 1 to build a profit-maximizing strategy, "the backtest," under the assumption that *the other market participants are not aware of her existence*. This is the key assumption of the model, as it captures the notion of myopia inherent in quantitative strategies: by definition, the backtest ignores the strategic response of the discretionary investor and the market maker to the trading strategy of the quant.

Next, we introduce a second differentially informed speculator, the QI, to the market. If QI is assumed to be fully strategic (as DI), the resulting economy (Economy 2) captures the effects of incremental informed speculation on the market akin to Foster and Viswanathan (1996). Finally, we introduce quantitative investing by allowing the QI to trade based on either a convex combination of the profit-maximizing strategy of a fully strategic quant in Economy 2 and the backtest (Economy 3) or full automation (i.e. exclusively the backtest; Economy 4). The effect of incremental informed speculation on market quality is identified as the change from Economy 1 to Economy 2, while the implications of quantitative investing are identified via changes from Economy 2 to Economy 3 or Economy 4.

Our primary finding is that the direction of the effect of greater automation by the quantitative investor on financial market quality is determined by the sign of the wedge between the trading intensity of a fully-automated quant (Economy 4) and a fully-discretionary quant (Economy 2). The sign of the wedge acts as a sufficient statistic for the net effect of the three primary parameters in the model: information precision of the DI, information precision of the QI, and the correlation of their signals. A positive wedge, i.e. when the automated quant trades more than the discretionary quant, generally results in positive effects on market quality. Greater automation leads to greater expected trading volume by the quant, generally acts as a deterrent from more trading by the discretionary investor (as he becomes concerned about excess information slippage), and results in a net increase in the expected trading volume for the speculative sector as a whole. Since this net increase is driven by the trading behavior of a QI who adheres less and less to a profit-maximizing strategy, profits for the speculator sector decrease, price informativeness and price volatility rise, market depth rises, and return volatility falls. However, a negative wedge between the trading intensity of a fully-automated and a fully-discretionary quant implies that the quant trades less aggressively with greater automation. Less aggressive trading creates opportunity for the discretionary investor to trade more aggressively at the expense of the quant. As such, the DI trades more with greater automation by the QI, makes greater profits, which more than offsets the declining profits for the quant, and ultimately results in a lower trading volume and greater profits for the speculative sector as a whole. Market quality suffers as market depth, price informativeness, and price volatility fall, while return volatility rises.

We are able to theoretically characterize the determinants of the sign of the wedge, which in turn directly maps to market quality outcomes. Greater (lower) relative signal precision for the quant (DI) and a lower signal correlation generally result in a positive wedge. However, the relationship ultimately depends on the initial level of signal precision for the discretionary investor. If discretionary investors have relatively low signal precision, e.g. as for fundamentally-oriented proprietary quant funds, greater information quality for the quant or lower signal correlation will generally lead to improvements in market quality. If instead the DI has a sufficiently precise signal, e.g. as for smart-beta ETFs, improvements in the quant's signal may lead to a deterioration in market quality.

We further find that the introduction of the quant, whether fully-discretionary or fully-automated, generally benefits market quality as price informativeness and price volatility rise while return volatility falls. However, the effect on market depth and the trading volume and profits of the speculative sector are more nuanced and are again driven by the wedge with discretionary trading. As expected, the DI is almost always worse off due to the introduction of the QI. However, with a positive wedge, the speculative sector as a whole trades more and makes greater profits, resulting in a decline in market depth. With a negative wedge, the speculator sector trades more and makes lower profits, resulting in increasing market depth.

These results contribute to an important theoretical literature on price formation in the financial markets. Holden and Subrahmanyam (1992) show how aggressive competition among perfectly informed speculators may lead to speedy incorporation of information into prices (the "rat race"). Foster and Viswanathan (1996) demonstrate that under imperfect information and signals with sufficiently low correlation, the "rat race" may occur only initially, while less aggressive trading, i.e. a "waiting game," may emerge in later rounds of trading. While our framework is static in nature, we capture both the "waiting game" and the "rat race" outcomes via the single-period trading intensity of the speculators in our game. Importantly, we allow for differential information quality, which interacts with signal correlation to dictate market quality outcomes.

Most literature on quantitative investing has focused on the effects of high frequency trading on market liquidity and price discovery. As pointed out by Kirilenko and Lo (2013) a key trade-off exists in greater automation. On the one hand, algorithmic trading lacks human emotion, potentially leading to greater market stability. On the other hand, quantitative investing may be more highly correlated between funds, resulting in greater market instability. To our knowledge, only Abis (2020) develops a theoretical framework for exploring the trading behavior of quantitative versus discretionary funds. The author builds on the mutual fund attention allocation model of Kacperczyk et al. (2016) and introduces quantitative investors as those who have unlimited learning capacity only in idiosyncratic shocks. However, investors are competitive (rather than strategic) in their model. We complement this approach by allowing for strategic trading, a realistic feature of most financial markets, while abstracting from idiosyncratic learning capacity. Several studies, including Khandani and Lo (2007), Khandani and Lo (2011), and Beggs et al. (2019), have documented the fact that quantitative funds, whether hedge funds or mutual funds, tend to not consider the behavior of others, consistent with our model's predictions, resulting in violent reversals in periods of unwinding. Many have shown that algorithmic trading, in particular high frequency trading, leads to improving price efficiency and liquidity (e.g. Hendershott et al. (2011), Hendershott and Riordan (2013), Brogaard et al. (2014), Zhang (2017), and Chakrabarty et al. (2019)), although acquisition of new information may deteriorate (Weller (2018)).

2.2 Model

We model a three-date financial market with a single risky asset. Participants include informed speculators, noise traders, and a market maker. At time $t = 0$, informed speculators observe their signals with regards to the payoff of the risky asset. At time $t = 1$, speculators and noise traders submit their demands for the risky asset and market markers set prices according to the aggregate order flow. At time $t = 2$, the risky asset payoff is realized. Agents have rational expectations in that the informed speculators and the market maker are aware of the model parameters and each others' price setting and trading behavior respectively. We describe each feature of the model below.

2.2.1 Model Setup

Economies

There will be four market environments, or economies, indexed by $j \in \{1, 2, 3, 4\}$, which will differ based on the presence and trading strategy of the quantitative investor.

The economies will be introduced in greater detail below.

Risky asset

The risky asset has an endogenously determined price P_j set by the market maker in the first period and an exogenous payoff v distributed according to $v \sim N(P_0, \sigma_v^2)$, realized in the second period.[2] The equilibrium price function is indexed by Economy j as it will depend on the presence and the trading strategy of the quantitative investor. The payoff distribution remains the same across all four economies.

Informed speculators

There are two informed speculators indexed by $i \in \{d, q\}$ with different trading behavior: a discretionary investor ($i = d$, "DI") and a quantitative investor ($i = q$, "QI"). Each informed speculator i receives a signal about the final payoff of the risky asset given by:

$$s_i = v + e_i, \tag{2.1}$$

where $e_i \sim N(0, \sigma_{e_i}^2)$, $\sigma_{ve_i} = 0$, $\rho = \sigma_{e_d e_q}/(\sigma_{e_d}\sigma_{e_q})$, and $\phi_i = \sigma_v^2/\sigma_{s_i}^2$ is a measure of information quality.

Discretionary investor The discretionary investor is risk-neutral and behaves akin to the Kyle (1985) insider, choosing a share amount x_{dj} to maximize period two expected profit:

$$x_{dj} = \arg\max_{\tilde{x}_{dj}} \mathbb{E}[\tilde{x}_{dj}(v - P_j)|s_d]. \tag{2.2}$$

[2] Henceforth $\sigma_x^2 = \tau_x^{-1}$ refers to the variance of x, τ_x refers to the precision of x, andσ_{xy} refers to the covariance between x and y.

The investor is strategic in that he accounts for the expected effect of his order on the price set by the market maker as well as the information revealed by his signal about the expected trading behavior of the quantitative investor. Within our framework, discretionary behavior is modeled via an investor's ability to take into account the expected behavior of others in the market when developing a trading strategy. A quant, in contrast, either fully or partially lacks this ability.

Quantitative investor The quantitative investor develops a profit-maximizing trading rule under the assumption that the discretionary investor and market maker are not aware of her existence. This strategy originates from the notion that many quantitative strategies are based on backtested signals, which by definition take as given the behavior of market participants. In other words, when developing trading strategies based on historical market data, quant traders are not able to incorporate the effect that their *presence* may have had on the strategic trading behavior of other market participants present at the time.[3] The backtested strategy, labeled x_b and introduced in greater detail below, is derived from a world in which only the discretionary investor is present.

We model the potentially myopic behavior of the quantitative investor in a reduced form manner. Specifically, we assume that the quant's optimal asset demand x_{qj} is given by the solution to the following optimization:

$$x_{qj} = \arg\max_{\tilde{x}_{qj}} \mathbb{E}[\underbrace{-\gamma(\tilde{x}_{qj} - x_b)^2}_{\text{backtest}} + \underbrace{(1-\gamma)\tilde{x}_{qj}(v - P_j)}_{\text{strategic speculation}} | s_q], \tag{2.3}$$

where $0 \leq \gamma \leq 1$.

Eq. (2.3) is based on the observation that a quant's welfare is at least partially de-

[3] This concept is known within the macroeconomics literature as the "Lucas Critique" (Lucas (1976))

termined by her adherence to a backtested, potentially well-known, trading approach (x_b). The degree of adherence, or "automation," is parametrized by γ. Such behavior may be explained by marketing reasons, and the related greater investing discipline and transparency. For example, quantitative fund strategies are less-likely to be affected by human emotion, and the associated discipline may appeal to investors thereby generating greater fund flows. Furthermore, the adherence to a particular well-known strategy may reduce the moral hazard issues associated with delegated portfolio management.

It is important to note that human judgement continues to play an important role in quantitative strategies. From the choice of investment strategy to its real-world implementation, human decisions are at the core of launching and maintaining the operations of quant funds. Quants may very well be strategic in their trading via, e.g., "feeler orders" to test market liquidity, scaling and turning strategies on and off depending on market conditions, and attending industry conferences to interact with other quant fund managers to understand the behavior of others. We model this "discretionary" aspect of quantitative investing in a reduced form via the loading $1-\gamma$ of the strategy on the traditional profit maximization objective of the Kyle insider.

An immediate consequence of Eq. (2.3) is that for $\gamma > 0$ the quant is not maximizing her profits. Implicit in this assumption are certain un-modeled gains that may arise due to the pursuit of a quantitative versus a discretionary strategy. For example, the lower reliance on human judgement may lower the risk of moral hazard, thereby reducing agency frictions. Furthermore, adherence to a very specific strategy may cater to a particular set of investors, thereby attracting greater fund flows. In addition, quantitative strategies may have a greater breadth due to their reliance on computing power, thereby potentially allocating to a much higher number of securities. The resulting benefits to diversification, although absent from the model, may also be captured by the backtested trading strategy.

Noise traders

Noise traders provide an exogenously given aggregate demand of the risky asset of $z \sim N(0, \sigma_z^2)$ such that $\sigma_{ze_d} = \sigma_{ze_q} = \sigma_{zv} = 0$. Some examples of noise demand include uninformed retail investors buying or selling shares for liquidity needs, informed asset managers buying or selling securities purely for hedging purposes, or corporations unexpectedly issuing or buying back stock. We will rely on $\eta = \sigma_z^2/\sigma_v^2$ as a normalized measure of noise trading in the market.

Market makers

Following Kyle (1985), market makers ("MM") are assumed to have risk-neutral preferences and operate in a competitive environment. Therefore, given aggregate order flow $\omega_j = x_{dj} + x_{qj} + z$, the market maker sets a price such that he breaks even in expectation:

$$P_j(\omega_j) = \mathbb{E}[v|\omega_j], \tag{2.4}$$

i.e. such that the equilibrium asset price is semi-strong from efficient. An underlying assumption in this framework is that the market maker has sufficient inventory (if selling) or liquidity (if buying) to satisfy the net demands of the traders. The market maker attempts to discern between the informed and uninformed order flow and to set prices accordingly.

2.2.2 Model Solution

Our objective is to identify the effect of growth in quantitative investing on financial market quality. Intuitively, this growth may be driven by new quant funds entering the market, incumbent discretionary funds switching to quantitative strategies, or both new entrants and incumbent strategy transitions. We focus primarily on the effects of new quant funds in the financial markets as we believe this to be the primary driver of the growth in quantitative investing. However, our framework can also speak to the effects of automation by incumbent discretionary funds and can be extended to jointly consider new quant funds and incumbent strategy transitions.

Growth in quantitative investing that is driven by new entrants in the financial market has a dual effect on market quality. First, market quality changes due to the incremental informed participants in the financial markets. Second, market quality is impacted by the quantitative investing approach of the new entrants. To speak directly to the effects of quantitative strategies we must untangle the two forces. We do so by considering market quality across four economies:

- Economy 1: DI
- Economy 2: DI + QI ($\gamma = 0$)
- Economy 3: DI + QI ($0 < \gamma < 1$)
- Economy 4: DI + QI ($\gamma = 1$)

Economy 1 resembles a world before quantitative investing was introduced. This is also the quant's backtesting environment, which allows her to back out the DI's trading rule and the MM's pricing function. The backtest will be a regression of prices on fundamentals, which coupled with observable order flow provides the quant with sufficient information to form her profit-maximizing strategy. The quant will then build her strategy such that profits are maximized under the assumption that the MM and the DI are not aware of her existence.

Economies 2-4 introduce the quantitative investor, but do so in phases in order to distinguish between the effects of an additional informed speculator and the speculator's backtested trading rule. First, Economy 2 isolates the effect of introducing a second informed speculator to the economy. Applying $\gamma = 0$ to Eq. (2.3), demonstrates that the QI's and DI's objective functions are identical. Therefore, Economy 2 is simply an extension of Kyle (1985) with two informed speculators, a static version of Foster and Viswanathan (1996) but with differential signal precision. Second, Economy 3 "turns on" the quant's strategy. Increasing γ leads to greater adherence to the backtested strategy implying greater "automation" by the quant. Finally, Economy 4 reflects a full-automated quant. In our baseline theoretical results we only consider Economies 1,2, and 4, however we also conduct a simulation exercise for intermediate values of γ in Economy 3.

We now proceed to sequentially solve for the DI's and QI's trading strategies and the MM's pricing rule in Economies 1-4. For each economy j we conjecture that the optimal trading strategy of investor i is linear in his / her signal and is given by

$$x_{ij}(s_i) = \alpha_{ij} + \beta_{ij} s_i, \tag{2.5}$$

and the optimal pricing rule of the market maker is linear in the aggregate order flow such that

$$P_j(\omega_j) = \mu_j + \lambda_j \omega_j. \tag{2.6}$$

We then derive the coefficients that satisfy the investors' and market makers' optimization functions given by Eq. (2.2), (2.3), and (3.4). We refer to β_{ij} as the "trading intensity" of speculator i as it reflects how aggressively the speculator reacts to an incremental unit of signal. We identify $1/\lambda_j$ as "market depth," our primary measure of liquidity: the inverse of the market maker's price sensitivity to an incremental share of order flow.

2.2.2.1 Economy 1

Economy 1 is a world with a single informed speculator (DI) who trades strategically to minimize information slippage and maximize profits. This is the traditional Kyle (1985) framework with the sole difference that the informed speculator observes a noisy signal $s_d = v + e_d$ instead of perfect information v about the risky asset payoff. The DI solves Eq. (2.2) while the MM solves Eq. (3.4) with $\omega_1 = x_{d1} + z$. Lemma II.1 demonstrates the existence and uniqueness of a linear strategy equilibrium.

Lemma II.1 (Equilibrium in Economy 1)**.** *Within the class of linear equilibria, there exists a unique linear trading rule for the DI and a unique linear pricing function for the MM specified by Eq.* (2.5) *and* (2.6) *respectively such that Eq.* (2.2) *and* (3.4) *are satisfied for all realizations of* v*,* e_d*, and* z*.*

Specifically, as derived in Appendix B.1, the equilibrium trading rule of the DI is

$$x_{d1}(s_d) = \sqrt{\phi_d \eta}(s_d - P_0), \tag{2.7}$$

and the pricing rule of the MM is

$$P_1(\omega_1) = P_0 + \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}\omega_1. \tag{2.8}$$

The DI is more aggressive in trading on his signal with more noise trading, a more precise signal, and a greater expected payoff. The MM sets prices such that the expected profits from providing liquidity to the noise traders exactly offset the expected losses from trading against the DI. Therefore, equilibrium liquidity, as measured by $1/\lambda_1 = 2\sqrt{\eta}/\sqrt{\phi_d}$, is decreasing in signal quality and is increasing in the relative amount of noise trading. Both are driven by the market maker's reaction to adverse selection: the more precise the DI's information or the greater the likelihood of informed trading, the more sensitive will the price be to order flow.

2.2.2.2 Quant's Backtest

A defining characteristic of backtested strategies is their reliance on historical data to identify a signal, which is predictive of future risk-adjusted returns. The quant relies on the strategy to map a real-time signal to a trading rule, which is implemented in the financial markets. By relying on historical data to generate a signal, the quant cannot know how other market participants would have reacted to her presence in the market. Even if similar strategies already exist in the market during the backtest, the quant's strategy will likely be incremental, implying that the backtest does not fully capture the strategic behavior of other market participants. It is this lack of strategic consideration that we focus on as the primary friction associated with quantitative strategies.

We model this inherent myopia of backtested strategies, by endowing the quant with a signal, but assuming that she builds the strategy under the assumption that the other market participants are not aware of her existence. In other words, the quant assumes that the behavior of other market participants does not change in response to her presence. Prior to launching her strategy, the quant uses historical data from Economy 1 to generate her optimal trading rule. One can think of this as years of trading data that is based on the data generating process of Economy 1, used as the backtest into the quant's trading approach. She gathers an extended time series of this data, and backs out the trading rule of the DI and the pricing strategy of the MM.

The DI's trading strategy, Eq. (2.7), and a simple manipulation of the MM's pricing rule, Eq. (2.8), demonstrate that:

$$P_1 = P_0 + \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}\omega_1$$

$$P_1 = P_0 - \frac{\phi_d P_0}{2} + \frac{\phi_d}{2}v + \frac{\phi_d}{2}e_d + \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}z$$

Let's assume that Economy 1 plays out over many periods, e.g days, $t = 1, \ldots, T$ and in each day t the following data is recorded: $\{P_{1t}, \omega_{1t}, v_t\}$. In other words, the QI is able to observe historical prices, trading volume, and fundamentals. The quant can then take this data and run the following ordinary least squares regressions:

$$P_{1t} = a + b\omega_{1t} + e_t$$

$$P_{1t} = c + dv_t + \epsilon_t$$

By construction, $e_t = 0 \perp \omega_1$ and $\varepsilon_t = \frac{\phi_d}{2}e_d + \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}z \perp v$ implying that the QI will be able to obtain unbiased estimates for $a = P_0$, $b = \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}$ and $d = \frac{\phi_d}{2}$. Assuming the QI has a sufficiently long data sample, she will deduce both the DI's signal precision ϕ_d and the normalized amount of noise trading η from the data. Together, these parameters allow her to identify the MM's pricing rule and the DI's trading strategy: we assume that the data sample is sufficiently long such that she obtains highly accurate estimates. This, in turn, allows her to construct an optimal, profit-maximizing strategy *under the assumption that the DI and MM are not aware of her existence:*

$$\begin{aligned} x_b &= \arg\max_{\tilde{x}_b} \mathbb{E}\Big[\tilde{x}_b\Big(v - P_1\big(x_{d1} + z + \tilde{x}_b\big)\Big)|s_q\Big] \\ &= \frac{1}{2}\sqrt{\frac{\eta}{\phi_d}}\left((2-\phi_d)\phi_q - \rho\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}\right)(s_q - P_0), \end{aligned} \tag{2.9}$$

derived in Appendix B.2. The notion of myopia is incorporated by the QI's reliance on the MM's pricing function $P_1(\omega_1)$ and the DI's demand $x_{d1}(s_d)$ from Economy 1. In other words, the QI assumes that the MM will treat her order flow as noise trading and the DI will not internalize the QI's presence.

Similar to the DI's Economy 1 trading strategy, the QI's backtest trades more aggressively with more noise trading and a greater expected payoff. For negative and sufficiently low positive signal correlation, the quant will always buy more shares with greater expected payoff as she's less concerned about moving the price in the same direction as the DI. However, with high signal correlation, the QI may trade against the DI since she becomes concerned about excess information slippage. Ultimately, the backtested strategy will depend not only on the quant's signal precision and level of noise trading, but also on the DI's signal precision and the signal correlation.

2.2.2.3 Economies 2-4

Armed with the backtested strategy of the quantitative investor, we proceed to derive the equilibrium trading and price setting behavior in Economies 2-4. Economy 2 will be the classic extension of Kyle (1985) to include two informed speculators with potentially correlated signals (e.g. Foster and Viswanathan (1996)) and potentially differential information quality. Economy 3 will incorporate the notion of quantitative investing as the QI will at least partially rely on the backtested strategy ($\gamma > 0$), while Economy 4 will include a full-automated quant ($\gamma = 1$). The trading behavior of the DI and the QI are determined by the solutions to Eq. (2.2) and (2.3) respectively, while the MM continues to set prices to break-even in expectation as specified in Eq. (3.4). Lemma II.2 demonstrates the existence and uniqueness of a linear strategy equilibrium.

Lemma II.2 (Equilibrium in Economies 2-4)**.** *Within the class of linear equilibria, there exist unique linear trading rules for the DI and QI and a unique linear pricing function for the MM specified by Eq.* (2.5) *and* (2.6) *respectively such that Eq.* (2.2)*,* (2.3), *and* (3.4) *are satisfied for all realizations of* v, e_d, e_q, *and* z.

As demonstrated in Appendix B.3 multiple equilibria exist in this problem. In part this is due to the two equilibria within the Kyle (1985) framework whereby a speculator can buy *less* on a stronger signal, which in equilibrium leads to a downward sloping supply curve for the MM. Within our framework, this interacts with the QI's objective function resulting in up to four equilibria. We focus on economically meaningful equilibria, which appear to be unique. Specifically, we restrict λ_j to be greater than zero: supply curves are upward sloping.

To gain intuition behind the equilibrium it is helpful to analyze the functional forms of the trading rules for the informed speculators and the pricing function of the market maker. For the DI, demand is given by:

$$\begin{aligned} x_{dj} &= \frac{\mathbb{E}[v|s_d] - \mu_j - \lambda_j \mathbb{E}[x_{qj}|s_d]}{2\lambda_j} \\ &= \underbrace{\frac{\mathbb{E}[v|s_d] - \mu_j}{2\lambda_j}}_{\text{Single strategic speculator demand}} - \underbrace{\frac{\mathbb{E}[x_{qj}|s_d]}{2}}_{\text{Adjustment for second speculator}}, \end{aligned} \tag{2.10}$$

as derived in Eq. (B.10) and (B.11). The DI's demand is the sum of his demand function as if he were acting alone (conditional on the market maker's pricing rule) and an adjustment for the expected trading behavior of the QI conditional on his information. The adjustment is equal to the expected price impact of the QI's expected trading activity ($\lambda_j \mathbb{E}[x_{qj}|s_d]$) scaled by $2\lambda_j$, which maps ex-ante expected profits to share demand. The adjustment term is still dependent on λ_j because the expected share demand of the QI is also (inversely) dependent on λ_j. Furthermore, the adjustment will depend on the QI's level of automation γ, known by both the MM and the DI, since it will factor into both the equilibrium pricing function of the MM via λ_j and the expectation of the demand of the QI via $\mathbb{E}[x_{qj}|s_d]$. Decomposing the numerator of the adjustment even further yields:

$$\mathbb{E}[x_{qj}|s_d] = \underbrace{\alpha_{qj} + \beta_{qj} P_0}_{\text{Second speculator's ex-ante expected trade}} \underbrace{+ \beta_{qj}\left(\phi_d + \rho\sqrt{\frac{\phi_d}{\phi_q}(1-\phi_d)(1-\phi_q)}\right)(s_d - P_0)}_{\text{Signal inference}}, \tag{2.11}$$

such that the DI's optimal trading strategy depends on both what he expects v to be and what he expects the QI to know about v and use in her trading strategy. The expected trading strategy of the quant based on the DI's information is the sum of the unconditional expectation of her trading strategy and an information adjustment based on the DI's signal. The signal inference will be jointly determined by the DI's signal, QI's trading intensity, the absolute and relative precision of QI's and DI's signal, and the signal correlation.

We can apply a similar decomposition, derived in Eq. (B.56), to understand the trading behavior of the quantitative investor:

$$x_{qj} = \frac{\gamma}{\gamma + (1-\gamma)\lambda_j} x_b + \frac{(1-\gamma)\lambda_j}{\gamma + (1-\gamma)\lambda_j} \frac{\mathbb{E}[v|s_q] - \mu_j - \lambda_j \mathbb{E}[x_{dj}|s_q]}{2\lambda_j}. \tag{2.12}$$

The exogenously assigned automation parameter γ will determine whether the model represents Economy 2 ($\gamma = 0$), Economy 3 ($\gamma > 0$), or Economy 4 ($\gamma = 1$).

2.2.2.4 Economy 2: $\gamma = 0$

With no automation the QI becomes a fully-strategic speculator akin to the discretionary investor. The forces driving the behavior of the DI specified above will be symmetric for the QI. Of particular note is the relationship between the trading intensity of the DI and the QI as derived in Eq. (B.21) since the aggregate trading intensity will determine various market quality measures introduced in the next section:

$$\begin{aligned}\beta_{d2} &= \frac{2\phi_d - \phi_d\phi_q - \rho\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}}{2\phi_q - \phi_d\phi_q - \rho\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}}\beta_{q2} \\ &\equiv k\beta_{q2}\end{aligned} \tag{2.13}$$

It is immediate that $\partial k/\partial\rho > (<)0 \Leftrightarrow \phi_d > (<)\phi_q$, which implies that greater signal correlation will cause the investor with an information advantage to trade more aggressively. The importance of signal correlation to trading behavior is consistent with the findings of Foster and Viswanathan (1996).

2.2.2.5 Economy 3: $0 < \gamma < 1$

For a partially automated quant, i.e. intermediate values of γ, the trading rule is an endogenously determined weighted average of the backtested strategy and the fully-strategic trading approach. As evident from Eq. (2.12), the relative weights are determined by λ_j, which, in turn, identifies the depth of the market. For a fixed level of liquidity λ_j by the market maker, greater automation (higher γ) results in a greater weight being placed on the back-tested strategy versus the strategic rule. For relatively liquid markets ($\lambda_j < 1$), greater γ leads to a more than a one-to-one weight increase on the backtested strategy (e.g. a 1% increase in γ leads to a more than 1% of incremental weight being placed on the back-tested rule). For relatively illiquid markets ($\lambda_j > 1$), greater γ leads to a less than one-to-one increase in the weight placed on the backtested strategy. Finally, for $\lambda_j = 1$, a greater γ results in exactly a one-to-one increase in the weight placed on x_b. The quant is more aggressive in implementing her backtest in more liquid environments.

2.2.2.6 Economy 4: $\gamma = 1$

A fully-automated quant adheres solely to the back-tested strategy x_b, which is developed under the assumption that the DI and the MM are not aware of her existence. However, both the DI and MM internalize the QI's expected behavior thereby rendering QI's strategy suboptimal from the perspective of profit-maximizing behavior.

2.2.2.7 Economy 2-4: Market Maker

For the market maker's pricing rule, first note that the aggregate order flow ω_j provides a noisy signal for v:

$$\omega_j = \alpha_{dj} + \alpha_{qj} + \underbrace{(\beta_{dj} + \beta_{qj})v}_{\text{Signal}} + \underbrace{\beta_{dj}e_d + \beta_{qj}e_q + z}_{\text{Noise}}, \tag{2.14}$$

where the informativeness of the order flow can be expressed as:

$$\text{Var}(\omega_j|v) = \beta_{dj}^2\sigma_{e_d}^2 + \beta_{qj}^2\sigma_{e_q}^2 + 2\beta_{dj}\beta_q j\sigma_{e_d e_q} + \sigma_z^2. \tag{2.15}$$

A lower conditional variance implies a more informative order flow. Lower trading intensity by either informed speculator, greater signal precision for either informed speculator, lower noise trading variance, and a lower signal covariance all lead to a more informative order flow. This in-turn translates into the following effects on market liquidity as given by $1/\lambda_j$:

$$\begin{aligned}
\lambda_j^{-1} &= \frac{\text{Var}(\omega_j)}{\text{Cov}(v, \omega_j)} \\
&= \beta_{dj} + \beta_{qj} + \frac{\text{Var}(\omega_j|v)}{(\beta_{dj} + \beta_{qj})\sigma_v^2} \\
&= \underbrace{(\beta_{dj} + \beta_{qj})}_{\text{Aggregate trading intensity}} \times \left(1 + \underbrace{\frac{\text{Var}(\omega_j|v)}{\text{Var}(\omega_j|\beta_{dj}e_d + \beta_{qj}e_q + z)}}_{\text{Noisiness of order flow}}\right)
\end{aligned} \tag{2.16}$$

Holding aggregate trading intensity of the speculators constant, a noisier order flow results in a more liquid price system. Market makers are less concerned about adverse selection and are therefore less sensitive to greater order flow. Less precise signals, greater signal covariance, and more noise trading all lead to a noisier system. Changes to trading intensity for either speculator may either increase or decrease the liquidity provision both through the noise-to-signal ratio and through the aggregate trading intensity measure.

2.3 Comparative Statics

We explore the effects of the introduction of and the greater automation by the quantitative investor on market quality. As customary in the literature (e.g. Vives (2008)), we focus on the following measures of market quality as well as certain important asset pricing measures for each Economy j:

1. Market depth: $1/\lambda_j$
2. Price informativeness: $\mathrm{Var}(v|P_j)^{-1}$
3. Price volatility: $\mathrm{Var}(P_j)$
4. Return volatility: $\mathrm{Var}(v - P_j)$
5. Expected trading volume for each informed speculator: $\mathbb{E}[|x_{ij}|]$
6. Expected trading volume for the speculator sector: $\mathbb{E}[|x_{dj} + x_{qj}|]$
7. Expected profits for each informed speculator: $\mathbb{E}[\pi_{ij}] = \mathbb{E}[x_{ij}(v - P_j)]$
8. Expected profits for the speculator sector: $\mathbb{E}[\pi_j] = \mathbb{E}[\pi_{dj}] + \mathbb{E}[\pi_{qj}]$

The closed-form solutions for each of the measures above utilizing the informed speculators' trading rules and the market maker's pricing function as derived in Appendices B.1 and B.3 are provided in Appendix B.4. Unfortunately, the theoretical framework does not allow for analytically tractable solutions for the trading rules and pricing functions. Therefore, we split the comparative statics analysis into two sections. The first is a special case of the generalized model, which assumes that the DI has perfect information, and only considers changes to market quality when the QI switches from full discretion ($\gamma = 0$) to full automation ($\gamma = 1$), thereby only exploring changes to market quality for Economies 1,2, and 4. Within this framework, we are able to theoretically demonstrate the effects of the introduction of and automation by the quant on market quality. The second section conducts a quantitative exercise for the generalized model, documenting changes in market quality due to the addition of and all levels of automation by the QI (Economies 1-4) across a large number of parameter draws. We first discuss the theoretical benchmark and conclude with the market quality outcomes of the generalized model.

2.3.1 Theoretical Benchmark

The benchmark case assumes the DI has perfect information ($\phi_d = 1$) and the unconditional expected payoff is zero. The latter assumption is without loss of generality, but the former needs further justification. Inherent in this assumption is the notion that traditional fundamental investors may have superior access to fundamental information via corporate management meetings, industry and market expertise, and deep-dive fundamental analysis. Quants, on the other hand, may look at rough proxies for fundamental value, such as historical multiples or price trends, and hope that on average, across a large sample of securities, such proxies will be indicative of future payoffs. Examples of such quant funds would be smart beta exchange traded funds and statistical arbitrage strategies nested within hedge funds. However, this category does not include fundamental investors who rely on big data as an input into their investment process, since such fundamental investors may actually have superior information. Examples of the latter would be fundamental funds relying on credit card transaction data to accurately estimate retailers' revenues or utilizing airfare pricing datasets to forecast airline revenues.

The introduction of quantitative investing has a dual effect on the financial markets. First, an incremental informed speculator is introduced to the economy. Second, the incremental speculator chooses to pursue an automated strategy. We attempt to distinguish between these two effects by first exploring the implication of adding a second informed strategic speculator to the market (Economy 1 to Economy 2), and, second, analyzing the impact of automation by the incremental strategic speculator (Economy 2 to Economy 4). Under the assumption that growth in quantitative investing is driven by new entrants in the market, the net effect of quantitative investing will be interpreted as the change in market quality between Economy 1 and Economy 4. However, if the growth is driven by strategy transitions of incumbent funds from discretionary to quantitive, the effects could be viewed through the changes from Economy 2 to Economy 4.

Proposition II.1 derived in Appendix B.6 highlights our main theoretical findings regarding the effects of quantitative investing on financial market quality. Figure 2.1 demonstrates the comparative statics graphically.

Proposition II.1 (Quantitative investing and financial market quality)**.** *Market quality improves due to the introduction of the quantitative investor, but the improvement is dampened by the quant's pursuit of the backtested trading rule. The shift from Economy 1 to Economy 2 results in greater trading volume by the speculator sector, more informative prices, greater market depth, and lower profits for the speculator sector. The* *transition from Economy 2 to Economy 4 leads to less aggressive trading and lower profits by the QI, more aggressive trading and greater profits by the DI, less informative prices, lower market depth, and greater profits for the speculator sector as a whole. Overall, market quality improves from Economy 1 to Economy 4.*

The introduction of a second strategic speculator to the economy, i.e. the transition from Economy 1 to Economy 2, unequivocally improves financial market quality. The result is consistent with the findings of Foster and Viswanathan (1996) for the case of uncorrelated information for multiple informed speculators. As can be seen from Fig. 2.1a, the introduction of a fully-discretionary quant is internalized by the DI, causing him to trade less aggressively relative to a market in which he's the only informed speculator due to concerns regarding incremental information slippage. However, the decline in the DI's trading, is more than offset by incremental trading by the QI, resulting in a net growth in expected trading volume by the speculator sector. This net growth is increasing in the QI's signal precision, as she trades more aggressively on her information. Fig. 2.1d highlights that the greater trading by the speculator sector leads to more informative prices relative to a benchmark world of a single informed speculator, which, in turn is translated into lower profits for the discretionary investor and the speculator sector as a whole (Fig. 2.1b). In equilibrium, the greater trading volume and more informative prices correspond to greater market depth (Fig. 2.1c), greater price volatility (Fig. 2.1e), and lower return volatility (Fig. 2.1f). The latter two outcomes are directly driven by prices that more accurately track fundamental values.

The primary contribution of the present framework is an analysis of the effects of automation by the QI on the behavior of the DI and the net effects to financial market quality. The quant's backtested rule is less aggressive (i.e. has lower trading intensity) than her trading strategy with full discretion. In part this is due to the foundation of the backtest, which assumes that the DI and the MM are not aware of the QI's existence: they treat the incremental order flow as originating from the noise traders. Once the presence of the quant is internalized by other market participants, the DI trades more cautiously, allowing a discretionary quant to trade more aggressively in response. On the other hand, an automated quant is, by definition, pursuing a strategy that is not profit maximizing. This opens the door for the discretionary investor to trade more aggressively on his information, but the net effect is less expected trading volume by the speculator sector as a whole relative to Economy 2. The decline in trading volume leads to less informative prices and greater profits for the DI and the speculator sector as a whole. The QI earns less in expectation in Economy 4 due to the sub-optimality of her strategy from a profit maximization perspective. Concurrent with the lower price informativeness, liquidity dries up in the form of lower market depth, price volatility declines, and return volatility increases. Overall, Economy 4 market quality decreases relative to Economy 2.

Despite the potentially detrimental effects of automation by the quant, as the shift from Economy 2 to Economy 4 would suggest, we emphasize the overall improvement to market quality from Economy 1 to Economy 4. The beneficial effects from the introduction of an incremental informed speculator to the economy dominate the potential inefficiency induced by the trading strategy they pursue. To the extent that quantitative investing growth is driven by new entrants in the market, our model would suggest overall improvements in market quality. However, if existing funds are transitioning from discretionary to quantitative strategies, market quality may be negatively impacted.

The presented framework makes strong assumptions in exchange for theoretical tractability. Next, we conduct a numerical exercise to describe the properties of a more general version of the model.

2.3.2 Generalized Model

In the generalized model, we do not restrict the DI to perfect information, allow for signal correlation between the informed speculators, and allow for partial automation (i.e. intermediate values of γ). As such, we are able to not only consider the relationships between the levels of market quality in Economies 1,2, and 4, but also to explore the effects of intermediate automation through Economy 3. After all, quantitative funds are ultimately deployed by humans, and inherent in this is a certain level of discretion. Humans may choose to strategically scale, turn off, and alter their investment strategies, even if such strategies full rely on backtested analysis. Furthermore, the framework allows for differential signal precision for the QI and the DI, thereby allowing us to explore the effects of both superior and inferior information for the quant. The latter feature of the model allows us to speak more generally to systematic quantitative strategies, whereby quants may actually have superior information given their ability to process large data to more accurately forecast future earnings.

As the simulation will highlight, the effects of quantitative investing on financial market quality will be determined by whether a fully-automated quant (Economy 4) trades more than a fully-discretionary quant (Economy 2). Since the quant's strategy is (loosely, i.e. Eq. (2.12)) a weighted average between the Economy 2 and the Economy 4 strategy, growth in automation (γ) will lead to more (less) aggressive trading by the QI if $\beta_{q4} > (<) \beta_{q2}$. The change in trading behavior due to greater automation is internalized by the discretionary investor and ultimately determines the effects of quantitative investing on financial market quality. Since the fully-automated quant is by definition not maximizing profits, to the extent that this sub-optimality is accompanied by less aggressive trading, the DI has the opportunity to exploit the greater resulting mis-pricing to his advantage by trading more aggressively without positive implications to market quality. If, however, greater automation leads to greater trading intensity by the quant, the DI is mostly unable to exploit this to his advantage as he's concerned about excess information leakage. The net effect in the latter scenario is an improvement in market quality. The simulation results, provided in Tables 2.1-2.3, are summarized in Conclusions II.1 and II.2.

Conclusion II.1 (Effect of the introduction of a fully-discretionary quant)**.** *The introduction of a fully-discretionary quantitative investor leads to improvements in market quality via more informative prices, greater price volatility, and lower return volatility. The effects on market depth, trading volume, and profits for the speculator sector depend on signal correlation and precision for the QI and DI. If the parameters are such that the fully discretionary quant trades more than the fully automated quant ($\beta_{q2} > \beta_{q4}$), trading volume for the speculator sector increases, profits mostly decrease, the DI trades less and makes less in profits, the QI makes positive profits, and market depth mostly rises. Conversely, if $\beta_{q2} \leq \beta_{q4}$, trading volume for the speculator sector mostly increases, profits always increase, the DI trades less and makes less in profits, the QI makes positive profits, and market depth always decreases.*

Conclusion II.2 (Effect of automation by the quantitative investor)**.** *The effect of greater automation γ by the QI on financial market quality depends on whether the fully-automated quant trades more than the fully-discretionary quant. If the fully-automated quant trades more ($\beta_{q4} > \beta_{q2}$), market quality broadly improves as market depth, price informativeness, price volatility are mostly rising and return volatility is falling in γ. Furthermore, the speculator sector as a whole trades more, profits are mostly lower, the effects on the DI's trading and profits can be either be increasing, decreasing, or U-shaped in γ, while the QI always trades more and her profits are increasing, decreasing or hump-shaped in γ. If the fully-automated quant trades less ($\beta_{q4} < \beta_{q2}$), market quality deteriorates via lower market depth, less informative prices, lower price volatility, and greater return volatility. The speculator sector as a whole trades less and makes* *greater profits as the DI trades more aggressively and makes greater profits, while the QI trades less and makes lower profits.*

To understand the intuition behind the main conclusions from the simulation, it is helpful to first separately consider the effects of signal correlation and relative information precision on market quality. Figures 2.2-2.3 demonstrate the effect of automation on market quality for different correlation structures for the generalized model within which the DI and QI have equal information precision ($\phi_d = \phi_q$) and other parameter values are held constant. The first notable pattern, highlighted in Fig. 2.2e, is the growth in the QI's trading intensity for lower correlation levels. This is a direct outcome from the QI's backtest, as described in Section 2.2.2.2, whereby the backtest trades more aggressively with lower correlation. A lower correlation allows the quant to trade more aggressively because she's less concerned about excess information slippage. For sufficiently low correlation levels, greater automation leads to a greater expected trading volume for the QI since the backtest has greater trading intensity than the fully-discretionary strategy. Conversely, for sufficiently high levels of correlation, greater automation leads to lower expected trading volume as more weight is placed on the backtest, which has lower trading intensity than the fully-discretionary approach. Formulaically, we have

$$\frac{\partial}{\partial \rho}\Big(\frac{\partial \mathbb{E}[|x_{q3}|]}{\partial \gamma}\Big) < 0, \tag{2.17}$$

for all $\gamma \in (0,1)$ such that for each $\hat{\gamma}$ there exists $\rho^*_{\hat{\gamma}} > 0$ whereby $\frac{\partial \mathbb{E}[|x_{q3}|]}{\partial \gamma}|_{\gamma=\hat{\gamma}} = 0$.

Pictured in Fig. 2.2c, the discretionary investor's reaction to the changes in signal correlation is more nuanced. On the one hand, a sufficiently negative signal correlation implies that he can trade more aggressively on his information since the QI is more likely to trade in the opposite direction thereby decreasing information slippage. On the other hand, the QI is also trading more aggressively, and is acting less strategically, thereby limiting the extent to which the DI can be more aggressive in his trading. The net effect of the two forces can lead to either an increase or a decrease in the trading volume of the DI. However, if the signals are sufficiently positively correlated, the QI will always trade less aggressively with greater automation, which opens the door for the DI to be more aggressive in his trading.

As demonstrated in Fig. 2.2a, the net effect of the trading behavior of the DI and the QI on the aggregate speculator trading volume is lower aggregate trading volume in response to lower correlation levels. This is intuitive, as the negative correlation leads to offsetting orders, causing a decrease in aggregate order flow. However, a more intricate outcome, is that the increase in total trading volume in response to greater γ is greatest for lower correlation levels. That is the change in the trading intensity of the QI in response to greater automation for varying correlation levels dominates the change in the trading intensity of the DI. Directionally, an equivalent relationship to Eq. (2.17) emerges:

$$\frac{\partial}{\partial\rho}\Big(\frac{\partial\mathbb{E}[|x_{d3}+x_{q3}|]}{\partial\gamma}\Big) < 0, \tag{2.18}$$

for all $\gamma \in (0,1)$ such that for each $\hat{\gamma}$ there exists $\rho^*_{\hat{\gamma}} > 0$ whereby $\frac{\partial\mathbb{E}[|x_{d3}+x_{q3}|]}{\partial\gamma}|_{\gamma=\hat{\gamma}} = 0$.

The trading patterns above always result in lower profits for the quant with greater automation. This is to be expected as greater automation implies a lower adherence to a profit-maximizing strategy, which, by definition, leads to lower profits. Furthermore, a lower correlation level results in greater profits as price impact is lower. For the discretionary investor, profits are rising with greater automation by the quant to the extent that he is able to trade more aggressively. However, the relationship with correlation levels is nontrivial as highlighted above. The speculator sector as a whole makes greater expected profits with more negatively correlated signals and aggregate profits are rising in automation if the quant is less aggressive in her trading.

Greater aggregate trading volume generally results in more informative prices, and price informativeness is falling in correlation. A lower correlation implies that the aggregate order flow is more reflective of the fundamental value, while higher correlation introduces more noise. Formulaically,

$$\frac{\partial}{\partial \rho}\Big(\frac{\partial \mathrm{Var}(v|P_3)^{-1}}{\partial \gamma}\Big) < 0, \tag{2.19}$$

for all $\gamma \in (0,1)$ such that for each $\hat{\gamma}$ there exists $\rho^*_{\hat{\gamma}} > 0$ whereby $\frac{\partial \mathrm{Var}(v|P_3)^{-1}}{\partial \gamma}|_{\gamma=\hat{\gamma}} = 0$. Therefore, price volatility exhibits the same pattern, while return volatility is the reverse. Ultimately, market depth is increasing with correlation, however the rate of increase due to greater automation is decreasing with ρ:

$$\frac{\partial}{\partial \rho}\Big(\frac{\partial \lambda_3^{-1}}{\partial \gamma}\Big) < 0, \tag{2.20}$$

for all $\gamma \in (0,1)$ such that for each $\hat{\gamma}$ there exists $\rho^*_{\hat{\gamma}} > 0$ whereby $\frac{\partial \lambda_3^{-1}}{\partial \gamma}|_{\gamma=\hat{\gamma}} = 0$.

The effect of differential signal quality, highlighted in Figures 2.4 and 2.5, is more straightforward. Great relative signal precision for the quant leads to more trading and more expected profits for the QI, less trading and less expected profits for the DI, and more aggregate trading and aggregate profits for the speculator sector. Greater automation generally dampens these effects, as the quant's trading activity leads to excess information slippage. As such:

$$
\begin{aligned}
&\frac{\partial}{\partial \phi_q}\Big(\frac{\partial \mathbb{E}[|x_{d3}+x_{q3}|]}{\partial \gamma}\Big) > 0 \\
&\frac{\partial}{\partial \phi_q}\Big(\frac{\partial \mathbb{E}[|\pi_{d3}+\pi_{q3}|]}{\partial \gamma}\Big) < 0
\end{aligned}
\tag{2.21}
$$

Greater signal precision for the quant leads to more informative prices, greater price volatility, and lower return volatility. Greater automation strengthens these effects as the quant becomes overly aggressive in her trading approach:

$$
\frac{\partial}{\partial \phi_q}\Big(\frac{\partial \mathrm{Var}(v|P_3)^{-1}}{\partial \gamma}\Big) > 0. \tag{2.22}
$$

Market depth is negatively impacted by the quant's greater signal precision but the effect is attenuated by greater automation levels:

$$
\frac{\partial}{\partial \phi_d}\Big(\frac{\partial \lambda_3^{-1}}{\partial \gamma}\Big) < 0. \tag{2.23}
$$

Given the differential effects of signal precision and information correlation on trading activity and market quality, at first sight it may seem complex to fully describe the effects of the parameter space (ϕ_d, ϕ_q, ρ) on market quality. One may conjecture that for sufficiently low correlation together with sufficiently high information advantage for the quant, market quality would generally benefit from greater automation and vice versa. However, understanding the implications when the parameter effects are offsetting, e.g. high correlation and high information advantage for the quant, appears inherently challenging.

Fortunately, as highlighted by the simulation results in Tables 2.1-2.3, a sufficient statistic defined for each triplet (ϕ_d, ϕ_q, ρ) nearly unambiguously determines the effect of the introduction of and greater automation by the quantitative investor on financial market quality. The sign of the wedge between the trading intensity of a fully-discretionary and a fully-automated quant, i.e. sign($\beta_{q4} - \beta_{q2}$), captures the net effect of the signal precision for each speculator and the signal correlation. As derived in Appendix B.7, the relationship between the trading intensity of a fully-automated and a fully-discretionary quant is jointly determined by (ϕ_d, ϕ_q, ρ), which allows us to fully characterize the parameter space and the effects of any triplet on the sign of the wedge.

We provide a classification of the parameter space via a contour plot illustrated in Fig. 2.6. Specifically, for each pair of signal precision parameters (ϕ_d, ϕ_q), we plot the maximum correlation level for all values below which the fully automated quant trades more than the fully-discretionary quant. It is immediate from the theoretical exercise in Appendix B.7 and apparent from the graph, that for a sufficiently high information advantage for the quant, specifically if $\phi_q > 3\phi_d$, the automated QI will always trade more aggressively than the discretionary QI regardless of how strong the signal correlation is. For a lower information advantage and a sufficiently imprecise signal for the DI, the quant can still absorb a positive correlation, however as the DI's signal increases in precision, we soon reach a point where despite an information advantage, the automated quant does not have a greater trading intensity than a discretionary one.

With an inferior information quality, the quant requires significantly lower correlation levels to have greater trading intensity with full automation. Greater signal precision for the discretionary investor brings down this break-even correlation even further, and, beyond a certain point, the fully automated quant never trades more aggressively than a fully discretionary quant. Ultimately, increases in the signal precision for the discretionary investor always lower the breakeven correlation level required for a more aggressive fully-automated quant. However, the effect of greater signal precision for the quantitative investor is twofold. For a sufficiently low signal precision for the DI (to the left of the $\rho = 0$ contour), greater ϕ_q leads to a *higher* breakeven correlation level. For a sufficiently high signal precision for the DI (to the right of the $\rho = 0$ contour), greater ϕ_q implies a *lower* breakeven correlation level.

We now connect the sign of the trading intensity wedge to the results of the simulation highlighted in Tables 2.1-2.3. A greater trading intensity results in a mostly increasing market depth, increasing price informativeness and volatility, and decreasing return volatility in the level of automation γ by the quant. These overall improvements in market quality are driven by greater expected trading volume for the quant, mostly lower volume for the DI, and greater volume for the speculator sector as a whole. The effects on the quant's profits are ambiguous, while the DI and the speculator sector as a whole mostly see declining profits. Greater automation by the quantitative investor has a positive impact on financial market quality for parameters, which imply greater trading intensity for a fully-automated quant.

In cases where the fully-automated quant trades less than the fully-discretionary quant, market quality generally deteriorates as the lower trading intensity opens the door for the DI to benefit at the expense of the QI. The simulation highlights that market depth, price informativeness, and price volatility unambiguously decrease, while return volatility increases with greater automation by the quant. This is driven by lower trading volume for the quant, greater trading volume for the DI, and lower trading volume for the speculator sector as a whole. Concurrently, profits for the quant unambiguously fall for the quant and increase for the discretionary investor and the speculator sector. In all, market quality suffers from the greater automation by the quant with a negative trading intensity wedge.

Given the differential effects of the trading intensity wedge on market quality, we consider the importance of Figure 2.6. The effects of quantitative investing on market quality will ultimately depend on where we believe we are in the graph. If we believe discretionary funds to have subpar information (e.g. $\rho < 0.5$), improvements in the quality of the information obtained by the quants should lead to greater market quality as market quality will increase for a greater set of potential correlation levels. However, if we believe that discretionary investors have strong information (e.g. $\rho > 0.6$), a more precise signal for the quant makes it less likely that automation improves market quality. More generally, if we believe that information quality is increasing for the speculator sector as a whole, i.e. for both the QI and the DI, the implications to market quality are not trivial since both the starting point and the relative improvement will dictate the effect on market quality.

2.4 Conclusion

Quantitative investment strategies are playing an increasingly important role in the financial markets. What was previously available primarily via hedge funds to a select clientele is now widely accessible for household investments via mutual funds and smart-beta ETFs. The effects of the growth in quantitative investing on the quality of the financial market is not immediate. Quantitative trading is inherently disciplined by backtesting. Precisely due to the reliance on backtesting, quantitative investing may not fully incorporate the strategic trading behavior of other market participants. We develop a model of strategic speculation that captures the potential myopia of quantitative funds via their reliance on a backtested trading strategy – i.e. one assuming by definition that other market participants are unaware of the quant's existence.

The introduction of the quantitative investor to our model broadly benefits market quality. However, greater automation by the quant (i.e. a greater reliance on the backtest) may have disparate effects on market quality. If greater adherence to backtesting results in greater trading intensity for the quant, market quality improves as the strategic market participants are unable to take advantage of the quant's myopia. Conversely, if greater automation leads to less trading by the quant, the fully-strategic speculators may take advantage of the quant's myopia, leading to worse market quality. Importantly, many of these effects depend on potentially observable fund characteristics, such as the relative precision and correlation of the information of discretionary and quant investors.

Our theoretical analysis yields numerous empirical implications. Given the importance of relative signal precision and signal correlation for the discretionary and quantitative investors in determining the effects of quantitative investing on market quality, it is essential to categorize quant and discretionary funds by signal quality. For example, smart-beta ETFs are generally deemed to be less informed than proprietary quant traders, and the correlation of their trading strategies may change over time with potentially material implications for the quality of the affected markets.

Thus, our analysis provides a framework for empirically investigating the market quality implications of the growth in quantitative investing with strategic interaction among investors. Applying the model to the data will shed light on the empirical effects of quantitative investing on the trading behavior of the financial markets.

2.5 Tables and Figures

	Incr.	Decr.	Hump	U	E1<E2	E1<E4	E2>E4	E3>E1 (% γ)	E3<E2 (% γ)	E3>E4 (% γ)	% Sim.
All	43.9	44.6	0.0	11.5	27.1	24.6	46.8	25.5	48.9	45.2	100.0
$\beta_{q4} > \beta_{q2}$	75.1	5.5	0.0	19.5	0.0	0.0	9.0	0.0	12.7	6.4	58.2
$\beta_{q4} \leq \beta_{q2}$	0.5	99.1	0.0	0.5	64.8	58.8	99.4	60.9	99.5	99.2	41.8

(a) Market depth

	Incr.	Decr.	Hump	U	E1<E2	E1<E4	E2>E4	E3>E1 (% γ)	E3<E2 (% γ)	E3>E4 (% γ)	% Sim.
All	57.1	42.9	0.0	0.0	100.0	100.0	42.9	100.0	42.9	42.9	100.0
$\beta_{q4} > \beta_{q2}$	98.0	2.0	0.0	0.0	100.0	100.0	2.0	100.0	2.0	2.0	58.2
$\beta_{q4} \leq \beta_{q2}$	0.1	99.9	0.0	0.0	100.0	100.0	99.9	100.0	99.9	99.9	41.8

(b) Price informativeness

	Incr.	Decr.	Hump	U	E1<E2	E1<E4	E2>E4	E3>E1 (% γ)	E3<E2 (% γ)	E3>E4 (% γ)	% Sim.
All	57.1	42.9	0.0	0.0	100.0	100.0	42.9	100.0	42.9	42.9	100.0
$\beta_{q4} > \beta_{q2}$	98.0	2.0	0.0	0.0	100.0	100.0	2.0	100.0	2.0	2.0	58.2
$\beta_{q4} \leq \beta_{q2}$	0.1	99.9	0.0	0.0	100.0	100.0	99.9	100.0	99.9	99.9	41.8

(c) Price volatility

	Incr.	Decr.	Hump	U	E1<E2	E1<E4	E2>E4	E3>E1 (% γ)	E3<E2 (% γ)	E3>E4 (% γ)	% Sim.
All	42.9	57.1	0.0	0.0	0.0	0.0	57.1	0.0	57.1	57.1	100.0
$\beta_{q4} > \beta_{q2}$	2.0	98.0	0.0	0.0	0.0	0.0	98.0	0.0	98.0	98.0	58.2
$\beta_{q4} \leq \beta_{q2}$	99.9	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.1	41.8

(d) Return volatility

Table 2.1: Simulation Results for Market Quality Measures

We use 100,000 parameter draws: $\rho \sim U(-1, 1), \phi_d \sim U(0, 1), \phi_q \sim U(0, 1), \eta \sim U(0, 2), P_0 = 0, \sigma_v^2 = 1$. For Economy 3, we evaluate market quality for 99 values of γ from 1% to 99% at 1% increments. Numbers represent the percent of simulations such that the condition in the header is satisfied (Col. 1-7) or the average fraction of γ's across all simulations such that the condition satisfied (Col. 8-10). "E" stands for Economy. First four columns document the shape of the market quality measure in Economy 3 for increasing γ.

	Incr.	Decr.	Hump	U	E1<E2	E1<E4	E2>E4	E3>E1 (%γ)	E3<E2 (%γ)	E3>E4 (%γ)	% Sim.
All	57.7	42.1	0.0	0.2	84.3	94.6	42.1	93.2	42.2	42.1	100.0
$\beta_{q4} > \beta_{q2}$	99.2	0.4	0.0	0.4	73.1	90.9	0.6	88.4	0.7	0.5	58.2
$\beta_{q4} \leq \beta_{q2}$	0.0	100.0	0.0	0.0	99.9	99.9	100.0	99.9	100.0	100.0	41.8

(a) Trading volume: speculator sector

	Incr.	Decr.	Hump	U	E1<E2	E1<E4	E2>E4	E3>E1 (%γ)	E3<E2 (%γ)	E3>E4 (%γ)	% Sim.
All	58.8	25.1	0.0	16.0	0.0	1.7	31.4	1.1	34.4	26.9	100.0
$\beta_{q4} > \beta_{q2}$	29.4	43.1	0.0	27.5	0.0	2.9	53.9	1.9	59.1	46.1	58.2
$\beta_{q4} \leq \beta_{q2}$	99.8	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.1	41.8

(b) Trading volume: discretionary investor

	Incr.	Decr.	Hump	U	E1<E2	E1<E4	E2>E4	E3>E1 (%γ)	E3<E2 (%γ)	E3>E4 (%γ)	% Sim.
All	57.1	42.9	0.0	0.0	100.0	100.0	42.9	100.0	42.9	42.9	100.0
$\beta_{q4} > \beta_{q2}$	98.0	2.0	0.0	0.0	100.0	100.0	2.0	100.0	2.0	2.0	58.2
$\beta_{q4} \leq \beta_{q2}$	0.1	99.9	0.0	0.0	100.0	100.0	99.9	100.0	99.9	99.9	41.8

(c) Trading volume: quantitative investor

Table 2.2: Simulation Results for the Speculator Sector Trading Intensity

We use 100,000 parameter draws: $\rho \sim U(-1,1), \phi_d \sim U(0,1), \phi_q \sim U(0,1), \eta \sim U(0,2), P_0 = 0, \sigma_v^2 = 1$. For Economy 3, we evaluate trading volume for 99 values of γ from 1% to 99% at 1% increments. Numbers represent the percent of simulations such that the condition in the header is satisfied (Col. 1-7) or the average fraction of γ's across all simulations such that the condition is satisfied (Col. 8-10). "E" stands for Economy. First four columns document the shape of the market quality measure in Economy 3 for increasing γ.

	Incr.	Decr.	Hump	U	E1<E2	E1<E4	E2>E4	E3>E1 (%γ)	E3<E2 (%γ)	E3>E4 (%γ)	% Sim.
All	44.6	43.9	11.5	0.0	72.9	75.4	53.2	74.5	51.1	54.8	100.0
$\beta_{q4} > \beta_{q2}$	5.5	75.1	19.5	0.0	100.0	100.0	91.0	100.0	87.3	93.6	58.2
$\beta_{q4} \leq \beta_{q2}$	99.1	0.5	0.5	0.0	35.2	41.2	0.6	39.1	0.5	0.8	41.8

(a) Expected profits: speculator sector

	Incr.	Decr.	Hump	U	E1<E2	E1<E4	E2>E4	E3>E1 (%γ)	E3<E2 (%γ)	E3>E4 (%γ)	% Sim.
All	57.0	33.0	0.0	10.0	2.8	7.6	37.9	6.3	39.5	34.3	100.0
$\beta_{q4} > \beta_{q2}$	26.2	56.6	0.0	17.2	4.7	12.9	65.1	10.7	67.8	58.9	58.2
$\beta_{q4} \leq \beta_{q2}$	100.0	0.0	0.0	0.0	0.1	0.1	0.0	0.1	0.0	0.0	41.8

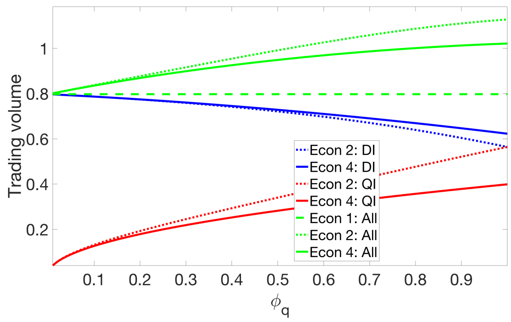
(b) Expected profits: discretionary investor

	Incr.	Decr.	Hump	U	E1<E2	E1<E4	E2>E4	E3>E1 (%γ)	E3<E2 (%γ)	E3>E4 (%γ)	% Sim.
All	13.1	62.4	24.5	0.0	100.0	98.3	80.9	98.9	76.1	85.2	100.0
$\beta_{q4} > \beta_{q2}$	22.4	35.4	42.2	0.0	100.0	97.1	67.3	98.1	59.1	74.7	58.2
$\beta_{q4} \leq \beta_{q2}$	0.1	99.9	0.0	0.0	100.0	100.0	99.9	100.0	99.9	99.9	41.8

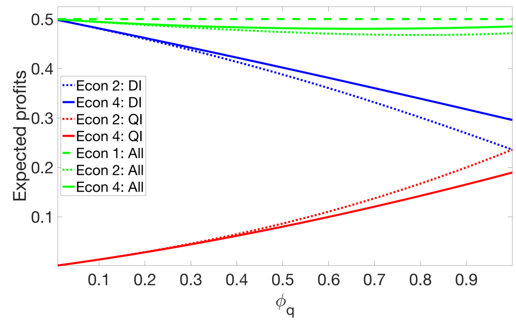
(c) Expected profits: quantitative investor

Table 2.3: Simulation Results for the Speculator Sector Profits

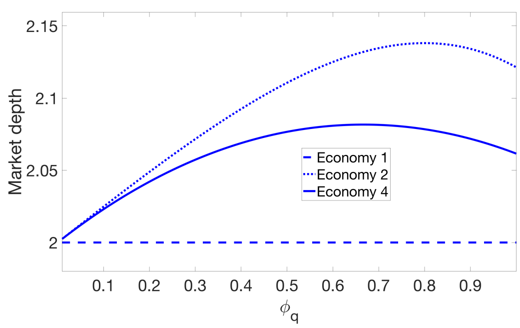
We use 100,000 parameter draws: $\rho \sim U(-1,1), \phi_d \sim U(0,1), \phi_q \sim U(0,1), \eta \sim U(0,2), P_0 = 0, \sigma_v^2 = 1$. For Economy 3, we evaluate expected profits for 99 values of γ from 1% to 99% at 1% increments. Numbers represent the percent of simulations such that the condition in the header is satisfied (Col. 1-7) or the average fraction of γ's across all simulations such that the condition is satisfied (Col. 8-10). "E" stands for Economy. First four columns document the shape of the market quality measure in Economy 3 for increasing γ.



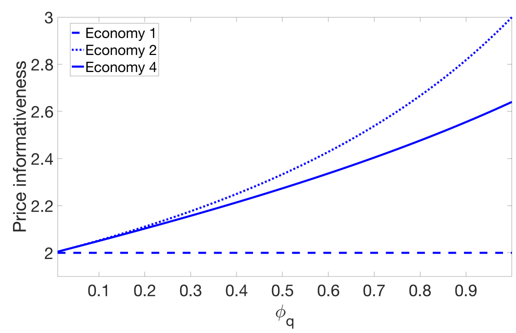
(a) Expected trading volume for speculator sector



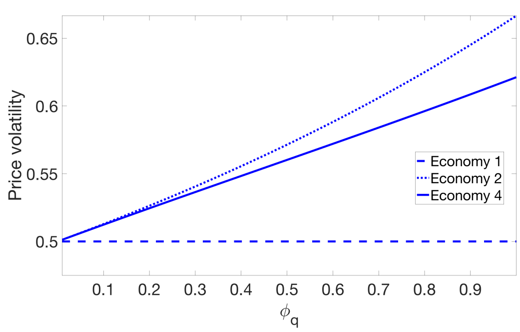
(b) Expected trading profits for speculator sector



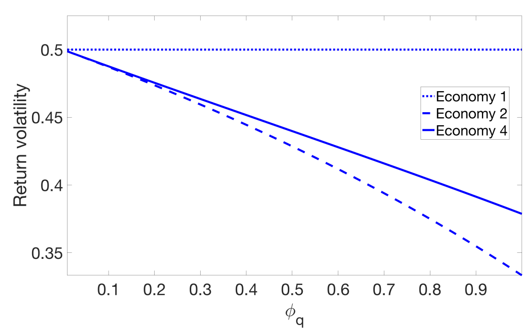
(c) Market depth in Economy j: $1/\lambda_j$



(d) Price informativeness in Economy j: $\mathrm{Var}(v|P_j)^{-1}$

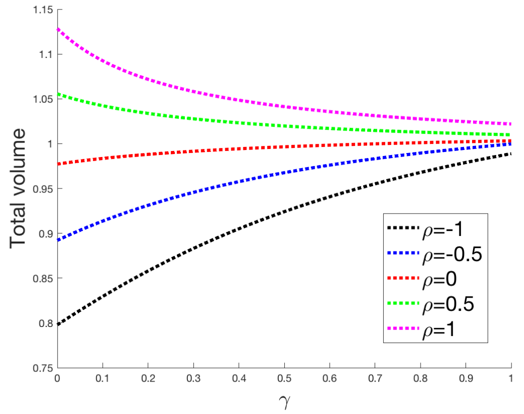


(e) Price volatility in Economy j: $\mathrm{Var}(P_j)$

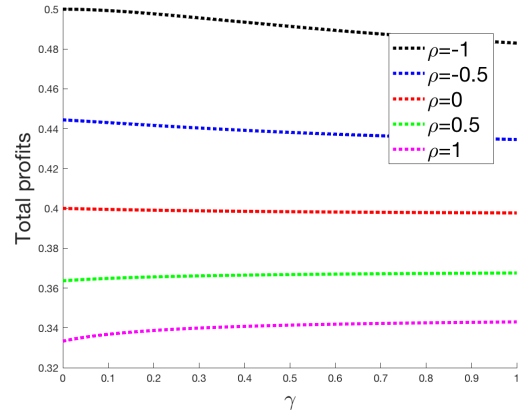


(f) Return volatility in Economy j: $\mathrm{Var}(v - P_j)$

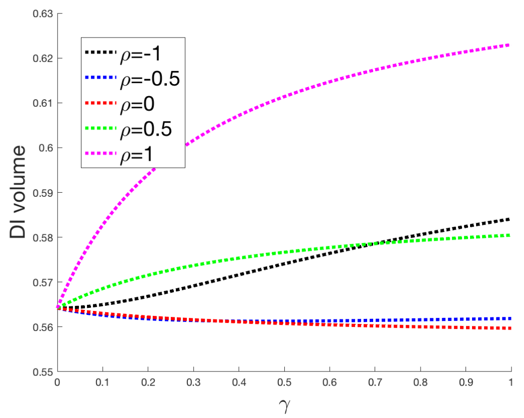
Figure 2.1: Effect of Quantitative Investing on Market Quality by Quant's Information Quality.



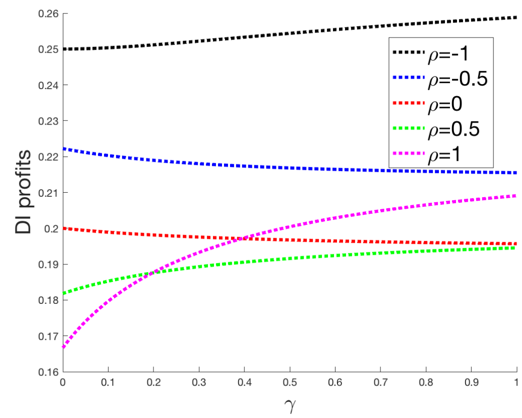
(a) Expected trading volume for speculator sector



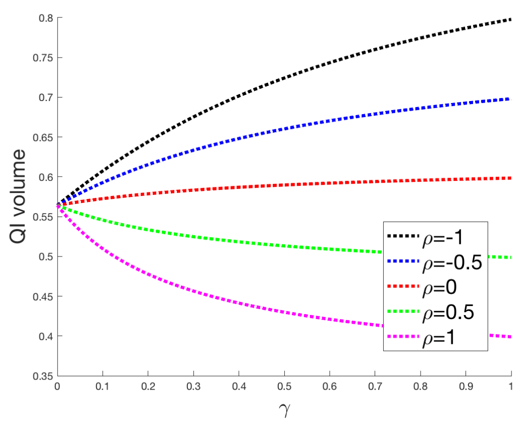
(b) Expected trading profits for speculator sector



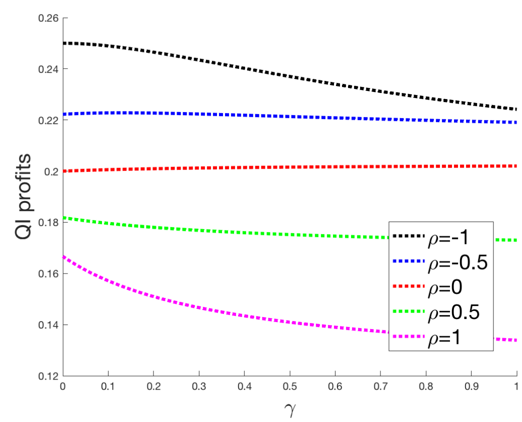
(c) Expected trading volume for DI



(d) Expected trading profits for DI

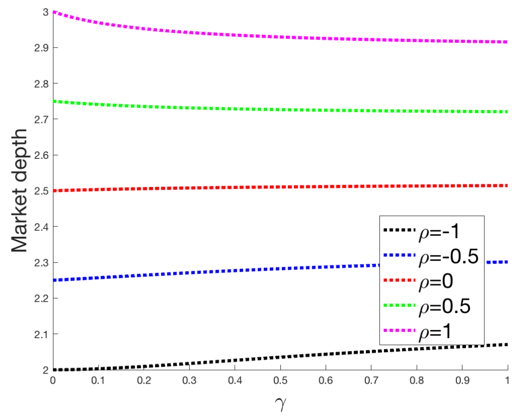


(e) Expected trading volume for QI

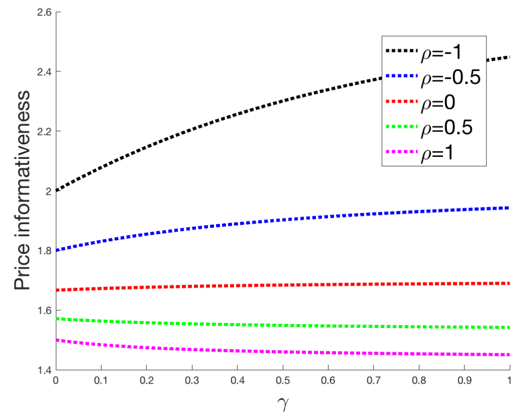


(f) Expected trading profits for QI

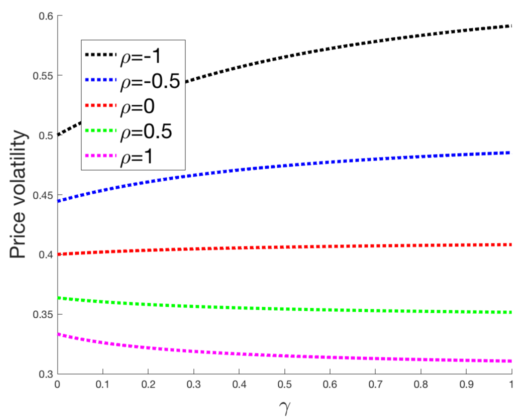
Figure 2.2: Effect of Signal Correlation on Speculator Sector Trading Volume and Profits.
Figures are constructed for the generalized model and assume equal information quality for both the discretionary investor and the quantitative investor: $\phi_d = \phi_q = 0.5$.



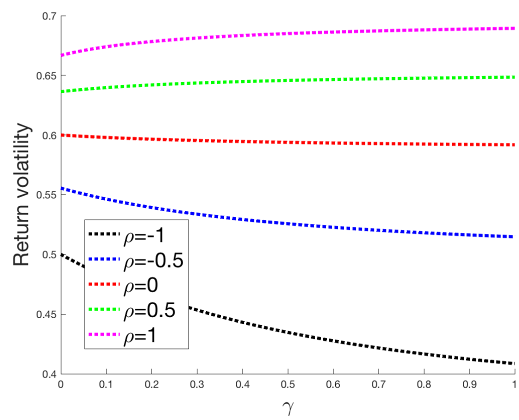
(a) Market depth



(b) Price informativeness

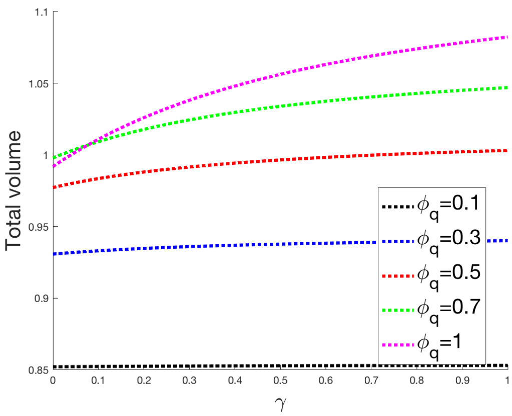


(c) Price volatility

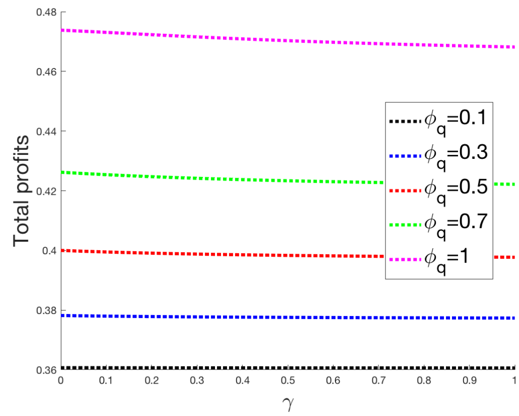


(d) Return volatility

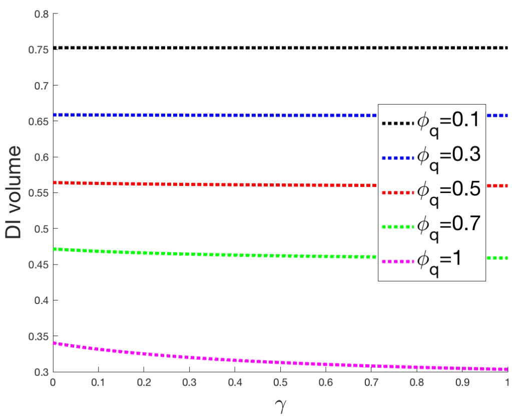
Figure 2.3: Effect of Signal Correlation on Market Quality.
Figures are constructed for the generalized model and assume equal information quality for both the discretionary investor and the quantitative investor: $\phi_d = \phi_q = 0.5$.



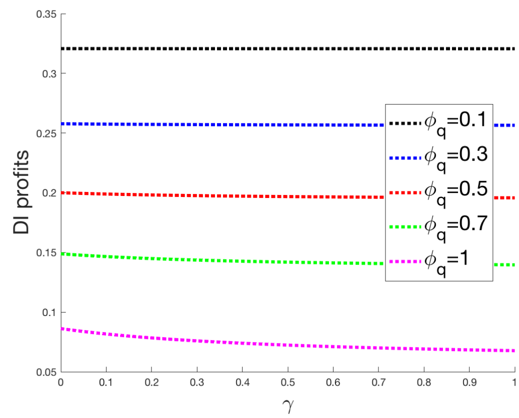
(a) Expected trading volume for speculator sector



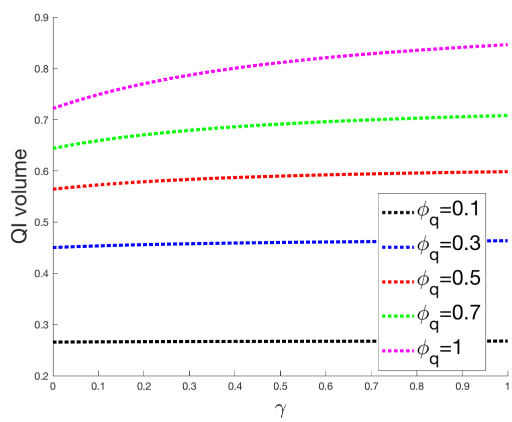
(b) Expected trading profits for speculator sector



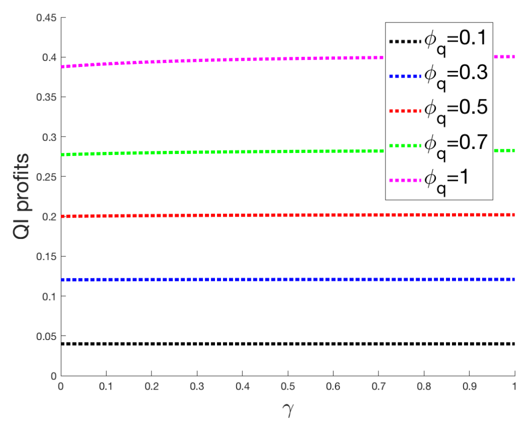
(c) Expected trading volume for DI



(d) Expected trading profits for DI

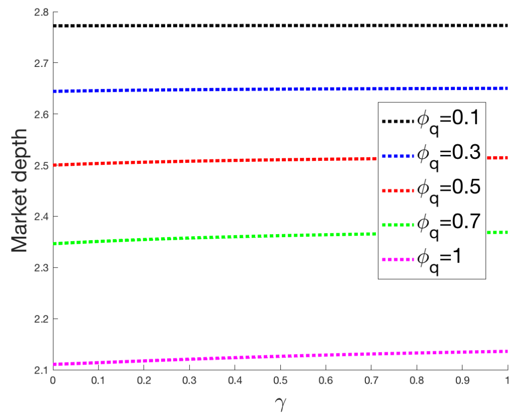


(e) Expected trading volume for QI

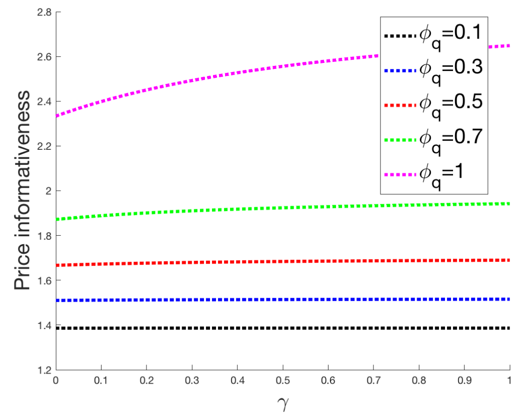


(f) Expected trading profits for QI

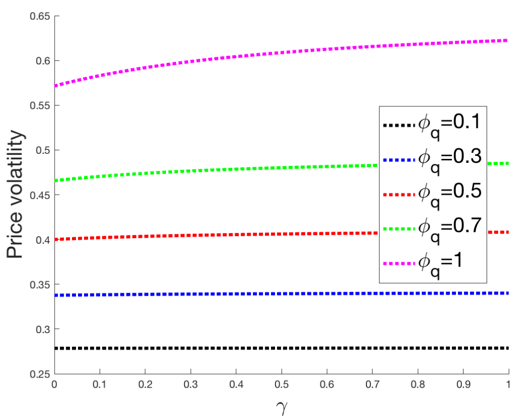
Figure 2.4: Effect of Relative Signal Precision on the Speculator Sector. Figures are constructed for the generalized model, assume uncorrelated signals for the discretionary and the quantitative investor, and allow for potentially inferior or superior information for the discretionary investor: $\phi_d = 0.5, \rho = 0$.



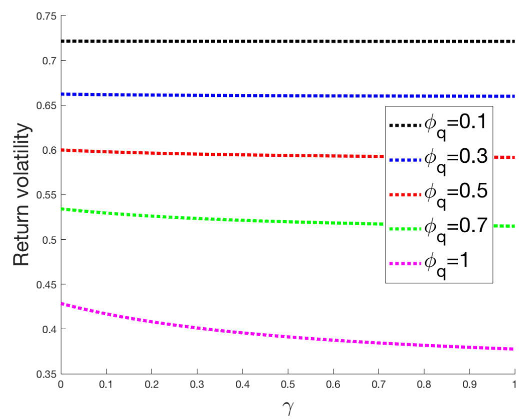
(a) Market depth



(b) Price informativeness



(c) Price volatility



(d) Return volatility

Figure 2.5: Effect of Relative Signal Precision on Market Quality

Figures are constructed for the generalized model, assume uncorrelated signals for the discretionary and the quantitative investor, and allow for potentially inferior or superior information for the discretionary investor: $\phi_d = 0.5, \rho = 0$.

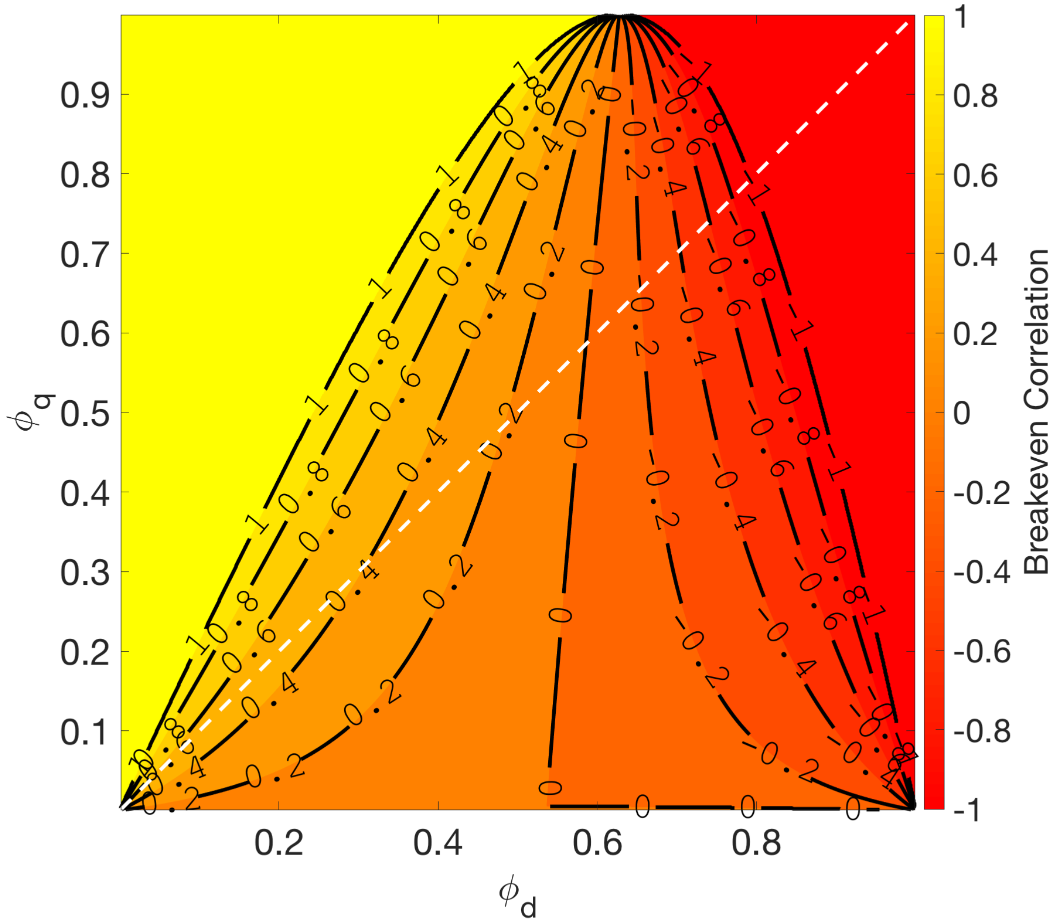


Figure 2.6: Breakeven Correlation for the Quant's Trading Wedge
The contour plot demonstrates the maximum signal correlation levels such that the fully discretionary quant (Economy 2) trades less than the fully automated quant (Economy 4) for varying levels of signal precision for the DI and QI. The white dotted line highlights equal signal precision for the DI and QI, with the region above signifying an informational advantage for the quant, and the region below an informational advantage for the discretionary investor.

CHAPTER III

Arbitrage as Camouflage

3.1 Introduction

The dramatic proliferation of exchange traded funds (ETFs) over the past few decades has greatly increased the interconnectedness of the financial markets. By design, ETFs track a basket of underlying securities, yet may deviate from their net asset value (NAV) as dictated by the exposure of the basket via the forces of supply and demand in the secondary markets. The associated law of one price violations are continuously exploited by various market players including authorized participants (APs) who are incentivized to ensure that ETF prices do not diverge from their NAV drastically and various arbitrageurs such as high frequency traders at hedge funds and other sophisticated investors. The resulting buying and selling in the synthetic security and its constituents, which attempts to ensure that the ETF tracks its underlying securities, may be viewed as a form of noise trading as it is ultimately agnostic to fundamental value. In the present work, I theoretically explore the effects of arbitrage trading due to law of one price violations on market quality in both the underlying and synthetic securities.

I first explore the implications of introducing arbitrage trading to a market that will otherwise see different prices of assets with identical payoffs. The markets for the underlying and synthetic assets are assumed to be initially segmented via different levels of noise trading, which in turn leads to price divergence. The introduction of an arbitrageur who ensures that prices of assets with identical payoffs converge, leads to an averaging of noise trading activity across the otherwise segmented markets. The arbitrageur buys (sells) in the market with excess noise supply (demand), implying that effective noise trading in each market is a weighted average (based on market depth) of noise trading in both markets. This smoothing diminishes the level of available camouflage (i.e. Kyle (1985)) via a lower effective noise trading volatility for the informed speculator and increases concerns of adverse selection for the market makers, resulting in a declining trading intensity for the informed speculator, lower market depth, and unchanged price informativeness from a world with law of one price violations. Further segmenting the markets by introducing asset-specific speculators attenuates the negative effects above, as the arbitrageur links the markets, thereby inducing competition among the informed speculators akin to Holden and Subrahmanyam (1992). The increased competition leads to greater trading intensity by the speculators, greater market depth set by the market makers, and more informative prices relative to a world with a single speculator and arbitrageur.

The analysis above assumes the existence of a market for both the underlying and the synthetic assets. The results are dramatically different if one were to study the *introduction* of a synthetic security. The arbitrageur connects the synthetic and the underlying markets and his activity results in an averaging of noise trading between the two markets. To the extent that the synthetic asset attracts sufficient noise trading, the trading intensity of the informed speculator will rise in the underlying market and the market maker will be able to provide greater market depth while price informativeness remains unchanged. The result is even more pronounced in the case of greater market segmentation via asset-specific speculators.

In the main analysis, I model arbitrage trading as an activity separate from market making. However, much of the leading theoretical literature on ETFs (e.g. Bhattacharya and O'Hara (2018)), models arbitrage activity via the price updating process of the market makers at an intermediate stage. The market makers will take all information available from order flow across markets, and will set prices of assets with identical payoffs equal conditional on the same information set. This assumption boils down to having the market maker also perform the function of an authorized participant, which need not always be the case institutionally. I directly compare the two approaches to arbitrage activity, and document the following differential implications to market quality.

A stand-alone arbitrageur is generally agnostic to fundamental payoffs as he cares only about price convergence, and therefore trades based on the relative order flow in both markets to ensure that prices equate. His presence allows the informed speculators to be more aggressive in their trading. However, the market makers become more concerned about adverse selection. Conversely, the price-updating market maker cares about adverse selection, and is therefore able to more accurately impute the asset payoff via information contained in the order flow of segmented markets. This reduces his fear of adverse selection, yet dampens the ability of the informed speculators to trade aggressively. The net effect of the two offsetting forces are identically informative prices for both the stand-alone arbitrageur and the price-updating market maker. However, the implications to various measures of market quality such as informed trading intensity and market depth are different, and are governed by the respective objective of the arbitrageur and the market maker.

My model provides various testable empirical implications. An immediate outcome is that the introduction of an ETF, which attracts significant noise trading volume, should result in greater informed trading in the underlying assets and greater market depth, and should not distort price informativeness. Exploiting the introduction of ETFs across various markets, one can rely on empirical measures of price informativeness (e.g. Bai et al. (2016)), probability of informed trading (e.g. Easley et al. (1996), Duarte and Young (2009)), liquidity (e.g Amihud (2002), Holden and Jacobsen (2014)), and adverse selection (e.g. Hasbrouck (1988), Hasbrouck (1991), Glosten and Harris (1988)) to identify whether empirically the implications of the model stand up. Furthermore, the nuanced market quality effects of stand-alone arbitrageurs versus market makers jointly acting as APs allow for a cross-sectional exploration of ETFs based on their exposure to stand-alone authorized participants versus market makers also performing AP functions. Greater exposure to the latter should result in more liquid prices with lower informed trading intensity, while exposure to the former should imply lower liquidity and greater informed trading intensity.

The present work contributes to the rapidly growing theoretical and empirical literature on ETFs. Surprisingly, research specifically tackling the question of strategic speculation in the presence of arbitrage trading and its impact on market quality has been limited. To my knowledge, the most closely related paper would be Shim (2020), who argues that arbitrage trading distorts asset prices as it trades based on mechanically assigned weights versus the true exposure of the underlying securities to factors. My model is complementary to his work: I focus on the effects of strategic speculation in the presence of arbitrage trading and abstract away from the inefficiencies associated with the weighting schemes of multi-security ETFs. Bhattacharya and O'Hara (2018) is another complementary paper analyzing the effects of ETFs on hard-to-access markets, whereby the no-arbitrage mechanism is embedded via market makers' short-term price adjustments. The distinguishing feature of the present framework is its explicit focus on the implications of arbitrage trading on financial market quality.

On the empirical side, much work has attempted to discern whether ETFs improve or impede price discovery in the underlying securities with conflicting results. On the one hand, Glosten et al. (2017) highlight that greater ETF ownership leads to more accurate incorporation of accounting information suggesting that price efficiency may be rising due to ETF ownership. On the other hand, in a study of Russell index reconstitutions, Coles et al. (2018) find that weak-form price efficiency deteriorates due to index funds and ETFs. Furthermore, Ben-David et al. (2018) find that greater ETF ownership leads to greater return volatility also implying lower price efficiency. Similarly, Israeli et al. (2017) report that ETF ownership results in decreasing price efficiency. Several papers, including Easley et al. (2020) and Huang et al. (2021) point out various advantages of ETFs via their ability to serve as hedging instruments and their optimal "activeness."

3.2 Model

I model a three-date financial market with two risky assets with identical payoffs. Participants include up to two perfectly informed speculators, up to one arbitrageur, noise traders, and up to two market makers. At time $t = 0$, informed speculators observe the payoff of the risky assets. At time $t = 1$, speculators, noise traders, and the arbitrageur (if present) submit their demands for the risky asset and market makers set prices according to the aggregate order flow. At time $t = 2$, the risky asset payoffs are realized. Agents have rational expectations in that the informed speculators, the market makers, and the arbitrageur are aware of the model parameters and each others' price setting and trading behavior respectively. I describe each feature of the model below.

3.2.1 Model Setup

Economies

There will be four market environments, or economies, indexed by $j \in \{b, a, s, m\}$, which will differ based on the number of informed speculators, the number of market makers, and the presence of an arbitrageur who will ensure that prices of assets with identical payoffs converge. Each economy will have two risky assets traded in separate markets with identical payoffs but potentially different prices due to differential noise trading. Economy b will be the benchmark economy and will feature a single informed speculator allocating to the two risky assets. Prices may diverge in Economy b due to differential levels of noise trading in the two markets. Economy a will attempt to isolate the effect of arbitrage trading while holding all else equal by introducing an arbitrageur to Economy b who will ensure that the prices of the two risky assets converge. Economy b is introduced primarily for theoretical purposes, to highlight the effects of a stand-alone arbitrageur. Economy s will identify the effect of arbitrage trading in the presence of segmented markets, whereby each risky asset will have its own informed speculator. Economy s will speak to the effects of arbitrage activity in markets in which one can plausibly assume separate speculators for the underlying and synthetic assets. Finally, Economy m will substitute the arbitrageur for a single market maker, who sets prices to be the same based on order flow in both underlying markets. Economy m will address institutionally-relevant markets in which the market maker in the underlying and synthetic securities jointly acts as an AP.

Risky assets

Each economy will feature two risky assets $i \in \{1, 2\}$ traded in different markets with identical payoffs v distributed according to $v \sim N(0, \sigma_v^2)$, realized in the second period. The identical payoffs are meant to resemble the relationship between an ETF and the underlying securities, which, absent frictions, should yield identical payoffs. The endogenously determined price P_{ij} set by the market maker in the first period is indexed by Economy j as it will depend on the presence of the arbitrageur and the number and trading behavior of the informed speculators. I label asset 1 as the "underlying" and asset 2 as the "synthetic" security without loss of generality.

Noise traders and segmented markets

To incorporate law of one price violations I assume segmented markets for the two assets. Specifically, I require that each Asset i trades only in Market i, which has an asset-specific level of noise trading z_i distributed according to $N(0, \sigma_{z_i}^2)$, such that the level of noise trading is orthogonal to the fundamental payoff ($\sigma_{vz_i} = 0$).[1] Some examples of noise demand include uninformed retail investors buying or selling shares for liquidity needs, informed asset managers buying or selling securities purely for hedging purposes, or corporations unexpectedly issuing or buying back stock. One can think of reasons why noise trading may differ between the synthetic and underlying securities. For example, ETF purchases and sales may be driven by household wealth shocks, or hedge fund hedging needs, while sales and purchases in the underlying may be driven by corporate secondary equity offerings and buy-backs. I will rely on $\eta_i = \sigma_{z_i}^2/\sigma_v^2$ as a normalized measure of noise trading in Market i. The differential levels of noise trading for the two markets will result in law of one price violations as assets will be functions of asset-specific order flow implying that assets with identical payoffs will have different prices.

[1] σ_{xy} refers to the covariance between x and y.

Informed speculators

Each economy will feature up to two informed speculators who observe the payoff v. Each Asset i in Economy j will have exactly one speculator investing in the asset to maximize profits. As in Kyle (1985), in placing her order, the speculator will balance her knowledge of the payoff with revealing her knowledge to the market maker via her demand. For Economy $j \in \{b, a\}$, where a single strategic speculator invests in both assets, her optimization is as follows:

$$[x_{1j}, x_{2j}] = \underset{\tilde{x}_{1j}, \tilde{x}_{2j}}{\arg\max}\, \mathbb{E}[\tilde{x}_{1j}(v - P_{1j}) + \tilde{x}_{2j}(v - P_{2j})|v]. \tag{3.1}$$

For Economy $j \in \{s, m\}$, which includes asset-specific speculators, whereby Speculator i is assigned to Asset i, Speculator i's demand will be:

$$x_{ij} = \underset{\tilde{x}_{ij}}{\arg\max}\, \mathbb{E}[\tilde{x}_{ij}(v - P_{ij})|v]. \tag{3.2}$$

Aribtrageur

The arbitrageur will be introduced in Economies a and s to eliminate law of one price violations. I will assume that the arbitrageur buys ε_j shares in Market 1 of Economy j and sells ε_j shares in Market 2 of Economy j. The quantity bought and sold will be endogenously determined, and, ultimately, the arbitrageur makes exactly zero in profits, which is consistent with a competitive arbitrage market. For Economy $j \in \{a, s\}$, the arbitrageur will submit the following orders in Markets 1 and 2 to ensure that prices converge:

$$P_{1j}(x_{1j} + z_1 + \varepsilon_j) = P_{2j}(x_{2j} + z_2 - \varepsilon_j) \tag{3.3}$$

My notion of arbitrageurs most closely resembles authorized participants (APs) in the financial markets, who are responsible for ensuring that ETF prices do not deviate drastically from the value of the underlying basket, and to accomplish this role are granted the ability to issue and buy back ETF securities exactly at NAV.

Market makers

Each Market i in Economy $j \in \{b, a, s\}$ will have two market makers, each dedicated to Asset i such that he only observes order flow in his Market i, and is assumed to have risk-neutral preferences and operate in a competitive environment. Economy m will feature a single market maker who will observe order flow and set prices both in Market 1 and Market 2. For Economy $j \in \{b, a, s\}$, given aggregate order flow ω_{ij}, the market maker sets a price such that he breaks even in expectation:

$$P_{ij}(\omega_{ij}) = \mathbb{E}[v|\omega_{ij}], \tag{3.4}$$

i.e. such that the equilibrium asset price is semi-strong from efficient. For Economy m, the market maker will take advantage of all available information in both Market 1 and Market 2, to set the same price in both markets:

$$P_{1m}(\omega_{1m}, \omega_{2m}) = P_{2m}(\omega_{1m}, \omega_{2m}) \equiv P_m(\omega_{1m}, \omega_{2m}) = \mathbb{E}[v|\omega_{1m}, \omega_{2m}]. \tag{3.5}$$

An underlying assumption in this framework is that the market maker has sufficient inventory (if selling) or liquidity (if buying) to satisfy the net demands of the traders. The market maker attempts to discern between the informed and uninformed order flow and to set prices accordingly.

The aggregate order flow will differ based on whether an arbitrageur is present or if there are two speculators. As such, order flow will be

$$\omega_{1j} = x_{1j} + z_1 + \mathbb{1}_{j \in \{a,s\}} \varepsilon \tag{3.6}$$

for the underlying asset, and

$$\omega_{2j} = x_{2j} + z_2 - \mathbb{1}_{j \in \{a,s\}} \varepsilon \tag{3.7}$$

for the synthetic asset. The arbitrageur's buying and selling in Markets 1 and 2 respectively will only take place in Economies a and s. For these Economies I assume that the arbitrageur and the market makers operate independently, which is consistent with some industry segmentation between market makers and authorized participants. Implicit in this assumption is the arbitrageur's ability to infer the pricing rule of the market maker and the demand functions of the informed speculators. Economy m relaxes this assumption, and allows the market maker to partake in both functions by allowing him to update prices based on the order flow in each market.

3.2.2 Model Solution

I aim to identify the effect of arbitrage trading activity on various measures of market quality such as trading intensity, liquidity, and price informativeness. As such, I begin with the simplest possible benchmark (Economy b), within which law of one price violations may occur and identify the relevant measures of market quality in this setting. Next, I introduce the arbitrageur in Economy a, who ensures that the law of one price holds and revisit the effects on market quality. In Economy s, I allow for potential segmentation between the informed speculators, assigning a market-specific informed speculator, and explore how the activity of the arbitrageur interacts with this segmentation. Finally, Economy m replaces the arbitrageur by allowing a single market maker to set prices, thereby automatically setting prices equal to each other.

For each Economy j I assume that the demand function of the speculator in Asset i is linear in her signal

$$x_{ij} = \beta_{ij} v, \tag{3.8}$$

and the market maker's pricing function is linear either in aggregate order flow for Economy $j \in \{b, a, s\}$

$$P_{ij} = \lambda_{ij} \omega_{ij}, \tag{3.9}$$

or a linear combination of the order flow from the two markets in Economy m

$$P_m = \lambda_{1m} \omega_{1m} + \lambda_{2m} \omega_{2m}. \tag{3.10}$$

I then derive the coefficients that satisfy the speculators', arbitrageur's, and market makers' optimization functions given by Eq. (3.1)-(3.5). I refer to β_{ij} as the "trading intensity" of the informed speculator in Asset i as it reflects how aggressively the speculator reacts to an incremental unit of signal. I identify $1/\lambda_{ij}$ as "market depth," my primary measure of liquidity: the inverse of the market maker's price sensitivity to an incremental share of order flow.

I solve the model for all Economies in Appendices C.1.1,C.1.2,C.1.3, and C.1.4, and summarize the solutions in Proposition III.1 below.

Proposition III.1 (Existence and Lack of Uniqueness of Linear Equilibrium)**.** *There exists a linear trading rule for the speculator(s), specified by Eq.* (3.8)*, and a linear pricing function for the market maker(s), specified by Eq.* (3.9) *or* (3.10)*, with positive coefficients, such that the speculator(s) maximize expected profits, specified by Eq.* (3.1) *or* (3.2)*, the market maker(s) set prices to break even, specified by Eq.* (3.4) *and* (3.5)*, and* $P_{1j} = P_{2j}$ *for* $j \in \{a, s, m\}$*, for all realizations of* v*,* z_1*, and* z_2*. Positive coefficients are uniquely identified for Economies* b *and* m*, and symmetric equilibria with positive coefficients are uniquely identified for Economies* a *and* s*, which have a continuum of linear equilibria.*

I now describe and summarize my findings for each Economy all of which are derived in Appendices C.1.1-C.1.4.

Benchmark: Economy b This is a simple extension of Kyle (1985) to include a second asset and partially segment the market for Asset 1 and Asset 2 via different levels of noise trading and asset-specific market makers. The segmentation is partial since there continues to be a single strategic speculator trading in both assets. The trading intensity for the informed speculator and each market maker's pricing slope for each Asset i has the standard form:

$$\beta_{ib} = \sqrt{\eta_i} \tag{3.11}$$

$$\lambda_{ib} = \frac{1}{2\sqrt{\eta_i}} \tag{3.12}$$

Both market depth ($1/\lambda_{ib}$) and trading intensity are increasing in the amount of normalized noise trading as the speculator has greater camouflage to trade more aggressively, while the market makers are less concerned about adverse selection. The two forces offset resulting in a constant (inverse) price informativeness:

$$\mathrm{Var}(v|P_{ib}) = \frac{1}{2}\sigma_v^2. \tag{3.13}$$

Introducing arbitrageur: Economy a I now introduce the arbitrageur who will buy ε shares in Market 1 and sell ε shares in Market 2 so that the prices in the two markets are equal. The change in market depth and trading intensity from Economy b will highlight the effect of arbitrage trading. As pointed out in Prop. III.1, there is a continuum of equilibria in this setting. This is intuitive since the speculator is active in both Markets 1 and 2, and prices will be equal in the two markets by design. Therefore, she will be indifferent between investing in either market and what will matter is the aggregate amount of informed trading and the aggregate market depth, both of which are uniquely identified for upwards sloping supply curves:

$$\beta_{1a} + \beta_{2a} = \sqrt{\eta_1 + \eta_2} \tag{3.14}$$

$$\frac{1}{\lambda_{1a}} + \frac{1}{\lambda_{2a}} = 2\sqrt{\eta_1 + \eta_2} \tag{3.15}$$

A similar pattern to Economy b holds, whereby aggregate market depth and aggregate trading intensity are increasing in the amount of noise trading. It is important to note that arbitrage trading links noise trading in the two markets via the activity of the arbitrageur, resulting in aggregate measures of noise trading (i.e. $\eta_1 + \eta_2$) as the defining equilibrium parameters. The market with greater (lower) market-depth adjusted noise trading will see the arb *sell* (*buy*) ε shares. Therefore, the effective level of noise trading in each market will be a weighted average of noise trading in both markets, thereby reducing its volatility in each.[2] Ultimately this prevents the informed speculator from trading as aggressively as in the benchmark economy as

$$\beta_{1a} + \beta_{2a} < \beta_{1b} + \beta_{2b},$$

[2]For further intuition consider the case where $\sigma^2_{z_1} = \sigma^2_{z_2} = \sigma^2_z$ and the realization of noise trading is such that $z_1 > z_2$. Here the model is fully symmetric and the arb will sell $\varepsilon = (z_1 - z_2)/2$ in Market 1 and buy $\varepsilon = (z_1 - z_2)/2$ in Market 2 resulting in equal effective noise trading of $(z_1 + z_2)/2$ in each market. The "averaging" of noise trading leads to its lower volatility in each market of $\sigma^2_z/2$.

and the market makers react to the reduced effective amount of noise trading by lowering market depth

$$\frac{1}{\lambda_{1a}} + \frac{1}{\lambda_{2a}} < \frac{1}{\lambda_{1b}} + \frac{1}{\lambda_{2b}}.$$

The net effect of lower camouflage and greater adverse selection on price informativeness is identical to Economy b as

$$\mathrm{Var}(v|P_{ia}) = \frac{1}{2}\sigma_v^2. \tag{3.16}$$

Overall, the introduction of arbitrage trading is fully internalized by all market participants allowing for both lower trading intensity and less liquid markets, while maintaining price informativeness.

Segmenting informed speculators: Economy s I now consider the realistic possibility that markets may not only be segmented for the market makers, but also for the informed speculators. As such, I introduce a second speculator, with Speculator 1 specializing in Asset 1 and Speculator 2 specializing in Asset 2, while keeping all other parameters of the model constant. In this framework, both speculators and market makers are segmented, with the only link between the two markets being the arbitrageur. The trading intensity of each Speculator i is identified as

$$\beta_{is} = \sqrt{\frac{\eta_1 + \eta_2}{2}}, \tag{3.17}$$

while the aggregate market depth is

$$\frac{1}{\lambda_{1s}} + \frac{1}{\lambda_{2s}} = 3\sqrt{\frac{\eta_1 + \eta_2}{2}}, \tag{3.18}$$

and if I assume a symmetric equilibrium, market depth for Asset i will be

$$\frac{1}{\lambda_{is}} = \frac{3}{2}\sqrt{\frac{\eta_1 + \eta_2}{2}}. \tag{3.19}$$

Segmenting the markets leads to greater aggregate trading intensity and greater market depth versus both the benchmark case and Economy a. Despite segmented markets each speculator no longer has a monopoly on her information as informed trading in one market will be translated to the other market via the activity of the arbitrageur. For example, more aggressive buying by one of the speculators will lead to more aggressive buying by the arb in the other market, therefore affecting the trading behavior of the other speculator. The introduction of competition among the speculators leads to improved market quality, which translates into more informative prices:

$$\mathrm{Var}(v|P_{is}) = \frac{1}{3}\sigma_v^2. \tag{3.20}$$

Single market maker: Economy m In the final step, I consider the implications of shutting down the arbitrage trading channel, and, instead, allowing market makers to observe order flow across the segmented markets (e.g. Bhattacharya and O'Hara (2018)). Since order flow in each market will be an incremental source of information, the market maker will condition on order flow across markets when setting prices in each market, which will result in equal prices. The trading intensity of each speculator reverts to the benchmark with law of one price violations

$$\beta_{im} = \sqrt{\eta_i}, \tag{3.21}$$

however market depth increases to

$$\frac{1}{\lambda_{im}} = 3\sqrt{\eta_i}, \tag{3.22}$$

as the market maker becomes less concerned about adverse selection given the incremental information via order flow. Although the aggregate trading intensity is lower than the fully segmented setting, the greater aggregate market depth coupled with a more precise signal for the market maker, results in an equal level of price informativeness to Economy s:

$$\mathrm{Var}(v|P_{im}) = \frac{1}{3}\sigma_v^2. \tag{3.23}$$

3.3 Discussion

Table 3.1 summarizes my main findings for market quality across all economies. The introduction of an arbitrageur to the benchmark economy has the effect of absorbing noise trading in each of the markets, since the market with the higher prices will have seen greater noise demand, while the market with the lower price will have seen lower noise demand. The arbitrageur steps in to correct these discrepancies, which implies that the effective level of noise trading in each market becomes a (market depth based) weighted average of noise trading in both markets. The averaging of noise trading leads to a decrease in camouflage, which is internalized by both the speculator and the market makers, leading to both lower trading intensity and lower market depth, which, ultimately, does not impact price informativeness relative to the benchmark.

It is important to note that the changes from Economy b to Economy a assume that both the underlying and synthetic securities previously existed, and that the level of noise trading has remained the same in both. However, if one is exploring the *introduction* of a synthetic security, which will bring incremental noise traders to the market, then the implications are drastically different. The arbitrageur will "transfer" the noise trading from the synthetic security to the underlying, implying that the strategic speculator will be able to trade more aggressively in the underlying as long as $3\eta_1 < \eta_2$, i.e. the synthetic market is sufficiently noisy. Furthermore, the market maker in the underlying will be able to provide greater liquidity due to lower adverse selection concerns. The informed speculator will also be able to participate in the synthetic security, resulting in unambiguously greater aggregate informed trading intensity and higher aggregate liquidity.

Segmenting the asset markets by assigning asset-specific speculators in Economy s improves market quality relative to the case of a single speculator in Economy a as the trading intensity of the informed speculators is higher in each asset and, in aggregate, market depth rises unequivocally, and prices become more informative as a result. The strategic speculator of Economy a no longer has a monopoly on all of the noise traders in both the underlying and synthetic securities. Due to the connectedness of the two markets via the arbitrageur, this economy is akin to a multi-speculator version of Kyle (1985) (e.g. Holden and Subrahmanyam (1992)), whereby the two speculators compete more aggressively than one.

With asset-specific speculators in Economy s, the threshold for improvements in market quality with the *introduction* of a synthetic security are lower than with a single speculator in Economy a. As long as $\eta_2 > \eta_1$, i.e. if the synthetic market brings more noise traders than already exist in the underlying, the informed speculator will be able to trade more aggressively on her information. The threshold for the market maker's improvement in liquidity for the underlying will also be lower at $\frac{23}{9}\eta_1 < \eta_2$.[3] Ultimately, regardless of parameter values, the introduction of a synthetic asset has

[3]Condition under which $\frac{1}{\lambda_{1s}} > \frac{1}{\lambda_{1b}} \Leftrightarrow \frac{3}{2}\sqrt{\frac{\eta_1+\eta_2}{2}} > 2\sqrt{\eta_1}$.

the effect of increasing competition for the incumbent speculator, resulting in more informative prices.

Finally, the substitution of a dedicated arbitrageur with a single market maker who observes order flow in both markets and accordingly uses all available information to set prices, has a subtle effect on market quality. On one hand, the trading intensity of the informed speculator simplifies to that of the benchmark economy, and improves relative to Economy s in Asset i if $\eta_i > \eta_{-i}$, that is if she belongs to the market with greater trading intensity. Aggregate trading intensity always decreases relative to Economy s. On the other hand, the market maker now observes incremental information about the asset payoff via access to the order flow in both markets, which unequivocally leads to greater market depth both relative to Economy s and the benchmark Economy b. The lower concerns of the market maker about adverse selection exactly offset the decline in camouflage available to the speculators due to the absence of arbitrage trading resulting in price informativeness identical to Economy s.

The differential outcomes to the informed speculators' trading intensity and the liquidity levels are driven by the differential objectives of the market maker and the arbitrageur. The arbitrageur only cares about having the prices converge, and, therefore, is agnostic to asset fundamentals. His trading is a form of noise trading, which acts as a link between the noise trading levels in the segmented markets and allows the speculators to trade more aggressively. The more aggressive trading causes aggregate liquidity to decline as market makers become more concerned about adverse selection. On the other hand, the single market maker is concerned about adverse selection in the underlying markets, and therefore uses both order flows to more accurately impute the fundamental value. This, in turn, has the effect of decreasing his concerns about adverse selection, which allows him to provide greater liquidity, but also decreases the aggregate trading intensity of the speculator sector. The opposite effects of changes to liquidity and trading intensity as one moves from Economy s to Economy m, exactly offset such that price informativeness does not change.

3.4 Conclusion

I present a stylized model to highlight the implications of arbitrage trading in ETFs and their underlying securities on financial market quality. On one hand, the introduction of arbitrage trading to segmented markets with otherwise diverging prices, results in a "smoothing" of noise trading across the markets, reducing the levels of camouflage available to the informed speculator and required for the market maker to combat adverse selection risks. On the other hand, the introduction of an ETF with a threshold level of noise trading leads to unambiguous improvements in the market quality of the underlying security as some of the incremental noise trading is transferred to the underlying market. Furthermore, I highlight the differential effects on market quality of stand-alone arbitrageurs and market makers jointly serving as authorized participants, with the former leading to greater informed trading intensity for the speculators and greater adverse selection for the market makers, and vice versa for the latter.

The present framework provides for ample testable empirical implications. The focus on a single-asset model in the underlying security, thereby abstracting away from the complexities associated with multi-asset ETFs, makes the model most applicable to commodity ETFs. For example, one can explore the effects of the introduction of the United States Oil Fund (USO) ETF on market quality measures in the futures and spot markets. Furthermore, given the differential impacts to market quality of stand-alone arbitrageurs and market makers also serving as authorized participants, one can exploit cross-sectional variation in ETFs' exposure to either category of arbitrageurs to analyze whether the market quality measures for the underlying securities are consistent with the model presented. The prominence and the pace of proliferation of ETFs underscore the importance of understanding their effects on the financial markets.

3.5 Tables and Figures

Table 3.1: Market Quality by Economy

Economy	β_{ij}	$1/\lambda_{ij}$	$\beta_{1j} + \beta_{2j}$	$1/\lambda_{1j} + 1/\lambda_{2j}$	$\text{Var}(v \mid P_{ij})$
b	$\sqrt{\eta_i}$	$2\sqrt{\eta_i}$	$\sqrt{\eta_1} + \sqrt{\eta_2}$	$2(\sqrt{\eta_1} + \sqrt{\eta_2})$	$\frac{1}{2}\sigma_v^2$
a	$\frac{1}{2}\sqrt{\eta_1 + \eta_2}$	$\sqrt{\eta_1 + \eta_2}$	$\sqrt{\eta_1 + \eta_2}$	$2\sqrt{\eta_1 + \eta_2}$	$\frac{1}{2}\sigma_v^2$
s	$\sqrt{\frac{\eta_1+\eta_2}{2}}$	$\frac{3}{2}\sqrt{\frac{\eta_1+\eta_2}{2}}$	$\sqrt{2(\eta_1 + \eta_2)}$	$\sqrt{\frac{9}{2}(\eta_1 + \eta_2)}$	$\frac{1}{3}\sigma_v^2$
m	$\sqrt{\eta_i}$	$3\sqrt{\eta_i}$	$\sqrt{\eta_1} + \sqrt{\eta_2}$	$3(\sqrt{\eta_1} + \sqrt{\eta_2})$	$\frac{1}{3}\sigma_v^2$

Asset-specific trading intensity and market depth for Economy a and market depth only for Economy s assume a symmetric equilibrium.

APPENDICES

APPENDIX A

Passive Investing: Derivations and Proofs

A.1 Proof of Lemma I.1

Case 1: $\tilde{\gamma}_i < 1$ and $\tilde{\lambda}_i > 0$ Recall that

$$S_i(\tilde{P}_i) = \tilde{\theta}_i - \frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\tilde{x}_i$$

which implies

$$\begin{aligned}
\mathbb{E}[S_i(\tilde{P}_i)] &= \mu_{\tilde{\theta}_i} \\
\mathrm{Var}(S_i(\tilde{P}_i)] &= \sigma^2_{\tilde{\theta}_i} + \frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2} \\
\mathrm{Cov}(S_i(\tilde{P}_i), \tilde{v}) &= \sigma^2_{\tilde{\theta}_i}
\end{aligned}$$

The sufficiency of $S_i(\tilde{P}_i)$ implies that

$$\mathbb{E}[\tilde{v}_i|S_i(\tilde{P}_i)] = \mu_{\tilde{\theta}_i} + \sigma^2_{\tilde{\theta}_i}\left(\sigma^2_{\tilde{\theta}_i} + \frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2}\right)^{-1}\left(\tilde{\theta}_i - \frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\tilde{x}_i - \mu_{\tilde{\theta}_i}\right) \quad \text{(A.1)}$$

$$\mathrm{Var}[\tilde{v}_i|S_i(\tilde{P}_i)] = \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \sigma^4_{\tilde{\theta}_i}\left(\sigma^2_{\tilde{\theta}_i} + \frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2}\right)^{-1} = \sigma^2_{\tilde{v}_i|\tilde{P}_i} \quad \text{(A.2)}$$

Recall that the market clearing condition is given by

$$(1-\tilde{\gamma}_i)\tilde{\lambda}_i \frac{\tau(\tilde{\theta}_i - \tilde{P}_i)}{\sigma^2_{\tilde{u}_i}} + (1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\frac{\tau(\mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i)}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} = \tilde{X}_i + \tilde{x}_i$$

Label

$$m_0 = \sigma^2_{\tilde{\theta}_i}\left(\sigma^2_{\tilde{\theta}_i} + \frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2}\right)^{-1} \tag{A.3}$$

Label the average risk-adjusted precision measure as:

$$\begin{aligned} m_1 &= \frac{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}{\sigma^2_{\tilde{u}_i}} + \frac{\tau(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \\ &= m_2 + m_3 \end{aligned}$$

The price function for asset i is:

$$\begin{aligned} \tilde{P}_i(\tilde{\theta}_i, \tilde{x}_i) &= \frac{m_3(1-m_0)\mu_{\tilde{\theta}_i}}{m_1} + \frac{m_2 + m_3 m_0}{m_1}\tilde{\theta}_i - \frac{m_2 + m_3 m_0}{m_1 m_2}\tilde{x}_i \\ &= q_1 + q_2\tilde{\theta}_i + q_3\tilde{x}_i \end{aligned}$$

Case 2: $\tilde{\lambda}_i = 0$ Recall that

$$S_i(\tilde{P}_i) = \tilde{x}_i$$

which implies

$$\begin{aligned} \mathbb{E}[S_i(\tilde{P}_i)] &= 0 \\ \mathrm{Var}(S_i(\tilde{P}_i)] &= \sigma^2_{\tilde{x}_i} \\ \mathrm{Cov}(S_i(\tilde{P}_i), \tilde{v}) &= 0 \end{aligned}$$

The price equation greatly simplifies due to the lack of fundamentals in prices:

$$\begin{aligned}\tilde{P}_i(\tilde{x}_i) &= \mu_{\tilde{\theta}_i} - \frac{\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)} - \frac{\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)}\tilde{x}_i \\ &= q_1 + q_3\tilde{x}_i\end{aligned}$$

Case 3: $\tilde{\gamma}_i = 1$ Assets do not clear because trader demand becomes zero and supply is stochastic. Therefore no equilibrium exists.

A.2 Properties of Expected Returns

I first exploit market clearing conditions to derive closed form solutions for unconditional expectations. By market clearing:

$$(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau\frac{\tilde{\theta}_i - \tilde{P}_i}{\sigma^2_{\tilde{u}_i}} + (1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau\frac{\mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} = \tilde{X}_i + \tilde{x}_i$$

$$\Downarrow$$

$$(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau\frac{\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i] + \mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i}{\sigma^2_{\tilde{u}_i}} + (1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau\frac{\mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} = \tilde{X}_i + \tilde{x}_i$$

$$\Downarrow$$

$$\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i]) + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)(\mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i) = \tilde{X}_i + \tilde{x}_i$$

The conditional return for any asset i:

$$\mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i = \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-1}\left(\tilde{X}_i + \tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right) \tag{A.4}$$

By the law of iterated expectations the unconditional return for any asset i is given by:

$$\mathbb{E}[\tilde{v}_i - \tilde{P}_i] = \tilde{X}_i\left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-1} \tag{A.5}$$

Given aggregate supplies and participation levels for the index and stock picking strategies, the unconditional risk-premium for the index and stock picking strategies are as follows.

Risk Premium: Index Asset ($i = 1$)

$$\mathbb{E}[\tilde{v}_1 - \tilde{P}_1] = \left(\frac{(1-\tilde{\gamma}_1)\tilde{\lambda}_1 \tau}{\sigma^2_{\tilde{u}_1}} + \frac{(1-\tilde{\gamma}_1)(1-\tilde{\lambda}_1)\tau}{\sigma^2_{\tilde{v}_1|\tilde{P}_1}} \right)^{-1} \tag{A.6}$$

Risk Premia: Stock Picking Strategies ($i > 1$)

$$\mathbb{E}[\tilde{v}_i - \tilde{P}_i] = 0 \tag{A.7}$$

Return Variance: All Assets I utilize the law of total variance and equation (A.4) to estimate the variance of the unconditional return:

$$\begin{aligned} &\mathrm{Var}(\tilde{v}_i - \tilde{P}_i) = \mathrm{Var}(\tilde{v}_i - \tilde{P}_i|\tilde{P}_i) + \mathrm{Var}(\mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i) \\ &= \sigma^2_{\tilde{v}_i|\tilde{P}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i \tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \right)^{-2} \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i \tau}{\sigma^2_{\tilde{u}_i}} (\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i]) \right) \end{aligned} \tag{A.8}$$

A.3 Proof of Proposition I.2

I am after changes in the ratio of unconditional expected utilities of participating and non-participating traders with respect to $\tilde{\gamma}_i$ in an information equilibrium. The ratios are defined by equation (1.24):

$$\begin{cases} \text{If } \tilde{c}_i \geq \frac{\tau}{2}\log(1+\tilde{n}_i) & \text{then } \frac{EU_{I_j^* R_j^*}}{EU_{I_j R_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i - \tilde{P}_i}}} \exp\left(\frac{\tilde{k}_i}{\tau}\right) \\ \text{If } \frac{\tau}{2}\log(1+\frac{\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{n}_i + \sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}\tilde{n}_i) \leq \tilde{c}_i < \frac{\tau}{2}\log(1+\tilde{n}_i) & \text{then } \frac{EU_{I_j^* R_j^*}}{EU_{I_j R_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i - \tilde{P}_i}}} \exp\left(\frac{\tilde{k}_i}{\tau}\right) \\ \text{If } \tilde{c}_i \leq \frac{\tau}{2}\log(1+\frac{\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{n}_i + \sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}\tilde{n}_i) & \text{then } \frac{EU_{I_j^* R_j^*}}{EU_{I_j R_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i - \tilde{P}_i}}} \exp\left(\frac{\tilde{k}_i}{\tau} + \frac{\tilde{c}_i}{\tau}\right) \end{cases} \tag{A.9}$$

The cases above specify the three levels of information equilibrium as outlined in equation (1.22):

$$\begin{cases} \text{Fully-uninformed equilibrium: if } \tilde{c}_i \geq \frac{\tau}{2}\log(1+\tilde{n}_i) \text{ then } \tilde{\lambda}_i = 0 \\ \text{Interior equilibrium: if } f_i(\tilde{\gamma}_i, \tilde{\lambda}_i) = 1 \text{ then } \tilde{\lambda}_i = \frac{\sigma_{\tilde{u}_i}\sigma_{\tilde{x}_i}}{(1-\tilde{\gamma}_i)\tau}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}} \\ \text{Fully-informed equilibrium: if } 1-\tilde{\gamma}_i < \frac{\sigma_{\tilde{u}_i}\sigma_{\tilde{x}_i}}{\tau}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}} \text{ then } \tilde{\lambda}_i = 1 \end{cases} \tag{A.10}$$

I proceed to derive closed form solutions of the ratios of expected utilities to participation (the function $g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)$) for the three cases specified in equation (A.10) and evaluate their derivatives with respect to $\tilde{\gamma}_i$. Since the ratio is always positive, I can equivalently evaluate the derivative for the function $g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)^2$. For the fully-uninformed and interior equilibrium:

$$\begin{aligned} g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)^2 &= \frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i|\tilde{P}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2} \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)} \exp\left(\frac{2\tilde{k}_i}{\tau}\right) \\ &= \frac{1}{1 + \sigma^{-2}_{\tilde{v}_i|\tilde{P}_i}\left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2} \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)} \exp\left(\frac{2\tilde{k}_i}{\tau}\right) \end{aligned} \tag{A.11}$$

For the fully-informed equilibrium:

$$g_i(\tilde{\gamma}_i,\tilde{\lambda}_i)^2 = \frac{\sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{v}_i|\tilde{P}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2} \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)} \exp\left(\frac{2\tilde{c}_i}{\tau} + \frac{2\tilde{k}_i}{\tau}\right) \tag{A.12}$$

A.3.1 Fully-Uninformed Equilibrium

In this equilibrium traders do not acquire information ($\tilde{\lambda}_i = 0$) implying that prices contain no information:

$$\sigma^2_{\tilde{v}_i|\tilde{P}_i} = \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} \tag{A.13}$$

Substituting the above into equation (A.11)

$$\begin{aligned}
g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)^2 &= \frac{1}{1+\sigma^{-2}_{\tilde{v}_i|\tilde{P}_i}\left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}+\frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2}\mathrm{Var}\left(\tilde{x}_i-\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i-\mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)} \exp\left(\frac{2\tilde{k}_i}{\tau}\right) \\
&= \frac{1}{1+\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}\sigma^2_{\tilde{x}_i}}{(1-\tilde{\gamma}_i)^2\tau^2}} \exp\left(\frac{2\tilde{k}_i}{\tau}\right) \\
&= \frac{1}{1+\frac{(\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i})\sigma^2_{\tilde{x}_i}}{(1-\tilde{\gamma}_i)^2\tau^2}} \exp\left(\frac{2\tilde{k}_i}{\tau}\right)
\end{aligned} \tag{A.14}$$

$$\Downarrow$$

$$\frac{\partial g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)^2}{\partial \tilde{\gamma}_i} < 0 \Leftrightarrow \frac{\partial g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)}{\partial \tilde{\gamma}_i} < 0 \text{ (since } g_i(\tilde{\gamma}_i, \tilde{\lambda}_i) > 0)$$

A.3.2 Interior Equilibrium

In an interior equilibrium, from equations (1.16),(1.17),(1.18),(A.1),(A.2),(1.15) conditional expectations are

$$\mathbb{E}[\tilde{v}_i|S_i(\tilde{P}_i)] = \mu_{\tilde{\theta}_i} + \sigma^2_{\tilde{\theta}_i}\left(\sigma^2_{\tilde{\theta}_i} + \frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2}\right)^{-1}\left(\tilde{\theta}_i - \frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\tilde{x}_i - \mu_{\tilde{\theta}_i}\right) \tag{A.15}$$

$$= \mu_{\tilde{\theta}_i} + \rho^2_{S(\tilde{P}_i),\tilde{\theta}_i}\left(\tilde{\theta}_i - \frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\tilde{x}_i - \mu_{\tilde{\theta}_i}\right) \tag{A.16}$$

$$\tag{A.17}$$

and conditional variances are

$$\begin{aligned}
\sigma^2_{\tilde{v}_i|\tilde{P}_i} &= \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \sigma^4_{\tilde{\theta}_i}\left(\sigma^2_{\tilde{\theta}_i} + \frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2}\right)^{-1} \\
&= \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \sigma^2_{\tilde{\theta}_i}\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i} \\
&= \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \frac{\sigma^2_{\tilde{\theta}_i}}{1+\tilde{m}_i} \\
&= \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i}
\end{aligned} \tag{A.18}$$

Utilizing equation (1.20) I also have (note that these identities only hold for the interior equilibrium):

$$\begin{aligned}\frac{\tilde{m}_i}{1+\tilde{m}_i} &= \frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}\\ \frac{1}{1+\tilde{m}_i} &= \frac{1+\tilde{n}_i-\exp(2\tilde{c}_i/\tau)}{\tilde{n}_i}\end{aligned} \tag{A.19}$$

Utilizing the above it can be shown that:

$$\begin{aligned}&\text{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right) =\\ &= \text{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}\left(\tilde{\theta}_i - \mu_{\tilde{\theta}_i} - \rho^2_{S(\tilde{P}_i),\tilde{\theta}_i}\big(\tilde{\theta}_i - \frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\tilde{x}_i - \mu_{\tilde{\theta}_i}\big)\right)\right)\\ &= \text{Var}\left((1-\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i})\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(1-\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i})\tilde{\theta}_i\right)\\ &= (1-\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i})^2\text{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}\tilde{\theta}_i\right)\\ &= (1-\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i})^2\left(\sigma^2_{\tilde{x}_i} + \frac{(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2\tau^2\sigma^2_{\tilde{\theta}_i}}{\sigma^4_{\tilde{u}_i}}\right)\\ &= (1-\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i})^2\sigma^2_{\tilde{x}_i}\left(\frac{1}{\tilde{m}_i}+1\right)\\ &= \sigma^2_{\tilde{x}_i}\left(1-\frac{1}{1+\tilde{m}_i}\right)^2\left(\frac{1}{\tilde{m}_i}+1\right)\\ &= \sigma^2_{\tilde{x}_i}\left(\frac{\tilde{m}_i}{1+\tilde{m}_i}\right)\\ &= \sigma^2_{\tilde{x}_i}\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}\end{aligned} \tag{A.20}$$

I can substitute equations (A.20),(A.18) and (1.22) into equation (A.11)

$$
\begin{aligned}
g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)^2 &= \frac{1}{1+\sigma_{\tilde{v}_i|\tilde{P}_i}^{-2}\left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma_{\tilde{u}_i}^2}+\frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma_{\tilde{v}_i|\tilde{P}_i}^2}\right)^{-2} \mathrm{Var}\left(\tilde{x}_i-\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma_{\tilde{u}_i}^2}(\tilde{\theta}_i-\mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)} \exp\left(\frac{2\tilde{k}_i}{\tau}\right) \\
&= \frac{1}{1+\left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma_{\tilde{u}_i}^2}\sigma_{\tilde{v}_i|\tilde{P}_i}+\frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2} \mathrm{Var}\left(\tilde{x}_i-\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma_{\tilde{u}_i}^2}(\tilde{\theta}_i-\mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)} \exp\left(\frac{2\tilde{k}_i}{\tau}\right) \\
&= \frac{1}{1+\left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma_{\tilde{u}_i}^2}\exp\left(\frac{\tilde{c}_i}{\tau}\right)\sigma_{\tilde{u}_i}+\frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\exp\left(\frac{\tilde{c}_i}{\tau}\right)\sigma_{\tilde{u}_i}}\right)^{-2} \mathrm{Var}\left(\tilde{x}_i-\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma_{\tilde{u}_i}^2}(\tilde{\theta}_i-\mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)} \exp\left(\frac{2\tilde{k}_i}{\tau}\right) \\
&= \frac{1}{1+\left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma_{\tilde{u}_i}}\left(\exp(\frac{\tilde{c}_i}{\tau})-\exp(-\frac{\tilde{c}_i}{\tau})\right)+\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma_{\tilde{u}_i}^{-1}-\tilde{\gamma}_i\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma_{\tilde{u}_i}^{-1}\right)^{-2} \sigma_{\tilde{x}_i}^2\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}} \\
&\times\exp\left(\frac{2\tilde{k}_i}{\tau}\right) \\
&= \frac{1}{1+\left(\sigma_{\tilde{x}_i}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1}-\frac{1}{\tilde{n}_i}}\left(\exp(\frac{\tilde{c}_i}{\tau})-\exp(-\frac{\tilde{c}_i}{\tau})\right)+\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma_{\tilde{u}_i}^{-1}-\tilde{\gamma}_i\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma_{\tilde{u}_i}^{-1}\right)^{-2} \sigma_{\tilde{x}_i}^2\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}} \\
&\times\exp\left(\frac{2\tilde{k}_i}{\tau}\right) \\
&\Downarrow
\end{aligned}
$$

$$
\frac{\partial g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)^2}{\partial \tilde{\gamma}_i} < 0 \Leftrightarrow \frac{\partial g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)}{\partial \tilde{\gamma}_i} < 0 \text{ (since } g_i(\tilde{\gamma}_i, \tilde{\lambda}_i) > 0) \tag{A.21}
$$

A.3.3 Fully-Informed Equilibrium

For the fully-informed equilibrium the following for conditional variance continues to hold:

$$
\begin{aligned}
\sigma_{\tilde{v}_i|\tilde{P}_i}^2 &= \sigma_{\tilde{\theta}_i}^2 + \sigma_{\tilde{u}_i}^2 - \sigma_{\tilde{\theta}_i}^4\left(\sigma_{\tilde{\theta}_i}^2 + \frac{\sigma_{\tilde{u}_i}^4\sigma_{\tilde{x}_i}^2}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2}\right)^{-1} \\
&= \sigma_{\tilde{\theta}_i}^2 + \sigma_{\tilde{u}_i}^2 - \sigma_{\tilde{\theta}_i}^2\rho_{S(\tilde{P}_i),\tilde{\theta}_i}^2 \\
&= \sigma_{\tilde{\theta}_i}^2 + \sigma_{\tilde{u}_i}^2 - \frac{\sigma_{\tilde{\theta}_i}^2}{1+\tilde{m}_i}
\end{aligned} \tag{A.22}
$$

Furthermore, equation (A.20) holds except for the final equality:

$$\begin{aligned}
&\mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right) = \\
&= \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}\left(\tilde{\theta}_i - \mu_{\tilde{\theta}_i} - \rho^2_{S(\tilde{P}_i),\tilde{\theta}_i}(\tilde{\theta}_i - \frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\tilde{x}_i - \mu_{\tilde{\theta}_i})\right)\right) \\
&= \mathrm{Var}\left((1-\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i})\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(1-\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i})\tilde{\theta}_i\right) \\
&= (1-\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i})^2\mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}\tilde{\theta}_i\right) \\
&= (1-\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i})^2\left(\sigma^2_{\tilde{x}_i} + \frac{(1-\tilde{\gamma}_i)^2\tilde{\lambda}_i^2\tau^2\sigma^2_{\tilde{\theta}_i}}{\sigma^4_{\tilde{u}_i}}\right) \\
&= (1-\rho^2_{S(\tilde{P}_i),\tilde{\theta}_i})^2\sigma^2_{\tilde{x}_i}\left(\frac{1}{\tilde{m}_i}+1\right) \\
&= \sigma^2_{\tilde{x}_i}\left(1-\frac{1}{1+\tilde{m}_i}\right)^2\left(\frac{1}{\tilde{m}_i}+1\right) \\
&= \sigma^2_{\tilde{x}_i}\left(\frac{\tilde{m}_i}{1+\tilde{m}_i}\right)
\end{aligned} \tag{A.23}$$

By equations (A.22), (A.23) and (A.12), I have the following

$$\begin{aligned}
g_i(\tilde{\gamma}_i,\tilde{\lambda}_i) &= \frac{\sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{v}_i|\tilde{P}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2}\mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)}\exp\left(\frac{\tilde{c}_i}{\tau}+\frac{\tilde{k}_i}{\tau}\right) \\
&= \frac{\sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \frac{\sigma^2_{\tilde{\theta}_i}}{1+\tilde{m}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tau}{\sigma^2_{\tilde{u}_i}}\right)^{-2}\sigma^2_{\tilde{x}_i}\left(\frac{\tilde{m}_i}{1+\tilde{m}_i}\right)}\exp\left(\frac{\tilde{c}_i}{\tau}+\frac{\tilde{k}_i}{\tau}\right)
\end{aligned} \tag{A.24}$$

By equation (1.16)

$$\begin{aligned}
m_i &= \left(\frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\right)^2\frac{\sigma^2_{\tilde{x}_i}}{\sigma^2_{\tilde{\theta}_i}} \\
&= \frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\sigma^2_{\tilde{\theta}_i}}
\end{aligned}$$

Therefore, I have

$$
\begin{aligned}
g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)^2 &= \frac{\sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \frac{\sigma^2_{\tilde{\theta}_i}}{1+\tilde{m}_i} + \frac{\tilde{m}_i^2\sigma^2_{\tilde{\theta}_i}}{1+\tilde{m}_i}} \exp\left(\frac{2\tilde{c}_i}{\tau} + \frac{2\tilde{k}_i}{\tau}\right) \\
&= \frac{\sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{u}_i} + \sigma^2_{\tilde{\theta}_i}\tilde{m}_i} \exp\left(\frac{2\tilde{c}_i}{\tau} + \frac{2\tilde{k}_i}{\tau}\right) \\
&= \frac{1}{1 + \tilde{n}_i\tilde{m}_i} \exp\left(\frac{2\tilde{c}_i}{\tau} + \frac{2\tilde{k}_i}{\tau}\right) \\
&\Downarrow \\
\frac{\partial g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)^2}{\partial \tilde{\gamma}_i} &< 0 \Leftrightarrow \frac{\partial g_i(\tilde{\gamma}_i, \tilde{\lambda}_i)}{\partial \tilde{\gamma}_i} < 0 \text{ (since } g_i(\tilde{\gamma}_i, \tilde{\lambda}_i) > 0)
\end{aligned}
\tag{A.25}
$$

For all three cases of information equilibria, I have demonstrated that the ratio of unconditional expected utility to participating vs. not-participating in asset i, given by equations (A.14), (A.33), and (A.25), is decreasing in $\tilde{\gamma}_i$, equivalently increasing in participation $(1 - \tilde{\gamma}_i)$. Since utilities are negative, this demonstrates strategic substitutability in participation: the more traders participate in the trading of an asset, the lower are the expected gains to participation. An immediate consequence are closed form solutions for equilibrium participation levels, demonstrated in Appendix A.4.

A.4 Closed-Form Solutions for Passive Share

A.4.1 Fully-Uninformed Equilibrium

By equation (A.14), after some algebra,

$$
\begin{aligned}
&\frac{1}{1 + \frac{(\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i})\sigma^2_{\tilde{x}_i}}{(1-\tilde{\gamma}_i)^2\tau^2}} \exp\left(\frac{2\tilde{k}_i}{\tau}\right) = 1 \\
&\Downarrow \\
&\tilde{\gamma}_i = 1 - \frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{n+1}{\exp(2\tilde{k}_i/\tau) - 1}}
\end{aligned}
\tag{A.26}
$$

A.4.2 Interior Equilibrium

By equation (A.33), after some algebra,

$$
\begin{aligned}
&\frac{1}{1+\left(\sigma_{\tilde{x}_i}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1}-\frac{1}{\tilde{n}_i}}\left(\exp(\frac{\tilde{c}_i}{\tau})-\exp(-\frac{\tilde{c}_i}{\tau})\right)+\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma_{\tilde{u}_i}^{-1}-\tilde{\gamma}_i\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma_{\tilde{u}_i}^{-1}\right)^{-2}\sigma_{\tilde{x}_i}^2\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}}\exp\left(\frac{2\tilde{k}_i}{\tau}\right)=1\\
&\Downarrow\\
&\tilde{\gamma}_i=1-\frac{1}{\sqrt{\exp(2k_i/\tau)-1}}\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\exp(\tilde{c}_i/\tau)\sqrt{\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}}+\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}}\sqrt{n+1-\exp(2\tilde{c}_i/\tau)}\\
&=1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}}\left(\frac{\exp(\tilde{c}_i/\tau)}{\sqrt{\exp(2k_i/\tau)-1}}-\sqrt{n+1-\exp(2\tilde{c}_i/\tau)}\right)
\end{aligned}
\tag{A.27}
$$

A.4.3 Fully-Informed Equilibrium

By equation (A.25), after some algebra,

$$
\begin{aligned}
&\frac{1}{1+\tilde{n}_i\tilde{m}_i}\exp\left(\frac{2\tilde{c}_i}{\tau}+\frac{2\tilde{k}_i}{\tau}\right)=1\\
&\Downarrow\\
&\tilde{\gamma}_i=1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{1}{\exp(2(k_i+\tilde{c}_i)/\tau)-1}}
\end{aligned}
\tag{A.28}
$$

A.4.4 Passive Share Closed-Form Solutions: Summary

The following summarizes the closed-form solutions for passive share (note that it's possible to have parameter values such that $\tilde{\gamma}_i$ falls below zero, therefore I take the maximum):

$$
\begin{cases}
\text{Case 1: if } \tilde{c}_i\geq\frac{\tau}{2}\log(1+\frac{\sigma_{\tilde{\theta}_i}^2}{\sigma_{\tilde{u}_i}^2}) \text{ then } \tilde{\gamma}_i=\max\left\{0,1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{n+1}{\exp(2k_i/\tau)-1}}\right\}\\
\text{Case 2: if } f_i(\tilde{\gamma}_i,\tilde{\lambda}_i)=1 \text{ then } \tilde{\gamma}_i=\max\left\{0,1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}}\left(\frac{\exp(\tilde{c}_i/\tau)}{\sqrt{\exp(2k_i/\tau)-1}}-\sqrt{n+1-\exp(2\tilde{c}_i/\tau)}\right)\right\}\\
\text{Case 3: if } \tilde{\gamma}_i>1-\frac{\sigma_{\tilde{u}_i}\sigma_{\tilde{x}_i}}{\tau}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1}-\frac{\sigma_{\tilde{u}_i}^2}{\sigma_{\tilde{\theta}_i}^2}} \text{ then } \tilde{\gamma}_i=\max\left\{0,1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{1}{\exp(2(k_i+\tilde{c}_{I_i})/\tau)-1}}\right\}
\end{cases}
\tag{A.29}
$$

Equivalently:

$$\begin{cases} \text{Case 1: if } \frac{1}{\exp(2\tilde{c}_i/\tau)-1} \leq \frac{1}{\tilde{n}_i} \text{ then } \tilde{\gamma}_i = \max\left\{0, 1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{n+1}{\exp(2k_i/\tau)-1}}\right\} \\ \text{Case 2: if } \frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\exp(2(k_i+\tilde{c}_i)/\tau)-1} < \frac{1}{\tilde{n}_i} < \frac{1}{\exp(2\tilde{c}_i/\tau)-1} \text{ then} \\ \tilde{\gamma}_i = \max\left\{0, 1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}}\left(\frac{\exp(\tilde{c}_i/\tau)}{\sqrt{\exp(2k_i/\tau)-1}} - \sqrt{n+1-\exp(2\tilde{c}_i/\tau)}\right)\right\} \\ \text{Case 3: if } \frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\exp(2(k_i+\tilde{c}_i)/\tau)-1} \geq \frac{1}{n} \text{ then } \tilde{\gamma}_i = \max\left\{0, 1-\frac{\sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i}}{\tau}\sqrt{\frac{1}{\exp(2(k_i+\tilde{c}_i)/\tau)-1}}\right\} \end{cases} \tag{A.30}$$

A.5 Proof of Lemma I.4

I am after the changes to the unconditional expected utility of trader j participating in stock-picking strategy i for a fixed $\tilde{\gamma}_i$ across the three information equilibria. I rely on the derivations in Appendix A.3 in the following.

Fully-Uninformed Equilibrium From equation (A.14)

$$\begin{gathered} EU^2_{I^*_j} = \frac{1}{1+\frac{(\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i})\sigma^2_{\tilde{x}_i}}{(1-\tilde{\gamma}_i)^2\tau^2}} \exp\left(-\frac{2W_0}{\tau}+\frac{2\tilde{k}_i}{\tau}\right) \\ \Downarrow \\ \frac{\partial EU^2_{I^*_j}}{\partial \tilde{c}_i} = 0 \Leftrightarrow \frac{\partial EU_{I^*_j}}{\partial \tilde{c}_i} = 0 \end{gathered} \tag{A.31}$$

Fully-Informed Equilibrium From equation (A.25)

$$\begin{gathered} EU^2_{I^*_j} = \frac{1}{1+\tilde{n}_i\tilde{m}_i} \exp\left(-\frac{2W_0}{\tau}+\frac{2\tilde{c}_i}{\tau}+\frac{2\tilde{k}_i}{\tau}\right) \\ \Downarrow \\ \frac{\partial EU^2_{I^*_j}}{\partial \tilde{c}_i} > 0 \Leftrightarrow \frac{\partial EU^2_{I^*_j}}{\partial \tilde{c}_i} < 0 \end{gathered} \tag{A.32}$$

Interior Equilibrium From equation (A.33)

$$
\begin{aligned}
&EU^2_{I^*_j} = \frac{1}{1+\left(\sigma_{\tilde{x}_i}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1}-\frac{1}{\tilde{n}_i}}\left(\exp(\frac{\tilde{c}_i}{\tau})-\exp(-\frac{\tilde{c}_i}{\tau})\right)+\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma^{-1}_{\tilde{u}_i}-\tilde{\gamma}_i\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma^{-1}_{\tilde{u}_i}\right)^{-2}\sigma^2_{\tilde{x}_i}\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}} \\
&\times\exp\left(-\frac{2W_0}{\tau}+\frac{2\tilde{k}_i}{\tau}\right) \\
&\Downarrow \\
&= \frac{1}{1+\frac{\sigma^2_{\tilde{x}_i}}{\tilde{n}_i}\left(\sigma_{\tilde{x}_i}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1}-\frac{1}{\tilde{n}_i}}\left(\exp(\frac{\tilde{c}_i}{\tau})-\exp(-\frac{\tilde{c}_i}{\tau})\right)+\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma^{-1}_{\tilde{u}_i}-\tilde{\gamma}_i\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma^{-1}_{\tilde{u}_i}\right)^{-2}\left(\frac{1}{\sqrt{\exp(2\tilde{c}_i/\tau)-1}}\right)^{-2}} \\
&\times\exp\left(-\frac{2W_0}{\tau}+\frac{2\tilde{k}_i}{\tau}\right) \\
&\Downarrow \\
&= \frac{1}{1+\frac{\sigma^2_{\tilde{x}_i}}{\tilde{n}_i}\left(\sigma_{\tilde{x}_i}\sqrt{\exp(-2\tilde{c}_i/\tau)\left(1+\frac{1}{\tilde{n}_i}\right)-\frac{1}{\tilde{n}_i}}+(1-\tilde{\gamma}_i)\tau\exp(-\frac{\tilde{c}_i}{\tau})\sigma^{-1}_{\tilde{u}_i}\right)^{-2}} \\
&\times\exp\left(-\frac{2W_0}{\tau}+\frac{2\tilde{k}_i}{\tau}\right) \\
&\Downarrow \\
&\frac{\partial EU^2_{I^*_j}}{\partial\tilde{c}_i}<0 \Leftrightarrow \frac{\partial EU^2_{I^*_j}}{\partial\tilde{c}_i}>0
\end{aligned}
\tag{A.33}
$$

A.6 Asset Pricing Implications of Changing Information and Participation Costs

A.6.1 Unconditional Expected Returns

The unconditional expected return for the any stock picking strategy is given by equation (A.7)

$$\mathbb{E}[\tilde{v}_i - \tilde{P}_i] = 0 \tag{A.34}$$

The unconditional expected return for the market index is given by equation (A.6)

$$\mathbb{E}[\tilde{v}_1 - \tilde{P}_1] = \left(\frac{\tilde{\lambda}_1\tau}{\sigma^2_{\tilde{u}_1}} + \frac{(1-\tilde{\lambda}_1)\tau}{\sigma^2_{\tilde{v}_1|\tilde{P}_1}}\right)^{-1} \tag{A.35}$$

I apply the three cases of information equilibria, as specified by the conditions of equation (1.22) for $i = 1$ ($\tilde{\gamma}_i = 0$):

$$\begin{cases} \text{If } \tilde{c}_i \geq \frac{\tau}{2}\log(1+\tilde{n}_i) & \text{then } \tilde{\lambda}_i = 0 \\ \text{If } f_i(\tilde{\gamma}_i, \tilde{\lambda}_i) = 1 & \text{then } \tilde{\lambda}_i = \frac{\sigma_{\tilde{u}_i}\sigma_{\tilde{x}_i}}{(1-\tilde{\gamma}_i)\tau}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}} \\ \text{If } \tilde{c}_i \leq \frac{\tau}{2}\log(1+\frac{\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{n}_i+\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}\tilde{n}_i) & \text{then } \tilde{\lambda}_i = 1 \end{cases} \tag{A.36}$$

and equation (1.30)

$$\begin{cases} \text{Fully-Uninformed Equilibrium:} & \sigma^2_{\tilde{v}_i|\tilde{P}_i} = \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} \\ \text{Interior Equilibrium:} & \sigma^2_{\tilde{v}_i|\tilde{P}_i} = \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} \\ \text{Fully-Informed Equilibrium:} & \sigma^2_{\tilde{v}_i|\tilde{P}_i} = \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \frac{\sigma^2_{\tilde{\theta}_i}}{1+\tilde{m}_i} = \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \frac{\sigma^2_{\tilde{\theta}_i}}{1+\frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2\sigma^2_{\tilde{\theta}_i}(1-\tilde{\gamma}_i)^2}} \end{cases} \tag{A.37}$$

It's immediate that for the first and third cases, changes to information costs have no effect on passive share. For the interior equilibrium (second case):

$$\begin{aligned} \frac{\partial \mathbb{E}[\tilde{v}_1 - \tilde{P}_1]}{\partial \tilde{c}_1} &= -\left(\frac{\tilde{\lambda}_1\tau}{\sigma^2_{\tilde{u}_1}} + \frac{(1-\tilde{\lambda}_1)\tau}{\sigma^2_{\tilde{v}_1|\tilde{P}_1}}\right)^{-2}\frac{\partial}{\partial \tilde{c}_1}\left(\frac{\tilde{\lambda}_1\tau}{\sigma^2_{\tilde{u}_1}} + \frac{(1-\tilde{\lambda}_1)\tau}{\sigma^2_{\tilde{v}_1|\tilde{P}_1}}\right) \\ &= -\left(\frac{\tilde{\lambda}_1\tau}{\sigma^2_{\tilde{u}_1}} + \frac{(1-\tilde{\lambda}_1)\tau}{\sigma^2_{\tilde{v}_1|\tilde{P}_1}}\right)^{-2}\left(\tau\frac{\partial\lambda_1}{\partial\tilde{c}_1}\left(\frac{1}{\sigma^2_{\tilde{u}_1}} - \frac{1}{\sigma^2_{\tilde{v}_1|\tilde{P}_1}}\right) - \frac{2}{\tau}(1-\tilde{\lambda}_1)\exp\left(-\frac{2\tilde{c}_i}{\tau}\right)\sigma^{-2}_{\tilde{u}_i}\tau\right) \\ &> 0 \end{aligned} \tag{A.38}$$

Where the inequality follows from the fact that in an interior equilibrium $\frac{1}{\sigma^2_{\tilde{u}_1}} > \frac{1}{\sigma^2_{\tilde{v}_1|\tilde{P}_1}}$ and $\frac{\partial\lambda_1}{\partial\tilde{c}_1} < 0$

A.6.2 Unconditional Return Variances

The unconditional variance of returns is given by equation (A.8):

$$\begin{aligned} &\mathrm{Var}(\tilde{v}_i - \tilde{P}_i) = \mathrm{Var}(\tilde{v}_i - \tilde{P}_i|\tilde{P}_i) + \mathrm{Var}(\mathbb{E}[\tilde{v}_i|\tilde{P}_i] - \tilde{P}_i) \\ &= \sigma^2_{\tilde{v}_i|\tilde{P}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2}\mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right) \end{aligned} \tag{A.39}$$

I rely on equations (1.30),(A.20),and (A.23) to derive the unconditional variance for both the index and stock picking strategies across three levels of equilibrium.

A.6.2.1 Unconditional Variance of Index Asset

Fully-Uninformed Equilibrium

$$\begin{aligned}
\mathrm{Var}(\tilde{v}_1 - \tilde{P}_1) =& \sigma^2_{\tilde{v}_1|\tilde{P}_1} + \left(\frac{(1-\gamma_1)\tilde{\lambda}_1\tau}{\sigma^2_{\tilde{u}_1}} + \frac{(1-\gamma_1)(1-\tilde{\lambda}_1)\tau}{\sigma^2_{\tilde{v}_1|\tilde{P}_1}}\right)^{-2} \mathrm{Var}\left(\tilde{x}_1 - \frac{(1-\gamma_1)\tilde{\lambda}_1\tau}{\sigma^2_{\tilde{u}_1}}(\tilde{\theta}_1 - \mathbb{E}[\tilde{v}_1|\tilde{P}_1])\right) \\
&= \sigma^2_{\tilde{\theta}_1} + \sigma^2_{\tilde{u}_1} + \left(\frac{\tau}{\sigma^2_{\tilde{\theta}_1} + \sigma^2_{\tilde{u}_1}}\right)^{-2} \sigma^2_{\tilde{x}_1} \\
&\Downarrow \\
\frac{\partial \mathrm{Var}(\tilde{v}_1 - \tilde{P}_1)}{\partial \tilde{c}_1} &= 0
\end{aligned} \tag{A.40}$$

Interior Equilibrium By equations (A.6) and (A.38):

$$\begin{aligned}
\mathrm{Var}(\tilde{v}_1 - \tilde{P}_1) &= \sigma^2_{\tilde{v}_1|\tilde{P}_1} + \left(\frac{\tilde{\lambda}_1\tau}{\sigma^2_{\tilde{u}_1}} + \frac{(1-\tilde{\lambda}_1)\tau}{\sigma^2_{\tilde{v}_1|\tilde{P}_1}}\right)^{-2} \mathrm{Var}\left(\tilde{x}_1 - \frac{\tilde{\lambda}_1\tau}{\sigma^2_{\tilde{u}_1}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_1|\tilde{P}_1])\right) \\
&= \exp\left(\frac{2\tilde{c}_1}{\tau}\right)\sigma^2_{\tilde{u}_i} + \mathbb{E}[\tilde{v}_1 - \tilde{P}_1]^2 \sigma^2_{\tilde{x}_1} \frac{\exp(2\tilde{c}_1/\tau) - 1}{n_1} \\
&\Downarrow \\
\frac{\partial \mathrm{Var}(\tilde{v}_1 - \tilde{P}_1)}{\partial \tilde{c}_1} &> 0
\end{aligned} \tag{A.41}$$

Fully-Informed Equilibrium

$$
\begin{aligned}
\mathrm{Var}(\tilde{v}_1-\tilde{P}_1) =&\sigma^2_{\tilde{v}_1|\tilde{P}_1}+\left(\frac{(1-\gamma_1)\tilde{\lambda}_1\tau}{\sigma^2_{\tilde{u}_1}}+\frac{(1-\gamma_1)(1-\tilde{\lambda}_1)\tau}{\sigma^2_{\tilde{v}_1|\tilde{P}_1}}\right)^{-2}\mathrm{Var}\left(\tilde{x}_1-\frac{(1-\gamma_1)\tilde{\lambda}_1\tau}{\sigma^2_{\tilde{u}_1}}(\tilde{\theta}_1-\mathbb{E}[\tilde{v}_1|\tilde{P}_1])\right)\\
&=\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i}-\frac{\sigma^2_{\tilde{\theta}_i}}{1+\tilde{m}_i}+\left(\frac{\tau}{\sigma^2_{\tilde{\theta}_1}+\sigma^2_{\tilde{u}_1}}\right)^{-2}\sigma^2_{\tilde{x}_i}\left(\frac{\tilde{m}_i}{1+\tilde{m}_i}\right)\\
&=\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i}-\frac{\sigma^2_{\tilde{\theta}_i}}{1+\frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2\sigma^2_{\tilde{\theta}_i}}}+\left(\frac{\tau}{\sigma^2_{\tilde{\theta}_1}+\sigma^2_{\tilde{u}_1}}\right)^{-2}\sigma^2_{\tilde{x}_i}\left(\frac{1}{1+\frac{\tau^2\sigma^2_{\tilde{\theta}_i}}{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}}\right) \qquad \text{(A.42)}\\
&\Downarrow\\
\frac{\partial\mathrm{Var}(\tilde{v}_1-\tilde{P}_1)}{\partial\tilde{c}_1}&=0
\end{aligned}
$$

A.6.3 Unconditional Variance of Stock Picking Strategies

Fully-Uninformed Equilibrium

$$
\begin{aligned}
\mathrm{Var}(\tilde{v}_i-\tilde{P}_i) =&\sigma^2_{\tilde{v}_i|\tilde{P}_i}+\left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}+\frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2}\mathrm{Var}\left(\tilde{x}_i-\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i-\mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)\\
&=\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i}+\left(\frac{\tau(1-\tilde{\gamma}_i)}{\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i}}\right)^{-2}\sigma^2_{\tilde{x}_i}\\
&\Downarrow\\
\frac{\partial\mathrm{Var}(\tilde{v}_i-\tilde{P}_i)}{\partial\tilde{c}_i}&=2\left(\frac{\tau(1-\tilde{\gamma}_i)}{\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i}}\right)^{-3}\sigma^2_{\tilde{x}_i}\frac{\partial\tilde{\gamma}_i}{\partial\tilde{c}_i} \qquad \text{(A.43)}\\
&=0\\
\frac{\partial\mathrm{Var}(\tilde{v}_i-\tilde{P}_i)}{\partial\tilde{k}_i}&=2\left(\frac{\tau(1-\tilde{\gamma}_i)}{\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i}}\right)^{-3}\sigma^2_{\tilde{x}_i}\frac{\partial\tilde{\gamma}_i}{\partial\tilde{k}_i}\\
&\geq 0
\end{aligned}
$$

Interior Equilibrium

$$\mathrm{Var}(\tilde{v}_i - \tilde{P}_i) = \sigma^2_{\tilde{v}_i|\tilde{P}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \right)^{-2} \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} (\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i]) \right)$$

$$= \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \right)^{-2} \sigma^2_{\tilde{x}_i} \frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}$$

$$= \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \right)^{-2} \sigma^2_{\tilde{x}_i} \frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}$$

$$= \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} + \left((1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau \left(\frac{1}{\sigma^2_{\tilde{u}_i}} - \frac{1}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \right) + \frac{(1-\tilde{\gamma}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \right)^{-2} \sigma^2_{\tilde{x}_i} \frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}$$

$$= \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} + \left(\sigma_{\tilde{u}_i}\sigma_{\tilde{x}_i} \sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}} \left(\frac{1}{\sigma^2_{\tilde{u}_i}} - \frac{1}{\exp(2\tilde{c}_i/\tau)\sigma^2_{\tilde{u}_i}} \right) \sqrt{\frac{\tilde{n}_i}{\exp(2\tilde{c}_i/\tau)-1}} \right.$$

$$\left. + \sigma_{\tilde{x}_i}\sigma_{\tilde{u}_i} \sqrt{\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}} \left(\frac{\exp(\tilde{c}_i/\tau)}{\sqrt{\exp(2\tilde{k}_i/\tau)-1}} - \sqrt{n+1-\exp(2\tilde{c}_i/\tau)} \right) \times \frac{1}{\exp(2\tilde{c}_i/\tau)\sigma^2_{\tilde{u}_i}} \sqrt{\frac{\tilde{n}_i}{\exp(2\tilde{c}_i/\tau)-1}} \right)^{-2} \sigma^2_{\tilde{x}_i}$$

$$= \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} + \left(\frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}} \sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}} \left(\frac{\exp(2\tilde{c}_i/\tau)-1}{\exp(2\tilde{c}_i/\tau)} \right) \sqrt{\frac{\tilde{n}_i}{\exp(2\tilde{c}_i/\tau)-1}} \right.$$

$$\left. + \frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}} \left(\frac{1}{\exp(\tilde{c}_i/\tau)\sqrt{\exp(2\tilde{k}_i/\tau)-1}} - \frac{\sqrt{n+1-\exp(2\tilde{c}_i/\tau)}}{\exp(2\tilde{c}_i/\tau)} \right) \right)^{-2} \sigma^2_{\tilde{x}_i}$$

$$= \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} + \left(\frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}} \sqrt{1 - \frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}} \left(\frac{1}{\exp(2\tilde{c}_i/\tau)} \right) \sqrt{\tilde{n}_i} \right.$$

$$\left. + \frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}} \left(\frac{1}{\exp(\tilde{c}_i/\tau)\sqrt{\exp(2\tilde{k}_i/\tau)-1}} - \frac{\sqrt{n+1-\exp(2\tilde{c}_i/\tau)}}{\exp(2\tilde{c}_i/\tau)} \right) \right)^{-2} \sigma^2_{\tilde{x}_i}$$

$$= \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} + \left(\frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}} \frac{\sqrt{n+1-\exp(2\tilde{c}_i/\tau)}}{\exp(2\tilde{c}_i/\tau)} + \frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}} \left(\frac{1}{\exp(\tilde{c}_i/\tau)\sqrt{\exp(2\tilde{k}_i/\tau)-1}} - \frac{\sqrt{n+1-\exp(2\tilde{c}_i/\tau)}}{\exp(2\tilde{c}_i/\tau)} \right) \right)^{-2} \sigma^2_{\tilde{x}_i}$$

$$= \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} + \left(\frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}} \frac{1}{\exp(\tilde{c}_i/\tau)\sqrt{\exp(2\tilde{k}_i/\tau)-1}} \right)^{-2} \sigma^2_{\tilde{x}_i}$$

$$= \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} + \sigma^2_{\tilde{u}_i} \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\left(\exp\left(\frac{2\tilde{k}_i}{\tau}\right) - 1 \right)$$

$$= \sigma^2_{\tilde{u}_i} \exp\left(\frac{2(\tilde{c}_i + \tilde{k}_i)}{\tau}\right)$$

$$\Downarrow$$

$$\frac{\partial \mathrm{Var}(\tilde{v}_i - \tilde{P}_i)}{\partial \tilde{c}_i} > 0 \text{ and } \frac{\partial \mathrm{Var}(\tilde{v}_i - \tilde{P}_i)}{\partial \tilde{k}_i} > 0$$

Fully-Informed Equilibrium

$$
\begin{aligned}
\mathrm{Var}(\tilde{v}_i - \tilde{P}_i) &= \sigma^2_{\tilde{v}_i|\tilde{P}_i} + \left(\tfrac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i \tau}{\sigma^2_{\tilde{u}_i}} + \tfrac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} \right)^{-2} \mathrm{Var}\left(\tilde{x}_i - \tfrac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i \tau}{\sigma^2_{\tilde{u}_i}} (\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i]) \right) \\
&= \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \frac{\sigma^2_{\tilde{\theta}_i}}{1 + \frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2 \sigma^2_{\tilde{\theta}_i}(1-\tilde{\gamma}_i)^2}} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\tau}}{\sigma^2_{\tilde{u}_i}} \right)^{-2} \sigma^2_{\tilde{x}_i} \left(\frac{\tilde{m}_i}{1+\tilde{m}_i} \right) \\
&= \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \frac{\sigma^2_{\tilde{\theta}_i}}{1 + \frac{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2 \sigma^2_{\tilde{\theta}_i}(1-\tilde{\gamma}_i)^2}} + \left(\frac{(1-\tilde{\gamma}_i)\tau}{\sigma^2_{\tilde{u}_i}} \right)^{-2} \sigma^2_{\tilde{x}_i} \left(1 + \frac{\tau^2 \sigma^2_{\tilde{\theta}_i}(1-\tilde{\gamma}_i)^2}{\sigma^4_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}} \right)^{-1} \qquad \text{(A.44)}
\end{aligned}
$$

$$\Downarrow$$

$$\frac{\partial \mathrm{Var}(\tilde{v}_i - \tilde{P}_i)}{\partial \tilde{c}_i} > 0 \text{ and } \frac{\partial \mathrm{Var}(\tilde{v}_i - \tilde{P}_i)}{\partial \tilde{k}_i} \geq 0$$

The derivative signs follow from the fact that all derivatives functions are positive, therefore the direction of the derivation is entirely determined by the sign on $\frac{\partial \tilde{\gamma}_i}{\partial \tilde{c}_i}$ and $\frac{\partial \tilde{\gamma}_i}{\partial \tilde{k}_i}$

A.7 Participation Levels with the Index Asset

I estimate the ratio of participating to not participating in asset i across three levels of information equilibria. The utility ratio is provided by equation (1.23):

$$\frac{EU_{I_j^* R_j^*}}{EU_{I_j R_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|I^*_{j_i}}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}} \exp\left(\frac{\mathbb{1}_{I^*_{j_i}}\tilde{c}_i}{\tau} + \frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i - \tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right) \tag{A.45}$$

where for various equilibrium levels, the ratio is given by:

$$\begin{cases} \text{Fully-Uninformed Equilibrium:} & \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}} \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right) \\ \text{Interior Equilibrium:} & \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}} \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right) \\ \text{Fully-Informed Equilibrium:} & \sqrt{\frac{\sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}} \exp\left(\frac{\tilde{c}_i}{\tau} + \frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right) \end{cases} \tag{A.46}$$

By equation (A.5), the risk premium for each asset is given by:

$$\mathbb{E}[\tilde{v}_i - \tilde{P}_i] = \tilde{X}_i\left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-1} \tag{A.47}$$

By the law of total variance we have:

$$\mathrm{Var}(\tilde{v}_i - \tilde{P}_i) = \sigma^2_{\tilde{v}_i|\tilde{P}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2} \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right) \tag{A.48}$$

which for assets with non-zero supply equals

$$\mathrm{Var}(\tilde{v}_i - \tilde{P}_i) = \sigma^2_{\tilde{v}_i|\tilde{P}_i} + \frac{\mathbb{E}[\tilde{v}_i - \tilde{P}_i]^2}{\tilde{X}_i^2}\mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right) \tag{A.49}$$

I apply the three cases of information equilibria, as specified by the conditions of equation (1.22):

$$\begin{cases} \text{Fully-Uninformed Equilibrium:} & \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right) = \sigma^2_{\tilde{x}_i} \\ \text{Interior Equilibrium:} & \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right) = \sigma^2_{\tilde{x}_i}\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i} \\ \text{Fully-Informed Equilibrium:} & \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right) = \sigma^2_{\tilde{x}_i}\left(\frac{\tilde{m}_i}{1+\tilde{m}_i}\right) \end{cases} \tag{A.50}$$

$$\begin{cases} \text{Fully-Uninformed Equilibrium:} & \tilde{c}_i \geq \frac{\tau}{2}\log(1+\tilde{n}_i) \text{ and } \tilde{\lambda}_i = 0 \\ \text{Interior Equilibrium:} & (1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau = \sigma_{\tilde{u}_i}\sigma_{\tilde{x}_i}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}} \\ \text{Fully-Informed Equilibrium:} & \tilde{c}_i \leq \frac{\tau}{2}\log(1+\frac{\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}{\tau^2(1-\tilde{\gamma}_i)^2\tilde{n}_i+\sigma^2_{\tilde{u}_i}\sigma^2_{\tilde{x}_i}}\tilde{n}_i) \text{ and } \tilde{\lambda}_i = 1 \end{cases} \tag{A.51}$$

and equation (1.30)

$$\begin{cases} \text{Fully-Uninformed Equilibrium:} & \sigma^2_{\tilde{v}_i|\tilde{P}_i} = \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} \\ \text{Interior Equilibrium:} & \sigma^2_{\tilde{v}_i|\tilde{P}_i} = \exp\left(\frac{2\tilde{c}_i}{\tau}\right)\sigma^2_{\tilde{u}_i} \\ \text{Fully-Informed Equilibrium:} & \sigma^2_{\tilde{v}_i|\tilde{P}_i} = \sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} - \frac{\sigma^2_{\tilde{\theta}_i}}{1+\tilde{m}_i} \end{cases} \tag{A.52}$$

where

$$m_i = \left(\frac{\sigma^2_{\tilde{u}_i}}{\tau(1-\tilde{\gamma}_i)\tilde{\lambda}_i}\right)^2\frac{\sigma^2_{\tilde{x}_i}}{\sigma^2_{\tilde{\theta}_i}} \tag{A.53}$$

$$n_i = \frac{\sigma^2_{\tilde{\theta}_i}}{\sigma^2_{\tilde{u}_i}} \tag{A.54}$$

Therefore we have

$$\begin{cases} \text{Fully-Uninformed Equilibrium:} & \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} = \frac{(1-\tilde{\gamma}_i)\tau}{\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i}} \\ \text{Interior Equilibrium:} & \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} = \frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}}\Big(1-\exp(-2\tilde{c}_i/\tau)\Big) \\ & + \frac{(1-\tilde{\gamma}_i)\tau}{\exp(2\tilde{c}_i/\tau)\sigma^2_{\tilde{u}_i}} \\ \text{Fully-Informed Equilibrium:} & \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}} = \frac{(1-\tilde{\gamma}_i)\tau}{\sigma^2_{\tilde{u}_i}} \end{cases}$$

Armed with the above, we can write the ratio for the utility change due to participation for any asset with non-zero supply as follows:

Fully-Uninformed Equilibrium:

$$\frac{EU_{I_j^* R_j^*}}{EU_{I_j R_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}} \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i - \tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right)$$

$$= \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i|\tilde{P}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2} \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)}}$$

$$\times \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i - \tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right)$$

$$= \sqrt{\frac{\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} + \frac{(\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i})^2}{(1-\tilde{\gamma}_i)^2\tau^2}\sigma^2_{\tilde{x}_i}}} \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i - \tilde{P}_i]^2}{2\left(\sigma^2_{\tilde{v}_i|\tilde{P}_i} + \frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{\bar{X}_i^2}\sigma^2_{\tilde{x}_i}\right)}\right)$$

$$= \sqrt{\frac{\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} + \frac{(\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i})^2}{(1-\tilde{\gamma}_i)^2\tau^2}\sigma^2_{\tilde{x}_i}}} \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{1/2}{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2} + \frac{\sigma^2_{\tilde{x}_i}}{\bar{X}_i^2}}\right)$$

$$= \sqrt{\frac{\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i}}{\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i} + \frac{(\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i})^2}{(1-\tilde{\gamma}_i)^2\tau^2}\sigma^2_{\tilde{x}_i}}} \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{1/2}{\frac{(1-\tilde{\gamma}_i)^2\tau^2}{\bar{X}_i^2(\sigma^2_{\tilde{\theta}_i} + \sigma^2_{\tilde{u}_i})} + \frac{\sigma^2_{\tilde{x}_i}}{\bar{X}_i^2}}\right)$$

By the chain rule both the square root term and the exponential term are decreasing in $\tilde{\gamma}_i$ implying that the ratio of expected utilities is decreasing as well.

Interior Equilibrium:

$$\frac{EU_{I_j^* R_j^*}}{EU_{I_j R_j}} = \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}} \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right)$$

$$= \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i|\tilde{P}_i} + \left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}} + \frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2} \mathrm{Var}\left(\tilde{x}_i - \frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i - \mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)}}$$

$$\times \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right)$$

$$= \sqrt{\frac{1}{1+\sigma^2_{\tilde{x}_i}\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}\exp(-2\tilde{c}_i/\tau)\sigma^{-2}_{\tilde{u}_i}\left(\frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}}\Big(1-\exp(-2\tilde{c}_i/\tau)\Big) + \frac{(1-\tilde{\gamma}_i)\tau}{\exp(2\tilde{c}_i/\tau)\sigma^2_{\tilde{u}_i}}\right)^{-2}}}$$

$$\times \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{2\left(\sigma^2_{\tilde{v}_i|\tilde{P}_i} + \frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{\tilde{X}_i^2}\sigma^2_{\tilde{x}_i}\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}\right)}\right)$$

$$= \sqrt{\frac{1}{1+\sigma^2_{\tilde{x}_i}\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}\exp(-2\tilde{c}_i/\tau)\sigma^{-2}_{\tilde{u}_i}\left(\frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}}\Big(1-\exp(-2\tilde{c}_i/\tau)\Big) + \frac{(1-\tilde{\gamma}_i)\tau}{\exp(2\tilde{c}_i/\tau)\sigma^2_{\tilde{u}_i}}\right)^{-2}}}$$

$$\times \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{1/2}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^{-2} + \frac{\sigma^2_{\tilde{x}_i}}{\tilde{X}_i^2}\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}}\right)$$

$$= \sqrt{\frac{1}{1+\sigma^2_{\tilde{x}_i}\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}\exp(-2\tilde{c}_i/\tau)\sigma^{-2}_{\tilde{u}_i}\left(\frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}}\Big(1-\exp(-2\tilde{c}_i/\tau)\Big) + \frac{(1-\tilde{\gamma}_i)\tau}{\exp(2\tilde{c}_i/\tau)\sigma^2_{\tilde{u}_i}}\right)^{-2}}}$$

$$\times \exp\left(\frac{\tilde{k}_i}{\tau} - \frac{1/2}{\frac{\exp(2\tilde{c}_i/\tau)\sigma^2_{\tilde{u}_i}}{\tilde{X}_i^2}\left(\frac{\sigma_{\tilde{x}_i}}{\sigma_{\tilde{u}_i}}\sqrt{\frac{1}{\exp(2\tilde{c}_i/\tau)-1} - \frac{1}{\tilde{n}_i}}\Big(1-\exp(-2\tilde{c}_i/\tau)\Big) + \frac{(1-\tilde{\gamma}_i)\tau}{\exp(2\tilde{c}_i/\tau)\sigma^2_{\tilde{u}_i}}\right)^2 + \frac{\sigma^2_{\tilde{x}_i}}{\tilde{X}_i^2}\frac{\exp(2\tilde{c}_i/\tau)-1}{\tilde{n}_i}}\right)$$

By the chain rule both the square root term and the exponential term are decreasing in $\tilde{\gamma}_i$ implying that the ratio of expected utilities is decreasing as well.

Fully-Informed Equilibrium:

$$
\begin{aligned}
\frac{EU_{I_j^* R_j^*}}{EU_{I_j R_j}} &= \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i-\tilde{P}_i}}}\exp\left(\frac{\tilde{k}_i}{\tau}+\frac{\tilde{c}_i}{\tau}-\frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right)\\
&= \sqrt{\frac{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}+\left(\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}+\frac{(1-\tilde{\gamma}_i)(1-\tilde{\lambda}_i)\tau}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}}\right)^{-2}\mathrm{Var}\left(\tilde{x}_i-\frac{(1-\tilde{\gamma}_i)\tilde{\lambda}_i\tau}{\sigma^2_{\tilde{u}_i}}(\tilde{\theta}_i-\mathbb{E}[\tilde{v}_i|\tilde{P}_i])\right)}}\times\exp\left(\frac{\tilde{k}_i}{\tau}+\frac{\tilde{c}_i}{\tau}-\frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{2\sigma^2_{\tilde{v}_i-\tilde{P}_i}}\right)\\
&= \sqrt{\frac{1}{1+\left(\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i}-\frac{\sigma^2_{\tilde{\theta}_i}}{1+\tilde{m}_i}\right)^{-1}\frac{\sigma^4_{\tilde{u}_i}}{(1-\tilde{\gamma}_i)^2\tau^2}\sigma^2_{\tilde{x}_i}\left(\frac{\tilde{m}_i}{1+\tilde{m}_i}\right)}}\times\exp\left(\frac{\tilde{k}_i}{\tau}+\frac{\tilde{c}_i}{\tau}-\frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{2\left(\sigma^2_{\tilde{v}_i|\tilde{P}_i}+\frac{\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^2}{\tilde{X}_i^2}\sigma^2_{\tilde{x}_i}\left(\frac{\tilde{m}_i}{1+\tilde{m}_i}\right)\right)}\right)\\
&= \sqrt{\frac{1}{1+\frac{\tilde{m}_i^2}{\tilde{m}_i(\sigma^4_{\tilde{\theta}_i}+\sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i})+\sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i}}}}\times\exp\left(\frac{\tilde{k}_i}{\tau}+\frac{\tilde{c}_i}{\tau}-\frac{1/2}{\sigma^2_{\tilde{v}_i|\tilde{P}_i}\mathbb{E}[\tilde{v}_i-\tilde{P}_i]^{-2}+\frac{\sigma^2_{\tilde{x}_i}}{\tilde{X}_i^2}\left(\frac{\tilde{m}_i}{1+\tilde{m}_i}\right)}\right)\\
&= \sqrt{\frac{1}{1+\frac{\tilde{m}_i^2}{\tilde{m}_i(\sigma^4_{\tilde{\theta}_i}+\sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i})+\sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i}}}}\times\exp\left(\frac{\tilde{k}_i}{\tau}+\frac{\tilde{c}_i}{\tau}-\frac{1/2}{\left(\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i}-\frac{\sigma^2_{\tilde{\theta}_i}}{1+\tilde{m}_i}\right)\frac{(1-\tilde{\gamma}_i)^2\tau^2}{\tilde{X}_i^2\sigma^4_{\tilde{u}_i}}+\frac{\sigma^2_{\tilde{x}_i}}{\tilde{X}_i^2}\left(\frac{\tilde{m}_i}{1+\tilde{m}_i}\right)}\right)\\
&= \sqrt{\frac{1}{1+\frac{\tilde{m}_i^2}{\tilde{m}_i(\sigma^4_{\tilde{\theta}_i}+\sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i})+\sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i}}}}\times\exp\left(\frac{\tilde{k}_i}{\tau}+\frac{\tilde{c}_i}{\tau}-\frac{1/2}{\left(\sigma^2_{\tilde{\theta}_i}+\sigma^2_{\tilde{u}_i}-\frac{\sigma^2_{\tilde{\theta}_i}}{1+\tilde{m}_i}\right)\frac{1}{\sigma^2_{\tilde{\theta}_i}\tilde{m}_i}\frac{\sigma^2_{\tilde{x}_i}}{\tilde{X}_i^2}+\frac{\sigma^2_{\tilde{x}_i}}{\tilde{X}_i^2}\left(\frac{\tilde{m}_i}{1+\tilde{m}_i}\right)}\right)\\
&= \sqrt{\frac{1}{1+\frac{\tilde{m}_i^2}{\tilde{m}_i(\sigma^4_{\tilde{\theta}_i}+\sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i})+\sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i}}}}\times\exp\left(\frac{\tilde{k}_i}{\tau}+\frac{\tilde{c}_i}{\tau}-\frac{1/2}{\frac{\sigma^2_{\tilde{x}_i}}{\tilde{X}_i^2}\left(\frac{1}{\tilde{m}_i}+\frac{1}{\tilde{n}_i\tilde{m}_i}-\frac{1}{\tilde{m}_i(1+\tilde{m}_i)}+\frac{\tilde{m}_i}{1+\tilde{m}_i}\right)}\right)\\
&= \sqrt{\frac{1}{1+\frac{\tilde{m}_i^2}{\tilde{m}_i(\sigma^4_{\tilde{\theta}_i}+\sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i})+\sigma^2_{\tilde{\theta}_i}\sigma^2_{\tilde{u}_i}}}}\times\exp\left(\frac{\tilde{k}_i}{\tau}+\frac{\tilde{c}_i}{\tau}-\frac{1/2}{\frac{\sigma^2_{\tilde{x}_i}}{\tilde{X}_i^2}\frac{\tilde{n}_i\tilde{m}_i^2+(\tilde{n}_i+1)\tilde{m}_i+1}{\tilde{n}_i\tilde{m}_i^2+\tilde{n}_i\tilde{m}_i}}\right)
\end{aligned}
$$

By the chain rule and the fact that $\frac{\partial \tilde{m}_i}{\partial \tilde{\gamma}_i} > 0$ both the square root term and the exponential term are decreasing in $\tilde{\gamma}_i$ implying that the ratio of expected utilities is decreasing as well.

A.8 Effects of Changing Information and Participation Costs on the Original Assets

The following portfolio weights, as given by the coefficients on the corresponding synthetic asset payoffs, map the synthetic assets with payoffs $\tilde{v}_1, \ldots, \tilde{v}_N$ back to the original assets with payoffs $v_1, \ldots, v_N$

$$\begin{cases} \text{If } i = 1: & v_1 = \frac{X_1 \sigma^2_{\theta_1}}{Nq} \tilde{v}_1 + \frac{X_1 \sigma^2_{\theta_1}}{N} \tilde{v}_2 \\ \text{If } 1 < i < N: & v_i = \frac{X_i \sigma^2_{\theta_i}}{Nq} \tilde{v}_1 - \sum_{l=2}^{i} \frac{X_i \sigma^2_{\theta_i}}{(N-l+2)(N-l+1)} \tilde{v}_l + \frac{X_i \sigma^2_{\theta_i}}{N-i+1} \tilde{v}_{i+1} \\ \text{If } i = N: & v_N = \frac{X_N \sigma^2_{\theta_N}}{Nq} \tilde{v}_1 - \sum_{l=2}^{N} \frac{X_N \sigma^2_{\theta_N}}{(N-l+2)(N-l+1)} \tilde{v}_l \end{cases} \tag{A.55}$$

The coefficients on the synthetic payoffs above correspond to the portfolio weights placed on each of the synthetic assets in order to arrive at the original asset. For each original asset i, I label the portfolio weights for the synthetic assets as $w_{i1}, w_{i2}, \ldots, w_{iN}$. According to equation (A.55), the weight placed on the index portfolio for each original asset i, i.e. its market beta, is:

$$w_{i1} = \frac{X_i \sigma^2_{\theta_i}}{Nq} \tag{A.56}$$

where q defines the relationship between expected supply and payoff uncertainty, and is assumed to be constant across assets. Therefore the original assets' market betas are increasing in their supply and in fundamental uncertainty and are decreasing in the number of risky assets. All original assets $i < N$ are negatively exposed to synthetic assets $1 < j < i+1$ and are positively exposed to synthetic asset $j = i+1$. Original asset $i = N$ is negatively exposed to all synthetic assets $j > 1$. For directional derivations of the effects of information and participation costs on variances and covariances only the signs on the portfolio weights matter, not the magnitudes.

Utilizing the notation above and the law of one price I can define various measures of expected returns, variances and covariances:

1. The variance of the return of the original asset i will be equal to:

$$\text{Var}(v_i - P_i) = \sum_{j=1}^{N} w_{ij}^2 \text{Var}(\tilde{v}_j - \tilde{P}_j) \tag{A.57}$$

2. The expected return of original asset i will be equal to:

$$\begin{aligned} \mathbb{E}[v_i - P_i] &= \sum_{j=1}^{N} w_{ij} \mathbb{E}[\tilde{v}_j - \tilde{P}_j] \\ &= w_{i1} \mathbb{E}[\tilde{v}_1 - \tilde{P}_1] \end{aligned} \tag{A.58}$$

3. The covariance of the return of original asset m with original asset n

$$\mathrm{Cov}(v_m - p_m, v_n - p_n) = \sum_{j=1}^{N} w_{mj} w_{nj} \mathrm{Var}(\tilde{v}_j - \tilde{P}_j) \tag{A.59}$$

4. The covariance of the return of original asset i with the market

$$\mathrm{Cov}(v_i - p_i, \tilde{v}_1 - \tilde{p}_1) = w_{i1} \mathrm{Var}(\tilde{v}_1 - \tilde{P}_1) \tag{A.60}$$

From the definitions above, the mapping from the synthetic assets to the original assets as given by equation (A.55), and the properties of expected returns and variances as highlighted in Lemmas I.5 and I.6, the following return properties arise.

Lemma A.1 (Effects of Information and Participation Costs on Return Properties)**.** *The effects of stock picking costs on the return properties of the original assets are as follows:*

1. *Expected returns: not affected by changes to stock picking information costs.*
2. *Return variances: change in the variance of a synthetic asset as specified by Lemma I.6, affects the variances of ALL original assets in the same direction.*
3. *Return covariances between original assets m and n, where $m < n$:*
   - *If $1 < m < n \leq N$: changes in information costs for synthetic assets $1 < j < m+1$ will have the same effect directionally as the change in variance as specified by Lemma I.6. Change in the information cost for synthetic asset $j = m+1$ will have the opposite effect directionally as the change in variance specified by Lemma I.6. Change in information costs for synthetic assets $j > m+1$ will have no effect on the covariance.*
   - *If $m = 1$: change in the information cost for synthetic asset $j = 2$ will have the opposite effect directionally as the change in variance specified by Lemma I.6. Change in information costs for synthetic assets $j > 2$ will have no effect on the covariance.*

The effects of index information costs on the return properties of the original assets are as follows:

1. *Expected returns: proportional to the change in the expected return for the index asset as specified by Lemma I.5 for ALL original assets. Coefficient of proportionality is given by the original asset's weight in the market index as specified by equation* (A.56).

2. *Return variances: change in the variance of the index asset as specified by Lemma I.6, affects the variances of ALL original assets in the same direction.*

3. *Return covariances: change in the variance of the index asset as specified by Lemma I.6, affects ALL of the pair-by-pair covariances of the original assets in the same direction.*

The effects of stock picking participation costs on the return properties of the original assets are as follows:

1. *Expected returns: not affected by changes to stock-picking participation costs.*

2. *Return variances: change in the variance of the synthetic asset as specified by Lemma I.6, affects the variances of ALL original assets in the same direction.*

3. *Return covariances between original assets m and n, where $m < n$:*

    - *If $1 < m < n \leq N$: changes in participation costs for synthetic assets $1 < j < m+1$ will have the same effect directionally as the change in variance as specified by Lemma I.6. Change in the participation cost for synthetic asset $j = m+1$ will have the opposite effect directionally as the change in variance specified by Lemma I.6. Change in participation costs for synthetic assets $j > m+1$ will have no effect on the covariance.*
    - *If $m = 1$: change in the participation cost for synthetic asset $j = 2$ will have the opposite effect directionally as the change in variance specified by Lemma I.6. Change in participation costs for synthetic assets $j > 2$ will have no effect on the covariance.*

A.9 Lemma A.2 with Proof

Lemma A.2 (Reducing Correlated Fundamentals to the Diagonal Case)**.** *Assume that $\Sigma_u = \sigma_u^2 \times I_N$ and Σ_θ is positive definite. Using the eigen-decomposition of Σ_θ, the payoff space can be re-spanned with portfolios of the underlying securities whose fundamental and noise payoffs are orthogonal.*

Consider the eigen-decomposition of the fundamental payoff matrix:

$$\Sigma_\theta = Q_\theta \Lambda_\theta Q'_\theta \tag{A.61}$$

Since Σ_θ is a covariance matrix and is thus symmetric, such a decomposition exists where $\Lambda_\theta = \text{diag}(\lambda_1, \ldots, \lambda_N)$ contains the eigenvalues of the Σ_θ and $Q_\theta = [\mathbf{q_1}, \ldots, \mathbf{q_N}]$ contains the corresponding weights on the assets underlying the variance-covariance matrix. The eigen-decomposition decomposes the N potentially correlated assets into N orthogonal portfolios, which span the original asset space. Any portfolio w_i of the original assets can be replicated using portfolio w_o of the orthogonal assets as follows:

$$w_i = Q_\theta w_o \Rightarrow Q'_\theta w_i = w_o \tag{A.62}$$

since $Q'_\theta Q_\theta = I_N$ under the original decomposition. I can thus restate the original problem in terms of the orthogonal assets. Returns are governed by:

$$Q'_\theta v = Q'_\theta \theta + Q'_\theta u \tag{A.63}$$

Distributional assumptions are

$$\begin{pmatrix} Q'_\theta \theta \\ Q'_\theta u \end{pmatrix} \sim N \left[\begin{pmatrix} Q'_\theta \mu_\theta \\ 0 \end{pmatrix}, \begin{pmatrix} Q'_\theta \Sigma_\theta Q_\theta & 0 \\ 0 & Q'_\theta \Sigma_u Q_\theta \end{pmatrix} \right] \tag{A.64}$$

Renaming the original variables to have subscripts "o" for their orthogonal versions, under the assumption that the noise matrix has the form $\Sigma_u = \sigma_u^2 I_N$ and by the eigen-value decomposition, I have:

$$\begin{pmatrix} \theta_o \\ u_o \end{pmatrix} \sim N \left[\begin{pmatrix} \mu_{\theta_o} \\ 0 \end{pmatrix}, \begin{pmatrix} \Lambda_\theta & 0 \\ 0 & \Sigma_u \end{pmatrix} \right] \tag{A.65}$$

where Λ_θ and Σ_u are diagonal with

$$v_o = \theta_o + u_o \tag{A.66}$$

The aggregate supply of the orthogonal portfolios, X_o is given by:

$$Q_\theta X_o = X \Rightarrow X_o = Q'_\theta X \tag{A.67}$$

Therefore, I can restate the problem with correlated fundamentals in terms of assets, which have uncorrelated fundamentals and noise payoffs.

APPENDIX B

Quants: Derivations and Proofs

B.1 Proof of Lemma II.1

We conjecture that the DI's demand and MM's price functions are governed by Eq. (2.5) and (2.6). Given rational expectations, the DI's problem simplifies to:

$$\begin{aligned}
x_d &= \arg\max_{\tilde{x}} \mathbb{E}[\tilde{x}(v - P_1)|s_d] \\
&= \arg\max_{\tilde{x}} \mathbb{E}[\tilde{x}(v - \mu_1 - \lambda_1\omega_1)|s_d] \\
&= \arg\max_{\tilde{x}} \mathbb{E}[\tilde{x}(v - \mu_1 - \lambda_1(\tilde{x} + z))|s_d] \\
&= \arg\max_{\tilde{x}} \mathbb{E}[\tilde{x}(v - \mu_1 - \lambda_1 z) - \lambda_1\tilde{x}^2|s_d] \\
&= \arg\max_{\tilde{x}} \left\{\tilde{x}(\mathbb{E}[v|s_d] - \mu_1) - \lambda_1\tilde{x}^2\right\} \\
&= \frac{\mathbb{E}[v|s_d] - \mu_1}{2\lambda_1} \\
&= \frac{P_0(1-\phi_d) + s_d\phi_d - \mu_1}{2\lambda_1} \\
&= \frac{P_0(1-\phi_d) - \mu_1}{2\lambda_1} + \frac{\phi_d}{2\lambda_1}s_d \\
&\equiv \alpha_{d1} + \beta_{d1}s_d
\end{aligned} \tag{B.1}$$

Similarly, the market-maker's problem becomes

$$\begin{aligned}
P_1(\omega_1) &= \mathbb{E}[v|\omega_1] \\
&= \mathbb{E}[v|x_d + z] \\
&= \mathbb{E}[v|\alpha_{d1} + \beta_{d1}(v + e_d) + z] \\
&= P_0 + \frac{\beta_{d1}\sigma_v^2}{\beta_{d1}^2(\sigma_v^2 + \sigma_{e_d}^2) + \sigma_z^2}(\omega_1 - \alpha_{d1} - \beta_{d1}P_0) \qquad \text{(B.2)} \\
&= P_0 - \frac{\beta_{d1}\sigma_v^2(\alpha_{d1} + \beta_{d1}P_0)}{\beta_{d1}^2(\sigma_v^2 + \sigma_{e_d}^2) + \sigma_z^2} + \frac{\beta_{d1}\sigma_v^2}{\beta_{d1}^2(\sigma_v^2 + \sigma_{e_d}^2) + \sigma_z^2}\omega_1 \\
&\equiv \mu_1 + \lambda_1\omega_1
\end{aligned}$$

We first solve for λ_b utilizing Eq. (B.1) and (B.2). From Eq. (B.1):

$$\begin{gathered}
\beta_{d1} = \frac{\phi_d}{2\lambda_1} \\
\Downarrow \\
\frac{1}{\lambda_1} = \frac{2\beta_{d1}}{\phi_d}
\end{gathered} \qquad \text{(B.3)}$$

From Eq. (B.2)

$$\begin{gathered}
\lambda_1 = \frac{\beta_{d1}\sigma_v^2}{\beta_{d1}^2(\sigma_v^2 + \sigma_{e_d}^2) + \sigma_z^2} \\
\Downarrow \\
\frac{1}{\lambda_1} = \frac{\beta_{d1}}{\phi_d} + \frac{\eta}{\beta_{d1}}
\end{gathered} \qquad \text{(B.4)}$$

Setting Eq. (B.3) into (B.4) equal

$$\begin{gathered}
\frac{\beta_{d1}}{\phi_d} + \frac{\eta}{\beta_{d1}} = \frac{2\beta_{d1}}{\phi_d} \\
\Downarrow \\
\frac{\eta}{\beta_{d1}} = \frac{\beta_{d1}}{\phi_d} \\
\Downarrow \\
\beta_{d1} = \pm\sqrt{\eta\phi_d} \\
\Downarrow \\
\lambda_1 = \pm\frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}
\end{gathered}$$

We will focus on economically meaningful equilibria, i.e. strategies that buy more with a stronger signal ($\beta_{d1} > 0$) and price functions that charge more with greater demand ($\lambda_1 > 0$). Therefore:

$$\beta_{d1} = \sqrt{\eta\phi_d} \qquad\qquad \lambda_1 = \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}$$

Next we solve for α_{d1} and μ_1. Utilizing Eq. (B.1)

$$\begin{aligned}
\alpha_{d1} &= \frac{P_0(1-\phi_d) - \mu_1}{2\lambda_1} \\
&\Downarrow \\
\mu_1 &= P_0(1-\phi_d) - 2\lambda_1\alpha_{d1}
\end{aligned}$$

Substituting into Eq. (B.2)

$$\begin{aligned}
P_0(1-\phi_d) - 2\lambda_1\alpha_{d1} &= P_0 - \frac{\beta_{d1}\sigma_v^2(\alpha_{d1} + \beta_{d1}P_0)}{\beta_{d1}^2(\sigma_v^2 + \sigma_{e_d}^2) + \sigma_z^2} \\
&\Downarrow \\
P_0\phi_d + 2\lambda_1\alpha_{d1} &= \frac{\beta_{d1}\sigma_v^2(\alpha_{d1} + \beta_{d1}P_0)}{\beta_{d1}^2(\sigma_v^2 + \sigma_{e_d}^2) + \sigma_z^2} \\
&\Downarrow \\
P_0\phi_d + 2\lambda_1\alpha_{d1} &= \frac{\beta_{d1}(\alpha_{d1} + \beta_{d1}P_0)}{\beta_{d1}^2/\phi_d + \eta} \\
&\Downarrow \\
P_0\phi_d + 2\lambda_1\alpha_{d1} &= \frac{\beta_{d1}(\alpha_{d1} + \beta_{d1}P_0)}{2\eta} \\
&\Downarrow \\
2P_0\eta\phi_d + 4\lambda_1\eta\alpha_{d1} &= \beta_{d1}\alpha_{d1} + \beta_{d1}^2P_0 \\
&\Downarrow \\
P_0\eta\phi_d &= (\beta_{d1} - 4\lambda_1\eta)\alpha_{d1} \\
&\Downarrow \\
P_0\eta\phi_d &= -\sqrt{\eta\phi_d}\,\alpha_{d1} \\
&\Downarrow \\
\alpha_{d1} &= -P_0\sqrt{\eta\phi_d} \\
&\Downarrow \\
\mu_1 &= P_0
\end{aligned}$$

The solution coefficients are:

$$
\begin{aligned}
\mu_1 &= P_0 \\
\lambda_1 &= \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}} \\
\alpha_{d1} &= -P_0\sqrt{\eta\phi_d} \\
\beta_{d1} &= \sqrt{\eta\phi_d}
\end{aligned}
\tag{B.5}
$$

B.2 Derivation of Backtest: Eq. (2.9)

The QI feeds the DI's trading strategy and the MM's pricing rule from Economy 1 given by Eq. (2.7) and (2.8) into her optimization problem:

$$
\begin{aligned}
x_b &= \arg\max_{\tilde{x}_b} \mathbb{E}\Big[\tilde{x}_b\Big(v - P_1\big(x_{d1} + z + \tilde{x}_b\big)\Big)|s_q\Big] \\
&= \arg\max_{\tilde{x}_b} \mathbb{E}\left[\tilde{x}_b\left(v - P_0 - \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}\Big(x_{d1} + z + \tilde{x}_b\Big)\right)|s_q\right] \\
&= \arg\max_{\tilde{x}_b} \mathbb{E}\left[\tilde{x}_b\left(v - P_0 - \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}\Big(\sqrt{\phi_d\eta}(s_d - P_0) + z + \tilde{x}_b\Big)\right)|s_q\right] \\
&= \arg\max_{\tilde{x}_b} \mathbb{E}\left[\tilde{x}_b\left(v - P_0 - \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}\Big(\sqrt{\phi_d\eta}(v + e_d - P_0) + z + \tilde{x}_b\Big)\right)|s_q\right] \\
&= \arg\max_{\tilde{x}_b} \left(\mathbb{E}[v|s_q] - P_0 - \frac{\phi_d}{2}\Big(\mathbb{E}[v|s_q] + \mathbb{E}[e_d|s_q] - P_0\Big)\right)\tilde{x}_b - \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}\tilde{x}_b^2 \\
&= \arg\max_{\tilde{x}_b} \left(\frac{2-\phi_d}{2}\mathbb{E}[v|s_q] - \frac{2-\phi_d}{2}P_0 - \frac{\phi_d}{2}\mathbb{E}[e_d|s_q]\right)\tilde{x}_b - \frac{1}{2}\sqrt{\frac{\phi_d}{\eta}}\tilde{x}_b^2 \\
&= \sqrt{\frac{\eta}{\phi_d}}\left(\frac{2-\phi_d}{2}\mathbb{E}[v|s_q] - \frac{2-\phi_d}{2}P_0 - \frac{\phi_d}{2}\mathbb{E}[e_d|s_q]\right)
\end{aligned}
$$

(...note that $\mathbb{E}[v|s_q]=(1-\phi_q)P_0+\phi_q s_q$ and $\mathbb{E}[e_d|s_q]=\mathbb{E}[e_d|v+e_q]=\sigma_{e_d e_q}/\sigma^2_{s_q}(s_q-P_0)$ $=\rho\sigma_{e_d}\sigma_{e_q}/\sigma^2_{s_q}(s_q-P_0)=\rho(\sigma_{e_d}/\sigma_{s_q})(\sigma_{e_q}/\sigma_{s_q})(s_q-P_0)=\rho\sqrt{\frac{\phi_q(1-\phi_d)(1-\phi_q)}{\phi_d}}(s_q-P_0)\ldots)$

$$
\begin{aligned}
&= \sqrt{\frac{\eta}{\phi_d}}\left(\frac{2-\phi_d}{2}\Big(-\phi_q P_0 + \phi_q s_q\Big) - \frac{\rho}{2}\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}(s_q - P_0)\right) \\
&= \frac{1}{2}\sqrt{\frac{\eta}{\phi_d}}\left((2-\phi_d)\phi_q - \rho\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}\right)(s_d - P_0)
\end{aligned}
$$

B.3 Proof of Lemma II.2

B.3.1 Economy 2. Fully-strategic Quant, $\gamma = 0$

This is the case of two strategic speculators $i \in \{d, q\}$. The problem is fully symmetric for both strategic speculators, therefore we solve it from the QI's perspective and reverse notation for the DI. In what follows, note that:

$$
\begin{aligned}
\mathbb{E}[v|s_q] &= (1-\phi_q)P_0 + \phi_q s_q \\
\mathbb{E}[v|s_d] &= (1-\phi_d)P_0 + \phi_d s_d \\
\mathbb{E}[e_d|s_q] &= \rho\sqrt{\frac{\phi_q(1-\phi_d)(1-\phi_q)}{\phi_d}}(s_q - P_0) \\
&\equiv \nu_q(s_q - P_0) \\
\mathbb{E}[e_q|s_d] &= \rho\sqrt{\frac{\phi_d(1-\phi_d)(1-\phi_q)}{\phi_q}}(s_d - P_0) \\
&\equiv \nu_d(s_d - P_0)
\end{aligned}
$$

The QI chooses demand x_q to maximize the following expectation:

$$x_{q2} = \arg\max_{\tilde{x}_q} \mathbb{E}[\tilde{x}_q(v - P_2)|s_q] \tag{B.6}$$

For both speculators we conjecture a demand function linear in the signal:

$$x_{i2}(s_i) = \alpha_{i2} + \beta_{i2} s_i \tag{B.7}$$

Noise traders supply z shares and market-makers set prices according to the aggregate order flow:

$$\omega_2 = x_{d2} + x_{q2} + z \tag{B.8}$$

The market makers in the economy are competitive and risk-neutral resulting and are assumed to have a pricing rule linear in the aggregate order flow:

$$
\begin{aligned}
P_2(\omega_2) &= \mathbb{E}[v|\omega_2] \\
&= \mu_2 + \lambda_2\omega_2
\end{aligned}
\tag{B.9}
$$

We first consider the QI's problem:

$$
\begin{aligned}
x_{q2} &= \arg\max_{\tilde{x}_{q2}} \mathbb{E}[\tilde{x}_{q2}(v - P_2(\omega_2))|s_q] \\
&= \arg\max_{\tilde{x}_{q2}} \mathbb{E}[\tilde{x}_{q2}(v - \mu_2 - \lambda_2\omega_2)|s_q] \\
&= \arg\max_{\tilde{x}_{q2}} \mathbb{E}[\tilde{x}_{q2}(v - \mu_2 - \lambda_2(x_{d2} + \tilde{x}_{q2} + z))|s_q] \\
&= \arg\max_{\tilde{x}_{q2}} \mathbb{E}[\tilde{x}_{q2}(v - \mu_2 - \lambda_2 x_{d2} - \lambda_2 z) - \lambda_2\tilde{x}_{q2}^2|s_q] \\
&= \arg\max_{\tilde{x}_{q2}} \{\tilde{x}_{q2}(\mathbb{E}[v|s_q] - \mu_2 - \lambda_2\mathbb{E}[x_{d2}|s_q]) - \lambda\tilde{x}_{q2}^2\} \\
&= \frac{\mathbb{E}[v|s_q] - \mu_2 - \lambda_2\mathbb{E}[x_{d2}|s_q]}{2\lambda_2} \\
&= \frac{\mathbb{E}[v|s_q] - \mu_2 - \lambda_2\alpha_{d2} - \lambda_2\beta_{d2}\mathbb{E}[s_d|s_q]}{2\lambda_2} \\
&= \frac{\mathbb{E}[v|s_q] - \mu_2 - \lambda_2\alpha_{d2} - \lambda_2\beta_{d2}\mathbb{E}[v + e_d|s_q]}{2\lambda_2} \\
&= \frac{(1 - \lambda_2\beta_{d2})\mathbb{E}[v|s_q] - \mu_2 - \lambda_2\alpha_{d2} - \lambda_2\beta_{d2}\mathbb{E}[e_d|s_q]}{2\lambda_2} \\
&= \frac{(1 - \lambda_2\beta_{d2})\big((1 - \phi_q)P_0 + \phi_q s_q\big) - \mu_2 - \lambda_2\alpha_{d2} - \lambda_2\beta_{d2}\big(\nu_q s_q - \nu_q P_0\big)\big)}{2\lambda_2} \\
&= \frac{(1 - \lambda_2\beta_{d2})(1 - \phi_q)P_0}{2\lambda_2} - \frac{\mu_2}{2\lambda_2} - \frac{\alpha_{d2}}{2} + \frac{\beta_{d2}\nu_q P_0}{2} + \frac{(1 - \lambda_2\beta_{d2})\phi_q - \lambda_2\beta_{d2}\nu_q}{2\lambda_2} s_q \\
&= \frac{(1 - \lambda_2\beta_{d2})(1 - \phi_q)P_0}{2\lambda_2} - \frac{\mu_2}{2\lambda_2} - \frac{\alpha_{d2}}{2} + \frac{\beta_{d2}\nu_q P_0}{2} + \left(\frac{\phi_q}{2\lambda_2} - \frac{\phi_q\beta_{d2}}{2} - \frac{\nu_q\beta_{d2}}{2}\right) s_q \\
&\equiv \alpha_{q2} + \beta_{q2}s_q
\end{aligned}
\tag{B.10}
$$

Similarly, for the DI:

$$
\begin{aligned}
x_{d2} &= \arg\max_{\tilde{x}_{d2}} \mathbb{E}[\tilde{x}_{d2}(v - P_2(\omega_2))|s_d] \\
&= \frac{(1 - \lambda_2\beta_{q2})(1 - \phi_d)P_0}{2\lambda_2} - \frac{\mu_2}{2\lambda_2} - \frac{\alpha_{q2}}{2} + \frac{\beta_{q2}\nu_d P_0}{2} + \left(\frac{\phi_d}{2\lambda_2} - \frac{\phi_d\beta_{q2}}{2} - \frac{\nu_d\beta_{q2}}{2}\right) s_d \\
&\equiv \alpha_{d2} + \beta_{d2}s_d
\end{aligned}
\tag{B.11}
$$

Finally, the competitive market maker sets prices according to:

$$
\begin{aligned}
P_2(\omega_2) &= \mathbb{E}[v|\omega_2] \\
&= \mathbb{E}[v|x_{d2} + x_{q2} + z] \\
&= \mathbb{E}[v|\alpha_{d2} + \beta_{d2}(v + e_d) + \alpha_{q2} + \beta_{q2}(v + e_q) + z] \\
&= P_0 + \frac{(\beta_{d2}+\beta_{q2})\sigma_v^2}{(\beta_{d2}+\beta_{q2})^2\sigma_v^2+\beta_{d2}^2\sigma_{e_d}^2+\beta_{q2}^2\sigma_{e_q}^2+2\beta_{d2}\beta_{q2}\sigma_{e_d e_q}+\sigma_z^2}(\omega_2-\alpha_{d2}-\alpha_{q2}-(\beta_{d2}+\beta_{q2})P_0) \\
&= P_0 - \frac{(\beta_{d2}+\beta_{q2})\sigma_v^2(\alpha_{d2}+\alpha_{q2}+(\beta_{d2}+\beta_{q2})P_0)}{(\beta_{d2}+\beta_{q2})^2\sigma_v^2+\beta_{d2}^2\sigma_{e_d}^2+\beta_{q2}^2\sigma_{e_q}^2+2\beta_{d2}\beta_{q2}\sigma_{e_d e_q}+\sigma_z^2} + \frac{(\beta_{d2}+\beta_{q2})\sigma_v^2}{(\beta_{d2}+\beta_{q2})^2\sigma_v^2+\beta_{d2}^2\sigma_{e_d}^2+\beta_{q2}^2\sigma_{e_q}^2+2\beta_{d2}\beta_{q2}\sigma_{e_d e_q}+\sigma_z^2}\omega_2 \\
&\equiv \mu_2 + \lambda_2\omega_2
\end{aligned} \tag{B.12}
$$

We have six equations and six unknowns:

$$
\mu_2 = P_0 - \frac{(\beta_{d2} + \beta_{q2})\sigma_v^2(\alpha_{d2} + \alpha_{q2} + (\beta_{d2} + \beta_{q2})P_0)}{(\beta_{d2} + \beta_{q2})^2\sigma_v^2 + \beta_{d2}^2\sigma_{e_d}^2 + \beta_{q2}^2\sigma_{e_q}^2 + 2\beta_{d2}\beta_{q2}\sigma_{e_d e_q} + \sigma_z^2} \tag{B.13}
$$

$$
\begin{aligned}
&\lambda_2 = \frac{(\beta_{d2} + \beta_{q2})\sigma_v^2}{(\beta_{d2} + \beta_{q2})^2\sigma_v^2 + \beta_{d2}^2\sigma_{e_d}^2 + \beta_{q2}^2\sigma_{e_q}^2 + 2\beta_{d2}\beta_{q2}\sigma_{e_d e_q} + \sigma_z^2} \\
&\left(\text{note that } \sigma_{e_d e_q}/\sigma_v^2 = \rho(\sigma_{e_d}/\sigma_v)(\sigma_{e_q}/\sigma_v) = \rho\sqrt{\frac{(1-\phi_d)(1-\phi_q)}{\phi_d\phi_q}} = \nu_q/\phi_q = \nu_d/\phi_d\right) \\
&\Downarrow \\
&\frac{1}{\lambda_2} = \beta_{d2}+\beta_{q2}+\frac{\beta_{d2}^2}{\beta_{d2}+\beta_{q2}}\frac{1-\phi_d}{\phi_d}+\frac{\beta_{q2}^2}{\beta_{d2}+\beta_{q2}}\frac{1-\phi_q}{\phi_q}+\frac{2\beta_{d2}\beta_{q2}}{\beta_{d2}+\beta_{q2}}\frac{\nu_q}{\phi_q}+\frac{\eta}{\beta_{d2}+\beta_{q2}}
\end{aligned} \tag{B.14}
$$

$$
\alpha_{d2} = \frac{(1 - \lambda_2\beta_{q2})(1 - \phi_d)P_0}{2\lambda_2} - \frac{\mu_2}{2\lambda_2} - \frac{\alpha_{q2}}{2} + \frac{\beta_{q2}\nu_d P_0}{2} \tag{B.15}
$$

$$
\beta_{d2} = \frac{\phi_d}{2\lambda_2} - \frac{\phi_d\beta_{q2}}{2} - \frac{\nu_d\beta_{q2}}{2} \tag{B.16}
$$

$$
\alpha_{q2} = \frac{(1 - \lambda_2\beta_{d2})(1 - \phi_q)P_0}{2\lambda_2} - \frac{\mu_2}{2\lambda_2} - \frac{\alpha_{d2}}{2} + \frac{\beta_{d2}\nu_q P_0}{2} \tag{B.17}
$$

$$
\beta_{q2} = \frac{\phi_q}{2\lambda_2} - \frac{\phi_q\beta_{d2}}{2} - \frac{\nu_q\beta_{d2}}{2} \tag{B.18}
$$

We first utilize the equations B.16 and B.18 to solve for $1/\lambda_2$ and set the results equal to each other. From equation B.16:

$$\begin{aligned} \beta_{d2} &= \frac{\phi_d}{2\lambda_2} - \frac{\phi_d \beta_{q2}}{2} - \frac{\nu_d \beta_{q2}}{2} \\ \Rightarrow \frac{1}{\lambda_2} &= \frac{2\beta_{d2}}{\phi_d} + \beta_{q2} + \frac{\nu_d \beta_{q2}}{\phi_d} \end{aligned} \tag{B.19}$$

Similarly, from equation B.18:

$$\begin{aligned} \beta_{q2} &= \frac{\phi_q}{2\lambda_2} - \frac{\phi_q \beta_{d2}}{2} - \frac{\nu_q \beta_{d2}}{2} \\ \Rightarrow \frac{1}{\lambda_2} &= \frac{2\beta_{q2}}{\phi_q} + \beta_{d2} + \frac{\nu_q \beta_{d2}}{\phi_q} \end{aligned} \tag{B.20}$$

Equating equations B.19 and B.20:

$$\begin{aligned} \frac{2\beta_{d2}}{\phi_d} + \beta_{q2} + \frac{\nu_d \beta_{q2}}{\phi_d} &= \frac{2\beta_{q2}}{\phi_q} + \beta_{d2} + \frac{\nu_q \beta_{d2}}{\phi_q} \\ \Rightarrow \beta_{d2} &= \frac{2\phi_d - \phi_d\phi_q - \nu_d\phi_q}{2\phi_q - \phi_d\phi_q - \nu_q\phi_d}\beta_{q2} \\ &\equiv k\beta_{q2} \end{aligned} \tag{B.21}$$

Substituting β_{d2} from equation B.21 into equation B.14 and into equation B.19 and setting the two equal to each other, we obtain:

$$\beta_{d2}+\beta_{q2}+\frac{\beta_{d2}^2}{\beta_{d2}+\beta_{q2}}\frac{1-\phi_d}{\phi_d}+\frac{\beta_{q2}^2}{\beta_{d2}+\beta_{q2}}\frac{1-\phi_q}{\phi_q}+\frac{2\beta_{d2}\beta_{q2}}{\beta_{d2}+\beta_{q2}}\frac{\nu_q}{\phi_q}+\frac{\eta}{\beta_{d2}+\beta_{q2}}=\frac{2\beta_{d2}}{\phi_d}+\beta_{q2}+\frac{\nu_d\beta_{q2}}{\phi_d}$$

$$\Downarrow$$

$$\frac{\eta}{k+1}\frac{1}{\beta_{q2}}=\frac{2k}{\phi_d}\beta_{q2}+\frac{\nu_d}{\phi_d}\beta_{q2}-k\beta_{q2}-\frac{k^2}{k+1}\frac{1-\phi_d}{\phi_d}\beta_{q2}-\frac{1}{k+1}\frac{1-\phi_q}{\phi_q}\beta_{q2}-\frac{2k}{k+1}\frac{\nu_q}{\phi_q}\beta_{q2}$$

$$\Downarrow$$

$$\frac{\eta}{\beta_{q2}}=\beta_{q2}\left(\frac{2k(k+1)}{\phi_d}+\frac{(k+1)\nu_d}{\phi_d}-k(k+1)-\frac{k^2(1-\phi_d)}{\phi_d}-\frac{1-\phi_q}{\phi_q}-\frac{2k\nu_q}{\phi_q}\right)$$

$$\Downarrow$$

$$\frac{\eta}{\beta_{q2}}=\beta_{q2}\frac{2k(k+1)\phi_q+(k+1)\nu_d\phi_q-k(k+1)\phi_d\phi_q-k^2(1-\phi_d)\phi_q-(1-\phi_q)\phi_d-2k\nu_q\phi_d}{\phi_d\phi_q}$$

$$\Downarrow$$

$$\beta_{q2}=\pm\sqrt{\frac{\eta\phi_d\phi_q}{2k(k+1)\phi_q+(k+1)\nu_d\phi_q-k(k+1)\phi_d\phi_q-k^2(1-\phi_d)\phi_q-(1-\phi_q)\phi_d-2k\nu_q\phi_d}}$$

$$\Downarrow(\text{ note that } \nu_d\phi_q=\nu_q\phi_d) \tag{B.22}$$

$$\beta_{q2}=\pm\sqrt{\frac{\eta\phi_d\phi_q}{\phi_q k^2+(2\phi_q-\nu_d\phi_q-\phi_d\phi_q)k+\nu_d\phi_q-\phi_d+\phi_d\phi_q}}$$

$$\Downarrow$$

$$\beta_{d2}=\pm\sqrt{\frac{\eta\phi_d\phi_q k^2}{\phi_q k^2+(2\phi_q-\nu_d\phi_q-\phi_d\phi_q)k+\nu_d\phi_q-\phi_d+\phi_d\phi_q}}$$

$$\Downarrow$$

$$\frac{1}{\lambda_2}=\frac{2\beta_{d2}}{\phi_d}+\beta_{q2}+\frac{\nu_d\beta_{q2}}{\phi_d}=\beta_{q2}\left(\frac{2k}{\phi_d}+1+\frac{\nu_d}{\phi_d}\right)=\frac{2k+\phi_d+\nu_d}{\phi_d}\beta_{q2}$$

$$\Downarrow$$

$$\lambda_2=\pm\frac{\phi_d}{2k+\phi_d+\nu_d}\sqrt{\frac{\phi_q k^2+(2\phi_q-\nu_d\phi_q-\phi_d\phi_q)k+\nu_d\phi_q-\phi_d+\phi_d\phi_q}{\eta\phi_d\phi_q}}$$

We have solved for β_{q2}, β_{d2}, and λ_2. We now proceed to solve for α_{q2},α_{d2}, and μ_2.

Subtracting equations B.15 and B.17:

$$\begin{aligned}
\alpha_{d2} - \alpha_{q2} &= \frac{(1-\lambda_2\beta_{q2})(1-\phi_d)P_0}{2\lambda_2} - \frac{\mu_2}{2\lambda_2} - \frac{\alpha_{q2}}{2} + \frac{\beta_{q2}\nu_d P_0}{2} \\
&\quad - \frac{(1-\lambda_2\beta_{d2})(1-\phi_q)P_0}{2\lambda_2} + \frac{\mu_2}{2\lambda_2} + \frac{\alpha_{d2}}{2} - \frac{\beta_{d2}\nu_q P_0}{2} \\
&= \frac{(1-\lambda_2\beta_{q2})(1-\phi_d)P_0}{\lambda_2} + \beta_{q2}\nu_d P_0 - \frac{(1-\lambda_2\beta_{d2})(1-\phi_q)P_0}{\lambda_2} - \beta_{d2}\nu_q P_0 \\
&= \frac{1}{\lambda_2}P_0(\phi_q - \phi_d) + \beta_{q2}P_0(\nu_d + \phi_d - 1) - \beta_{d2}P_0(\nu_q + \phi_q - 1) \\
&= \frac{1}{\lambda_2}P_0(\phi_q - \phi_d) + \beta_{q2}P_0(\nu_d + \phi_d - 1 - k\nu_q - k\phi_q + k) \\
&= \beta_{q2}P_0\left(\frac{(2k+\phi_d+\nu_d)(\phi_q-\phi_d)}{\phi_d} + \nu_d + \phi_d - 1 - k\nu_q - k\phi_q + k\right) \\
\Rightarrow \alpha_{d2} &= \alpha_{q2} + \beta_{q2}P_0\left(\frac{(2k+\phi_d+\nu_d)(\phi_q-\phi_d)}{\phi_d} + \nu_d + \phi_d - 1 - k\nu_q - k\phi_q + k\right)
\end{aligned} \tag{B.23}$$

Similarly, adding equations B.15 and B.17:

$$\begin{aligned}
\alpha_{d2}+\alpha_{q2} &= \frac{(1-\lambda_2\beta_{q2})(1-\phi_d)P_0}{2\lambda_2} - \frac{\mu_2}{2\lambda_2} - \frac{\alpha_{q2}}{2} + \frac{\beta_{q2}\nu_d P_0}{2} + \frac{(1-\lambda_2\beta_{d2})(1-\phi_q)P_0}{2\lambda_2} - \frac{\mu_2}{2\lambda_2} - \frac{\alpha_{d2}}{2} + \frac{\beta_{d2}\nu_q P_0}{2} \\
&= \frac{(1-\lambda_2\beta_{q2})(1-\phi_d)P_0}{3\lambda_2} + \frac{(1-\lambda_2\beta_{d2})(1-\phi_q)P_0}{3\lambda_2} + \frac{\beta_{q2}\nu_d P_0}{3} + \frac{\beta_{d2}\nu_q P_0}{3} - \frac{2\mu_2}{3\lambda_2}
\end{aligned} \tag{B.24}$$

From Eq. (B.13) and (B.14), we have:

$$\mu_2 = P_0 - \lambda_2(\alpha_{d2} + \alpha_{q2} + (\beta_{d2} + \beta_{q2})P_0) \tag{B.25}$$

Substituting $\alpha_{d2} + \alpha_{q2}$ from Eq. (B.24) and solving for μ_2:

$$\begin{aligned}
\mu_2 &= P_0 - \lambda_2\left(\frac{(1-\lambda_2\beta_{q2})(1-\phi_d)P_0}{3\lambda_2} + \frac{(1-\lambda_2\beta_{d2})(1-\phi_q)P_0}{3\lambda_2} + \frac{\beta_{q2}\nu_d P_0}{3} + \frac{\beta_{d2}\nu_q P_0}{3} - \frac{2\mu_2}{3\lambda_2} + (k+1)\beta_{q2}P_0\right) \\
3\mu_2 &= 3P_0 - (1-\lambda_2\beta_{q2})(1-\phi_d)P_0 - (1-\lambda_2\beta_{d2})(1-\phi_q)P_0 - \beta_{q2}\lambda_2\nu_d P_0 - \beta_{d2}\lambda_2\nu_q P_0 + 2\mu_2 - 3(k+1)\beta_{q2}\lambda_2 P_0 \\
\mu_2 &= 3P_0 - (1-\lambda_2\beta_{q2})(1-\phi_d)P_0 - (1-\lambda_2\beta_{d2})(1-\phi_q)P_0 - \beta_{q2}\lambda_2\nu_d P_0 - \beta_{d2}\lambda_2\nu_q P_0 - 3(k+1)\beta_{q2}\lambda_2 P_0
\end{aligned} \tag{B.26}$$

We have now solved for μ_2 with previously calculated λ_2,β_{q2},β_{d2}, and can add Eq. (B.23) and (B.24) and divide by 2 to solve for α_{d2}:

$$\begin{aligned}
\alpha_{d2} &= \beta_{q2}P_0\left(\frac{(2k+\phi_d+\nu_d)(\phi_q-\phi_d)}{\phi_d} + \nu_d + \phi_d - 1 - k\nu_q - k\phi_q + k\right) \\
&\quad + \frac{(1-\lambda_2\beta_{q2})(1-\phi_d)P_0}{3\lambda_2} + \frac{(1-\lambda_2\beta_{d2})(1-\phi_q)P_0}{3\lambda_2} + \frac{\beta_{q2}\nu_d P_0}{3} + \frac{\beta_{d2}\nu_q P_0}{3} - \frac{2\mu_2}{3\lambda_2}
\end{aligned} \tag{B.27}$$

Finally, using Eq. (B.23), we obtain:

$$\alpha_{q2} = \alpha_{d2} - \beta_{q2} P_0 \Bigg(\frac{(2k + \phi_d + \nu_d)(\phi_q - \phi_d)}{\phi_d} + \nu_d + \phi_d - 1 - k\nu_q - k\phi_q + k \Bigg) \tag{B.28}$$

B.3.1.1 Economy 2: Solution Summary

To summarize the solution, we first solve for β_{q2},β_{d2},λ_2

$$\beta_{q2} = \pm \sqrt{\frac{\eta \phi_d \phi_q}{\phi_q k^2 + (2\phi_q - \nu_d \phi_q - \phi_d \phi_q)k + \nu_d \phi_q - \phi_d + \phi_d \phi_q}} \tag{B.29}$$

$$\beta_{d2} = k\beta_{q2} \tag{B.30}$$

$$\lambda_2 = \frac{\phi_d}{2k + \phi_d + \nu_d} \beta_{q2}^{-1}, \tag{B.31}$$

where

$$k = \frac{2\phi_d - \phi_d\phi_q - \nu_d\phi_q}{2\phi_q - \phi_d\phi_q - \nu_q\phi_d}$$
$$\nu_q = \rho \sqrt{\frac{\phi_q(1 - \phi_d)(1 - \phi_q)}{\phi_d}}$$
$$\nu_d = \frac{\phi_d}{\phi_q} \nu_q$$

and use the solutions for μ_2,α_{q2},α_{d2}

$$\mu_2 = 3P_0 - (1 - \lambda_2\beta_{q2})(1 - \phi_d)P_0 - (1 - \lambda_2\beta_{d2})(1 - \phi_q)P_0 - \beta_{q2}\lambda_2\nu_d P_0 - \beta_{d2}\lambda_2\nu_q P_0 - 3(k+1)\beta_{q2}\lambda_2 P_0 \tag{B.32}$$

$$\begin{aligned} \alpha_{d2} = {} & \beta_{q2} P_0 \left(\frac{(2k+\phi_d+\nu_d)(\phi_q-\phi_d)}{\phi_d} + \nu_d + \phi_d - 1 - k\nu_q - k\phi_q + k \right) \\ & + \frac{(1-\lambda_2\beta_{q2})(1-\phi_d)P_0}{3\lambda_2} + \frac{(1-\lambda_2\beta_{d2})(1-\phi_q)P_0}{3\lambda_2} + \frac{\beta_{q2}\nu_d P_0}{3} + \frac{\beta_{d2}\nu_q P_0}{3} - \frac{2\mu_2}{3\lambda_2} \end{aligned} \tag{B.33}$$

$$\alpha_{q2} = \alpha_{d2} - \beta_{q2} P_0 \Bigg(\frac{(2k + \phi_d + \nu_d)(\phi_q - \phi_d)}{\phi_d} + \nu_d + \phi_d - 1 - k\nu_q - k\phi_q + k \Bigg) \tag{B.34}$$

B.3.2 Economy 4: Fully-automated Quant, $\gamma = 1$

In this case the QI is fully-automated and applies the trading rule dictated by Equation (2.9) without taking into consideration the behavior of the DI:

$$
\begin{aligned}
x_{q4} &= x_b \\
&= \frac{1}{2}\sqrt{\frac{\eta}{\phi_d}}\left((2-\phi_d)\phi_q - \rho\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}\right)(s_q - P_0) \\
&= -\frac{1}{2}\sqrt{\frac{\eta}{\phi_d}}\left((2-\phi_d)\phi_q - \rho\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}\right)P_0 \\
&+ \frac{1}{2}\sqrt{\frac{\eta}{\phi_d}}\left((2-\phi_d)\phi_q - \rho\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}\right)s_q \\
&\equiv \alpha_b + \beta_b s_q
\end{aligned}
\tag{B.35}
$$

The DI pursues a strategy similar to Economy 2, specified by Eq. (B.11), but with different inputs for the QI's trading rule and the MM's pricing function:

$$
\begin{aligned}
x_{d4} &= \arg\max_{\tilde{x}_{d4}} \mathbb{E}[\tilde{x}_{d4}(v - P_4(\omega_4))|s_d] \\
&= \frac{(1-\lambda_4\beta_b)(1-\phi_d)P_0}{2\lambda_4} - \frac{\mu_4}{2\lambda_4} - \frac{\alpha_b}{2} + \frac{\beta_b\nu_d P_0}{2} + \left(\frac{\phi_d}{2\lambda_4} - \frac{\phi_d\beta_b}{2} - \frac{\nu_d\beta_b}{2}\right)s_d \\
&\equiv \alpha_{d4} + \beta_{d4}s_d
\end{aligned}
\tag{B.36}
$$

Similarly, the MM's pricing rule is comparable to Economy 2, as given by Eq. (B.37), with different inputs for the DI's and QI's strategies:

$$
\begin{aligned}
P_4(\omega_4) &= \mathbb{E}[v|\omega_4] \\
&= \mathbb{E}[v|x_{d4} + x_{q4} + z] \\
&= \mathbb{E}[v|\alpha_{d4} + \beta_{d4}(v+e_d) + \alpha_b + \beta_b(v+e_q) + z] \\
&= P_0 - \frac{(\beta_{d4}+\beta_b)\sigma_v^2(\alpha_{d4}+\alpha_b+(\beta_{d4}+\beta_b)P_0)}{(\beta_{d4}+\beta_b)^2\sigma_v^2+\beta_{d4}^2\sigma_{e_d}^2+\beta_b^2\sigma_{e_q}^2+2\beta_{d4}\beta_b\sigma_{e_de_q}+\sigma_z^2} \\
&+ \frac{(\beta_{d4}+\beta_b)\sigma_v^2}{(\beta_{d4}+\beta_b)^2\sigma_v^2+\beta_{d4}^2\sigma_{e_d}^2+\beta_b^2\sigma_{e_q}^2+2\beta_{d4}\beta_b\sigma_{e_de_q}+\sigma_z^2}\omega_4 \\
&\equiv \mu_4 + \lambda_4\omega_4
\end{aligned}
\tag{B.37}
$$

We have four equations and four unknowns:

$$
\mu_4 = P_0 - \frac{(\beta_{d4}+\beta_b)\sigma_v^2(\alpha_{d4}+\alpha_b+(\beta_{d4}+\beta_b)P_0)}{(\beta_{d4}+\beta_b)^2\sigma_v^2+\beta_{d4}^2\sigma_{e_d}^2+\beta_b^2\sigma_{e_q}^2+2\beta_{d4}\beta_b\sigma_{e_de_q}+\sigma_z^2}
\tag{B.38}
$$

$$\lambda_4 = \frac{(\beta_{d4}+\beta_b)\sigma_v^2}{(\beta_{d4}+\beta_b)^2\sigma_v^2+\beta_{d4}^2\sigma_{e_d}^2+\beta_b^2\sigma_{e_q}^2+2\beta_{d4}\beta_b\sigma_{e_de_q}+\sigma_z^2}$$
$$\left(\text{note that } \sigma_{e_de_q}/\sigma_v^2=\rho(\sigma_{e_d}/\sigma_v)(\sigma_{e_q}/\sigma_v)=\rho\sqrt{\frac{(1-\phi_d)(1-\phi_q)}{\phi_d\phi_q}}=\nu_q/\phi_q=\nu_d/\phi_d\right) \tag{B.39}$$
$$\Downarrow$$
$$\frac{1}{\lambda_4}=\beta_{d4}+\beta_b+\frac{\beta_{d4}^2}{\beta_{d4}+\beta_b}\frac{1-\phi_d}{\phi_d}+\frac{\beta_b^2}{\beta_{d4}+\beta_b}\frac{1-\phi_q}{\phi_q}+\frac{2\beta_{d4}\beta_b}{\beta_{d4}+\beta_b}\frac{\nu_q}{\phi_q}+\frac{\eta}{\beta_{d4}+\beta_b}$$

$$\alpha_{d4} = \frac{(1-\lambda_4\beta_b)(1-\phi_d)P_0}{2\lambda_4} - \frac{\mu_4}{2\lambda_4} - \frac{\alpha_b}{2} + \frac{\beta_b\nu_dP_0}{2} \tag{B.40}$$

$$\beta_{d4} = \frac{\phi_d}{2\lambda_4} - \frac{\phi_d\beta_b}{2} - \frac{\nu_d\beta_b}{2} \tag{B.41}$$

From Eq. (B.61)

$$\begin{aligned}\beta_{d4} &= \frac{\phi_d}{2\lambda_4} - \frac{\phi_d\beta_b}{2} - \frac{\nu_d\beta_b}{2}\\ \Rightarrow \frac{1}{\lambda_4} &= \frac{2}{\phi_d}\beta_{d4}+\beta_b+\frac{\nu_d}{\phi_d}\beta_b\end{aligned} \tag{B.42}$$

Setting Eq. (B.42) and (B.59) equal:

$$\begin{aligned}&\beta_{d4}+\beta_b+\frac{\beta_{d4}^2}{\beta_{d4}+\beta_b}\frac{1-\phi_d}{\phi_d}+\frac{\beta_b^2}{\beta_{d4}+\beta_b}\frac{1-\phi_q}{\phi_q}+\frac{2\beta_{d4}\beta_b}{\beta_{d4}+\beta_b}\frac{\nu_q}{\phi_q}+\frac{\eta}{\beta_{d4}+\beta_b}=\frac{2}{\phi_d}\beta_{d4}+\beta_b+\frac{\nu_d}{\phi_d}\beta_b\\
&\Downarrow\\
&\beta_{d4}(\beta_{d4}+\beta_b)+\frac{1-\phi_d}{\phi_d}\beta_{d4}^2+\frac{1-\phi_q}{\phi_q}\beta_b^2+\frac{2\beta_b\nu_q}{\phi_q}\beta_{d4}+\eta-(\frac{2}{\phi_d}\beta_{d4}+\frac{\nu_d}{\phi_d}\beta_b)(\beta_{d4}+\beta_b)=0\\
&\Downarrow\\
&\left(-\frac{1}{\phi_d}\right)\beta_{d4}^2+\left(\beta_b+\frac{2\beta_b\nu_q}{\phi_q}-\frac{2\beta_b}{\phi_d}-\frac{\nu_d\beta_b}{\phi_d}\right)\beta_{d4}+\left(\frac{1-\phi_q}{\phi_q}\beta_b^2+\eta-\frac{\nu_d\beta_b^2}{\phi_d}\right)=0\\
&\Downarrow\left(\nu_q/\phi_q=\nu_d/\phi_d\right)\\
&\left(-\frac{1}{\phi_d}\right)\beta_{d4}^2+\beta_b\left(1+\frac{\nu_q}{\phi_q}-\frac{2}{\phi_d}\right)\beta_{d4}+\left(\frac{1-\phi_q}{\phi_q}\beta_b^2+\eta-\frac{\nu_d\beta_b^2}{\phi_d}\right)=0\\
&\Downarrow\\
&\beta_{d4}=\beta_b\left(\frac{\phi_d}{2}+\frac{\nu_d}{2}-1\right)\pm\sqrt{\beta_b^2\left(\frac{\phi_d^2}{4}+\frac{\nu_d^2}{4}+1+\frac{\phi_d\nu_d}{2}-2\phi_d-2\nu_d+\frac{\phi_d}{\phi_q}\right)+\phi_d\eta}\end{aligned} \tag{B.43}$$

We can now substitute β_{d3} from above to solve for market depth:

$$\begin{aligned}
\frac{1}{\lambda_4} &= \frac{2}{\phi_d}\beta_{d4} + \beta_b + \frac{\nu_d}{\phi_d}\beta_b \\
&= \beta_b\Big(2 + \frac{2\nu_d}{\phi_d} - \frac{2}{\phi_d}\Big) \pm \frac{2}{\phi_d}\sqrt{\beta_b^2\Big(\frac{\phi_d^2}{4} + \frac{\nu_d^2}{4} + 1 + \frac{\phi_d\nu_d}{2} - 2\phi_d - 2\nu_d + \frac{\phi_d}{\phi_q}\Big) + \phi_d\eta} \\
&= \beta_b\Big(2 + \frac{2\nu_d}{\phi_d} - \frac{2}{\phi_d}\Big) \pm \sqrt{\beta_b^2\Big(1 + \frac{\nu_d^2}{\phi_d^2} + \frac{4}{\phi_d^2} + \frac{2\nu_d}{\phi_d} - \frac{8}{\phi_d} - \frac{8\nu_d}{\phi_d^2} + \frac{4}{\phi_d\phi_q}\Big) + \frac{4\eta}{\phi_d}}
\end{aligned} \tag{B.44}$$

From Eq. (B.58) and (B.59), we have:

$$\mu_4 = P_0 - \lambda_4(\alpha_{d4} + \alpha_b + (\beta_{d4} + \beta_b)P_0) \tag{B.45}$$

Substituting Eq. (B.60) into the above

$$\mu_4 = P_0 - \lambda_4\Bigg(\frac{(1-\lambda_4\beta_b)(1-\phi_d)P_0}{2\lambda_4} - \frac{\mu_4}{2\lambda_4} - \frac{\alpha_b}{2} + \frac{\beta_b\nu_d P_0}{2} + \alpha_b + (\beta_{d4} + \beta_b)P_0\Bigg) \tag{B.46}$$

$$= 2P_0 - (1-\lambda_4\beta_b)(1-\phi_d)P_0 - \lambda_4\alpha_b - \lambda_4\beta_b\nu_d P_0 - 2\lambda_4(\beta_{d4} + \beta_b)P_0 \tag{B.47}$$

Finally, we have all of the inputs to solve for α_{d3} given by Eq. (B.60)

$$\alpha_{d4} = \frac{(1-\lambda_4\beta_b)(1-\phi_d)P_0}{2\lambda_4} - \frac{\mu_4}{2\lambda_4} - \frac{\alpha_b}{2} + \frac{\beta_b\nu_d P_0}{2} \tag{B.48}$$

B.3.2.1 Economy 4: Solution Summary

To summarize, we first identify the QI's trading rule:

$$\alpha_b = -\frac{1}{2}\sqrt{\frac{\eta}{\phi_d}}\left((2-\phi_d)\phi_q - \rho\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}\right)P_0 \tag{B.49}$$

$$\beta_b = \frac{1}{2}\sqrt{\frac{\eta}{\phi_d}}\left((2-\phi_d)\phi_q - \rho\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}\right) \tag{B.50}$$

We then use this as an input into the DI's trading intensity

$$\beta_{d4} = \beta_b\Big(\frac{\phi_d}{2} + \frac{\nu_d}{2} - 1\Big) \pm \sqrt{\beta_b^2\Big(\frac{\phi_d^2}{4} + \frac{\nu_d^2}{4} + 1 + \frac{\phi_d\nu_d}{2} - 2\phi_d - 2\nu_d + \frac{\phi_d}{\phi_q}\Big) + \phi_d\eta}, \tag{B.51}$$

which feeds into the equilibrium market depth:

$$\frac{1}{\lambda_4} = \beta_b\Big(2 + \frac{2\nu_d}{\phi_d} - \frac{2}{\phi_d}\Big) \pm \sqrt{\beta_b^2\Big(1 + \frac{\nu_d^2}{\phi_d^2} + \frac{4}{\phi_d^2} + \frac{2\nu_d}{\phi_d} - \frac{8}{\phi_d} - \frac{8\nu_d}{\phi_d^2} + \frac{4}{\phi_d\phi_q}\Big) + \frac{4\eta}{\phi_d}}$$
$$\Downarrow$$
$$\lambda_4 = \left(\beta_b\Big(2 + \frac{2\nu_d}{\phi_d} - \frac{2}{\phi_d}\Big) \pm \sqrt{\beta_b^2\Big(\phi_d^2 + \frac{\nu_d^2}{\phi_d^2} + \frac{4}{\phi_d^2} + \frac{2\nu_d}{\phi_d} - \frac{8}{\phi_d} - \frac{8\nu_d}{\phi_d^2} + \frac{4}{\phi_d\phi_q}\Big) + \frac{4\eta}{\phi_d}}\right)^{-1}. \tag{B.52}$$

We can now solve for

$$\mu_4 = 2P_0 - (1 - \lambda_4\beta_b)(1 - \phi_d)P_0 - \lambda_4\alpha_b - \lambda_4\beta_b\nu_d P_0 - 2\lambda_4(\beta_{d4} + \beta_b)P_0, \tag{B.53}$$

which gives us

$$\alpha_{d4} = \frac{(1 - \lambda_4\beta_b)(1 - \phi_d)P_0}{2\lambda_4} - \frac{\mu_4}{2\lambda_4} - \frac{\alpha_b}{2} + \frac{\beta_b\nu_d P_0}{2}. \tag{B.54}$$

B.3.3 Economy 3: Partially-automated Quant, $0 < \gamma < 1$

The DI's optimal demands are formulaically identical to Economy 2:

$$\begin{aligned} x_{d3} &= \arg\max_{\tilde{x}_{d3}} \mathbb{E}[\tilde{x}_{d3}(v - P_3(\omega_3))|s_d] \\ &= \frac{(1 - \lambda_3\beta_{q3})(1 - \phi_d)P_0}{2\lambda_3} - \frac{\mu_3}{2\lambda_3} - \frac{\alpha_{q3}}{2} + \frac{\beta_{q3}\nu_d P_0}{2} + \left(\frac{\phi_d}{2\lambda_3} - \frac{\phi_d\beta_{q3}}{2} - \frac{\nu_d\beta_{q3}}{2}\right)s_d \\ &\equiv \alpha_{d3} + \beta_{d3}s_d \end{aligned} \tag{B.55}$$

The QI now solves the following objective function for this optimal demand x_{q3} where x_b is given by Eq. (2.9):

$$\begin{aligned}
x_{q3} &= \arg\min_{\tilde{x}_{q3}} \mathbb{E}[\gamma(\tilde{x}_{q3}-x_b)^2+(1-\gamma)\tilde{x}_{q3}(P_3(\omega_3)-v)|s_q] \\
&= \arg\min_{\tilde{x}_{q3}} \mathbb{E}[\gamma(\tilde{x}_{q3}-x_b)^2+(1-\gamma)\tilde{x}_{q3}(\mu_3+\lambda_3(x_{d3}+\tilde{x}_{q3}+z)-v)|s_q] \\
&= \arg\min_{\tilde{x}_{q3}} \mathbb{E}[\gamma\tilde{x}_{q3}^2-2\gamma\tilde{x}_{q3}x_b+\gamma x_b^2+(1-\gamma)\tilde{x}_{q3}(\mu_3+\lambda_3 x_{d3}+\lambda_3 z-v)+(1-\gamma)\lambda_3\tilde{x}_{q3}^2|s_q] \\
&= \arg\min_{\tilde{x}_{q3}} \mathbb{E}[(\gamma+(1-\gamma)\lambda_3)\tilde{x}_{q3}^2-2\gamma\tilde{x}_{q3}x_b+\gamma x_b^2+(1-\gamma)\tilde{x}_{q3}(\mu_3+\lambda_3 x_{d3}+\lambda_3 z-v)|s_q] \\
&= \arg\min_{\tilde{x}_{q3}} \mathbb{E}\left[\Big(\gamma+(1-\gamma)\lambda_3\Big)\tilde{x}_{q3}^2+\Big((1-\gamma)(\mu_3+\lambda_3 x_{d3}+\lambda_3 z-v)-2\gamma x_b\Big)\tilde{x}_{q3}+\gamma x_b^2|s_q\right] \\
&= \arg\min_{\tilde{x}_{q3}} \mathbb{E}\left\{\Big(\gamma+(1-\gamma)\lambda_3\Big)\tilde{x}_{q3}^2+\Big((1-\gamma)(\mu_3+\lambda_3\mathbb{E}[x_{d3}|s_q]-\mathbb{E}[v|s_q])-2\gamma x_b\Big)\tilde{x}_{q3}+\gamma x_b^2|s_q\right] \\
&= \frac{2\gamma x_b+(1-\gamma)(\mathbb{E}[v|s_q]-\mu_3-\lambda_3\mathbb{E}[x_{d3}|s_q])}{2\Big(\gamma+(1-\gamma)\lambda_3\Big)} \\
&= \frac{\gamma}{\gamma+(1-\gamma)\lambda_3}x_b+\frac{(1-\gamma)\lambda_3}{\gamma+(1-\gamma)\lambda_3}\frac{\mathbb{E}[v|s_q]-\mu_3-\lambda_3\mathbb{E}[x_{d3}|s_q]}{2\lambda_3} \\
&= \frac{\gamma}{\gamma+(1-\gamma)\lambda_3}(\alpha_b+\beta_b s_q) \\
&+\frac{(1-\gamma)\lambda_3}{\gamma+(1-\gamma)\lambda_3}\left(\frac{(1-\lambda_3\beta_{d3})(1-\phi_q)P_0}{2\lambda_3}-\frac{\mu_3}{2\lambda_3}-\frac{\alpha_{d3}}{2}+\frac{\beta_{d3}\nu_q P_0}{2}+\left(\frac{\phi_q}{2\lambda_3}-\frac{\phi_q\beta_{d3}}{2}-\frac{\nu_q\beta_{d3}}{2}\right)s_q\right) \\
&= \frac{2\gamma\alpha_b+(1-\gamma)\Big((1-\lambda_3\beta_{d3})(1-\phi_q)P_0-\mu_3\Big)-(1-\gamma)\lambda_3\alpha_{d3}+(1-\gamma)\lambda_3\beta_{d3}\nu_q P_0}{2(\gamma+(1-\gamma)\lambda_3)} \\
&+\frac{2\gamma\beta_b+(1-\gamma)\phi_q-(1-\gamma)\lambda_3(\phi_q+\nu_q)\beta_{d3}}{2(\gamma+(1-\gamma)\lambda_3)}s_q \\
&\equiv \alpha_{q3}+\beta_{q3}
\end{aligned} \tag{B.56}$$

The market maker also has a formulaically identical pricing rule to Economy 2:

$$\begin{aligned}
P_3(\omega_3) &= \mathbb{E}[v|\omega_3] \\
&= P_0-\frac{(\beta_{d3}+\beta_{q3})\sigma_v^2(\alpha_{d3}+\alpha_{q3}+(\beta_{d3}+\beta_{q3})P_0)}{(\beta_{d3}+\beta_{q3})^2\sigma_v^2+\beta_{d3}^2\sigma_{e_d}^2+\beta_{q3}^2\sigma_{e_q}^2+2\beta_{d3}\beta_{q3}\sigma_{e_d e_q}+\sigma_z^2}+\frac{(\beta_{d3}+\beta_{q3})\sigma_v^2}{(\beta_{d3}+\beta_{q3})^2\sigma_v^2+\beta_{d3}^2\sigma_{e_d}^2+\beta_{q3}^2\sigma_{e_q}^2+2\beta_{d3}\beta_{q3}\sigma_{e_d e_q}+\sigma_z^2}\omega_3 \\
&\equiv \mu_3+\lambda_3\omega_3
\end{aligned} \tag{B.57}$$

We have six equations and six unknowns:

$$\begin{aligned}
\mu_3 &= P_0-\frac{(\beta_{d3}+\beta_{q3})\sigma_v^2(\alpha_{d3}+\alpha_{q3}+(\beta_{d3}+\beta_{q3})P_0)}{(\beta_{d3}+\beta_{q3})^2\sigma_v^2+\beta_{d3}^2\sigma_{e_d}^2+\beta_{q3}^2\sigma_{e_q}^2+2\beta_{d3}\beta_{q3}\sigma_{e_d e_q}+\sigma_z^2} \\
&\Downarrow \\
\mu_3 &= P_0-\lambda_3(\alpha_{d3}+\alpha_{q3}+(\beta_{d3}+\beta_{q3})P_0)
\end{aligned} \tag{B.58}$$

$$
\begin{aligned}
\lambda_3 &= \frac{(\beta_{d3}+\beta_{q3})\sigma_v^2}{(\beta_{d3}+\beta_{q3})^2\sigma_v^2+\beta_{d3}^2\sigma_{e_d}^2+\beta_{q3}^2\sigma_{e_q}^2+2\beta_{d3}\beta_{q3}\sigma_{e_d e_q}+\sigma_z^2} \\
&\left(\text{note that } \sigma_{e_d e_q}/\sigma_v^2=\rho(\sigma_{e_d}/\sigma_v)(\sigma_{e_q}/\sigma_v)=\rho\sqrt{\tfrac{(1-\phi_d)(1-\phi_q)}{\phi_d\phi_q}}=\nu_q/\phi_q=\nu_d/\phi_d\right) \\
&\Downarrow \\
\frac{1}{\lambda_3} &= \beta_{d3}+\beta_{q3}+\frac{\beta_{d3}^2}{\beta_{d3}+\beta_{q3}}\frac{1-\phi_d}{\phi_d}+\frac{\beta_{q3}^2}{\beta_{d3}+\beta_{q3}}\frac{1-\phi_q}{\phi_q}+\frac{2\beta_{d3}\beta_{q3}}{\beta_{d3}+\beta_{q3}}\frac{\nu_q}{\phi_q}+\frac{\eta}{\beta_{d3}+\beta_{q3}}
\end{aligned}
\tag{B.59}
$$

$$
\alpha_{d3} = \frac{(1-\lambda_3\beta_{q3})(1-\phi_d)P_0}{2\lambda_3} - \frac{\mu_3}{2\lambda_3} - \frac{\alpha_{q3}}{2} + \frac{\beta_{q3}\nu_d P_0}{2} \tag{B.60}
$$

$$
\beta_{d3} = \frac{\phi_d}{2\lambda_3} - \frac{\phi_d\beta_{q3}}{2} - \frac{\nu_d\beta_{q3}}{2} \tag{B.61}
$$

$$
\alpha_{q3} = \frac{2\gamma\alpha_b+(1-\gamma)\Big((1-\lambda_3\beta_{d3})(1-\phi_q)P_0-\mu_3\Big)-(1-\gamma)\lambda_3\alpha_{d3}+(1-\gamma)\lambda_3\beta_{d3}\nu_q P_0}{2(\gamma+(1-\gamma)\lambda_3)} \tag{B.62}
$$

$$
\beta_{q3} = \frac{2\gamma\beta_b+(1-\gamma)\phi_q-(1-\gamma)\lambda_3(\phi_q+\nu_q)\beta_{d3}}{2(\gamma+(1-\gamma)\lambda_3)} \tag{B.63}
$$

We first use Eq. (B.61) and (B.63) to solve for λ_3 and set the results equal:

$$
\begin{aligned}
\beta_{d3} &= \frac{\phi_d}{2\lambda_3} - \frac{\phi_d\beta_{q3}}{2} - \frac{\nu_d\beta_{q3}}{2} \\
\Rightarrow \frac{1}{\lambda_3} &= \frac{2\beta_{d3}}{\phi_d} + \beta_{q3} + \frac{\nu_d\beta_{q3}}{\phi_d} \\
\beta_{q3} &= \frac{2\gamma\beta_b+(1-\gamma)\phi_q-(1-\gamma)\lambda_3(\phi_q+\nu_q)\beta_{d3}}{2(\gamma+(1-\gamma)\lambda_3)} \\
\Rightarrow \frac{1}{\lambda_3} &= \frac{2(1-\gamma)\beta_{q3}+(1-\gamma)(\phi_q+\nu_q)\beta_{d3}}{2\gamma\beta_b-2\gamma\beta_{q3}+(1-\gamma)\phi_q} \\
\Rightarrow \frac{2}{\phi_d}\beta_{d3}+\Big(1+\frac{\nu_d}{\phi_d}\Big)\beta_{q3} &= \frac{2(1-\gamma)\beta_{q3}+(1-\gamma)(\phi_q+\nu_q)\beta_{d3}}{2\gamma\beta_b-2\gamma\beta_{q3}+(1-\gamma)\phi_q} \\
\Rightarrow \beta_{d3} &= \frac{2\gamma\left(1+\frac{\nu_d}{\phi_d}\right)\beta_{q3}^2+\left(2(1-\gamma)-2\gamma\beta_b\left(1+\frac{\nu_d}{\phi_d}\right)-(1-\gamma)\phi_q\left(1+\frac{\nu_d}{\phi_d}\right)\right)\beta_{q3}}{-\frac{4\gamma}{\phi_d}\beta_{q3}+\frac{4\gamma\beta_b}{\phi_d}+\frac{2(1-\gamma)\phi_q}{\phi_d}-(1-\gamma)(\phi_q+\nu_q)} \\
&\equiv \frac{a\beta_{q3}^2+b\beta_{q3}}{c\beta_{q3}+d}
\end{aligned}
\tag{B.64}
$$

Next, we use Eq. (B.61) to solve for $1/\lambda_3$ and set the result equal to Eq. (B.59):

$$\begin{aligned}
&\beta_{d3} = \frac{\phi_d}{2\lambda_3} - \frac{\phi_d \beta_{q3}}{2} - \frac{\nu_d \beta_{q3}}{2} \\
&\Rightarrow \frac{1}{\lambda_3} = \frac{2\beta_{d3}}{\phi_d} + \beta_{q3} + \frac{\nu_d \beta_{q3}}{\phi_d} \\
&\Rightarrow \beta_{d3}+\beta_{q3}+\frac{\beta_{d3}^2}{\beta_{d3}+\beta_{q3}}\frac{1-\phi_d}{\phi_d}+\frac{\beta_{q3}^2}{\beta_{d3}+\beta_{q3}}\frac{1-\phi_q}{\phi_q}+\frac{2\beta_{d3}\beta_{q3}}{\beta_{d3}+\beta_{q3}}\frac{\nu_q}{\phi_q}+\frac{\eta}{\beta_{d3}+\beta_{q3}}=\frac{2\beta_{d3}}{\phi_d}+\beta_{q3}+\frac{\nu_d\beta_{q3}}{\phi_d} \\
&\Rightarrow \beta_{d3}^2 + \underbrace{\left(2 - \phi_d - \nu_d\right)}_{\equiv e} \beta_{q3}\beta_{d3} + \underbrace{\left(\nu_d + \phi_d - \frac{\phi_d}{\phi_q}\right)}_{\equiv f} \beta_{q3}^2 - \underbrace{\phi_d \eta}_{\equiv g} = 0 \\
&\Rightarrow (a\beta_{q3}^2 + b\beta_{q3})^2 + e\beta_{q3}(c\beta_{q3}+d)(a\beta_{q3}^2 + b\beta_{q3}) + f\beta_{q3}^2(c\beta_{q3}+d)^2 - g(c\beta_{q3}+d)^2 = 0 \\
&\Rightarrow (a^2+ace+c^2f)\beta_{q3}^4+(2ab+bce+ade+2cdf)\beta_{q3}^3+(b^2+bde+d^2f-c^2g)\beta_{q3}^2-2cdg\beta_{q3}-d^2g=0
\end{aligned} \tag{B.65}$$

This is quartic, which has a closed-form solution for β_{q3} resulting in a maximum of four roots. For each root, we can then utilize Eq. (C.20) to solve for β_{d3} and, in-turn, use Eq. B.59 to solve for λ. We have solved for β_{d3}, β_{q3}, and λ_3.

We proceed to solve for α_{d3}, α_{q3}, and μ_3. This is a simple three variable linear system of first degree. First we relabel Eq. (B.58), (B.60), and (B.62).

$$\begin{aligned}
\alpha_{d3} &= \frac{(1-\lambda_3\beta_{q3})(1-\phi_d)P_0}{2\lambda_3} - \frac{\mu_3}{2\lambda_3} - \frac{\alpha_{q3}}{2} + \frac{\beta_{q3}\nu_d P_0}{2} \\
&= \underbrace{\frac{(1-\lambda_3\beta_{q3})(1-\phi_d)P_0 + \beta_{q3}\lambda_3\nu_d P_0}{2\lambda_3}}_{\equiv h} - \underbrace{\frac{1}{2\lambda_3}}_{\equiv j}\mu_3 - \frac{\alpha_{q3}}{2} \\
&= h - j\mu_3 - \frac{\alpha_{q3}}{2}
\end{aligned} \tag{B.66}$$

$$\begin{aligned}
\alpha_{q3} &= \frac{2\gamma\alpha_b + (1-\gamma)\big((1-\lambda_3\beta_{d3})(1-\phi_q)P_0 - \mu_3\big) - (1-\gamma)\lambda_3\alpha_{d3} + (1-\gamma)\lambda_3\beta_{d3}\nu_q P_0}{2(\gamma+(1-\gamma)\lambda_3)} \\
&= \underbrace{\frac{2\gamma\alpha_b + (1-\gamma)(1-\lambda_3\beta_{d3})(1-\phi_q)P_0 + (1-\gamma)\lambda_3\beta_{d3}\nu_q P_0}{2(\gamma+(1-\gamma)\lambda_3)}}_{\equiv k} \\
&\quad - \underbrace{\frac{1-\gamma}{2(\gamma+(1-\gamma)\lambda_3)}}_{\equiv l}\mu_3 - \underbrace{\frac{(1-\gamma)\lambda_3}{2(\gamma+(1-\gamma)\lambda_3)}}_{m}\alpha_{d3} \\
&= k - l\mu_3 - m\alpha_{d3}
\end{aligned} \tag{B.67}$$

$$
\begin{aligned}
\mu_3 &= P_0 - \lambda_3(\alpha_{d3} + \alpha_{q3} + (\beta_{d3} + \beta_{q3})P_0) \\
&= \underbrace{P_0 - \lambda_3(\beta_{d3} + \beta_{q3})P_0}_{\equiv q} - \lambda_3\alpha_{d3} - \lambda_3\alpha_{q3} \\
&= q - \lambda_3\alpha_{d3} - \lambda_3\alpha_{q3}
\end{aligned} \tag{B.68}
$$

Solving the system of equations, we first obtain:

$$
\alpha_{q3} = \frac{(k - lq)(1 - j\lambda_3) + l\lambda_3(h - jq) - m(h - jq)}{(1 - l\lambda_3)(1 - j\lambda_3) - l\lambda_3(j\lambda_3 - \frac{1}{2}) + m(j\lambda_3 - \frac{1}{2})}. \tag{B.69}
$$

We use the solution above to solve for:

$$
\alpha_{d3} = \frac{h - jq + (j\lambda_3 - \frac{1}{2})\alpha_{q3}}{1 - j\lambda_3}, \tag{B.70}
$$

and, finally, for

$$
\mu_3 = q - \lambda_3\alpha_{d3} - \lambda_3\alpha_{q3}. \tag{B.71}
$$

B.3.3.1 Economy 3: Solution Summary

We first obtain β_{q3} from the solution to

$$
(a^2+ace+c^2f)\beta_{q3}^4+(2ab+bce+ade+2cdf)\beta_{q3}^3+(b^2+bde+d^2f-c^2g)\beta_{q3}^2-2cdg\beta_{q3}-d^2g=0
$$

where

$$
\begin{aligned}
a &= 2\gamma\big(1 + \frac{\nu_d}{\phi_d}\big) \\
b &= 2(1-\gamma) - 2\gamma\beta_b\big(1 + \frac{\nu_d}{\phi_d}\big) - (1-\gamma)\phi_q\big(1 + \frac{\nu_d}{\phi_d}\big) \\
c &= -\frac{4\gamma}{\phi_d} \\
d &= \frac{4\gamma\beta_b}{\phi_d} + \frac{2(1-\gamma)\phi_q}{\phi_d} - (1-\gamma)(\phi_q + \nu_q) \\
e &= 2 - \phi_d - \nu_d \\
f &= \nu_d + \phi_d - \frac{\phi_d}{\phi_q} \\
g &= \phi_d\eta
\end{aligned}
$$

and use the solution to obtain

$$\beta_{d3} = \frac{a\beta_{q3}^2 + b\beta_{q3}}{c\beta_{q3} + d}$$

and

$$\lambda_3 = \frac{\beta_{d3} + \beta_{q3}}{(\beta_{d3} + \beta_{q3})^2 + \beta_{d3}^2 \frac{1-\phi_d}{\phi_d} + \beta_{q3}^2 \frac{1-\phi_q}{\phi_q} + 2\beta_{d3}\beta_{q3}\frac{\nu_q}{\phi_q} + \eta}$$

Next we use the solutions above to obtain

$$\alpha_{q3} = \frac{(k - lq)(1 - j\lambda_3) + l\lambda_3(h - jq) - m(h - jq)}{(1 - l\lambda_3)(1 - j\lambda_3) - l\lambda_3(j\lambda_3 - \frac{1}{2}) + m(j\lambda_3 - \frac{1}{2})}$$

where

$$h = \frac{(1 - \lambda_3\beta_{q3})(1 - \phi_d)P_0 + \beta_{q3}\lambda_3\nu_d P_0}{2\lambda_3}$$

$$j = \frac{1}{2\lambda_3}$$

$$k = \frac{2\gamma\alpha_b + (1 - \gamma)(1 - \lambda_3\beta_{d3})(1 - \phi_q)P_0 + (1 - \gamma)\lambda_3\beta_{d3}\nu_q P_0}{2(\gamma + (1 - \gamma)\lambda_3)}$$

$$l = \frac{1 - \gamma}{2(\gamma + (1 - \gamma)\lambda_3)}$$

$$m = \frac{(1 - \gamma)\lambda_3}{2(\gamma + (1 - \gamma)\lambda_3)}$$

$$q = P_0 - \lambda_3(\beta_{d3} + \beta_{q3})P_0$$

and use the solution to find

$$\alpha_{d3} = \frac{h - jq + (j\lambda_3 - \frac{1}{2})\alpha_{q3}}{1 - j\lambda_3},$$

and

$$\mu_3 = q - \lambda_3\alpha_{d3} - \lambda_3\alpha_{q3}.$$

B.4 Derivations for Measures of Market Quality

B.4.1 Market Depth

We take the inverse of λ_j as derived for Economies 1-4 in Appendices B.1 and B.3.

B.4.2 Price Informativeness

Economy 1

$$\begin{aligned}
P_1(\omega_1) &= \mu_1 + \lambda_1\omega_1 \\
&= \mu_1 + \lambda_1(x_{d1} + z) \\
&= \mu_1 + \lambda_1(\alpha_{d1} + \beta_{d1}(v + e_d) + z) \\
&= \mu_1 + \lambda_1(\alpha_{d1} + \beta_{d1}v + \beta_{d1}e_d + z) \\
&= \mu_1 + \lambda_1\alpha_{d1} + \lambda_1\beta_{d1}v + \lambda_1\beta_{d1}e_d + \lambda_1 z
\end{aligned} \tag{B.72}$$

Therefore,

$$\begin{aligned}
\mathrm{Var}(v|P_1) &= \mathrm{Var}(v) - \mathrm{Cov}(v, P_1)\mathrm{Var}(P_1)^{-1}\mathrm{Cov}(P_1, v) \\
&= \sigma_v^2 - \frac{\mathrm{Cov}(v, \mu_1 + \lambda_1\alpha_{d1} + \lambda_1\beta_{d1}v + \lambda_1\beta_{d1}e_d + \lambda_1 z)^2}{\mathrm{Var}(\mu_1 + \lambda_1\alpha_{d1} + \lambda_1\beta_{d1}v + \lambda_1\beta_{d1}e_d + \lambda_1 z)} \\
&= \sigma_v^2 - \frac{\lambda_1^2\beta_{d1}^2\sigma_v^4}{\lambda_1^2\beta_{d1}^2\sigma_v^2 + \lambda_1^2\beta_{d1}^2\sigma_{e_d}^2 + \lambda_1^2\sigma_z^2} \\
&= \sigma_v^2 \frac{\beta_{d1}^2\sigma_{e_d}^2 + \sigma_z^2}{\beta_{d1}^2\sigma_v^2 + \beta_{d1}^2\sigma_{e_d}^2 + \sigma_z^2} \\
&\Downarrow \\
\mathrm{Var}(v|P_1)^{-1} &= \frac{1}{\sigma_v^2}\frac{\beta_{d1}^2 + \beta_{d1}^2\frac{1-\phi_d}{\phi_d} + \eta}{\beta_{d1}^2\frac{1-\phi_d}{\phi_d} + \eta}
\end{aligned}$$

Economies 2-4 For $j \in \{2, 3, 4\}$

$$\begin{aligned}
P_j(\omega_j) &= \mu_j + \lambda_j\omega_j \\
&= \mu_j + \lambda_j(x_{dj} + x_{qj} + z) \\
&= \mu_j + \lambda_j(\alpha_{dj} + \beta_{dj}s_d + \alpha_{qj} + \beta_{qj}s_q + z) \\
&= \mu_j + \lambda_j(\alpha_{dj} + \beta_{dj}(v + e_d) + \alpha_{qj} + \beta_{qj}(v + e_q) + z) \\
&= \mu_j + \lambda_j\alpha_{dj} + \lambda_j\alpha_{qj} + \lambda_j(\beta_{dj} + \beta_{qj})v + \lambda_j\beta_{dj}e_d + \lambda_j\beta_{qj}e_q + \lambda_j z
\end{aligned} \tag{B.73}$$

Therefore,

$$\begin{aligned}
&\mathrm{Var}(v|P_j) = \mathrm{Var}(v) - \mathrm{Cov}(v,P_j)\mathrm{Var}(P_j)^{-1}\mathrm{Cov}(P_j,v)\\
&= \sigma_v^2 - \frac{\mathrm{Cov}(v,\mu_j + \lambda_j\alpha_{dj} + \lambda_j\alpha_{qj} + \lambda_j(\beta_{dj}+\beta_{qj})v + \lambda_j\beta_{dj}e_d + \lambda_j\beta_{qj}e_q + \lambda_j z)^2}{\mathrm{Var}(\mu_j + \lambda_j\alpha_{dj} + \lambda_j\alpha_{qj} + \lambda_j(\beta_{dj}+\beta_{qj})v + \lambda_j\beta_{dj}e_d + \lambda_j\beta_{qj}e_q + \lambda_j z)}\\
&= \sigma_v^2 - \frac{\lambda_j^2(\beta_{dj}+\beta_{qj})^2\sigma_v^4}{\lambda_j^2(\beta_{dj}+\beta_{qj})^2\sigma_v^2 + \lambda_j^2\beta_{dj}^2\sigma_{e_d}^2 + \lambda_j^2\beta_{qj}^2\sigma_{e_q}^2 + 2\lambda_j^2\beta_{dj}\beta_{qj}\sigma_{e_de_q} + \lambda_j^2\sigma_z^2}\\
&= \frac{\sigma_v^2(\lambda_j^2\beta_{dj}^2\sigma_{e_d}^2 + \lambda_j^2\beta_{qj}^2\sigma_{e_q}^2 + 2\lambda_j^2\beta_{dj}\beta_{qj}\sigma_{e_de_q} + \lambda_j^2\sigma_z^2)}{\lambda_j^2(\beta_{dj}+\beta_{qj})^2\sigma_v^2 + \lambda_j^2\beta_{dj}^2\sigma_{e_d}^2 + \lambda_j^2\beta_{qj}^2\sigma_{e_q}^2 + 2\lambda_j^2\beta_{dj}\beta_{qj}\sigma_{e_de_q} + \lambda_j^2\sigma_z^2}\\
&= \sigma_v^2\frac{\beta_{dj}^2\frac{1-\phi_d}{\phi_d} + \beta_{qj}^2\frac{1-\phi_q}{\phi_q} + 2\beta_{dj}\beta_{qj}\frac{\nu_q}{\phi_q} + \eta}{(\beta_{dj}+\beta_{qj})^2 + \beta_{dj}^2\frac{1-\phi_d}{\phi_d} + \beta_{qj}^2\frac{1-\phi_q}{\phi_q} + 2\beta_{dj}\beta_{qj}\frac{\nu_q}{\phi_q} + \eta}\\
&\Downarrow\\
&\mathrm{Var}(v|P_j)^{-1} = \frac{1}{\sigma_v^2}\frac{(\beta_{dj}+\beta_{qj})^2 + \beta_{dj}^2\frac{1-\phi_d}{\phi_d} + \beta_{qj}^2\frac{1-\phi_q}{\phi_q} + 2\beta_{dj}\beta_{qj}\frac{\nu_q}{\phi_q} + \eta}{\beta_{dj}^2\frac{1-\phi_d}{\phi_d} + \beta_{qj}^2\frac{1-\phi_q}{\phi_q} + 2\beta_{dj}\beta_{qj}\frac{\nu_q}{\phi_q} + \eta}
\end{aligned}$$

B.4.3 Price Volatility

From equation B.73:

$$\begin{aligned}
\mathrm{Var}(P_j) &= \mathrm{Var}(\mu_j + \lambda_j\alpha_{dj} + \lambda_j\alpha_{qj} + \lambda_j(\beta_{dj}+\beta_{qj})v + \lambda_j\beta_{dj}e_d + \lambda_j\beta_{qj}e_q + \lambda_j z)\\
&= \lambda_j^2\Big((\beta_{dj}+\beta_{qj})^2\sigma_v^2 + \beta_{dj}^2\sigma_{e_d}^2 + \beta_{qj}^2\sigma_{e_q}^2 + 2\beta_{dj}\beta_{qj}\sigma_{e_de_q} + \sigma_z^2\Big)\\
&= \lambda_j^2\sigma_v^2\Big((\beta_{dj}+\beta_{qj})^2 + \beta_{dj}^2\frac{1-\phi_d}{\phi_d} + \beta_{qj}^2\frac{1-\phi_q}{\phi_q} + 2\beta_{dj}\beta_{qj}\frac{\nu_q}{\phi_q} + \eta\Big)
\end{aligned}$$

B.4.4 Risk Premium

$$\begin{aligned}
\mathbb{E}[v-P_j] &= \mathbb{E}\big[v - \big(\mu_j + \lambda_j\alpha_{dj} + \lambda_j\alpha_{qj} + \lambda_j(\beta_{dj}+\beta_{qj})v + \lambda_j\beta_{dj}e_d + \lambda_j\beta_{qj}e_q + \lambda_j z\big)\big]\\
&= P_0 - \mu_j - \lambda_j\alpha_{dj} - \lambda_j\alpha_{qj} - \lambda_j(\beta_{dj}+\beta_{qj})P_0
\end{aligned}$$

Return Volatility From equation B.73:

$$\begin{aligned}
&\mathrm{Var}(v-P) = \mathrm{Var}(v - \mu_j - \lambda_j\alpha_{dj} - \lambda_j\alpha_{qj} - \lambda_j(\beta_{dj}+\beta_{qj})v - \lambda_j\beta_{dj}e_d - \lambda_j\beta_{qj}e_q - \lambda_j z)\\
&= \mathrm{Var}\big((1-\lambda_j(\beta_{dj}+\beta_{qj}))v - \lambda_j\beta_{dj}e_d - \lambda_j\beta_{qj}e_q - \lambda_j z\big)\\
&= (1-\lambda_j(\beta_{dj}+\beta_{qj}))^2\sigma_v^2 + \lambda_j^2\beta_{dj}^2\sigma_{e_d}^2 + \lambda_j^2\beta_{qj}^2\sigma_{e_q}^2 + 2\lambda_j^2\beta_{dj}\beta_{qj}\sigma_{e_de_q} + \lambda_j^2\sigma_z^2\\
&= \sigma_v^2\Big((1-\lambda_j(\beta_{dj}+\beta_{qj}))^2 + \lambda_j^2\beta_{dj}^2\frac{1-\phi_d}{\phi_d} + \lambda_j^2\beta_{qj}^2\frac{1-\phi_q}{\phi_q} + 2\lambda_j^2\beta_{dj}\beta_{qj}\frac{\nu_q}{\phi_q} + \lambda_j^2\eta\Big)
\end{aligned}$$

Expected Trading Volume: Aggregate We are after $\mathbb{E}[|x_{dj}+x_{qj}|]$, which is the expected value of the absolute value of a normal random variable. Otherwise known as the "folded normal distribution," given a normally distributed random variable $X \sim N(\mu, \sigma^2)$, the expected value of $|X|$ is given by:

$$\mathbb{E}[|X|] = \sigma\sqrt{\frac{2}{\pi}}\exp\left(-\frac{\mu^2}{2\sigma^2}\right) + \mu\Big(1 - 2\Phi\big(-\frac{\mu}{\sigma}\big)\Big) \tag{B.74}$$

where Φ is the standard normal cumulative distribution function. We proceed to derive the distribution of $x_1 + x_2$.

$$\begin{aligned}
x_{dj} + x_{qj} = &\underbrace{\alpha_{dj} + \alpha_{qj} + (\beta_{dj} + \beta_{qj})v}_{\sim N\left(\alpha_{dj}+\alpha_{qj}+(\beta_{dj}+\beta_{qj})P_0, (\beta_{dj}+\beta_{qj})^2\sigma_v^2\right)} + \underbrace{\beta_{dj}e_d + \beta_{qj}e_q}_{\sim N\left(0, \beta_{dj}^2\sigma_{e_d}^2+\beta_{qj}^2\sigma_{e_q}^2+2\beta_{dj}\beta_{qj}\sigma_{e_d e_q}\right)} \\
&\Downarrow \\
x_{dj} + x_{qj} \sim N\big(&\underbrace{\alpha_{dj} + \alpha_{qj} + (\beta_{dj} + \beta_{qj})P_0}_{\equiv \mu_{x_{dj}+x_{qj}}}, \underbrace{(\beta_{dj} + \beta_{qj})^2\sigma_v^2 + \beta_{dj}^2\sigma_{e_d}^2 + \beta_{qj}^2\sigma_{e_q}^2 + 2\beta_{dj}\beta_{qj}\sigma_{e_d e_q}}_{\equiv \sigma^2_{x_{dj}+x_{qj}}}\big)
\end{aligned}$$

Substituting into equation B.74

$$\mathbb{E}[|x_{dj}+x_{qj}|] = \sigma_{x_{dj}+x_{qj}}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{\mu^2_{x_{dj}+x_{qj}}}{2\sigma^2_{x_{dj}+x_{qj}}}\right) + \mu_{x_{dj}+x_{qj}}\Big(1 - 2\Phi\big(-\frac{\mu_{x_{dj}+x_{qj}}}{\sigma_{x_{dj}+x_{qj}}}\big)\Big)$$

Expected Trading Volume: by Trader We are after $\mathbb{E}[|x_{ij}|]$, which is the expected value of the absolute value of a normal random variable. We utilize equation B.74 to derive the expectation of this random variable. The distribution of x_{ij} is:

$$\begin{aligned}
x_{ij} &= \alpha_{ij} + \beta_{ij}s_i \\
&\Downarrow \\
x_{ij} &\sim N\big(\underbrace{\alpha_{ij} + \beta_{ij}P_0}_{\equiv \mu_{x_{ij}}}, \underbrace{\beta_{ij}^2(\sigma_v^2 + \sigma_{e_i}^2)}_{\equiv \sigma^2_{x_{ij}}}\big)
\end{aligned}$$

Substituting into equation B.74

$$\mathbb{E}[|x_{ij}|] = \sigma_{x_{ij}}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{\mu^2_{x_{ij}}}{2\sigma^2_{x_{ij}}}\right) + \mu_{x_{ij}}\Big(1 - 2\Phi\big(-\frac{\mu_{x_{ij}}}{\sigma_{x_{ij}}}\big)\Big)$$

Expected Profits: By Trader and Aggregate Expected profits of trader i in Economy j (π_{ij}) are equal to:

$$\begin{aligned}
&\mathbb{E}[\pi_{ij}] = \mathbb{E}[x_{ij}(v - P)] \\
&= \mathbb{E}[(\alpha_{ij} + \beta_{ij}s_i)(v - \mu_j - \lambda_j\omega_j)] \\
&= \mathbb{E}[(\alpha_{ij} + \beta_{ij}s_i)(v - \mu_j - \lambda_j(x_{dj} + x_{qj} + z))] \\
&= \mathbb{E}[(\alpha_{ij} + \beta_{ij}(v + e_i))(v - \mu_j - \lambda_j(\alpha_{dj} + \beta_{dj}s_d + \alpha_{qj} + \beta_{qj}s_q + z))] \\
&= \mathbb{E}[(\alpha_{ij} + \beta_{ij}v + \beta_{ij}e_i)(v - \mu_j - \lambda_j(\alpha_{dj} + \beta_{dj}(v + e_d) + \alpha_{qj} + \beta_{qj}(v + e_q) + z))] \\
&= \mathbb{E}[(\alpha_{ij} + \beta_{ij}v + \beta_{ij}e_i) \\
&(v - \mu_j - \lambda_j\alpha_{dj} - \lambda_j\beta_{dj}v - \lambda_j\beta_{dj}e_d - \lambda_j\alpha_{dj} - \lambda_j\beta_{qj}v - \lambda_j\beta_{qj}e_q - \lambda_j z)] \\
&= \mathbb{E}[(\alpha_{ij} + \beta_{ij}v + \beta_{ij}e_i) \\
&((1 - \lambda_j\beta_{dj} - \lambda_j\beta_{qj})v - \mu_j - \lambda_j(\alpha_{dj} + \alpha_{qj}) - \lambda_j\beta_{dj}e_d - \lambda_j\beta_{qj}e_q - \lambda_j z)] \\
&= \alpha_{ij}(1 - \lambda_j\beta_{dj} - \lambda_j\beta_{qj})P_0 - \alpha_{ij}\mu_j - \lambda_j\alpha_{ij}(\alpha_{dj} + \alpha_{qj}) + \beta_{ij}(1 - \lambda_j\beta_{dj} - \lambda_j\beta_{qj})\mathbb{E}[v^2] \\
&- \beta_{ij}\mu_j P_0 - \beta_{ij}\lambda_j(\alpha_{dj} + \alpha_{qj})P_0 - \lambda_j\beta_{ij}^2\mathbb{E}[e_i^2] - \lambda_j\beta_{dj}\beta_{qj}\mathbb{E}[e_d e_q]
\end{aligned}$$

We have the following moments given to us by model parameters:

$$\begin{aligned}
\sigma_v^2 &= \mathbb{E}[v^2] - \mathbb{E}[v]^2 = \mathbb{E}[v^2] - P_0^2 \Rightarrow \mathbb{E}[v^2] = \sigma_v^2 + P_0^2 \\
\sigma_{e_d}^2 &= \mathbb{E}[e_d^2] - \mathbb{E}[e_d]^2 = \mathbb{E}[e_d^2] \\
\sigma_{e_q}^2 &= \mathbb{E}[e_q^2] - \mathbb{E}[e_q]^2 = \mathbb{E}[e_q^2] \\
\sigma_{e_d e_q} &= \mathbb{E}[e_d e_q] - \mathbb{E}[e_d]\mathbb{E}[e_q] = \mathbb{E}[e_d e_q]
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathbb{E}[\pi_{ij}] = &\alpha_{ij}(1 - \lambda_j\beta_{dj} - \lambda_j\beta_{qj})P_0 - \alpha_{ij}\mu_j - \lambda_j\alpha_{ij}(\alpha_{dj} + \alpha_{qj}) \\
&+ \beta_{ij}(1 - \lambda_j\beta_{dj} - \lambda_j\beta_{qj})(\sigma_v^2 + P_0^2) \\
&- \beta_{ij}\mu_j P_0 - \beta_{ij}\lambda_j(\alpha_{dj} + \alpha_{qj})P_0 - \lambda_j\beta_{ij}^2\sigma_{e_i}^2 - \lambda_j\beta_{dj}\beta_{qj}\sigma_{e_d e_q}
\end{aligned}$$

Finally,

$$\begin{aligned}
&\mathbb{E}[\pi_{dj} + \pi_{qj}] = \mathbb{E}[\pi_{dj}] + \mathbb{E}[\pi_{qj}] \\
&= (\alpha_{dj} + \alpha_{qj})(1 - \lambda_j\beta_{dj} - \lambda_j\beta_{qj})P_0 - (\alpha_{dj} + \alpha_{qj})\mu_j - \lambda_j(\alpha_{dj} + \alpha_{qj})^2 \\
&+ (\beta_{dj} + \beta_{qj})(1 - \lambda_j\beta_{dj} - \lambda_j\beta_{qj})(\sigma_v^2 + P_0^2) \\
&- (\beta_{dj} + \beta_{qj})\mu_j P_0 - (\beta_{dj} + \beta_{qj})\lambda_j(\alpha_{dj} + \alpha_{qj})P_0 - \lambda_j(\beta_{dj}^2\sigma_{e_d}^2 + \beta_{qj}^2\sigma_{e_q}^2) \\
&- 2\lambda_j\beta_{dj}\beta_{qj}\sigma_{e_d e_q}
\end{aligned}$$

B.5 Strategy Correlation

Strategy Correlation

$$
\begin{aligned}
\rho_{x_{dj}x_{qj}} &= \frac{\mathrm{Cov}(x_{dj}, x_{qj}))}{\sqrt{\mathrm{Var}(x_{dj})\mathrm{Var}(x_{qj})}} \\
&= \frac{\mathrm{Cov}(\alpha_{dj} + \beta_{dj}s_d, \alpha_{qj} + \beta_{qj}s_q)}{\sqrt{\mathrm{Var}(\alpha_{dj} + \beta_{dj}s_d)\mathrm{Var}(\alpha_{qj} + \beta_{qj}s_q)}} \\
&= \frac{\mathrm{Cov}(\alpha_{dj} + \beta_{dj}v + \beta_{dj}e_d, \alpha_{qj} + \beta_{qj}v + \beta_{qj}e_q)}{\sqrt{\mathrm{Var}(\alpha_{dj} + \beta_{dj}v + \beta_{dj}e_d)\mathrm{Var}(\alpha_{qj} + \beta_{qj}v + \beta_{qj}e_q)}} \\
&= \frac{\beta_{dj}\beta_{qj}\sigma_v^2 + \beta_{dj}\beta_{qj}\sigma_{e_d e_q}}{|\beta_{dj}\beta_{qj}|\sqrt{(\sigma_v^2 + \sigma_{e_d}^2)(\sigma_v^2 + \sigma_{e_q}^2)}} \\
&= \mathrm{sign}\left(\beta_{dj}\beta_{qj}\right)\frac{1 + \rho\sqrt{\frac{(1-\phi_d)(1-\phi_q)}{\phi_d\phi_q}}}{\sqrt{\frac{1}{\phi_d\phi_q}}} \\
&= \mathrm{sign}\left(\beta_{dj}\beta_{qj}\right)\left(\sqrt{\phi_d\phi_q} + \rho\sqrt{(1-\phi_d)(1-\phi_q)}\right)
\end{aligned}
$$

B.6 Comparative Statics for the Benchmark Case: Proposition II.1

The simplified version assumes the following model parameters: $P_0 = 0, \phi_d = 1, \rho = \nu_q = \nu_d = 0, \sigma_v^2 = 1$. The model solution across the four economies is as follows:

- Backtest:

$$
\begin{aligned}
\alpha_b &= 0 \\
\beta_b &= \frac{1}{2}\sqrt{\eta}\phi_q
\end{aligned}
$$

- Economy 1:

$$
\begin{aligned}
\alpha_{d1} = \alpha_{q1} = \mu_1 = \beta_{q2} &= 0 \\
\beta_{d1} &= \sqrt{\eta} \\
\lambda_1 &= \frac{1}{2}\sqrt{\frac{1}{\eta}}
\end{aligned}
$$

- Economy 2:

$$\alpha_{d2} = \alpha_{q2} = \mu_2 = 0$$
$$\beta_{d2} = \sqrt{\eta}\sqrt{\frac{(2-\phi_q)^2}{\phi_q^2 - 3\phi_q + 4}}$$
$$\beta_{q2} = \sqrt{\eta}\sqrt{\frac{\phi_q^2}{\phi_q^2 - 3\phi_q + 4}}$$
$$\lambda_2 = \sqrt{\frac{1}{\eta}}\sqrt{\frac{\phi_q^2 - 3\phi_q + 4}{(4-\phi_q)^2}}$$

- Economy 4 (with only the selected β_{d4} is market depth positive):

$$\alpha_{d4} = \alpha_{q4} = \mu_4 = 0$$
$$\beta_{d4} = -\frac{1}{4}\sqrt{\eta}\phi_q + \frac{1}{4}\sqrt{\eta}\sqrt{-3\phi_q^2 + 4\phi_q + 16}$$
$$\beta_{q4} = \beta_b = \frac{1}{2}\sqrt{\eta}\phi_q$$
$$\lambda_4 = \frac{2}{\sqrt{\eta}}\sqrt{\frac{1}{-3\phi_q^2 + 4\phi_q + 16}}$$

B.6.1 Market Depth

We first show that market depth increases with the introduction of the quantitative investor, both with and without automation:

$$\frac{1}{\lambda_1} < \frac{1}{\lambda_2}$$
$$\Updownarrow$$
$$\lambda_1 > \lambda_2$$
$$\Updownarrow$$
$$\frac{1}{2}\sqrt{\frac{1}{\eta}} > \sqrt{\frac{1}{\eta}}\sqrt{\frac{\phi_q^2 - 3\phi_q + 4}{(4-\phi_q)^2}}$$
$$\Updownarrow$$
$$3\phi_q\left(\frac{4}{3} - \phi_q\right) > 0$$

where the latter inequality holds for $\phi_q \in (0, 1]$

$$\frac{1}{\lambda_1} < \frac{1}{\lambda_4}$$
$$\Updownarrow$$
$$\lambda_1 > \lambda_4$$
$$\Updownarrow$$
$$\frac{1}{2}\sqrt{\frac{1}{\eta}} > \frac{2}{\sqrt{\eta}}\sqrt{\frac{1}{-3\phi_q^2 + 4\phi_q + 16}}$$
$$\Updownarrow$$
$$3\phi_q\left(\frac{4}{3} - \phi_q\right) > 0$$

where the latter inequality holds for $\phi_q \in (0, 1]$

We now show that market depth decreases due to automation by quant.

$$\frac{1}{\lambda_2} > \frac{1}{\lambda_4}$$
$$\Updownarrow$$
$$\lambda_2 < \lambda_4$$
$$\Updownarrow$$
$$\sqrt{\frac{1}{\eta}}\sqrt{\frac{\phi_q^2 - 3\phi_q + 4}{(4 - \phi_q)^2}} < \frac{2}{\sqrt{\eta}}\sqrt{\frac{1}{-3\phi_q^2 + 4\phi_q + 16}}$$
$$\Updownarrow$$
$$\frac{\phi_q^2 - 3\phi_q + 4}{(4 - \phi_q)^2} < \frac{4}{-3\phi_q^2 + 4\phi_q + 16}$$
$$\Updownarrow \text{(note that } -3\phi_q^2 + 4\phi_q + 16 > 0 \text{ for } \phi_q \in (0, 1])$$
$$f(\phi_q) \equiv -3\phi_q^2 + 13\phi_q - 12 < 0$$

where the latter inequality holds for $\phi_q \in (0, 1]$ since the roots of $f(\phi_q)$ are $\phi_q^* \approx 1.33$ and $\phi_q^{**} = 3$.

Finally, we demonstrate that the gap in market depth between Economy 2 and

Economy 4 increases with greater signal precision for the quant:

$$
\begin{aligned}
&\frac{\partial}{\partial \phi_q}\frac{\lambda_4}{\lambda_2} \vee 0\\
&\Updownarrow\\
&\frac{\partial}{\partial \phi_q}\frac{2(4-\phi_q)}{\sqrt{(-3\phi_q^2+4\phi_q+16)(\phi_q^2-3\phi_q+4)}} \vee 0\\
&\Updownarrow\\
&\frac{\partial}{\partial \phi_q}\frac{(4-\phi_q)^2}{(-3\phi_q^2+4\phi_q+16)(\phi_q^2-3\phi_q+4)} \vee 0\\
&\Updownarrow\\
&(2\phi_q-8)(-3\phi_q^4+13\phi_q^3-8\phi_q^2-32\phi_q+64)-(\phi_q^2-8\phi_q+16)(-12\phi_q^3+39\phi_q^2-16\phi_q-32)\vee 0\\
&\Updownarrow\\
&f(\phi_q) \equiv 6\phi_q^4 - 85\phi_q^3 + 400\phi_q^2 - 720\phi_q + 384 \vee 0
\end{aligned}
$$

We first note that $f(0) = 384, f(1) = -15$. We now demonstrate that $f(\phi_q)$ is monotone for $\phi_q \in (0, 1]$, implying a single root for $\phi_q \in (0, 1]$. Note the following derivatives of $f(\phi_q)$:

$$
\begin{aligned}
f'(\phi_q) &= 24\phi_q^3 - 255\phi_q^2 + 800\phi_q - 720\\
f''(\phi_q) &= 72\phi_q^2 - 510\phi_q + 800
\end{aligned}
$$

The two roots of $f''(\phi_q) = 0$ are $\approx 2.3, \approx 4.7$, which implies that $f''(\phi_q) > 0$ for $\phi_q \in (0, 1]$. This in turn means that $f'(\phi_q)$ is increasing for $\phi_q \in (0, 1]$, which coupled with $f'(0) = -720, f'(1) = -151$ demonstrates that $f''(\phi_q) < 0$ for $\phi_q \in (0, 1]$. The latter implies that $f(\phi_q)$ is decreasing for $\phi_q \in (0, 1]$, which proves that there exists $\phi_q^* \in (0, 1]$ such that $f(\phi_q) > 0$ for $\phi_q \in (0, \phi_q^*)$ and $f(\phi_q) < 0$ for $\phi_q \in (\phi_q^*, 1]$. Therefore, the percent decrease in market depth due to automation is hump-shaped in the precision of the QI's signal. The peak of the relative decline in market depth is around $\phi_q^* \approx 0.91$.

B.6.2 Price Informativeness

$$
\begin{aligned}
\mathrm{Var}(v|P_j)^{-1} &= \frac{1}{\sigma_v^2}\frac{(\beta_{dj}+\beta_{qj})^2+\beta_{dj}^2\frac{1-\phi_d}{\phi_d}+\beta_{qj}^2\frac{1-\phi_q}{\phi_q}+2\beta_{dj}\beta_{qj}\frac{\nu_q}{\phi_q}+\eta}{\beta_{dj}^2\frac{1-\phi_d}{\phi_d}+\beta_{qj}^2\frac{1-\phi_q}{\phi_q}+2\beta_{dj}\beta_{qj}\frac{\nu_q}{\phi_q}+\eta} \\
&= \frac{(\beta_{dj}+\beta_{qj})^2+\beta_{qj}^2\frac{1-\phi_q}{\phi_q}+\eta}{\beta_{qj}^2\frac{1-\phi_q}{\phi_q}+\eta} \\
&= 1+\frac{(\beta_{dj}+\beta_{qj})^2}{\beta_{qj}^2\frac{1-\phi_q}{\phi_q}+\eta}
\end{aligned}
$$

We first show that price informativeness decreases due to automation by quant.

$$
\begin{aligned}
&\mathrm{Var}(v|P_2)^{-1} > \mathrm{Var}(v|P_4)^{-1} \\
&\Updownarrow \\
&\frac{(\beta_{d2}+\beta_{q2})^2}{\beta_{q2}^2\frac{1-\phi_q}{\phi_q}+\eta} > \frac{(\beta_{d4}+\beta_{q4})^2}{\beta_{q4}^2\frac{1-\phi_q}{\phi_q}+\eta} \\
&\Updownarrow \\
&\frac{\left(\sqrt{\eta}\sqrt{\frac{(2-\phi_q)^2}{\phi_q^2-3\phi_q+4}}+\sqrt{\eta}\sqrt{\frac{\phi_q^2}{\phi_q^2-3\phi_q+4}}\right)^2}{\left(\sqrt{\eta}\sqrt{\frac{\phi_q^2}{\phi_q^2-3\phi_q+4}}\right)^2\frac{1-\phi_q}{\phi_q}+\eta} > \frac{\left(-\frac{1}{4}\sqrt{\eta}\phi_q+\frac{1}{4}\sqrt{\eta}\sqrt{-3\phi_q^2+4\phi_q+16}+\frac{1}{2}\sqrt{\eta}\phi_q\right)^2}{\left(\frac{1}{2}\sqrt{\eta}\phi_q\right)^2\frac{1-\phi_q}{\phi_q}+\eta} \\
&\Updownarrow \\
&\frac{\left(\sqrt{\frac{(2-\phi_q)^2}{\phi_q^2-3\phi_q+4}}+\sqrt{\frac{\phi_q^2}{\phi_q^2-3\phi_q+4}}\right)^2}{\left(\sqrt{\frac{\phi_q^2}{\phi_q^2-3\phi_q+4}}\right)^2\frac{1-\phi_q}{\phi_q}+1} > \frac{\left(\frac{1}{4}\phi_q+\frac{1}{4}\sqrt{-3\phi_q^2+4\phi_q+16}\right)^2}{\left(\frac{1}{2}\phi_q\right)^2\frac{1-\phi_q}{\phi_q}+1} \\
&\Updownarrow \\
&\frac{2}{2-\phi_q} > \frac{-\phi_q^2+2\phi_q+8+\phi_q\sqrt{-3\phi_q^2+4\phi_q+16}}{-2\phi_q^2+2\phi_q+8} \\
&\Updownarrow \text{ note that } -2\phi_q^2+2\phi_q+8>0 \text{ for } \phi_q\in(0,1] \\
&-\phi_q^3+8\phi_q+16 > (2-\phi_q)\sqrt{-3\phi_q^4+4\phi_q^3+16\phi_q^2} \\
&\Updownarrow \\
&f(\phi_q) \equiv 4\phi_q^6-16\phi_q^5-4\phi_q^4+16\phi_q^3+256\phi_q+256>0
\end{aligned}
$$

The latter inequality holds for the following reasons. First, note that $f(0)=256$ and $f(1)=512$. Since $f(\phi_q)$ is continuous, if $f(\phi_q)$ is either monotone or hump-shaped for

$\phi_q \in (0,1]$ then $f(\phi_q) > 0$ in this interval. We now take derivatives of $f(\phi_q)$ and work backwards to demonstrate that $f(\phi_q)$ is monotone over the interval.

$$
\begin{aligned}
f'(\phi_q) &= 24\phi_q^5 - 80\phi_q^4 - 16\phi_q^3 + 48\phi_q^2 + 256 \\
f''(\phi_q) &= 120\phi_q^4 - 320\phi_q^3 - 46\phi_q^2 + 96\phi_q \\
f'''(\phi_q) &= 480\phi_q^3 - 960\phi_q^2 - 92\phi_q + 96 \\
f''''(\phi_q) &= 1440\phi_q^2 - 1920\phi_q - 92
\end{aligned}
$$

First note that the zeros of $f''''(\phi_q)$ are $\approx -.05$ and ≈ 1.38 implying that $f''''(\phi_q) < 0$ for $\phi_q \in (0,1]$. Since $f'''(0) = 96$ and $f'''(1) = -476$ and $f'''(\phi_q)$ is decreasing for $\phi_q \in (0,1]$, there exists $\phi_q^* \in (0,1]$ such that $f'''(\phi_q) > 0$ for $\phi_q \in (0,\phi_q^*)$ and $f'''(\phi_q) < 0$ for $\phi_q \in (\phi_q^*,1]$. This implies that $f''(\phi_q)$ is hump-shaped in the interval $\phi_q \in (0,1]$, which coupled with $f''(0) = 0$ and $f''(1) = -150$ suggests that there exists $\phi_q^{**} \in (0,1]$ such that $f''(\phi_q) > 0$ for $\phi_q \in (0,\phi_q^{**})$ and $f''(\phi_q) < 0$ for $\phi_q \in (\phi_q^{**},1]$. Therefore $f'(\phi_q)$ is also hump-shaped in the interval $\phi_q \in (0,1]$. Since $f'(0) = 256$ and $f'(1) = 232$ and $f'(\phi_q)$ is hump-shaped in the interval $\phi_q \in (0,1]$, $f'(\phi_q) > 0$ in the interval. Coupled with $f(1) > f(0) > 0$, $f(\phi_q) > 0$ for $\phi_q \in (0,1]$.

We now show that price informativeness is always greater with the quant, both fully-discretionary and fully-automated.

$$
\begin{gathered}
\mathrm{Var}(v|P_2)^{-1} > \mathrm{Var}(v|P_1)^{-1} \\
\Updownarrow \\
\frac{(\beta_{d2} + \beta_{q2})^2}{\beta_{q2}^2 \frac{1-\phi_q}{\phi_q} + \eta} > \frac{\beta_{d1}^2}{\eta} \\
\Updownarrow \\
\frac{2}{2 - \phi_q} > 1
\end{gathered}
$$

The latter inequality always holds for $\phi_q \in (0, 1]$.

$$\begin{aligned}
\mathrm{Var}(v|P_4)^{-1} &> \mathrm{Var}(v|P_1)^{-1} \\
&\Updownarrow \\
\frac{(\beta_{d4} + \beta_{q4})^2}{\beta_{q4}^2 \frac{1-\phi_q}{\phi_q} + \eta} &> \frac{\beta_{d1}^2}{\eta} \\
&\Updownarrow \\
\frac{-\phi_q^2 + 2\phi_q + 8 + \phi_q \sqrt{-3\phi_q^2 + 4\phi_q + 16}}{-2\phi_q^2 + 2\phi_q + 8} &> 1 \\
&\Updownarrow \\
\sqrt{-3\phi_q^2 + 4\phi_q + 16} &> -\phi_q
\end{aligned}$$

The latter inequality always holds for $\phi_q \in (0, 1]$.

B.6.3 Return Volatility

$$\mathrm{Var}(v - P_j) = \sigma_v^2 \Big((1 - \lambda_j(\beta_{dj} + \beta_{qj}))^2 + \lambda_j^2 \beta_{dj}^2 \frac{1 - \phi_d}{\phi_d} + \lambda_j^2 \beta_{qj}^2 \frac{1 - \phi_q}{\phi_q} + 2\lambda_j^2 \beta_{dj} \beta_{qj} \frac{\nu_q}{\phi_q} + \lambda_j^2 \eta \Big)$$

We first demonstrate the return volatility increases with automation by the quant.

$$
\begin{gathered}
\operatorname{Var}(v-P_2)<\operatorname{Var}(v-P_4)\\
\Updownarrow\\
(1-\lambda_2(\beta_{d2}+\beta_{q2}))^2+\lambda_2^2\beta_{q2}^2\frac{1-\phi_q}{\phi_q}+\lambda_2^2\eta<(1-\lambda_4(\beta_{d4}+\beta_{q4}))^2+\lambda_4^2\beta_{q4}^2\frac{1-\phi_q}{\phi_q}+\lambda_4^2\eta\\
\Updownarrow\\
\left(1-\sqrt{\tfrac{1}{\eta}}\sqrt{\tfrac{\phi_q^2-3\phi_q+4}{(4-\phi_q)^2}}\left(\sqrt{\eta}\sqrt{\tfrac{(2-\phi_q)^2}{\phi_q^2-3\phi_q+4}}+\sqrt{\eta}\sqrt{\tfrac{\phi_q^2}{\phi_q^2-3\phi_q+4}}\right)\right)^2\\
+\left(\sqrt{\tfrac{1}{\eta}}\sqrt{\tfrac{\phi_q^2-3\phi_q+4}{(4-\phi_q)^2}}\sqrt{\eta}\sqrt{\tfrac{\phi_q^2}{\phi_q^2-3\phi_q+4}}\right)^2\tfrac{1-\phi_q}{\phi_q}+\left(\sqrt{\tfrac{1}{\eta}}\sqrt{\tfrac{\phi_q^2-3\phi_q+4}{(4-\phi_q)^2}}\right)^2\eta\\
<\\
\left(1-\tfrac{2}{\sqrt{\eta}}\sqrt{\tfrac{1}{-3\phi_q^2+4\phi_q+16}}\left(-\tfrac{1}{4}\sqrt{\eta}\phi_q+\tfrac{1}{4}\sqrt{\eta}\sqrt{-3\phi_q^2+4\phi_q+16}+\tfrac{1}{2}\sqrt{\eta}\phi_q\right)\right)^2\\
+\left(\tfrac{2}{\sqrt{\eta}}\sqrt{\tfrac{1}{-3\phi_q^2+4\phi_q+16}}\tfrac{1}{2}\sqrt{\eta}\phi_q\right)^2\tfrac{1-\phi_q}{\phi_q}+\left(\tfrac{2}{\sqrt{\eta}}\sqrt{\tfrac{1}{-3\phi_q^2+4\phi_q+16}}\right)^2\eta\\
\Updownarrow\\
\frac{\phi_q^2-6\phi_q+8}{\phi_q^2-8\phi_q+16}<\frac{1}{2}-\frac{1}{2}\sqrt{\frac{\phi_q^2}{-3\phi_q^2+4\phi_q+16}}\\
\Updownarrow\\
\frac{1}{\phi_q-4}>\sqrt{\frac{1}{-3\phi_q^2+4\phi_q+16}}\\
\Updownarrow\\
\phi_q-4<\sqrt{-3\phi_q^2+4\phi_q+16}
\end{gathered}
$$

The latter inequality always holds for $\phi_q\in(0,1]$ since the left-hand-side is negative and the right-hand-side is positive within the interval.

We now demonstrate that return volatility decreases with the introduction of the quantitative investor, both with and without automation.

$$\begin{gathered}\mathrm{Var}(v - P_2) < \mathrm{Var}(v - P_1) \\ \Updownarrow \\ \frac{\phi_q^2 - 6\phi_q + 8}{\phi_q^2 - 8\phi_q + 16} < \frac{1}{2} \\ \Updownarrow \\ \phi_q(\phi_q - 4) < 0\end{gathered}$$

The latter inequality always holds for $\phi_q \in (0, 1]$ since the left-hand-side is always negative in the interval.

$$\begin{gathered}\mathrm{Var}(v - P_4) < \mathrm{Var}(v - P_1) \\ \Updownarrow \\ \frac{1}{2} - \frac{1}{2}\sqrt{\frac{\phi_q^2}{-3\phi_q^2 + 4\phi_q + 16}} < \frac{1}{2} \\ \Updownarrow \\ -\phi_q\sqrt{\frac{1}{-3\phi_q^2 + 4\phi_q + 16}} < 0\end{gathered}$$

The latter inequality always holds for $\phi_q \in (0, 1]$ since the left-hand-side is always negative in the interval.

B.6.4 Price Volatility

We will rely on derivations in the return volatility analysis by noting that

$$\begin{aligned}\mathrm{Var}(v - P_j) &= \sigma_v^2 + \mathrm{Var}(P_j) - 2\mathrm{Cov}(v, P_j) \\ \mathrm{Var}(P_j) &= \mathrm{Var}(v - P_j) - \sigma_v^2 + 2\lambda_j(\beta_{dj} + \beta_{qj})\end{aligned}$$

We first demonstrate that price volatility increases with automation by the quant.

$$\begin{gathered}
\mathrm{Var}(P_2) > \mathrm{Var}(P_4) \\
\Updownarrow \\
\frac{\phi_q^2 - 6\phi_q + 8}{\phi_q^2 - 8\phi_q + 16} + \frac{4}{4-\phi_q} > \frac{1}{2} - \frac{1}{2}\sqrt{\frac{\phi_q^2}{-3\phi_q^2 + 4\phi_q + 16}} + 1 + \sqrt{\frac{\phi_q^2}{-3\phi_q^2 + 4\phi_q + 16}} \\
\Updownarrow \\
\frac{\phi_q}{4-\phi_q} > \sqrt{\frac{\phi_q^2}{-3\phi_q^2 + 4\phi_q + 16}} \\
\Updownarrow \\
4 - \phi_q < \sqrt{-3\phi_q^2 + 4\phi_q + 16} \\
\Updownarrow \\
\phi_q(\phi_q - 3) < 0
\end{gathered}$$

The latter inequality always holds for $\phi_q \in (0,1]$ since the left-hand-side is always negative in the interval.

We now demonstrate that price volatility increases with the introduction of the quantitative investor, both with and without automation.

$$\begin{gathered}
\mathrm{Var}(P_2) > \mathrm{Var}(P_1) \\
\Updownarrow \\
\frac{\phi_q^2 - 6\phi_q + 8}{\phi_q^2 - 8\phi_q + 16} + \frac{4}{4-\phi_q} > \frac{1}{2} + 1 \\
\Updownarrow \\
\phi_q(4 - \phi_q) > 0
\end{gathered}$$

The latter inequality always holds for $\phi_q \in (0,1]$ since the left-hand-side is always positive in the interval.

$$\begin{gathered}
\mathrm{Var}(P_4) > \mathrm{Var}(P_1) \\
\Updownarrow \\
\frac{1}{2} - \frac{1}{2}\sqrt{\frac{\phi_q^2}{-3\phi_q^2 + 4\phi_q + 16}} + 1 + \sqrt{\frac{\phi_q^2}{-3\phi_q^2 + 4\phi_q + 16}} > \frac{1}{2} + 1 \\
\Updownarrow \\
\phi_q\sqrt{\frac{1}{-3\phi_q^2 + 4\phi_q + 16}} > 0
\end{gathered}$$

The latter inequality always holds for $\phi_q \in (0, 1]$ since the left-hand-side is always positive in the interval.

B.6.5 Trading Intensity and Expected Trading Volume

Quantitative investor The QI does not trade in Economy 1. She trades more aggressively as a fully-discretionary than as a fully-automated investor:

$$\begin{gathered}
\beta_{q2} > \beta_{q4} \\
\Updownarrow \\
\sqrt{\eta}\sqrt{\frac{\phi_q^2}{\phi_q^2 - 3\phi_q + 4}} > \frac{1}{2}\sqrt{\eta}\phi_q \\
\Updownarrow \\
\phi_q^2 - 3\phi_q < 0 \\
\Updownarrow \\
\phi_q(\phi_q - 3) < 0,
\end{gathered}$$

where the latter inequality holds for $\phi_q \in (0, 1]$.

A higher trading intensity implies a greater expected trading volume and vice versa, immediate from Eq. (B.74), with $\mu_{x_{qj}} = 0$ and $\sigma_{x_{qj}} = \beta_{qj}\sigma_{s_q}$. Therefore the QI trades less as she turns on her strategy.

Discretionary investor We first show that β_{d4} is always positive:

$$\begin{aligned}
\beta_{d4} &> 0 \\
&\Updownarrow \\
-\frac{1}{4}\sqrt{\eta}\phi_q + \frac{1}{2}\sqrt{\eta}\sqrt{\phi_q - \frac{3}{4}\phi_q^2 + 4} &> 0 \\
&\Updownarrow \\
\frac{1}{2}\sqrt{\eta}\sqrt{\phi_q - \frac{3}{4}\phi_q^2 + 4} &> \frac{1}{4}\sqrt{\eta}\phi_q \\
&\Updownarrow \\
-\frac{3}{16}\phi_q^2 + \frac{1}{4}\phi_q + 1 &> \frac{1}{16}\phi_q^2 \\
&\Updownarrow \\
0 &> \phi_q^2 - \phi_q - 4
\end{aligned}$$

The latter inequality holds for $\phi_q \in (0, 1]$ since the right-hand-side is an upwards parabola with roots at ≈ -1.56 and ≈ 2.56.

Next, we show that the strategic speculator trades less once the QI is introduced in the economy both with and without automation:

$$
\begin{aligned}
\beta_{d2} &< \beta_{d1} \\
&\Updownarrow \\
\sqrt{\eta}\sqrt{\frac{(2-\phi_q)^2}{\phi_q^2 - 3\phi_q + 4}} &< \sqrt{\eta} \\
&\Updownarrow \\
(2-\phi_q)^2 &< \phi_q^2 - 3\phi_q + 4 \\
&\Updownarrow \\
-\phi_q &< 0
\end{aligned}
$$

$$
\begin{aligned}
\beta_{d4} &< \beta_{d1} \\
&\Updownarrow \\
-\frac{1}{4}\sqrt{\eta}\phi_q + \frac{1}{2}\sqrt{\eta}\sqrt{\phi_q - \frac{3}{4}\phi_q^2 + 4} &< \sqrt{\eta} \\
&\Updownarrow \\
-3\phi_q^2 + 4\phi_q + 16 &< (\phi_q + 4)^2 \\
&\Updownarrow \\
-4\phi_q(\phi_q + 1) &< 0
\end{aligned}
$$

where the latter inequalities hold because $\phi_q \in (0, 1]$.

Finally, we show that the DI always trades more aggressively in Economy 4 than in Economy 2:

$$
\begin{gathered}
\beta_{d4} > \beta_{d2} \\
\Updownarrow \\
-\frac{1}{4}\sqrt{\eta}\phi_q + \frac{1}{2}\sqrt{\eta}\sqrt{\phi_q - \frac{3}{4}\phi_q^2 + 4} > \sqrt{\eta}\sqrt{\frac{(2-\phi_q)^2}{\phi_q^2 - 3\phi_q + 4}} \\
\Updownarrow \\
\sqrt{-3\phi_q^2 + 4\phi_q + 16} - \phi_q > \frac{8 - 4\phi_q}{\sqrt{\phi_q^2 - 3\phi_q + 4}} \\
\Updownarrow \\
-2\phi_q^2 + 4\phi_q + 16 - \sqrt{-12\phi_q^4 + 16\phi_q^3 + 64\phi_q^2} > \frac{16\phi_q^2 - 64\phi_q + 64}{\phi_q^2 - 3\phi_q + 4} \\
\Updownarrow \\
(-2\phi_q^2 + 4\phi_q + 16)(\phi_q^2 - 3\phi_q + 4) - 16\phi_q^2 + 64\phi_q - 64 > (\phi_q^2 - 3\phi_q + 4)\sqrt{-12\phi_q^4 + 16\phi_q^3 + 64\phi_q^2} \\
\Updownarrow \\
-\phi_q^3 + 5\phi_q^2 - 10\phi_q + 16 > (\phi_q^2 - 3\phi_q + 4)\sqrt{-3\phi_q^2 + 4\phi_q + 16} \\
\Updownarrow \\
f(\phi_q) \equiv \phi_q^4 - 8\phi_q^3 + 26\phi_q^2 - 44\phi_q + 33 > 0
\end{gathered}
$$

We demonstrate that the latter inequality holds for $\phi_q \in [0, 1]$ by proof by contradiction. First, $f(0) = 33$, $f(1) = 8$. Therefore, for the inequality to not hold it must be the case that $f(\phi_q)$ has a local minimum in $\phi_q \in (0, 1)$. However, this, in turn, implies that $f'(\phi_q)$ has a a value of zero in the interval $\phi_q \in (0, 1)$, i.e.:

$$f'(\phi_q) = 4\phi_q^3 - 24\phi_q^2 + 52\phi_q - 44 = 0$$

However, this does not hold since, $f'(0) = -44, f'(1) = -12$, which in-turn implies that $f''(\phi_q)$ must have a local minimum for $\phi_q \in (0, 1)$. Since

$$f''(\phi_q) = 12\phi_q^2 - 48\phi_q + 52 > 0,$$

$f'(\phi_q)$ cannot have a zero in $\phi_q \in (0, 1)$, which implies that $f(\phi_q)$ is always greater than zero.

A higher trading intensity implies a greater expected trading volume and vice versa, immediate from Eq. (B.74), with $\mu_{x_{dj}} = 0$ and $\sigma_{x_{dj}} = \beta_{dj}\sigma_v$. Therefore, the DI trades more in expectation with automation by the QI.

Speculator sector First, we show that the speculator sector as a whole trades more with the addition of the QI:

$$\begin{gathered}\beta_{d4}+\beta_{q4}>\beta_{d1}\\ \Updownarrow\\ -\frac{1}{4}\sqrt{\eta}\phi_q+\frac{1}{2}\sqrt{\eta}\sqrt{\phi_q-\frac{3}{4}\phi_q^2+4}+\frac{1}{2}\sqrt{\eta}\phi_q>\sqrt{\eta}\\ \Updownarrow\\ \sqrt{\phi_q-\frac{3}{4}\phi_q^2+4}>-\frac{1}{2}\phi_q\end{gathered}$$

$$\begin{gathered}\beta_{d2}+\beta_{q2}>\beta_{d1}\\ \Updownarrow\\ \sqrt{\eta}\sqrt{\frac{(2-\phi_q)^2}{\phi_q^2-3\phi_q+4}}+\sqrt{\eta}\sqrt{\frac{\phi_q^2}{\phi_q^2-3\phi_q+4}}>\sqrt{\eta}\\ \Updownarrow\\ \frac{4}{\phi_q^2-3\phi_q+4}>1\\ \Updownarrow\\ 0>\phi_q(\phi_q-3)\end{gathered}$$

$$\beta_{d4} + \beta_{q4} < \beta_{d2} + \beta_{q2}$$
$$\Updownarrow$$
$$-\frac{1}{4}\sqrt{\eta}\phi_q + \frac{1}{2}\sqrt{\eta}\sqrt{\phi_q - \frac{3}{4}\phi_q^2 + 4} + \frac{1}{2}\sqrt{\eta}\phi_q < \sqrt{\eta}\sqrt{\frac{(2-\phi_q)^2}{\phi_q^2 - 3\phi_q + 4}} + \sqrt{\eta}\sqrt{\frac{\phi_q^2}{\phi_q^2 - 3\phi_q + 4}}$$
$$\Updownarrow$$
$$\frac{1}{4}(\sqrt{-3\phi_q^2 + 4\phi_q + 16} + \phi_q) < \sqrt{\frac{(2-\phi_q)^2}{\phi_q^2 - 3\phi_q + 4}} + \sqrt{\frac{\phi_q^2}{\phi_q^2 - 3\phi_q + 4}}$$
$$\Updownarrow$$
$$-\phi_q^2 + 2\phi_q + 8 + \phi_q\sqrt{-3\phi_q^2 + 4\phi_q + 16} < \frac{32}{\phi_q^2 - 3\phi_q + 4}$$
$$\Updownarrow$$
$$(\phi_q^3 - 3\phi_q^2 + 4\phi_q)\sqrt{-3\phi_q^2 + 4\phi_q + 16} < (\phi_q^2 - 3\phi_q + 4)(\phi_q^2 - 2\phi_q - 8) + 32$$
$$\Updownarrow$$
$$f(\phi_q) \equiv -\phi_q^5 + 8\phi_q^4 - 22\phi_q^3 + 8\phi_q^2 + 71\phi_q - 96 < 0$$

We prove the latter inequality for $\phi_q \in (0, 1]$ via proof by contradiction. Let's first consider the first three derivatives of $f(\phi_q)$ and work backwards:

$$f'(\phi_q) = -5\phi_q^4 + 32\phi_q^3 - 66\phi_q^2 + 16\phi_q + 71$$
$$f''(\phi_q) = -20\phi_q^3 + 96\phi_q^2 - 132\phi_q + 16$$
$$f'''(\phi_q) = -60\phi_q^2 + 192\phi_q - 132$$

Note that the two roots of $f'''(\phi_q) = 0$ are $\phi_q^* = 1$ and $\phi_q^{**} = 2.2$. Since $f'''(\phi_q)$ is a downwards parabola, $f'''(\phi_q) < 0$ for $\phi_q \in (0, 1)$. This implies that $f''(\phi_q)$ is decreasing for all $\phi_q \in (0, 1)$. Note that $f''(0) = 16$ and $f''(1) = -40$. Hence, there exists an $\phi_q^0 \in (0, 1)$ such that $f'(\phi_q)$ is increasing from $\phi_q \in (0, \phi_q^0)$ and decreasing from $\phi_q \in (\phi_q^0, 1)$, while reaching a local maxima at ϕ_q^0. Together with the fact that $f'(0) = 71$ and $f'(1) = 48$, $f'(\phi_q) > 0$ for $\phi_q \in (0, 1)$. Finally, since $f(0) = -96$ and $f(1) = -32$ and the function is monotone and increasing across the interval, $f'(\phi_q) < 0$ for all $\phi_q \in (0, 1)$. QED.

The speculator sector as a whole trades less with automation by the quant, which follows from Eq. (B.74), with $\mu_{x_{dj}+x_{qj}} = 0$ and $\sigma_{x_{dj}+x_{qj}} = \sqrt{(\beta_{dj} + \beta_{qj})^2\sigma_v^2 + \beta_{qj}^2\sigma_{e_q}^2}$, since both aggregate trading intensity $\beta_{dj} + \beta_{qj}$ and the QI's trading intensity β_{qj} decrease as the QI turns on the strategy (i.e. from Economy 2 to Economy 4).

B.6.6 Speculator Profits

$$\mathbb{E}[\pi_{ij}] = \beta_{ij}(1 - \lambda_j(\beta_{dj} + \beta_{qj})) - \lambda_j \beta_{ij}^2 \frac{1-\phi_i}{\phi_i}$$

In what follows, note that:

$$\begin{aligned}
\beta_{d2} + \beta_{q2} &= \sqrt{\eta} \frac{2}{\sqrt{\phi_q^2 - 3\phi_q + 4}} \\
\beta_{d4} + \beta_{q4} &= \sqrt{\eta} \frac{1}{4} \left(\phi_q + \sqrt{-3\phi_q^2 + 4\phi_q + 16} \right) \\
1 - \lambda_2(\beta_{d2} + \beta_{q2}) &= \frac{2 - \phi_q}{4 - \phi_q} \\
1 - \lambda_4(\beta_{d4} + \beta_{q4}) &= \frac{1}{2} \left(1 - \frac{\phi_q}{\sqrt{-3\phi_q^2 + 4\phi_q + 16}} \right) \\
\lambda_2 \beta_{q2}^2 \frac{1-\phi_q}{\phi_q} &= \sqrt{\eta} \frac{-\phi_q^2 + \phi_q}{(4-\phi_q)\sqrt{\phi_q^2 - 3\phi_q + 4}} \\
\lambda_4 \beta_{q4}^2 \frac{1-\phi_q}{\phi_q} &= \sqrt{\eta} \frac{1}{2} \frac{-\phi_q^2 + \phi_q}{\sqrt{-3\phi_q^2 + 4\phi_q + 16}}
\end{aligned} \tag{B.75}$$

Quantitative investor

$$\begin{aligned}
\mathbb{E}[\pi_{q2}] &= \beta_{q2}(1 - \lambda_2(\beta_{d2} + \beta_{q2})) - \lambda_2 \beta_{q2}^2 \frac{1-\phi_q}{\phi_q} \\
\mathbb{E}[\pi_{q4}] &= \beta_{q4}(1 - \lambda_4(\beta_{d4} + \beta_{q4})) - \lambda_4 \beta_{q4}^2 \frac{1-\phi_q}{\phi_q}
\end{aligned}$$

The quant's profits decrease with automation by design, since she switches from a profit-maximizing strategy.

Discretionary investor

$$\begin{aligned}
\mathbb{E}[\pi_{d1}] &= \beta_{d1}(1 - \lambda_1 \beta_{d1}) \\
\mathbb{E}[\pi_{d2}] &= \beta_{d2}(1 - \lambda_2(\beta_{d2} + \beta_{q2})) \\
\mathbb{E}[\pi_{d4}] &= \beta_{d4}(1 - \lambda_4(\beta_{d4} + \beta_{q4}))
\end{aligned}$$

First, we show that the profits of the DI increase with the automation by the QI:

$$\mathbb{E}[\pi_{d2}] < \mathbb{E}[\pi_{d4}]$$
$$\Updownarrow$$
$$\beta_{d2}(1-\lambda_2(\beta_{d2}+\beta_{q2})) < \beta_{d4}(1-\lambda_4(\beta_{d4}+\beta_{q4}))$$

We have already demonstrated that $\beta_{d2} < \beta_{d4}$. It is sufficient to show that $0 < 1-\lambda_2(\beta_{d2}+\beta_{q2}) < 1-\lambda_4(\beta_{d4}+\beta_{q4})$.

$$1-\lambda_2(\beta_{d2}+\beta_{q2}) < 1-\lambda_4(\beta_{d4}+\beta_{q4})$$
$$\Updownarrow$$
$$0 < \frac{2-\phi_q}{4-\phi_q} < \frac{1}{2}\left(1-\phi_q\sqrt{\frac{1}{-3\phi_q^2+4\phi_q+16}}\right)$$
$$\Updownarrow$$
$$\frac{1}{4-\phi_q} > \sqrt{\frac{1}{-3\phi_q^2+4\phi_q+16}}$$
$$\Updownarrow$$
$$(4-\phi_q)^2 < -3\phi_q^2+4\phi_q+16$$
$$\Updownarrow$$
$$\phi_q(\phi_q-3) < 0$$

where the latter inequality holds for $\phi_q \in (0,1]$. Therefore $\mathbb{E}[\pi_{d2}] < \mathbb{E}[\pi_{d4}]$.

Next, we show that the profits of the DI decrease with the introduction of the QI, both with and without automation. It is sufficient to show that profits in Economy 4 are lower, since, as demonstrated above, Economy 2 profits are greater than Economy 4 profits.

$$\mathbb{E}[\pi_{d4}] < \mathbb{E}[\pi_{d1}]$$
$$\Updownarrow$$
$$\beta_{d4}(1-\lambda_4(\beta_{d4}+\beta_{q4})) < \beta_{d1}(1-\lambda_1\beta_{d1})$$

We have already demonstrated that $\beta_{d4} < \beta_{d1}$. It remains to be shown that

$$
\begin{gathered}
1-\lambda_4(\beta_{d4}+\beta_{q4}) < 1-\lambda_1\beta_{d1} \\
\Updownarrow \\
\frac{1}{2}\left(1-\phi_q\sqrt{\frac{1}{-3\phi_q^2+4\phi_q+16}}\right) < \frac{1}{2} \\
\Updownarrow \\
0 < \phi_q\sqrt{\frac{1}{-3\phi_q^2+4\phi_q+16}}
\end{gathered}
$$

where the latter inequality holds for $\phi_q \in (0,1]$.

Speculator sector

$$
\begin{aligned}
\mathbb{E}[\pi_{d1}+\pi_{q1}] &= \beta_{d1}(1-\lambda_1\beta_{d1}) \\
\mathbb{E}[\pi_{d2}+\pi_{q2}] &= (\beta_{d2}+\beta_{q2})(1-\lambda_2(\beta_{d2}+\beta_{q2}))-\lambda_2\beta_{q2}^2\frac{1-\phi_q}{\phi_q} \\
\mathbb{E}[\pi_{d4}+\pi_{q4}] &= (\beta_{d4}+\beta_{q4})(1-\lambda_4(\beta_{d4}+\beta_{q4}))-\lambda_4\beta_{q4}^2\frac{1-\phi_q}{\phi_q}
\end{aligned}
$$

First, we show that profits for the entire speculator sector increase with automation by the quantitative investor. From Eq. (B.75):

$$
\begin{gathered}
\mathbb{E}[\pi_{d2}]+\mathbb{E}[\pi_{q2}] < \mathbb{E}[\pi_{d4}]+\mathbb{E}[\pi_{q4}] \\
\Updownarrow \\
(\beta_{d2}+\beta_{q2})(1-\lambda_2(\beta_{d2}+\beta_{q2}))-\lambda_2\beta_{q2}^2\frac{1-\phi_q}{\phi_q} < (\beta_{d4}+\beta_{q4})(1-\lambda_4(\beta_{d4}+\beta_{q4}))-\lambda_4\beta_{q4}^2\frac{1-\phi_q}{\phi_q} \\
\Updownarrow \\
\sqrt{\eta}\frac{2}{\sqrt{\phi_q^2-3\phi_q+4}}\frac{2-\phi_q}{4-\phi_q}-\sqrt{\eta}\frac{-\phi_q^2+\phi_q}{(4-\phi_q)\sqrt{\phi_q^2-3\phi_q+4}} \\
< \\
\sqrt{\eta}\frac{1}{4}\left(\phi_q+\sqrt{-3\phi_q^2+4\phi_q+16}\right)\frac{1}{2}\left(1-\frac{\phi_q}{\sqrt{-3\phi_q^2+4\phi_q+16}}\right)-\sqrt{\eta}\frac{1}{2}\frac{-\phi_q^2+\phi_q}{\sqrt{-3\phi_q^2+4\phi_q+16}} \\
\Updownarrow \\
\frac{\sqrt{\phi_q^2-3\phi_q+4}}{4-\phi_q} < \frac{2}{\sqrt{-3\phi_q^2+4\phi_q+16}} \\
\Updownarrow \\
-3\phi_q^2+13\phi_q-12 < 0
\end{gathered}
$$

The latter inequality holds for $\phi_q \in (0, 1]$ because the two roots of the quadratic are 4/3 and 3 and the parabola is concave.

We now demonstrate that the profits for the speculator sector as a whole decrease with the introduction of the quant, either with or without automation. It is sufficient to show that profits are lower for Economy 4, since they are greater than Economy 2, as just demonstrated.

$$\begin{aligned}
\mathbb{E}[\pi_{d4}] + \mathbb{E}[\pi_{q4}] &< \mathbb{E}[\pi_{d1}] + \mathbb{E}[\pi_{q1}] \\
&\Updownarrow \\
(\beta_{d4} + \beta_{q4})(1 - \lambda_4(\beta_{d4} + \beta_{q4})) - \lambda_4 \beta_{q4}^2 \frac{1 - \phi_q}{\phi_q} &< \beta_{d1}(1 - \lambda_1 \beta_{d1}) \\
&\Updownarrow \\
\frac{2}{\sqrt{-3\phi_q^2 + 4\phi_q + 16}} &< \frac{1}{2} \\
&\Updownarrow \\
0 &< \phi_q(1 - \frac{3}{4}\phi_q)
\end{aligned}$$

where the latter inequality holds for $\phi_q \in (0, 1]$.

B.7 Breakeven Correlation

We are after the conditions under which the fully automated quant (Economy 4) trades more aggressively than the fully discretionary quant (Economy 2). From Eq. (B.29) and (B.50), we have:

$$
\begin{aligned}
&\beta_{q4} > \beta_{q2} \\
&\Updownarrow \\
&\tfrac{1}{2}\sqrt{\tfrac{\eta}{\phi_d}}\left((2-\phi_d)\phi_q-\rho\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}\right)>\sqrt{\tfrac{\eta\phi_d\phi_q}{\phi_q k^2+(2\phi_q-\nu_d\phi_q-\phi_d\phi_q)k+\nu_d\phi_q-\phi_d+\phi_d\phi_q}} \\
&\Updownarrow \text{ (label } w=\sqrt{\phi_d\phi_q(1-\phi_d)(1-\phi_q)}) \\
&\frac{1}{2}\sqrt{\frac{\eta}{\phi_d}}(2\phi_q-\phi_d\phi_q-\rho w)>\sqrt{\frac{\eta\phi_d\phi_q}{\phi_q k^2+\phi_d}} \\
&\Updownarrow \\
&\frac{1}{2}\sqrt{\frac{\eta}{\phi_d}}(2\phi_q-\phi_d\phi_q-\rho w)>\sqrt{\frac{\eta\phi_d\phi_q}{\phi_q\left(\frac{2\phi_d-\phi_d\phi_q-\rho w}{2\phi_q-\phi_d\phi_q-\rho w}\right)^2+\phi_d}} \\
&\Updownarrow \\
&\frac{1}{4\phi_d}>\frac{\phi_d\phi_q}{\phi_q(2\phi_d-\phi_d\phi_q-\rho w)^2+\phi_d(2\phi_q-\phi_d\phi_q-\rho w)^2} \\
&\Updownarrow \\
&\phi_q(2\phi_d-\phi_d\phi_q-\rho w)^2+\phi_d(2\phi_q-\phi_d\phi_q-\rho w)^2>4\phi_d^2\phi_q \\
&\Updownarrow \\
&f(\rho)\equiv w^2(\phi_d+\phi_q)\rho^2+2w\phi_d\phi_q(\phi_d+\phi_q-4)\rho+\phi_d\phi_q^2(\phi_d\phi_q-8\phi_d+\phi_d^2+4)>0
\end{aligned}
$$

To identify parameter values for which the inequality above holds, we solve $f(\rho)=0$ analytically to obtain:

$$
\rho^*_\pm=\sqrt{\frac{\phi_d\phi_q}{(1-\phi_d)(1-\phi_q)}}\left(-1+\frac{4}{\phi_d+\phi_q}\pm\frac{2\sqrt{3-\phi_q/\phi_d}}{\phi_d+\phi_q}\right) \tag{B.76}
$$

Since $f(\rho)$ is an upward parabola and $\rho\in[-1,1]$, the inequality above holds for the following conditions:

1. $\forall\rho\in[-1,1]$ if $\phi_q>3\phi_d$ or $\rho^*_->1$ or $\rho^*_+<-1$
2. $\forall\rho\in[\rho^*_+,1]$ if $\rho^*_-\le -1$ and $\rho^*_+\in(0,1)$
3. $\forall\rho\in[-1,\rho^*_-]$ if $\rho^*_+\ge 1$ and $\rho^*_-\in(0,1)$
4. $\forall\rho\in[0,\rho^*_-]\cup[\rho^*_+,1]$ if $\rho^*_-\in(0,1)$ and $\rho^*_+\in(0,1)$
5. Never holds if $\rho^*_-<-1$ and $\rho^*_+>1$

Given the range of $\phi_d\in(0,1)$ and $\phi_q\in(0,1)$, a numerical analysis demonstrates that only conditions 1, 3, and 5 are possible.

APPENDIX C

Arbitrage as Camouflage: Derivations and Proofs

C.1 Model Solution

C.1.1 Economy b: No Arbitrageur

I first solve the model with law of one price violations. I assume that the functional forms for the price and supply functions are as follows (no intercept terms are necessary since all random variables are mean zero):

$$\begin{aligned} x_{1b} &= \beta_{1b} v \\ x_{2b} &= \beta_{2b} v \\ P_{1b} &= \lambda_{1b}\omega_{1b} \\ P_{2b} &= \lambda_{2b}\omega_{2b} \end{aligned}$$

For the strategic speculator:

$$\begin{aligned} [x_{1b}, x_{2b}] &= \arg\max_{\tilde{x}_1,\tilde{x}_2} \mathbb{E}[\tilde{x}_1(v - P_1(\omega_1)) + \tilde{x}_2(v - P_2(\omega_2))|v] \\ &= \arg\max_{\tilde{x}_1,\tilde{x}_2} \mathbb{E}[\tilde{x}_1(v - \lambda_{1b}\omega_1) + \tilde{x}_2(v - \lambda_{2b}\omega_2)|v] \\ &= \arg\max_{\tilde{x}_1,\tilde{x}_2} \mathbb{E}[\tilde{x}_1(v - \lambda_{1b}(\tilde{x}_1 + z_1)) + \tilde{x}_2(v - \lambda_{2b}(\tilde{x}_2 + z_2))|v] \\ &= \arg\max_{\tilde{x}_1,\tilde{x}_2} \left(\tilde{x}_1 v - \lambda_{1b}\tilde{x}_1^2 + \tilde{x}_2 v - \lambda_{2b}\tilde{x}_2^2\right) \\ &\equiv \arg\max_{\tilde{x}_1,\tilde{x}_2} f_b(\tilde{x}_1, \tilde{x}_2) \end{aligned} \tag{C.1}$$

First order condition #1:

$$\begin{aligned}
\frac{\partial}{\partial \tilde{x}_1}\left(\tilde{x}_1 v - \lambda_{1b}\tilde{x}_1^2 + \tilde{x}_2 v - \lambda_{2b}\tilde{x}_2^2\right) &= 0 \\
v - 2\lambda_{1b}\tilde{x}_1 &= 0 \\
&\Downarrow \\
x_{1b} = \frac{1}{2\lambda_{1b}} v \equiv \beta_{1b} v
\end{aligned} \tag{C.2}$$

First order condition #2:

$$\begin{aligned}
\frac{\partial}{\partial \tilde{x}_1}\left(\tilde{x}_1 v - \lambda_{1b}\tilde{x}_1^2 + \tilde{x}_2 v - \lambda_{2b}\tilde{x}_2^2\right) &= 0 \\
v - 2\lambda_{2b}\tilde{x}_2 &= 0 \\
&\Downarrow \\
x_{2b} = \frac{1}{2\lambda_{2b}} v \equiv \beta_{2b} v
\end{aligned} \tag{C.3}$$

Finally, second order conditions would require that:

$$\begin{aligned}
\frac{\partial^2}{\partial \tilde{x}_1^2}\left(\tilde{x}_1 v - \lambda_1\tilde{x}_1^2 + \tilde{x}_2 v - \lambda_2\tilde{x}_2^2\right) &= -2\lambda_1 < 0 \\
\frac{\partial^2}{\partial \tilde{x}_2^2}\left(\tilde{x}_1 v - \lambda_1\tilde{x}_1^2 + \tilde{x}_2 v - \lambda_2\tilde{x}_2^2\right) &= -2\lambda_2 < 0 \\
\frac{\partial^2 f_b}{\partial \tilde{x}_1^2}\frac{\partial^2 f_b}{\partial \tilde{x}_2^2} - \frac{\partial^2 f_b}{\partial \tilde{x}_1 \tilde{x}_2}\frac{\partial^2 f_b}{\partial \tilde{x}_2 \tilde{x}_1} &= 4\lambda_1\lambda_2 > 0
\end{aligned} \tag{C.4}$$

Therefore, to satisfy the second order conditions, I will require $\lambda_{1b} > 0$ and $\lambda_{2b} > 0$. For market maker 1:

$$\begin{aligned}
P_{1b}(\omega_{1b}) &= \mathbb{E}[v|\omega_{1b}] \\
&= \mathbb{E}[v|x_1 + z_1] \\
&= \mathbb{E}[v|\beta_{1b} v + z_1] \\
&= \frac{\beta_{1b}\sigma_v^2}{\beta_{1b}^2\sigma_v^2 + \sigma_{z_1}^2}\omega_1 \\
&\equiv \lambda_{1b}\omega_{1b}
\end{aligned} \tag{C.5}$$

For market maker 2:

$$
\begin{aligned}
P_{2b}(\omega_{2b}) &= \mathbb{E}[v|\omega_{2b}] \\
&= \mathbb{E}[v|x_2 + z_2] \\
&= \mathbb{E}[v|\beta_{2b}v + z_2] \\
&= \frac{\beta_{2b}\sigma_v^2}{\beta_{2b}^2\sigma_v^2 + \sigma_{z_2}^2}\omega_2 \\
&\equiv \lambda_{2b}\omega_{2b}
\end{aligned}
\tag{C.6}
$$

I have four equations and four unknowns:

$$
\begin{aligned}
\beta_{1b} &= \frac{1}{2\lambda_{1b}} \\
\lambda_{1b} &= \frac{\beta_{1b}\sigma_v^2}{\beta_{1b}^2\sigma_v^2 + \sigma_{z_1}^2} \\
\beta_{2b} &= \frac{1}{2\lambda_{2b}} \\
\lambda_{2b} &= \frac{\beta_{2b}\sigma_v^2}{\beta_{2b}^2\sigma_v^2 + \sigma_{z_2}^2},
\end{aligned}
$$

which is the classic Kyle (1985) setting yielding the following solutions:

$$
\begin{aligned}
\beta_{1b} &= \sqrt{\frac{\sigma_{z_1}^2}{\sigma_v^2}} \equiv \sqrt{\eta_1} \\
\lambda_{1b} &= \frac{1}{2}\sqrt{\frac{\sigma_v^2}{\sigma_{z_1}^2}} = \frac{1}{2\sqrt{\eta_1}} \\
\beta_{2b} &= \sqrt{\frac{\sigma_{z_2}^2}{\sigma_v^2}} \equiv \sqrt{\eta_2} \\
\lambda_{2b} &= \frac{1}{2}\sqrt{\frac{\sigma_v^2}{\sigma_{z_2}^2}} = \frac{1}{2\sqrt{\eta_2}}.
\end{aligned}
$$

Price informativeness for each Asset i would be measured as:

$$\begin{aligned}\text{Var}(v|P_{ib}) &= \text{Var}(v|\lambda_{ib}(x_{ib}+z_i)) \\ &= \text{Var}\big(v|\lambda_{ib}(\beta_{ib}v+z_i)\big) \\ &= \text{Var}\Big(v|\frac{1}{2}v+\frac{1}{2\sqrt{\eta_i}}z_i\Big) \\ &= \frac{1}{2}\sigma_v^2\end{aligned} \tag{C.7}$$

C.1.2 Economy a: With Arbitrageur

I now solve the model without law of one price violations. Once again, I assume that the strategic speculator's trading signal and the market makers' pricing rule is linear in the payoff and order flow respectively:

$$\begin{aligned}x_{1a} &= \beta_{1a}v \\ x_{2a} &= \beta_{2a}v \\ P_{1a} &= \lambda_{1a}\omega_{1a} \\ P_{2a} &= \lambda_{2a}\omega_{2a}\end{aligned}$$

I first derive the buy and sell order ε_a that will ensure that prices converge:

$$\begin{aligned}P_{1a} &= P_{2a} \\ &\Downarrow \\ \lambda_{1a}\omega_{1a} &= \lambda_{2a}\omega_{2a} \\ &\Downarrow \\ \lambda_1(x_{1a}+z_1+\varepsilon_a) &= \lambda_{2a}(x_{2a}+z_2-\varepsilon_a) \\ &\Downarrow \\ \varepsilon_a &= -\frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}x_{1a}+\frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}x_{2a}-\frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_1+\frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_2 \\ &\Downarrow \\ \varepsilon_a &= -\frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\beta_{1a}v+\frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\beta_{2a}v-\frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_1+\frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_2\end{aligned} \tag{C.8}$$

For the strategic speculator:

$$
\begin{aligned}
[x_{1a}, x_{2a}] &= \arg\max_{\tilde{x}_{1a},\tilde{x}_{2a}} \mathbb{E}[\tilde{x}_{1a}(v - P_{1a}(\omega_{1a})) + \tilde{x}_{2a}(v - P_{2a}(\omega_{2a}))|v] \\
&= \arg\max_{\tilde{x}_{1a},\tilde{x}_{2a}} \mathbb{E}[\tilde{x}_{1a}(v - \lambda_{1a}\omega_{1a}) + \tilde{x}_{2a}(v - \lambda_{2a}\omega_{2a})|v] \\
&= \arg\max_{\tilde{x}_{1a},\tilde{x}_{2a}} \mathbb{E}[\tilde{x}_{1a}(v - \lambda_{1a}(\tilde{x}_{1a} + z_1 + \varepsilon_a)) + \tilde{x}_{2a}(v - \lambda_{2a}(\tilde{x}_{2a} + z_2 - \varepsilon_a))|v] \\
&= \arg\max_{\tilde{x}_{1a},\tilde{x}_{2a}} \mathbb{E}\Bigg[\tilde{x}_{1a}\Big(v - \lambda_{1a}\Big(\tilde{x}_{1a} + z_1 - \tfrac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{1a} + \tfrac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{2a} - \tfrac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_1 + \tfrac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_2\Big)\Big) \\
&+ \tilde{x}_{2a}\Big(v - \lambda_{2a}\Big(\tilde{x}_{2a} + z_2 + \tfrac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{1a} - \tfrac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{2a} + \tfrac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_1 - \tfrac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_2\Big)\Big)|v\Bigg] \\
&= \arg\max_{\tilde{x}_{1a},\tilde{x}_{2a}} \Bigg(\tilde{x}_{1a}\Big(v - \lambda_{1a}\big(\tilde{x}_{1a} - \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{1a} + \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{2a}\big)\Big) \\
&+ \tilde{x}_{2a}\Big(v - \lambda_{2a}\big(\tilde{x}_{2a} + \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{1a} - \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{2a}\big)\Big)\Bigg) \\
&\equiv \arg\max_{\tilde{x}_{1a},\tilde{x}_{2a}} g(\tilde{x}_{1a}, \tilde{x}_{2a})
\end{aligned} \tag{C.9}
$$

First order condition #1:

$$
\begin{aligned}
\frac{\partial g(\tilde{x}_{1a}, \tilde{x}_{2a})}{\partial \tilde{x}_{1a}} &= \frac{\partial}{\partial \tilde{x}_{1a}}\Bigg(\tilde{x}_{1a}\Big(v - \frac{2\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{2a}\Big) - \tilde{x}_{1a}^2\Big(\lambda_{1a} - \frac{\lambda_{1a}^2}{\lambda_{1a}+\lambda_{2a}}\Big)\Bigg) \\
&= \Big(v - \frac{2\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{2a}\Big) - \frac{2\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{1a} = 0 \\
&\Downarrow \\
x_{1a} &= \frac{\lambda_{1a}+\lambda_{2a}}{2\lambda_{1a}\lambda_{2a}}v - \tilde{x}_{2a} = \frac{\lambda_{1a}+\lambda_{2a} - 2\lambda_{1a}\lambda_{2a}\beta_{2a}}{2\lambda_{1a}\lambda_{2a}}v \equiv \beta_{1a}v
\end{aligned} \tag{C.10}
$$

Similarly, first order condition #2:

$$
\begin{aligned}
\frac{\partial g(\tilde{x}_{1a}, \tilde{x}_{2a})}{\partial \tilde{x}_{2a}} &= \frac{\partial}{\partial \tilde{x}_{2a}}\Bigg(\tilde{x}_{2a}\Big(v - \frac{2\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{1a}\Big) - \tilde{x}_{2a}^2\Big(\lambda_{2a} - \frac{\lambda_{2a}^2}{\lambda_{1a}+\lambda_{2a}}\Big)\Bigg) \\
&= \Big(v - \frac{2\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{1a}\Big) - \frac{2\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\tilde{x}_{2a} = 0 \\
&\Downarrow \\
x_{2a} &= \frac{\lambda_{1a}+\lambda_{2a}}{2\lambda_{1a}\lambda_{2a}}v - \tilde{x}_{1a} = \frac{\lambda_{1a}+\lambda_{2a} - 2\lambda_{1a}\lambda_{2a}\beta_{1a}}{2\lambda_{1a}\lambda_{2a}}v \equiv \beta_{2a}v
\end{aligned} \tag{C.11}
$$

Finally, second order conditions would require that:

$$\begin{aligned}
\frac{\partial^2 g(\tilde{x}_{1a}, \tilde{x}_{2a})}{\partial \tilde{x}_{1a}^2} &= -\frac{2\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}} < 0 \\
\frac{\partial^2 g(\tilde{x}_{1a}, \tilde{x}_{2a})}{\partial \tilde{x}_{2a}^2} &= -\frac{2\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}} < 0 \\
\frac{\partial^2 g(\tilde{x}_{1a}, \tilde{x}_{2a})}{\partial \tilde{x}_{1a}^2}\frac{\partial^2 g(\tilde{x}_{1a}, \tilde{x}_{2a})}{\partial \tilde{x}_{2a}^2} - \frac{\partial^2 g(\tilde{x}_{1a}, \tilde{x}_{2a})}{\partial \tilde{x}_{1a}\tilde{x}_{2a}}\frac{\partial^2 g(\tilde{x}_{1a}, \tilde{x}_{2a})}{\partial \tilde{x}_{2a}\tilde{x}_{1a}} &= \frac{4\lambda_{1a}^2\lambda_{2a}^2}{(\lambda_{1a}+\lambda_{2a})^2} - \frac{4\lambda_{1a}^2\lambda_{2a}^2}{(\lambda_{1a}+\lambda_{2a})^2} \not> 0
\end{aligned} \tag{C.12}$$

Since the latter condition is not satisfied, I will demonstrate that I have found a non-unique local maximum by focusing on symmetric equilibria for which $\lambda_{1a} = \lambda_{2a} = \lambda_a$. In this case:

$$g(\tilde{x}_{1a}, \tilde{x}_{2a}) = -\frac{1}{2}\lambda(\tilde{x}_{1a} + \tilde{x}_{2a})^2 + (\tilde{x}_{1a} + \tilde{x}_{2a})v,$$

where, if I assume that $\lambda_a > 0$, then any cross-section of $g(\tilde{x}_{1a}, \tilde{x}_{2a})$, such that $\tilde{x}_{1a} + \tilde{x}_{2a} = c$, will be a downward parabola. Therefore, $g(\tilde{x}_{1a}, \tilde{x}_{2a})$ is a downwards pointing parabolic cylinder, implying that infinitely many solutions exist to the optimization. As long as

$$\tilde{x}_{1a} + \tilde{x}_{2a} = \frac{\lambda_{1a}+\lambda_{2a}}{2\lambda_{1a}\lambda_{2a}}v, \tag{C.13}$$

the speculator is maximizing her profits.

For market maker 1:

$$\begin{aligned}
P_{1a}(\omega_{1a}) &= \mathbb{E}[v|\omega_{1a}] \\
&= \mathbb{E}[v|x_{1a} + z_1 + \varepsilon_a] \\
&= \mathbb{E}[v|\beta_{1a}v + z_1 - \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\beta_{1a}v + \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\beta_{2a}v - \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_1 + \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_2] \\
&= \mathbb{E}[v|\Big(\beta_{1a} - \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\beta_{1a} + \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\beta_{2a}\Big)v + \Big(1 - \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\Big)z_1 + \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_2] \\
&= \mathbb{E}[v|\frac{(\beta_{1a}+\beta_{2a})\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}v + \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_1 + \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_2] \\
&= \frac{\frac{(\beta_{1a}+\beta_{2a})\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\sigma_v^2}{\frac{(\beta_{1a}+\beta_{2a})^2\lambda_{2a}^2}{(\lambda_{1a}+\lambda_{2a})^2}\sigma_v^2 + \frac{\lambda_{2a}^2}{(\lambda_{1a}+\lambda_{2a})^2}\sigma_{z_1}^2 + \frac{\lambda_{2a}^2}{(\lambda_{1a}+\lambda_{2a})^2}\sigma_{z_2}^2}\omega_{1a} \\
&= \frac{(\beta_{1a}+\beta_{2a})(\lambda_{1a}+\lambda_{2a})\lambda_{2a}\sigma_v^2}{(\beta_{1a}+\beta_{2a})^2\lambda_{2a}^2\sigma_v^2 + \lambda_{2a}^2(\sigma_{z_1}^2 + \sigma_{z_2}^2)}\omega_{1a} \\
&\equiv \lambda_{1a}\omega_{1a}
\end{aligned} \tag{C.14}$$

For market maker 2:

$$\begin{aligned}
P_{2a}(\omega_{2a}) &= \mathbb{E}[v|\omega_{2a}] \\
&= \mathbb{E}[v|x_{2a} + z_2 - \varepsilon_a] \\
&= \mathbb{E}[v|\beta_{2a}v + z_2 + \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\beta_{1a}v - \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\beta_{2a}v + \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_1 - \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_2] \\
&= \mathbb{E}[v|\Big(\beta_{2a} + \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\beta_{1a} - \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\beta_{2a}\Big)v + \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_1 + \Big(1 - \frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\Big)z_2] \\
&= \mathbb{E}[v|\frac{(\beta_{1a}+\beta_{2a})\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}v + \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_1 + \frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_2] \\
&= \frac{\frac{(\beta_{1a}+\beta_{2a})\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}\sigma_v^2}{\frac{(\beta_{1a}+\beta_{2a})^2\lambda_{1a}^2}{(\lambda_{1a}+\lambda_{2a})^2}\sigma_v^2 + \frac{\lambda_{1a}^2}{(\lambda_{1a}+\lambda_{2a})^2}\sigma_{z_1}^2 + \frac{\lambda_{1a}^2}{(\lambda_{1a}+\lambda_{2a})^2}\sigma_{z_2}^2}\omega_{2a} \\
&= \frac{(\beta_{1a}+\beta_{2a})(\lambda_{1a}+\lambda_{2a})\lambda_{1a}\sigma_v^2}{(\beta_{1a}+\beta_{2a})^2\lambda_{1a}^2\sigma_v^2 + \lambda_{1a}^2(\sigma_{z_1}^2+\sigma_{z_2}^2)}\omega_{2a} \\
&\equiv \lambda_{2a}\omega_{2a}
\end{aligned} \tag{C.15}$$

I have four equations and four unknowns:

$$\beta_{1a} = \frac{\lambda_{1a} + \lambda_{2a} - 2\lambda_{1a}\lambda_{2a}\beta_{2a}}{2\lambda_{1a}\lambda_{2a}} \tag{C.16}$$

$$\beta_{2a} = \frac{\lambda_{1a} + \lambda_{2a} - 2\lambda_{1a}\lambda_{2a}\beta_{1a}}{2\lambda_{1a}\lambda_{2a}} \tag{C.17}$$

$$\lambda_{1a} = \frac{(\beta_{1a}+\beta_{2a})(\lambda_{1a}+\lambda_{2a})\lambda_{2a}\sigma_v^2}{(\beta_{1a}+\beta_{2a})^2\lambda_{2a}^2\sigma_v^2 + \lambda_{2a}^2(\sigma_{z_1}^2+\sigma_{z_2}^2)} \tag{C.18}$$

$$\lambda_{2a} = \frac{(\beta_{1a}+\beta_{2a})(\lambda_{1a}+\lambda_{2a})\lambda_{1a}\sigma_v^2}{(\beta_{1a}+\beta_{2a})^2\lambda_{1a}^2\sigma_v^2 + \lambda_{1a}^2(\sigma_{z_1}^2+\sigma_{z_2}^2)} \tag{C.19}$$

First, note that Eq. (C.16) and (C.17) are equivalent, yielding:

$$\beta_{1a} + \beta_{2a} = \frac{\lambda_{1a} + \lambda_{2a}}{2\lambda_{1a}\lambda_{2a}} \tag{C.20}$$

Next, note that Eq. (C.18) and (C.19) are equivalent, yielding:

$$1 = (\beta_{1a}+\beta_{2a})\frac{\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}} + \frac{1}{\beta_{1a}+\beta_{2a}}\frac{\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}\frac{\sigma_{z_1}^2+\sigma_{z_2}^2}{\sigma_v^2} \tag{C.21}$$

Substituting Eq. (C.20) into Eq. (C.21) I obtain:

$$
\begin{aligned}
1 &= \frac{1}{(\beta_{1a}+\beta_{2a})^2}\frac{\sigma_{z_1}^2+\sigma_{z_2}^2}{\sigma_v^2} \\
&\Downarrow \\
\beta_{1a}+\beta_{2a} &= \pm\sqrt{\eta_1+\eta_2} \\
&\Downarrow \text{(assuming symmetric equilibria)} \\
\lambda_a \equiv \lambda_{1a} = \lambda_{2a} = &= \pm\frac{1}{\sqrt{\eta_1+\eta_2}}
\end{aligned}
\tag{C.22}
$$

I can derive the inverse of price informativeness for Market 1 as:

$$
\begin{aligned}
\mathrm{Var}(v|P_{1a}) &= \mathrm{Var}(v|\lambda_{1a}(x_{1a}+z_1+\varepsilon_a)) \\
&= \mathrm{Var}\left(v\middle|\lambda_{1a}\Big(\frac{(\beta_{1a}+\beta_{2a})\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}v+\frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_1+\frac{\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}z_2\Big)\right) \\
&= \mathrm{Var}\left(v\middle|\frac{(\beta_{1a}+\beta_{2a})\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}v+\frac{\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}(z_1+z_2)\right) \\
&= \mathrm{Var}\left(v\middle|\frac{1}{2}v+\frac{1}{2(\beta_{1a}+\beta_{2a})}(z_1+z_2)\right) \\
&= \frac{1}{2}\sigma_v^2
\end{aligned}
\tag{C.23}
$$

I can derive the inverse of price informativeness for Market 2 as:

$$
\begin{aligned}
\mathrm{Var}(v|P_{2a}) &= \mathrm{Var}(v|\lambda_{2a}(x_{2a}+z_2-\varepsilon_a)) \\
&= \mathrm{Var}\left(v\middle|\lambda_{2a}\Big(\frac{(\beta_{1a}+\beta_{2a})\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}v+\frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_1+\frac{\lambda_{1a}}{\lambda_{1a}+\lambda_{2a}}z_2\Big)\right) \\
&= \mathrm{Var}\left(v\middle|\frac{(\beta_{1a}+\beta_{2a})\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}v+\frac{\lambda_{1a}\lambda_{2a}}{\lambda_{1a}+\lambda_{2a}}(z_1+z_2)\right) \\
&= \mathrm{Var}\left(v\middle|\frac{1}{2}v+\frac{1}{2(\beta_{1a}+\beta_{2a})}(z_1+z_2)\right) \\
&= \frac{1}{2}\sigma_v^2
\end{aligned}
\tag{C.24}
$$

C.1.3 Economy s: With Fully Segmented Markets and Arbitrageur

The derivation of ε_s and functional form assumptions are the same as above:

$$\begin{aligned}
x_{1s} &= \beta_{1s} v \\
x_{2s} &= \beta_{2s} v \\
P_{1s} &= \lambda_{1s}\omega_{1s} \\
P_{2s} &= \lambda_{2s}\omega_{2s}
\end{aligned}$$

By law of one price it must be that:

$$\begin{aligned}
P_{1s} &= P_{2s} \\
&\Downarrow \\
\lambda_{1s}\omega_{1s} &= \lambda_{2s}\omega_{2s} \\
&\Downarrow \\
\lambda_{1s}(x_{1s} + z_1 + \varepsilon_s) &= \lambda_{2s}(x_{2s} + z_2 - \varepsilon_s) \\
&\Downarrow \\
\varepsilon_s &= -\frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}} x_{1s} + \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}} x_{2s} - \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}} z_1 + \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}} z_2 \\
&\Downarrow \\
\varepsilon_s &= -\frac{\lambda_{1s}}{\lambda_{s1}+\lambda_{2s}} \beta_{1s} v + \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}} \beta_{2s} v - \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}} z_1 + \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}} z_2
\end{aligned} \tag{C.25}$$

For Speculator 1:

$$\begin{aligned}
x_{1s} &= \arg\max_{\tilde{x}_{1s}} \mathbb{E}[\tilde{x}_{1s}(v - P_{1s}(\omega_{1s}))|v] \\
&= \arg\max_{\tilde{x}_{1s}} \mathbb{E}[\tilde{x}_{1s}(v - \lambda_{1s}\omega_{1s})|v] \\
&= \arg\max_{\tilde{x}_{1s}} \mathbb{E}[\tilde{x}_{1s}(v - \lambda_{1s}(\tilde{x}_{1s} + z_1 + \varepsilon_s))|v] \\
&= \arg\max_{\tilde{x}_{1s}} \mathbb{E}[\tilde{x}_{1s}\Big(v - \lambda_{1s}\Big(\tilde{x}_{1s} + z_1 - \tfrac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}\tilde{x}_{1s} + \tfrac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\beta_{2s} v - \tfrac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}} z_1 + \tfrac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}} z_2\Big)\Big)|v] \qquad \text{(C.26)} \\
&= \arg\max_{\tilde{x}_{1s}} \Big\{\tilde{x}_{1s}\Big(v - \frac{\lambda_{1s}\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\beta_{2s} v\Big) - \tilde{x}_{1s}^2\Big(\lambda_{1s} - \frac{\lambda_{1s}^2}{\lambda_{1s}+\lambda_{2s}}\Big)\Big\} \\
&= \arg\max_{\tilde{x}_{1s}} \Big\{\tilde{x}_{1s}\frac{\lambda_{1s}+\lambda_{2s}-\lambda_{1s}\lambda_{2s}\beta_{2s}}{\lambda_{1s}+\lambda_{2s}} v - \tilde{x}_{1s}^2 \frac{\lambda_{1s}\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\Big\}
\end{aligned}$$

First order condition for Speculator 1:

$$\begin{aligned}
\tilde{x}_{1s}\frac{2\lambda_{1s}\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}} &= \frac{\lambda_{1s}+\lambda_{2s}-\lambda_{1s}\lambda_{2s}\beta_{2s}}{\lambda_{1s}+\lambda_{2s}}v \\
&\Downarrow \\
x_{1s} &= \frac{\lambda_{1s}+\lambda_{2s}-\lambda_{1s}\lambda_{2s}\beta_{2s}}{2\lambda_{1s}\lambda_{2s}}v \\
&\equiv \beta_{1s}v
\end{aligned} \tag{C.27}$$

For Speculator 2:

$$\begin{aligned}
x_{2s} &= \arg\max_{\tilde{x}_{2s}} \mathbb{E}[\tilde{x}_{2s}(v-P_{2s}(\omega_{2s}))|v] \\
&= \arg\max_{\tilde{x}_{2s}} \mathbb{E}[\tilde{x}_{2s}(v-\lambda_{2s}\omega_{2s})|v] \\
&= \arg\max_{\tilde{x}_{2s}} \mathbb{E}[\tilde{x}_{2s}(v-\lambda_{2s}(\tilde{x}_{2s}+z_2-\varepsilon_s))|v] \\
&= \arg\max_{\tilde{x}_{2s}} \mathbb{E}[\tilde{x}_{2s}\Big(v-\lambda_{2s}\Big(\tilde{x}_{2s}+z_2+\tfrac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}\beta_{1s}v-\tfrac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\tilde{x}_{2s}+\tfrac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}z_1-\tfrac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}z_2\Big)\Big)|v] \\
&= \arg\max_{\tilde{x}_{2s}} \Big\{\tilde{x}_{2s}\Big(v-\frac{\lambda_{1s}\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\beta_{1s}v\Big)-\tilde{x}_{2s}^2\Big(\lambda_{2s}-\frac{\lambda_{2s}^2}{\lambda_{1s}+\lambda_{2s}}\Big)\Big\} \\
&= \arg\max_{\tilde{x}_{2s}} \Big\{\tilde{x}_{2s}\frac{\lambda_{1s}+\lambda_{2s}-\lambda_{1s}\lambda_{2s}\beta_{1s}}{\lambda_{1s}+\lambda_{2s}}v-\tilde{x}_{2s}^2\frac{\lambda_{1s}\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\Big\}
\end{aligned} \tag{C.28}$$

First order condition for Speculator 2:

$$\begin{aligned}
\tilde{x}_{2s}\frac{2\lambda_{1s}\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}} &= \frac{\lambda_{1s}+\lambda_{2s}-\lambda_{1s}\lambda_{2s}\beta_{1s}}{\lambda_{1s}+\lambda_{2s}}v \\
&\Downarrow \\
x_{2s} &= \frac{\lambda_{1s}+\lambda_{2s}-\lambda_{1s}\lambda_{2s}\beta_{1s}}{2\lambda_{1s}\lambda_{2s}}v \\
&\equiv \beta_{2s}v
\end{aligned} \tag{C.29}$$

For market maker 1:

$$\begin{aligned}
P_{1s}(\omega_{1s}) &= \mathbb{E}[v|\omega_{1s}] \\
&= \mathbb{E}[v|x_{1s} + z_1 + \varepsilon_s] \\
&= \mathbb{E}[v|\beta_{1s}v + z_1 - \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}\beta_{1s}v + \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\beta_{2s}v - \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}z_1 + \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}z_2] \\
&= \mathbb{E}[v|\Big(\beta_{1s} - \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}\beta_{1s} + \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\beta_{2s}\Big)v + \Big(1 - \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}\Big)z_1 + \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}z_2] \\
&= \mathbb{E}[v|\frac{(\beta_{1s}+\beta_{2s})\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}v + \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}z_1 + \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}z_2] \\
&= \frac{\frac{(\beta_{1s}+\beta_{2s})\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\sigma_v^2}{\frac{(\beta_{1s}+\beta_{2s})^2\lambda_{2s}^2}{(\lambda_{1s}+\lambda_{2s})^2}\sigma_v^2 + \frac{\lambda_{2s}^2}{(\lambda_{1s}+\lambda_{2s})^2}\sigma_{z_1}^2 + \frac{\lambda_{2s}^2}{(\lambda_{1s}+\lambda_{2s})^2}\sigma_{z_2}^2}\omega_{1s} \\
&= \frac{(\beta_{1s}+\beta_{2s})(\lambda_{1s}+\lambda_{2s})\lambda_{2s}\sigma_v^2}{(\beta_{1s}+\beta_{2s})^2\lambda_{2s}^2\sigma_v^2 + \lambda_{2s}^2(\sigma_{z_1}^2+\sigma_{z_2}^2)}\omega_{1s} \\
&\equiv \lambda_{1s}\omega_{1s}
\end{aligned} \tag{C.30}$$

For market maker 2:

$$\begin{aligned}
P_{2s}(\omega_{2s}) &= \mathbb{E}[v|\omega_{2s}] \\
&= \mathbb{E}[v|x_{2s} + z_2 - \varepsilon_s] \\
&= \mathbb{E}[v|\beta_{2s}v + z_2 + \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}\beta_{1s}v - \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\beta_{2s}v + \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}z_1 - \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}z_2] \\
&= \mathbb{E}[v|\Big(\beta_{2s} + \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}\beta_{1s} - \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\beta_{2s}\Big)v + \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}z_1 + \Big(1 - \frac{\lambda_{2s}}{\lambda_{1s}+\lambda_{2s}}\Big)z_2] \\
&= \mathbb{E}[v|\frac{(\beta_{1s}+\beta_{2s})\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}v + \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}z_1 + \frac{\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}z_2] \\
&= \frac{\frac{(\beta_{1s}+\beta_{2s})\lambda_{1s}}{\lambda_{1s}+\lambda_{2s}}\sigma_v^2}{\frac{(\beta_{1s}+\beta_{2s})^2\lambda_{1s}^2}{(\lambda_{1s}+\lambda_{2s})^2}\sigma_v^2 + \frac{\lambda_{1s}^2}{(\lambda_{1s}+\lambda_{2s})^2}\sigma_{z_1}^2 + \frac{\lambda_{1s}^2}{(\lambda_{1s}+\lambda_{2s})^2}\sigma_{z_2}^2}\omega_{2s} \\
&= \frac{(\beta_{1s}+\beta_{2s})(\lambda_{1s}+\lambda_{2s})\lambda_{1s}\sigma_v^2}{(\beta_{1s}+\beta_{2s})^2\lambda_{1s}^2\sigma_v^2 + \lambda_{1s}^2(\sigma_{z_1}^2+\sigma_{z_2}^2)}\omega_{2s} \\
&\equiv \lambda_{2s}\omega_{2s}
\end{aligned} \tag{C.31}$$

I have four equations and four unknowns:

$$\beta_{1s} = \frac{\lambda_{1s} + \lambda_{2s} - \lambda_{1s}\lambda_{2s}\beta_{2s}}{2\lambda_{1s}\lambda_{2s}} \tag{C.32}$$

$$\beta_{2s} = \frac{\lambda_{1s} + \lambda_{2s} - \lambda_{1s}\lambda_{2s}\beta_{1s}}{2\lambda_{1s}\lambda_{2s}} \tag{C.33}$$

$$\lambda_{1s} = \frac{(\beta_{1s} + \beta_{2s})(\lambda_{1s} + \lambda_{2s})\lambda_{2s}\sigma_v^2}{(\beta_{1s} + \beta_{2s})^2\lambda_{2s}^2\sigma_v^2 + \lambda_{2s}^2(\sigma_{z_1}^2 + \sigma_{z_2}^2)} \tag{C.34}$$

$$\lambda_{2s} = \frac{(\beta_{1s} + \beta_{2s})(\lambda_{1s} + \lambda_{2s})\lambda_{1s}\sigma_v^2}{(\beta_{1s} + \beta_{2s})^2\lambda_{1s}^2\sigma_v^2 + \lambda_{1s}^2(\sigma_{z_1}^2 + \sigma_{z_2}^2)} \tag{C.35}$$

First, taking equations C.32 and C.33, solving for $\lambda_1 + \lambda_2$ and setting the results equal, gets us:

$$\begin{gathered} 2\beta_{1s} + \beta_{2s} = \beta_{1s} + 2\beta_{2s} \\ \Downarrow \\ \beta_{1s} = \beta_{2s} \equiv \beta_s \end{gathered} \tag{C.36}$$

Plugging β into Eq. (C.32):

$$\frac{\lambda_{1s} + \lambda_{2s}}{\lambda_{1s}\lambda_{2s}} = 3\beta_s \tag{C.37}$$

Note that Eq. (C.34) and (C.35) are equivalent. Let's use the first one and substitute Eq. (C.37):

$$\begin{aligned} &1 = (\beta_{1s} + \beta_{2s})\frac{\lambda_{1s}\lambda_{2s}}{\lambda_{1s} + \lambda_{2s}} + \frac{1}{\beta_{1s} + \beta_{2s}}\frac{\lambda_{1s}\lambda_{2s}}{\lambda_{1s} + \lambda_{2s}}\frac{\sigma_{z_1}^2 + \sigma_{z_2}^2}{\sigma_v^2} \\ &\Downarrow \\ &\frac{1}{3} = \frac{1}{6\beta^2}(\eta_1 + \eta_2) \\ &\Downarrow \\ &\beta_s = \pm\sqrt{\frac{\eta_1 + \eta_2}{2}} \end{aligned} \tag{C.38}$$

Using Eq. (C.37), there are infinitely many solutions. Assume a symmetric equilibrium $\lambda_{1s} = \lambda_{2s} = \lambda_s$ to obtain:

$$\begin{aligned} \frac{2}{\lambda_s} &= \pm 3\sqrt{\frac{\eta_1 + \eta_2}{2}} \\ &\Downarrow \\ \lambda_s &= \pm \frac{2}{3}\sqrt{\frac{2}{\eta_1 + \eta_2}} \end{aligned} \tag{C.39}$$

I can derive the inverse of price informativeness as:

$$\begin{aligned} \mathrm{Var}(v|P_{1s}) &= \mathrm{Var}(v|\lambda_{1s}(x_{1s} + z_1 + \varepsilon_s)) \\ &= \mathrm{Var}\left(v\middle|\lambda_{1s}\Big(\frac{(\beta_{1s} + \beta_{2s})\lambda_{2s}}{\lambda_{1s} + \lambda_{2s}}v + \frac{\lambda_{2s}}{\lambda_{1s} + \lambda_{2s}}z_1 + \frac{\lambda_{2s}}{\lambda_{1s} + \lambda_{2s}}z_2\Big)\right) \\ &= \mathrm{Var}\left(v\middle|\frac{(\beta_{1s} + \beta_{2s})\lambda_{1s}\lambda_{2s}}{\lambda_{1s} + \lambda_{2s}}v + \frac{\lambda_{1s}\lambda_{2s}}{\lambda_{1s} + \lambda_{2s}}(z_1 + z_2)\right) \\ &= \mathrm{Var}\left(v\middle|\frac{2}{3}v + \frac{1}{3}\sqrt{\frac{2\sigma_v^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2}}(z_1 + z_2)\right) \\ &= \frac{1}{3}\sigma_v^2, \end{aligned} \tag{C.40}$$

which will be identical for Asset 2, since their prices are the same by construction.

C.1.4 Economy m: With Segmented Speculators and Single Market Maker

In this case, there is no arbitrageur, and, instead, the market maker uses order flow in both markets to set prices, which will be the same since they reflect the full information set of the market maker. I will have the following demand and price functions

$$x_{1m} = \beta_{1m}v \tag{C.41}$$

$$x_{2m} = \beta_{2m}v \tag{C.42}$$

$$P_m = \lambda_{1m}\omega_{1m} + \lambda_{2m}\omega_{2m}. \tag{C.43}$$

For Speculator 1:

$$\begin{aligned}
x_{1m} &= \arg\max_{\tilde{x}_{1m}} \mathbb{E}[\tilde{x}_{1m}(v - P_m(\omega_{1m}, \omega_{2m}))|v] \\
&= \arg\max_{\tilde{x}_{1m}} \mathbb{E}[\tilde{x}_{1m}(v - \lambda_{1m}\omega_{1m} - \lambda_{2m}\omega_{2m})|v] \\
&= \arg\max_{\tilde{x}_{1m}} \mathbb{E}[\tilde{x}_{1m}(v - \lambda_{1m}(\tilde{x}_{1m} + z_1) - \lambda_{2m}(\tilde{x}_{2m} + z_2))|v] \\
&= \arg\max_{\tilde{x}_{1m}} \mathbb{E}[\tilde{x}_{1m}(v - \lambda_{1m}(\tilde{x}_{1m} + z_1) - \lambda_{2m}(\beta_{2m}v + z_2))|v] \\
&= \arg\max_{\tilde{x}_{1m}} \left\{\tilde{x}_{1m}(v - \lambda_{2m}\beta_{2m}v) - \lambda_{1m}\tilde{x}_{1m}^2\right\}
\end{aligned} \tag{C.44}$$

First order condition for Speculator 1:

$$\begin{aligned}
2\lambda_{1m}\tilde{x}_{1m} &= v - \lambda_{2m}\beta_{2m}v \\
&\Downarrow \\
x_{1m} &= \frac{1 - \lambda_{2m}\beta_{2m}}{2\lambda_{1m}}v \\
&\equiv \beta_{1m}v
\end{aligned} \tag{C.45}$$

Similarly, for Speculator 2:

$$\begin{aligned}
x_{2m} &= \arg\max_{\tilde{x}_{2m}} \mathbb{E}[\tilde{x}_{2m}(v - P_m(\omega_{1m}, \omega_{2m}))|v] \\
&= \arg\max_{\tilde{x}_{2m}} \mathbb{E}[\tilde{x}_{2m}(v - \lambda_{1m}\omega_{1m} - \lambda_{2m}\omega_{2m})|v] \\
&= \arg\max_{\tilde{x}_{2m}} \mathbb{E}[\tilde{x}_{2m}(v - \lambda_{1m}(\tilde{x}_{1m} + z_1) - \lambda_{2m}(\tilde{x}_{2m} + z_2))|v] \\
&= \arg\max_{\tilde{x}_{2m}} \mathbb{E}[\tilde{x}_{2m}(v - \lambda_{1m}(\beta_{1m}v + z_1) - \lambda_{2m}(\tilde{x}_{2m} + z_2))|v] \\
&= \arg\max_{\tilde{x}_{2m}} \left\{\tilde{x}_{2m}(v - \lambda_{1m}\beta_{1m}v) - \lambda_{2m}\tilde{x}_{2m}^2\right\}
\end{aligned} \tag{C.46}$$

First order condition for Speculator 2:

$$\begin{aligned}
2\lambda_{2m}\tilde{x}_{2m} &= v - \lambda_{1m}\beta_{1m}v \\
&\Downarrow \\
x_{2m} &= \frac{1 - \lambda_{1m}\beta_{1m}}{2\lambda_{2m}}v \\
&\equiv \beta_{2m}v
\end{aligned} \tag{C.47}$$

The market maker breaks even in expectation:

$$\begin{aligned}
P_m &= \mathbb{E}[v|\omega_{1m}, \omega_{2m}] \\
&= \mathrm{Cov}(v, [\omega_{1m}, \omega_{2m}]) \begin{bmatrix} \sigma^2_{\omega_{1m}} & \sigma_{\omega_{1m}\omega_{2m}} \\ \sigma_{\omega_{1m}\omega_{2m}} & \sigma^2_{\omega_{2m}} \end{bmatrix}^{-1} \begin{bmatrix} \omega_{1m} \\ \omega_{2m} \end{bmatrix} \\
&(\text{using } \omega_{1m} = \beta_{1m} v + z_1, \omega_{2m} = \beta_{2m} v + z_2) \qquad (C.48) \\
&= \frac{\beta_{1m}\sigma_v^2\sigma_{z_2}^2}{\beta_{1m}^2\sigma_v^2\sigma_{z_2}^2 + \beta_{2m}^2\sigma_v^2\sigma_{z_1}^2 + \sigma_{z_1}^2\sigma_{z_2}^2}\omega_{1m} + \frac{\beta_{2m}\sigma_v^2\sigma_{z_1}^2}{\beta_{1m}^2\sigma_v^2\sigma_{z_2}^2 + \beta_{2m}^2\sigma_v^2\sigma_{z_1}^2 + \sigma_{z_1}^2\sigma_{z_2}^2}\omega_{2m} \\
&\equiv \lambda_{1m}\omega_{1m} + \lambda_{2m}\omega_{2m}
\end{aligned}$$

I have four equations and four unknowns:

$$\beta_{1m} = \frac{1 - \lambda_{2m}\beta_{2m}}{2\lambda_{1m}} \qquad (C.49)$$

$$\beta_{2m} = \frac{1 - \lambda_{1m}\beta_{1m}}{2\lambda_{2m}} \qquad (C.50)$$

$$\lambda_{1m} = \frac{\beta_{1m}\sigma_v^2\sigma_{z_2}^2}{\beta_{1m}^2\sigma_v^2\sigma_{z_2}^2 + \beta_{2m}^2\sigma_v^2\sigma_{z_1}^2 + \sigma_{z_1}^2\sigma_{z_2}^2} \qquad (C.51)$$

$$\lambda_{2m} = \frac{\beta_{2m}\sigma_v^2\sigma_{z_1}^2}{\beta_{1m}^2\sigma_v^2\sigma_{z_2}^2 + \beta_{2m}^2\sigma_v^2\sigma_{z_1}^2 + \sigma_{z_1}^2\sigma_{z_2}^2} \qquad (C.52)$$

I first substitute Eq. (C.50) into Eq. (C.49) and vice versa to obtain:

$$\frac{1}{\lambda_{1m}} = 3\beta_{1m} \qquad (C.53)$$

$$\frac{1}{\lambda_{2m}} = 3\beta_{2m} \qquad (C.54)$$

$$(C.55)$$

Next I take the reciprocal of Eq. (C.51) and (C.52):

$$\frac{1}{\lambda_{1m}} = \beta_{1m} + \frac{\beta_{2m}^2}{\beta_{1m}}\frac{\eta_1}{\eta_2} + \frac{\eta_1}{\beta_{1m}} \qquad (C.56)$$

$$\frac{1}{\lambda_{2m}} = \beta_{2m} + \frac{\beta_{1m}^2}{\beta_{2m}}\frac{\eta_2}{\eta_1} + \frac{\eta_2}{\beta_{2m}} \qquad (C.57)$$

$$(C.58)$$

Set the above equal to Eq. (C.53) and (C.54), to obtain:

$$\begin{aligned} 2\beta_{1m}^2 &= \beta_{2m}^2 \frac{\eta_1}{\eta_2} + \eta_1 \\ 2\beta_{2m}^2 &= \beta_{1m}^2 \frac{\eta_2}{\eta_1} + \eta_2, \end{aligned} \tag{C.59}$$

which yields:

$$\beta_{1m} = \pm\sqrt{\eta_1} \tag{C.60}$$

$$\beta_{2m} = \pm\sqrt{\eta_2} \tag{C.61}$$

$$\lambda_{1m} = \pm\frac{1}{3\sqrt{\eta_1}} \tag{C.62}$$

$$\lambda_{2m} = \pm\frac{1}{3\sqrt{\eta_2}} \tag{C.63}$$

I derive the inverse of price informativeness as:

$$\begin{aligned} \mathrm{Var}(v|P_m) &= \mathrm{Var}(v|\lambda_{1m}(x_{1m} + z_1) + \lambda_{2m}(x_{2m} + z_2)) \\ &= \mathrm{Var}(v|\lambda_{1m}(\beta_{1m}v + z_1) + \lambda_{2m}(\beta_{2m}v + z_2)) \\ &= \mathrm{Var}\left(v\middle|\frac{2}{3}v + \frac{1}{3\sqrt{\eta_1}}z_1 + \frac{1}{3\sqrt{\eta_2}}z_2\right) \\ &= \frac{1}{3}\sigma_v^2 \end{aligned} \tag{C.64}$$

BIBLIOGRAPHY

BIBLIOGRAPHY

Abis, S. (2020). Man vs. Machine: Quantitative and Discretionary Equity Management. *Working Paper*.

Admati, A. R. (1985). A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica 53*(3), 629–658.

Admati, A. R. and P. Pfleiderer (1987). Viable allocations of information in financial markets. *Journal of Economic Theory 43*, 76–115.

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

Bai, J., T. Philippon, and A. Savov (2016). Have financial markets become more informative? *Journal of Financial Economics 122*, 625–654.

Baker, M., L. Litov, J. A. Wachter, and J. Wurgle (2010). Can Mutual Fund Managers Pick Stocks? Evidence from Their Trades Prior to Earnings Announcement. *Journal of Financial and Quantitative Analysis 45*(5), 1111–1131.

Banerjee, S., J. Davis, and N. Gondhi (2018). When Transparency Improves, Must Prices Reflect Fundamentals Better? *Review of Financial Studies 31*(6), 2377–2414.

Beggs, W., J. Brogaard, and A. Hill-Kleespie (2019). Quantitative Investing and Market Instability. *Working Paper*.

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

Berk, J. B. and R. C. Green (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy 112*(6), 1269–1295.

Berk, J. B. and J. H. van Binsbergen (2015). Measuring Skill in the Mutual Fund Industry. *Journal of Financial Economics 118*(1), 1–20.

Bhattacharya, A. and M. O'Hara (2018). Can ETFs Increase Market Fragility? Effect of Information Linkages in ETF Markets. *Working Paper*.

Bhattacharya, N., Y. J. Cho, and J. B. Kim (2018). Leveling the Playing Field between Large and Small Institutions: Evidence from the SEC?s XBRL Mandate. *The Accounting Review 93*(5), 51–71.

Biais, B., P. Bossaerts, and C. Spatt (2010). Equilibrium Asset Pricing and Portfolio Choice Under Asymmetric Information. *Review of Financial Studies 23*(4), 1503–1543.

Blankespoor, E., B. P. Miller, and H. D. White (2014). Initial evidence on the market impact of the XBRL mandate. *Review of Accounting Studies 19*, 1468–1503.

Boehmer, E. and E. K. Kelley (2009). Institutional investors and the informational efficiency of prices. *Review of Financial Studies 22*(9), 3563–3594.

Bond, P., A. Edmans, and I. Goldstein (2012). The Real Effects of Financial Markets. *Annual Review of Financial Economics 4*, 339–360.

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

Brogaard, J., T. Hendershott, and R. Riordan (2014). High-Frequency Trading and Price Discovery. *Review of Financial Studies 27*(8), 2267–2306.

Brogaard, J., H. Nguyen, T. J. Putnins, and E. Wu (2019). What moves stock prices? The role of news, noise, and information. *Working Paper*.

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

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance 52*(1), 57–82.

Chakrabarty, B., P. C. Moulton, and X. Wang (2019). Attention Effects in a High-Frequency World. *Working Paper*.

Coles, J. L., D. Heath, and M. C. Ringgenberg (2018). On Index Investing. *Working Paper*.

Cong, L. W. and D. Xu (2016). Rise of Factor Investing: Asset Prices, Informational Efficiency, and Security Design. *Working Paper*.

Coval, J. D. and T. J. Moskowitz (2001). The Geography of Investment: Informed Trading and Asset Prices. *Journal of Political Economy 109*(4), 811–841.

Cremers, M., M. A. Ferreira, P. Matos, and L. Starks (2016). Indexing and active fund management: International evidence. *Journal of Financial Economics 120*, 539–560.

Davila, E. and C. Parlatore (2019). Identifying price informativeness. *Working Paper*.

Diamond, D. W. and R. E. Verrecchia (1981). Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics 9*, 221–235.

Dong, Y., O. Z. Li, Y. Lin, and C. Ni (2016). Does Information-Processing Cost Affect Firm-Specific Information Acquisition? Evidence from XBRL Adoption. *Journal of Financial and Quantitative Analysis 51*(2), 435–462.

Doshi, H., R. Elkamhi, and M. Simutin (2015). Managerial Activeness and Mutual Fund Performance. *Review of Asset Pricing Studies 5*(2), 156–184.

Duarte, J. and L. Young (2009). Why is PIN priced? *Journal of Financial Economics 91*, 119–138.

Dugast, J. and T. Foucault (2018). Data abundance and asset price informativeness. *Journal of Financial Economics 130*, 367–391.

Easley, D., N. M. Kiefer, M. O'Hara, and J. B. Paperman (1996). Liquidity, Information, and Infrequently Traded Stocks. *The Journal of Finance 51*(4), 1405–1436.

Easley, D., D. Michayluk, M. O?Hara, and T. J. Putnins (2020). The Active World of Passive Investing. *Working Paper*.

Fama, E. F. and K. R. French (2010). Luck versus Skill in the Cross-Section of Mutual Fund Returns. *The Journal of Finance 65*(5), 1915–1947.

Farboodi, M., A. Matray, and L. Veldkamp (2018). Where Has All the Big Data Gone? *Working Paper*.

Foster, F. D. and S. Viswanathan (1996). Strategic Trading When Agents Forecast the Forecasts of Others. *The Journal of Finance 51*(4), 1437–1478.

French, K. R. (2008). Presidential Address: The Cost of Active Investing. *The Journal of Finance 63*(4), 1537–1573.

Froot, K., N. Kang, G. Ozik, and R. Sadka (2017). What do measures of real-time corporate sales say about earnings surprises and post-announcement returns? *Journal of Financial Economics 125*(1), 143–162.

Garleanu, N. and L. H. Pedersen (2017). Efficiently inefficient markets for assets and assets under management. *The Journal of Finance 78*(4), 1663–1712.

Garleanu, N. and L. H. Pedersen (2020). Active and Passive Investing. *Working Paper*.

Glosten, L. R. and L. E. Harris (1988). Estimating the components of the bid / ask spread. *Journal of Financial Economics 21*, 123–142.

Glosten, L. R., S. Nallareddy, and Y. Zou (2017). ETF Activity and Informational Efficiency of Underlying Securities. *Working Paper*.

Gorton, G. B. and G. G. Pennacchi (1993). Security Baskets and Index-linked Securities. *66*(1), 1–28.

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

Hasbrouck, J. (1988). Trades, quotes, inventories, and information. *Journal of Financial Economics*, 229–252.

Hasbrouck, J. (1991). Measuring the Information Content of Stock Trades. *The Journal of Finance 46*(1), 179–207.

Hellwig, M. F. (1980). On the aggregation of information in competitive markets. *Journal of Economic Theory 22*, 447–498.

Hendershott, T., C. M. Jones, and A. J. Menkveld (2011). Does Algorithmic Trading Improve Liquidity? *The Journal of Finance 66*(1), 1–33.

Hendershott, T. and R. Riordan (2013). Algorithmic Trading and the Market for Liquidity. *Journal of Financial and Quantitative Analysis 48*(4), 1001–1024.

Holden, C. W. and S. Jacobsen (2014). Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions. *The Journal of Finance 69*(4), 1747–1785.

Holden, C. W. and A. Subrahmanyam (1992). Long-Lived Private Information and Imperfect Competition. *The Journal of Finance 47*(1), 247–270.

Huang, S., M. O?Hara, and Z. Zhong (2021). Innovation and Informed Trading: Evidence from Industry ETFs. *Review of Financial Studies 34*(3), 1280–1316.

Investment Company Institute (2020). *2020 Investment Company Fact Book*. Investment Company Institute.

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

Jaskowski, M. and D. A. Rettl (2018). Information Acquisition Costs and Credit Spreads. *Working Paper*.

Jensen, M. C. (1968). The Performance of Mutual Funds in the Period 1945-1964. *The Journal of Finance 23*.

Kacperczyk, M., S. V. Nieuwerburgh, and L. Veldkamp (2016). A Rational Theory of Mutual Funds' Attention Allocation. *Econometrica 84*(2), 571–626.

Kacperczyk, M., J. Nosal, and S. Sundaresan (2018). Market power and price informativeness. *Working Paper*.

Kacperczyk, M., C. Sialm, and L. Zheng (2005). On the Industry Concentration of Actively Managed Equity Mutual Funds. *The Journal of Finance 60*(4), 1983–2011.

Katona, Z., M. O. Painter, P. N. Patatoukas, and J. Zeng (2019). On the Capital Market Consequences of Alternative Data: Evidence from Outer Space. *Working Paper*.

Khandani, A. E. and A. W. Lo (2007). What Happened to the Quants in August 2007? *Journal of Investment Management 5*(4), 29–78.

Khandani, A. E. and A. W. Lo (2011). What Happened to the Quants in August 2007? Evidence From Factors and Transactions Data. *Journal of Financial Markets 14*, 1–46.

Kim, J. W., J.-H. Lim, and W. G. No (2012). The Effect of First Wave Mandatory XBRL Reporting across the Financial Information Environment. *Journal of Information Systems 26*(1), 127–153.

Kirilenko, A. A. and A. W. Lo (2013). Moore's Law versus Murphy's Law: Algorithmic Trading and Its Discontents. *Journal of Economic Perspectives 27*(2), 51–72.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica 53*(6), 1315–1336.

Lou, D. (2012). A Flow-Based Explanation for Return Predictability. *Review of Financial Studies 25*(12), 3457–3489.

Lucas, R. E. (1976). Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy 1*, 19–46.

Martineau, C. (2020). The Evolution of Financial Market Efficiency: Evidence from Earnings Announcements. *Working Paper*.

Mihet, R. (2020). Financial Innovation and the Inequality Gap. *Working Paper*.

Moon, H. R. and P. C. B. Phillips (2000). Estimation of Autoregressive Roots near Unity Using Panel Data. *Econometric Theory 16*(6), 927–997.

Nieuwerburgh, S. V. and L. Veldkamp (2009). Information immobility and home bias puzzle. *The Journal of Finance 64*(3), 1187–1215.

Nieuwerburgh, S. V. and L. Veldkamp (2010). Information Acquisition and Under-Diversification. *77*, 779–805.

O'Hara, M. (2003). Presidential address: Liquidity and price discovery. *The Journal of Finance 58*(4), 1335–1354.

Pástor, L., R. F. Stambaugh, and L. A. Taylor (2015). Scale and Skill in Active Management. *Journal of Financial Economics 116*(1), 23–45.

Pedersen, L. H. (2018). Sharpening the arithmetic of active management. *Financial Analysts Journal 75*(1), 21–36.

Peress, J. (2005). Information vs. Entry Costs: What Explains U.S. Stock Market Evolution? *Journal of Financial and Quantitative Analysis 40*(3), 563–594.

Sammon, M. (2020). Passive Ownership and Price Informativeness. *Working Paper*.

SEC (2009). *Interactive Data to Improve Financial Reporting*. Securities and Exchange Commission.

Shim, J. J. (2020). Arbitrage Comovement. *Working Paper*.

Sirri, E. R. and P. Tufano (1998). Costly Search and Mutual Fund Flows. *The Journal of Finance 53*(5), 1589–1622.

Subrahmanyam, A. (1991). A theory of trading in stock index futures. *Review of Financial Studies*.

Veldkamp, L. L. (2011). *Information Choice in Macroeconomics and Finance*. Princeton University Press.

Verrecchia, R. E. (1982). Information aggregation in a noisy rational expectations economy. *Econometrica 50*(6), 1415–1430.

Vives, X. (2008). Information and learning in markets: The impact of market microstructure. *Princeton University Press*, 120–122.

Weller, B. M. (2018). Does Algorithmic Trading Reduce Information Acquisition? *Review of Financial Studies 31*(6), 2184–2226.

Zhang, S. (2017). Need for Speed: Hard Information Processing in a High-frequency World. *Working Paper*.

Zhu, C. (2019). Big Data as a Governance Mechanism. *Working Paper*.

Zhu, M. (2018). Informative fund size, managerial skill, and investor rationality. *Journal of Financial Economics 130*(1), 114–134.

Zhu, Q. (2020). The Missing New Funds. *Management Science*, 1–12.