Essays on Informational Efficiency of the Financial Markets

Ran Liu

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2025

Abstract

Essays on Informational Efficiency of the Financial Markets

Ran Liu

This dissertation studies the effects of informational frictions and information acquisition on asset prices and investor decision-making.

Chapter 1 build models for capital market opening based on the noisy rational expectations equilibrium (NREE) framework. To reconcile the mixed empirical evidence on whether domestic investors have an advantage over foreign investors in domestic stock market tradings, I distinguish *soft information*, which is qualitative and cannot be transmitted if the receiver is not physically present, and *hard information*, which can be stored in the form of *data*, purchased, and transmitted regardless of the presence of the receiver. The models feature a mismatch of data processing technology level and availability of soft information. Furthermore, the partial opening model also features a mismatch between the accuracy of hard information and foreign investability of shares. The *equivalent data amount* and the *information spillover rate* governs domestic and foreign institutional investors' data choices, respectively. Sufficiently, when the unit precision of hard information for foreign investable shares directly from processing data for them exceeds that indirectly acquired through processing data for foreign non-investable shares, foreign institutional investors would exhaust data processing capacity on foreign investable shares. The proportion of foreign institutional investors and their technology level both need to be large enough to overcome the dilution of soft information. Given that the signal-to-noise ratio increases, if the domestic investors learn more about foreign

investable shares or the proportion of foreign investors is larger than a threshold, then the average precision of dividend innovation for foreign investable shares increases. In the full opening model, the average posterior precision of dividend innovation changes in the same direction as the signal-to-noise ratio for all shares.

The case of China is characterized by a very small proportion of foreign investors, a large proportion of domestic individual investors, a large disparity between domestic and foreign data processing technology, and high precision of soft information. Foreign institutional investors would allocate their data processing capacity in both type of shares, and the signal-to-noise ratio for foreign investable shares decreases. Domestic institutional investors are also induced to learn slightly less about those shares, and thus the average posterior precision of dividend innovation decreases. In the full opening counterfactual, the signal-to-noise raios for both type of shares increase.

Chapter 2 fills a gap in the theory literature of NREE models by studying the effects of learning about future demand shocks. Learning about future demand shocks makes price more informative in the future but less informative now. Ex-ante utility improves through what I term as the *future uncertainty channel*: on the aggregate level, fundamental information acquisition decreases, which makes dividends more risky and leads to higher risk premium; on the individual level, the investor is more certain about the asset payoff in the future due to the signal about future demand shocks she acquires today. As a result, her ex-ante utility improves.

Chapter 3 zooms in on the learning process and highlights a specific subprocess—the perception process. Agents take the same sensory information as inputs, direct them through different *perceptual sets* and thus form different *perceptual information*, which is then used to update posterior beliefs. Biased perceptual sets thus lead to persistent disagreement among agents, while *set revision* corrects any perceptual bias and allows agreement to be reached. Depending on the intensity of *perceptual shocks*, agents' sensitivity to perceptual shocks and their initial *entrenchment level*s, there could be overreaction or underreaction in the price to sensory information in the markets. When overreaction happens, the trading volume also increases.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Words cannot express my gratitude to my advisor Laura for her invaluable guidance and feedback over the years. I also could not have undertaken this journey without the support of my other committee members, Neng and Paolo, who generously provided their knowledge and expertise, as well as Jane and Simon, who kindly offered their help when I was in need.

I want to express my gratitude to participants of various seminars and conferences for their insightful feedback.

I am also grateful to my classmates and cohort members at Columbia Business School, especially my cubicle mates for regularly exchanging comments on research projects and organizing the presentation practice sessions. Among them I extend special thanks to my friend Kunjue, for the countless meals we had together over which we discussed our projects and her kind support in my every aspect of my life.

And my biggest thanks go to my family for all the support they have always shown me. Their belief in me has kept my spirits and motivation high during this process.

Dedication

In loving memory of my grandfather, who witnessed the start but not the end of this chapter of my life.

Introduction

Financial markets aggregate and transmit information held by individual traders (Hayek, 2009). A big-picture question is how well prices reflect information and what frictions best explain apparent deviations from market efficiency (Fama, 1991; Brunnermeier and Brunnermeier, 2001). This dissertation studies the effects of informational frictions and information acquisition on asset prices and investor decision-making.

For information choice problems in financial markets noisy rational expectations equilibrium (NREE) models have been an established framework. These models usually introduce exogenous noise trades, alternatively termed as shocks to the demand of the risky assets (demand shocks), to prevent their prices from fully revealing private information observed by investors (Grossman and Stiglitz, 1980; Hellwig, 1980; Diamond and Verrecchia, 1981; Verrecchia, 1982). These shocks are orthogonal to dividend innovations (also called fundamental shocks), and do not affect the fundamental value, or liquidation value, of the risky assets. It is now well-understood that noise trades can be micro-founded by hedging needs (Liu and Wang, 2016; Medrano and Vives, 2004), liquidity needs (Han et al., 2016), or sentiment (Shleifer and Summers, 1990). In Chapters 1 and 2, I follow this tradition and adopt the NREE framework to discuss issues of information efficiency in market liberalization and when future demand shocks are learnable.

Continuous-time finance models also plays an important role in models financial markets. Merton (1969, 1971, 1973) pioneered the use of continuous-time models in financial economics. Portfolio choices have been heavily studied in this framework (Karatzas et al., 1986; Cox and Huang, 1989; Basak, 1995). More recently, scholars have applied this tool to study more issues in

the financial markets (e.g., Chen et al., 2024). In Chapter 3, I adopt a continuous time framework, where information is filtered by Liptser-Shiryaev filtering Liptser and Shiryaev (2001).

Around the topics of informational efficiency in financial markets, two themes run through the dissertation. The first theme is how frictions in information processing affects the financial markets. In Chapter 1, I take a further step than Farboodi and Veldkamp (2020), where they distinguish data and information; I distinguish soft and hard information, with the latter stored and sold in the form of data, and the former generally requires the presence of the receiver and cannot be stored and distribute to a distant buyer. This is particularly meaningful for emerging market liberation. The old wisdom suggests that stock markets reevaluate when emerging economies liberalize since liberation substantially reduce the cost of capital (Henry, 2000, 2003; Bekaert and Harvey, 2000; Kim and Singal, 2000) and the systematic risk associated with holding investable securities, thus increasing risk sharing (Chari and Henry, 2004). In recent years, with the rapid-developing information technology, the new insight of the information channel plays has been shed light on (Griffin et al., 2010; Chan and Kwok, 2017). I argue that there are 2 mismatches that create market inefficiency. Firstly, the mismatch between the efficiency of hard information and foreign investability with markets being partially opened. When the *information spillover rate* is sufficiently high, foreign investors will only learn about shares not eligible for foreign investment and thus the processed hard information is not aggregated into the equilibrium prices of those shares. Secondly, the mismatch of data processing technology and availability of soft information: agents with the best data processing technology cannot acquire soft information. On the one hand, foreign institutional investors are more experienced in investing (Grinblatt and Keloharju, 2000). On the other hand, local investors have data advantage (Choe et al., 2005). The *equivalent data amount* and the *information spillover rate* governs domestic and foreign institutional investors' data choices, respectively. Sufficiently, when the unit precision of hard information for foreign investable shares directly from processing data for them exceeds that indirectly acquired through processing data for foreign non-investable shares, foreign institutional investors would exhaust data processing capacity on foreign investable shares.

It also happens in the limiting cases where the base precision level for the hard information on foreign non-investable shares are sufficiently large or close to zero. In the cases where foreign institutional investors only learn about foreign investable shares, their entry does not guarantee an increase in the signal-to-noise ratio of foreign investable shares since soft information is diluted. The proportion of foreign institutional investors and their technology level both need to be large enough to overcome the abovementioned dilution. Given that the signal-to-noise ratio increases, if the domestic investors learn more about foreign investable shares or the proportion of foreign investors is larger than a threshold, then the average precision of dividend innovation for foreign investable shares increases. In the full opening model, the average posterior precision of dividend innovation changes in the same direction as the signal-to-noise ratio for all shares.

The case of China is characterized by a very small proportion of foreign investors, a large proportion of domestic individual investors, a large disparity between domestic and foreign data processing technology, and high precision of soft information. Foreign institutional investors would allocate their data processing capacity in both type of shares, and the signal-to-noise ratio for foreign investable shares decreases. Domestic institutional investors are also induced to learn slightly less about those shares, and thus the average posterior precision of dividend innovation decreases. In the full opening counterfactual, the signal-to-noise raios for both type of shares increase.

Chapter 3 takes a behavioral approach and studies cognitive frictions during the learning process. The issue of persistent disagreement among agents have been tackled in the Bayesian framework (Kominers et al., 2018; Nimark and Sundaresan, 2019), as well as from a behavioral point of view, for example curse of knowledge (Cheng and Hsiaw, 2022), limited working memory (Koçak, 2018), wishes to agree with information source (Chauvin, 2023), double updating (Fryer et al., 2019), and etc. I highlight the perception process and biasness of perceptual sets. Individuals process the same *sensory information* through different perceptual sets (Weiten, 2001) and thus form different *perceptual information*, that is individuals obtain different internal representations of the same external inputs via the perception process (Schacter et al., 2014), and then

combine them with the prior belief to form posterior beliefs. In the spirit of entrenchment for beliefs (Hansson, 2017), I introduce a notion of set entrenchment so that a perceptual set with high entrenchment is less subject to revision. When perceptual shocks occur, however, the enacted perceptual sets would suddenly change and the entrenchment level would be reset to the default value, that is agents become entrenched with their new perceptual sets. With perceptual biases, persistent disagreement among agents exits even with sufficient information. Shifts from disagreement to agreement are caused by perceptual sets revisions induced by perceptual shocks. The magnitude and speed of the shift are affected by agents' experience and sensitivity to perceptual shocks. Perceptual set entrenchment and revision of perceptual sets in the face of huge shocks allow room for both overreaction and underreaction to sensory information from the markets. When overreaction happens, there is an increase in the trading volume of the risky asset.

The second theme is the effects of learning information of different nature. In particular, Chapter 2 deals with how learning about future demand shocks affects investors and the financial markets. In the literature, the effects of current fundamental shocks (Verrecchia, 1982; Van Nieuwerburgh and Veldkamp, 2010; Goldstein and Yang, 2015, 2019) and the current demand shocks (Medrano and Vives, 2004; Ganguli and Yang, 2009) are widely studied. However, research on the effects of learning about future shocks is extremely scarce. Cai (2017) deals with learning about future fundamental shocks, but the existing literature has not included future demand shocks into theoretical research. Chapter 3 fills this gap. Price volatility increases due to heterogeneous private information. Learning about future demand shocks makes price more informative in the future but less informative now. Ex-ante utility improves through what I term as the *future uncertainty channel*: on the aggregate level, fundamental information acquisition decreases, which makes dividends more risky and leads to higher risk premium; on the individual level, the investor is more certain about the asset payoff in the future due to the signal about future demand shocks she acquires today. As a result, her ex-ante utility improves.

Chapter 1: Data Processing Technology, Soft Information and Capital Market Opening

1.1 Introduction

There is widespread old wisdom that domestic investors have an advantage over foreign investors in domestic stock market tradings, which has been used to explain anomalies such as the home bias and herding. The empirical evidence, however, is mixed. On the one hand, using Korean data, Choe et al. (2005) find that domestic investors have an edge over foreign investors due in large part to the fact that prices move more against foreign investors than against domestic investors before trade, consistent with the claim that foreign investors are less informed. Dvořák (2005) applies the same approach to Indonesia and concurs that foreign investors are at an information disadvantage. On the other hand, Grinblatt and Keloharju (2000) find that in Finland the portfolios of foreign investors outperform those of domestic households, even after controlling for behavior differences. Hau (2001) find that in German traders located outside Germany in non-German-speaking cities show lower proprietary trading profit.

While the empirical evidence around this topic is ample, I try to add to the sparse theoretical work. To reconcile the mixed empirical results on the performance of foreign investors versus domestic investors, hereinafter, I distinguish soft information from hard information. The distinction between soft and hard information first arose in the banking literature, and then expanded to other areas of finance. Although the distinction has long been established, a clear and precise definition has not been provided by the literature. Liberti and Petersen (2019) argue that rather than two distinct classifications, a continuum along wich information can be classified should be thought of. They provide 3 dimensions, namely numbers versus text, the unimportance of context, and separation of information collection and decision-making. Gao and Zeng (2022) de-

fine soft information as qualitative information that is difficult to convey or easily convertible to values, which often requires physical presence to be acquired and evaluated. This information could be related to intangible features of the firm, such as corporate culture, ethics, employee loyalty, etc. It could also include small pieces of information that, when considered individually, are difficult to value. Moreover, critical to the goal of this study, a common feature of this type of information would be the lack of standard disclosure by the firms.

For the purpose of this research, I define soft information similar to Gao and Zeng (2022) but more restrictive. *Soft information* is qualitative information that cannot be converted to values or be transmitted if the receiver is not physically present. *Hard information*, on the other hand, can be physically stored on paper or an electronic device, in the form of *data*, can be purchased, exclusively or not, from data sources and transmitted regardless of the presence of receivers. See Figure 1.1 for a diagrammatic representation. In the context of foreign and domestic investors, it is reasonable to assume that only domestic investors hold soft information; foreign investors are shunned out because of distance, linguistic or cultural barriers. Both foreign and domestic investors can acquire data, i.e. encoded and stored hard information. Data can come in many different forms, quantitative or textual, precise or imprecise, direct or indirect about the target variable. For example, imprecise textual data include the textual descriptions in the company's news release (...slightly improves, ...relatively stable, etc.). Quantitative data is usually more precise, e.g. accounting numbers in a company's annual report (net income, total assets, long-term debt and etc.), daily trading prices and volumes. Acquired data are then processed by investors via certain *data processing technology* back into information. This information extracting process, however, can induce information loss, depending on software, experience, expertise and etc. This loss of information is compatible of the findings that investors under-react, or react with a significant delay, to value-relevant information from earnings announcements (Bernard and Thomas, 1989, 1990; Kothari, 2001; Chordia and Shivakumar, 2005; Truong, 2011). It is reasonable to assume that institutional investors from markets of the same sophistication level would employ similar data processing technology.

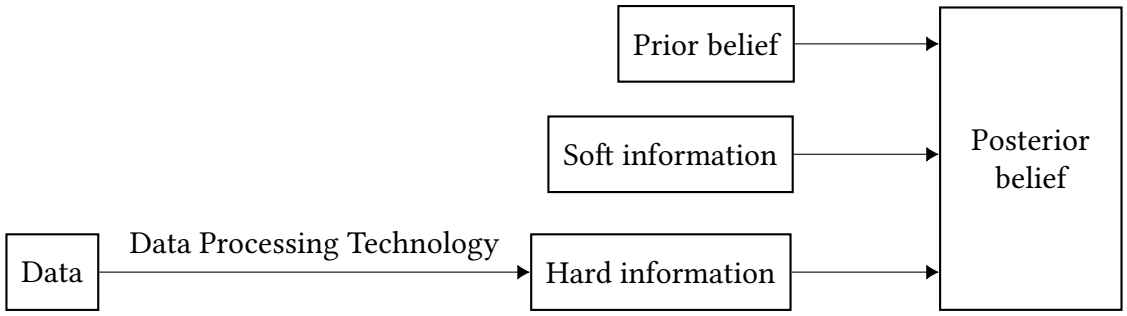


Figure 1.1. Data, hard information, and soft information

Data is not equal to information; it is a physical or digital format of storage for hard information. It can be purchased and transmitted regardless of the presence of receivers. The process of converting data back to hard information is abstracted as *data processing technology*. Depending on the precision of data processing technology, this process can be lossy or lossless. Soft information is qualitative information that cannot be converted to values or transmitted if the receiver is not physically present. Posterior beliefs are formed by combining the prior belief, soft information and hard information.

If we look back at the empirical evidence, for example in Hau (2001), since it is reasonable to assume that the data processing technology that professional investors in German employs is at the same level as foreign professional investors, the lower proprietary trading profit earned by traders located outside Germany in non-German-speaking cities is compatible with a lack of soft information due to distance and linguistic barriers.

This mismatch of better data processing technology and availability of soft information is especially interesting for emerging markets. On the one hand, emerging markets are more opaque and less efficient than developed markets, which usually means that the soft information held by domestic investors gives them larger informational advantage. On the other hand, investors in the emerging markets are likely less sophisticated than investors in the global markets, so global investors would have an advantage over them on data processing. In other words, when faced with the same data source, the ability of investors in the emerging markets to restore hard information is lower than that of investors in the global markets.

When emerging economies liberalize fully or partially, their stock markets go through revaluations due to substantial reductions of the cost of capital (Henry, 2000, 2003; Bekaert and Harvey, 2000; Kim and Singal, 2000). The source of reductions of the cost of capital is further decomposed into a fall in the risk-free rate and increased risk-sharing (Chari and Henry, 2004; Chan and Kwok,

2017). Since the fall in the risk-free rate introduces a common shock to all firms in the emerging economy, it should be the same regardless of whether the shares of the firm is investable or not by global investors after the liberalization. Increased risk-sharing, however, is experienced by foreign-investable shares. The relevant market portfolio has shifted from the domestic portfolio to the global busket, substantially reducing the systematic risks.

With the rapid-developing information technology, some new insight of the information channel has been shed light on. For one thing, foreign institutional investors are more experienced in investing (Grinblatt and Keloharju, 2000). Thus market liberalization introduces more informed traders to the emerging market. Chan and Kwok (2017) consider some information issues that may play a role in affecting the efficiency of price revaluation under liberalization. To the extent that institutional investors tend to be informed traders, a more prominent presense of informed traders exists in stocks with a high institutional ownership. These stocks will be more effective in incorporating firm-specific information. For another, an improved information environment, as measured by analysts coverage, number of analysts and analyst rating in Chan and Kwok (2017) and analyst coverage, number of forecasts, and frequency of analyst revisions in Griffin et al. (2010), allows information to be more effectively incorporated. My information channel stresses the mismatch of the best data processing technology and the availability of soft information.

Model To study the role which the information channel plays in capital market opening with the presence of imperfect data processing technology and soft information, I build a model in the noisy rational expectations equilibrium framework (NREE). For information choice problems in financial markets, NREE models are widely used (Grossman and Stiglitz, 1980; Hellwig, 1980; Diamond and Verrecchia, 1981; Verrecchia, 1982). Specifically, I propose a model of a 2-period competitive CARA-normal economy with domestic institutional investors, domestic individual investors and foreign institutional investors. In each period, institutional investors first decide how to allocate limited data processing capacity between foreign investable and uninvestable

shares. Through certain data processing technology, investors get a private signal about the interested target variable dividend innovation. Then they make subsequent investment decisions based on the information they extract from acquired data, soft information if investors are domestic, and the equilibrium asset price. To fit the evolution of emerging markets, I develop two versions of the model, partial opening and full opening.

The model features 2 mismatches that create information inefficiency. Firstly, the above-mentioned mismatch between better data processing technology and availability of soft data. While the entry of foreign institutional investors dilute the density of soft information, they also bring into the market more hard information from processed data. This mismatch exists in both partial and full opening models. Secondly, the mismatch between efficiency of hard information and foreign investability in partial opening. Although foreign investors are not allowed to invest in ineligible shares, it is still possible in their best interest to acquire and process data points for them; intuitively, in equilibrium the marginal gain in posterior precision of divided innovation for foreign investable shares directly from learning about them and indirectly from learning about ineligible shares should be equal. When the *information spillover rate* is sufficiently high, foreign investors will only learn about shares ineligible for foreign investments. Information inefficiency is created since the acquired hard information is not aggregated into the equilibrium price of shares ineligible for foreign investments.

The *equivalent data amount* and the *information spillover rate* governs domestic and foreign institutional investors' data choices, respectively. In the partial opening model, the equivalent data amount to the posterior precision of dividend information of foreign non-investable shares at full data processing capacity is still less than that without processing data for foreign investable shares, domestic institutional investors would only acquire data on foreign non-investable shares. In the opposite case, if the equivalent data amount without processing data for foreign non-investable shares is larger than that at full data processing capacity for foreign investable shares, they would only acquire data on foreign investable shares. If the information spillover rate without acquiring data for foreign non-investable shares is smaller than 1, then foreign insti-

tutional investors would only learn about foreign investable shares. If, however, the information spillover rate at full data capacity for foreign non-investable shares is larger than 1, then they would only learn about foreign non-investable shares. Sufficiently, when the unit precision of hard information for foreign investable shares directly from processing data for them exceeds that indirectly acquired through processing data for foreign non-investable shares, foreign institutional investors would exhaust data processing capacity on foreign investable shares. It also happens in the limiting cases where the base precision level for the hard information on foreign non-investable shares are sufficiently large or close to zero. Partial market opening influences the market equilibrium through the signal-to-noise ratios, the weighted average of the posterior variance of the dividend innovation, and the unit demand level. In the cases where foreign institutional investors only learn about foreign investable shares, their entry does not guarantee an increase in the signal-to-noise ratio of foreign investable shares since soft information is diluted. The proportion of foreign institutional investors and their technology level both need to be large enough to overcome the abovementioned dilution. Given that the signal-to-noise ratio increases, if the domestic investors learn more about foreign investable shares or the proportion of foreign investors is larger than a threshold, then the average precision of dividend innovation for foreign investable shares increases. In the full opening model, the average posterior precision of dividend innovation changes in the same direction as the signal-to-noise ratio for all shares.

This model is extended to include an active choice of technology level by Manager. In this chapter, technology is defined abstractly and inclusively: it includes everything that the investors need to convert data to information about the target variable, in this case dividend innovation, that can directly aid portfolio choice. Models of endogenous technical change fall into two broad categories. On the one hand, there are models of "invention" (e.g., Romer 1990; Segerstrom, Anant, and Dinopoulous 1990; Grossman and Helpman 1991), in which technical change is the outcome of costly and deliberative research aimed at the development of new technologies. On the other hand, there are models of "learning by doing" (e.g., Arrow 1962; Lucas 1988), in which technical change is the by-product of experience gained in the production of goods. In contrast to

the learning-by-doing literature where innovation is a by-product of production, in this chapter technological innovation is an active choice by the investors.

This model can be extended to include voluntary disclosure. Tsang et al. (2019) found that in their sample of 32 non-U.S. countries, foreign institutional investments lead to improved voluntary disclosure, and this effect is more pronounced when foreign institutional investors are unfamiliar with the firm's home country, have longer investment horizons, and are from countries with stronger investor protection and disclosure requirements than the firm's home country. Consistent with their findings, Ruan et al. (2021) discovered that after the implementation of the Shanghai-Hong Kong Stock Connect program, the quality of internationally tradable companies' public information disclosure, proxied by the proportion of commonly used words and the degree of text uncertainty, significantly improved.

Stock Connect Studies To illustrate the model, I provide a numerical example of the Shanghai-Hong Kong Stock Connect Program in China launched on November 17, 2014, which is an important step China takes to open its capital markets to the outside world. I look at specifically how the north-bound trades affect investors and the market.

I chose the China case for an illustration of my model mainly for two reasons. Firstly, China has become one of the fastest growing economies and capital markets in the world. As of the end of 2022, SSE is the third largest stock exchange in terms of market capitalization, trailing after NYSE and Nasdaq, while SZSE ranks the sixth. The total market capitalization of the two exchanges is at $11.4 trillion, accounting for 11.29% of the world total. The sheer size and enormous growth opportunities of the Chinese security markets have grabbed significant attention and participation from international investors in recent years. In addition, the gradual opening of the Chinese capital markets to the world also results in a significant share of the Chinese shares being held by international investors. The highlighted significance of the Chinese capital markets therefore increasingly calls for thorough studies of information efficiency in these markets, which is the focus of my research.

Secondly, China has characteristics that are conformant to the assumptions on soft information and technological difference in my model. China's socioeconomic structure and ideology differ from those of the majority of countries. The outside world, particularly countries outside of East Asia and Southeast Asia, is still largely unfamiliar with its dynamic culture, which influences every facet of its economy and value systems. Cultural and linguistic barriers contribute to the abundance of soft information that is exclusively available to local investors. Furthermore, it is also reasonable to assume that Chinese investors translate data less accurately into hard information. On the one hand, the trading activities of small retail investors have historically dominated the Chinese stock markets. The Shanghai Stock Exchange estimates that in China, retail trading makes up over 80% of total trading volume, compared to just 10% in the U.S. (Jones et al., 2025). While institutional investors are more adept at analyzing and acting upon financial data, small retail investors are less experienced (Hand, 1990). A less effective integration of accounting earnings information into stock prices is thus expected in the Chinese stock markets. On the other hand, post-earnings-announcement drift anomaly in the Chinese equity market documented by Truong (2011) also contributes to the validity of this assumption. A more transparent information environment is linked to foreign ownership, and foreign investors react to new information faster than domestic investors.

When fitted to the model, the case of China is characterized by a very small proportion of foreign investors, a large proportion of domestic individual investors, a large disparity between domestic and foreign data processing technology, and high precision of soft information. Foreign institutional investors would allocate their data processing capacity in both type of shares, and the signal-to-noise ratio for foreign investable shares decreases. Domestic institutional investors are also induced to learn slightly less about those shares, and thus the average posterior precision of dividend innovation decreases. In the full opening counterfactual, the signal-to-noise raios for both type of shares increase.

This chapter proceeds as follows. In Section 2, I first present present the partial-opening model, discuss the characteristics of its equilibrium, and then show the impacts of partial market

opening. In Section 3, I provide a numerical example of the Shanghai-Hong Kong Stock Connect programs. In Section 4, I expand the model to full market opening and shows the prospects of the China case. In Section 5, technology advancement is discussed. I conclude in Section 6.

1.2 Partial-Opening Model

When emerging markets open to the international markets, most economies (Korea, Mexico, Greece, to mention a few) first implement a partial opening and liberate the market step by step.

1.2.1 Model setup

Consider a 2-period economy. In the first period, the domestic economy is closed. In the beginning of period 2, the domestic financial market partially opens up to foreign investors.

Assets There is a domestic safe asset with rate of return r. The equities of domestic firms are originally only traded in domestic markets. I model them as two aggregate risky securities: type-I security that is eligible for international trade after partial market opening, and type-II risky asset that is not allowed to be traded by international investors even after partial market opening is in place. Period-t dividend $\widetilde{\boldsymbol{d}}_t$ follows an AR(1) process, $\widetilde{\boldsymbol{d}}_t = \boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_{t-1} + \widetilde{\boldsymbol{y}}_t$, where the dividend growth rate $\boldsymbol{G} \equiv diag((g_1, g_2)')$, the dividend shock $\widetilde{\boldsymbol{y}}_t \overset{i.i.d.}{\sim} N(0, \boldsymbol{\Sigma}_y)$. The dividend is paid out at the end of each period. Denote the ex-dividend aggregate firm value by $\widetilde{\boldsymbol{v}}_t$, then $\widetilde{\boldsymbol{v}}_t = \boldsymbol{\mu}_v + (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\widetilde{\boldsymbol{d}}_t$, where $\boldsymbol{\mu}_v = \frac{r}{r-1}(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu}$.

Investors There is a continuum of investors. Among them there are three types: λ_I domestic institutional investors, λ_J domestic individual investors, and λ_K foreign institutional investors. For simplicity, I use $l, l \in \{i, j, k\}$ to refer to an investor of any type, and $L, L \in \{I, J, K\}$ to refer to any type of investors. Investors L are endowed with initial wealth w_L. Domestic investors start to invest in the market from Period 1, while foreign investors only enter the market after the market becomes open to international investors in Period 2. Individual investors form their portfolios only based on security prices, while institutional investors acquire and process data,

obtaining private signals about the dividend innovation. All domestic investors are additionally endowed with soft information about the dividend innovation.

In Period 1, the securities are only traded by domestic investors. In the beginning of Period 2, the securities market of the domestic economy opens up to foreign investors. The initial wealth w_K for foreign investors is entirely invested in domestic securities markets; their wealth that is invested in foreign economies is not modeled here. All investors have CARA utility over their terminal wealth, with a risk-aversion coefficient ρ, that is $U_{lt} = -e^{-\rho w_{l2+}}$, $l \in \{i, j, k\}$[1]. They make portfolio choices $\boldsymbol{q}_{lt}$ in each period to maximize their expected utility conditional on their respective information set and subject to a budget constraint, which are

$$w_{l2} = rw_{l1} + \boldsymbol{q}'_{l2}(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r) \tag{1.1}$$

$$w_{l1} = rw_L + \boldsymbol{q}'_{l1}(\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1 r) \tag{1.2}$$

in Period 2 and Period 1, respectively.

Information structure In this economy, in each period before making the portfolio choice, institutional investors can acquire data to generate private signals about current dividend innovation $\widetilde{\boldsymbol{y}}_t$. In addition, domestic investors are endowed with soft information[2] about $\widetilde{\boldsymbol{y}}_t$. The equilibrium price aggregates all private information.

At different stages of decision, investors have different information sets. At the beginning of Period 1 when domestic institutional investor i is to make data acquisition decisions, she does not have any information other than the prior and her soft information $\boldsymbol{\xi}_{i1}$. Denote this information set by $\mathcal{I}_{i1^-} = \{\boldsymbol{\xi}_{i1}\}$. Let $\boldsymbol{s}_{lt}$, $l \in \{i, k\}$, denote the private signals generated from data, and $\boldsymbol{s}_{pt}$ denote the price signal. Later when she selects portfolio, the choice is based on the equilibrium price, the soft information, and the private signal generated from the data she acquired, so her

[1]The superscript + refers to the end of a period when dividend is paid out.

[2]Intuitively, domestic institutional investors should have more soft information than domestic individual investors. However, since it is not a focus of this chapter and does not change the structure of the model nor the results, I make the simplifying assumption that domestic investors all have the same amount of soft information.

information set is $\mathcal{I}_{i1} = \{\boldsymbol{s}_{p1}, \boldsymbol{s}_{i1}, \boldsymbol{\xi}_{i1}\}$. Domestic individual investor j does not acquire and process data, and makes her portfolio choice based only on her endowed soft information and the equilibrium price, so $\mathcal{I}_{j1} = \{\boldsymbol{s}_{p1}, \boldsymbol{\xi}_{j1}\}$.

In Period 2, the domestic market is open to foreign institutional investors. When making data acquisition decisions, domestic institutional investor l, $l = i, j$, conditions on $\mathcal{I}_{l2^-} = \mathcal{I}_{l1^+} \cup \{\boldsymbol{\xi}_{l2}\}$. Foreign institutional investor k does not have relevant information other than the prior. Progressing to the portfolio choice stage, investors base their decisions on $\mathcal{I}_{i2} = \begin{cases} \mathcal{I}_{l2^-} \cup \{\boldsymbol{s}_{p2}, \boldsymbol{s}_{l2}\}, & \text{for } l = i, k \\ \mathcal{I}_{l2^-} \cup \{\boldsymbol{s}_{p2}\}, & \text{for } i = j \end{cases}$. Let me discuss first the generation of price signals and then how acquired data is processed into private signals.

The equilibrium price is a linear function of the signals, and since signals are normally distributed, we can express the price as a linear function of shocks:

$$\boldsymbol{p}_t = \boldsymbol{A}_t + \boldsymbol{B}_t \boldsymbol{d}_{t-1} + \boldsymbol{C}_t \widetilde{\boldsymbol{y}}_t + \boldsymbol{D}_t \widetilde{\boldsymbol{x}}_t. \tag{1.3}$$

Define the price-to-current-noise ratio to be $\boldsymbol{\zeta}_t \equiv \boldsymbol{D}_t^{-1} \boldsymbol{C}_t$. A transformation of the equilibrium price $\boldsymbol{p}_t$ can generate an unbiased signal about current dividend shock. In Period t, investors extract a signal $\boldsymbol{s}_{pt} = \widetilde{\boldsymbol{y}}_t + \boldsymbol{\zeta}_t^{-1} \widetilde{\boldsymbol{x}}_t$, and the precision of the price signal is $\boldsymbol{\Omega}_{pt} = \boldsymbol{\zeta}_t' \boldsymbol{\Omega}_x \boldsymbol{\zeta}_t$.

Throughout this chapter, I specifically make the distinction between data and information. While data are physically stored on paper or in an electronic device, and can be purchased, exclusively or not, from data sources, information is the signal about an unknown random variable, in this case the dividend innovation $\widetilde{\boldsymbol{y}}_t$, resulting from processing the data. Data can come in different forms, including, for example, accounting numbers in a company's annual report (net income, total assets, long-term debt and etc.), daily trading price and volume, textual descriptions in the company's news release (...slightly improves, ...relatively stable, etc.). Through certain data processing technology, investors get a private signal about the interested target variable dividend innovation. Their data source and acquired amount of data could be different, and their data processing technology could also differ.

Specifically in the model, institutional investors can acquire a total of $\bar{M}$ data points from two data sources, and the data acquisition costs are normalized to zero. Then investors will acquire data at full capacity $\bar{M}$. Denote the number of data points acquired by investor l in Period 2 for type-I and type-II risky asset is $M_{l2}^{(\mathrm{I})}$ and $M_{l2}^{(\mathrm{II})}$, respectively. Each data source corresponds to one type of firm, and each data point is a linear function of current dividend innovation. So the acquired data point $\widetilde{\eta}_{lt}^{(m)}$ is a function of type-I or type-II firm dividend innovation, that is $\widetilde{\eta}_{lt}^{(m)} = \begin{cases} \beta\widetilde{y}_t^{(\mathrm{I})} + \widetilde{e}_{mlt}^{(\mathrm{I})} & \text{type-I firm} \\ \beta\widetilde{y}_t^{(\mathrm{II})} + \widetilde{e}_{mlt}^{(\mathrm{II})} & \text{type-II firm} \end{cases}$, with $\beta > 1$ and $\widetilde{\boldsymbol{e}}_{lt} \overset{i.i.d.}{\sim} N(0, \boldsymbol{\Sigma}_e)$, with the variance matrix $\boldsymbol{\Sigma}_e$ diagonal. With a perfect data processing technology, one correctly knows the value of β. Thus each data point can be transformed into an unbiased data signal $s_{lt}^{(m)} = \frac{1}{\beta}\widetilde{\eta}_{lt}^{(m)}$. Thus based on all data points, the conditional variance of current dividend innovation is $\mathbb{V}ar(\widetilde{\boldsymbol{y}}_t \mid s_{lt}^{(1)}, \ldots, s_{lt}^{(M)}) = (\boldsymbol{\Omega}_y + \beta^2 \boldsymbol{M}_{lt}\boldsymbol{\Omega}_e)^{-1}$. Suppose without loss of generality that foreign institutional investors can process data perfectly, that is they can process data with the correct β, while domestic institutional investors only have imperfect data processing technology with which β is incorrectly assumed to be β_I, with $\beta_I < \beta$. Therefore, in the eyes of domestic institutional investors, the data signal is noisier, which will in turn affect their portfolio choices.

The above-generated information is what I term as *hard information*; it can be stored and purchased in the form of *data*, which is then back-engineered into hard information using certain data processing technology by all domestic and foreign data buyer. However, the other type of information, *soft information*, cannot be stored and distributed as data; it is only held by domestic investors. To model hard and soft information, I assume that domestic investors are endowed with an unbiased signal about the dividend innovation, that is $\widetilde{\boldsymbol{\xi}}_{it} = \widetilde{\boldsymbol{y}}_t + \widetilde{\boldsymbol{\epsilon}}_{it}$, with $\widetilde{\boldsymbol{\epsilon}}_{it} \overset{i.i.d.}{\sim} N(0, \boldsymbol{\Sigma}_\epsilon)$. Although domestic institutional investor i has an informational advantage with the soft information, she has a technological disadvantage over processing buyable data and is not able to implement the perfect data processing technology. Thus a priori it is unclear who has an overall information advantage. Different types of investors along these two dimensions are summarized in Table 1.1.

	Domestic[a]	Foreign
Institutional[b]	Price information[c] Imperfect data processing technology Soft information $(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{yt} + \beta_I^2 \boldsymbol{M}_{It}\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon)$	Price information Perfect data processing technology $(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{yt} + \beta^2 \boldsymbol{M}_{Kt}\boldsymbol{\Omega}_e)$
Individual	Price information Soft information $(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{yt} + \boldsymbol{\Omega}_\epsilon)$	Price information[d] $(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{yt})$

[a] Domestic investors have soft information about dividend innovation, with precision $\boldsymbol{\Omega}_\epsilon$.

[b] Institutional investors can acquire and process data under capacity $\bar{M}$. Foreign institutional investors are assumed to have perfect data processing technology β, while domestic institutional investors only have imperfect technology β_I, with $\beta_I < \beta$.

[c] Apart from the common prior, all investors observe the equilibrium price and extract from it a price signal with precision $\boldsymbol{\Omega}_{yt}$.

[d] Foreign individual investors do not enter the domestic market in the partial opening model; they enter in the full opening extension discussed in later sections. They are included in this table for a complete description of all investors of interest throughout this chapter.

Table 1.1: Data processing technology and information held by investors and their corresponding posterior precision about dividend innovations

1.2.2 Equilibrium characterization

Equilibrium definition An equilibrium consists of a price function $\boldsymbol{P}_t(\boldsymbol{d}_{t-1}, \widetilde{\boldsymbol{y}}_t, \widetilde{\boldsymbol{x}}_t) :\to \mathbb{R}^2$, a data acquisition strategy $(M_{it}^{(\mathrm{I})}, M_{it}^{(\mathrm{II})})$ for domestic institutional investor i, a data acquisition strategy $(M_{kt}^{(\mathrm{I})}, M_{kt}^{(\mathrm{II})})$ for foreign institutional investor k, and a trading strategy $\boldsymbol{q}_{lt}$ for investor l.

Definition 1. *A rational expectations equilibrium in this economy is a sequence of data choices* $\{M_{it}^{(\mathrm{I})}, M_{it}^{(\mathrm{II})}\}$ *and* $\{M_{k2}^{(\mathrm{I})}, M_{k2}^{(\mathrm{II})}\}$*, prices of the risky asset* $\{\boldsymbol{p}_t\}$*, and portfolio choices* $\{\boldsymbol{q}_{lt}\}$*,* $t = 1, 2$ *such that*

(a) At the beginning of Period t*, institutional investor* l*,* $l = \begin{cases} i & \text{if } t = 1 \\ i, k & \text{if } t = 2 \end{cases}$*, allocates* $\{M_{lt}^{(\mathrm{I})}, M_{lt}^{(\mathrm{II})}\}$ *data points between type-I and type-II firms to maximize ex-ante expected utility* $\mathbb{E}[-e^{-\rho \widetilde{w}_{l2^+}} \mid \mathcal{I}_{lt^-}]$*, subject to data processing capacity* $\bar{M}$*;*

(b) At the portfolio choice stage of Period 1*, domestic investor* l $(l \in \{i, j\})$ *invests* $\boldsymbol{q}_{l1}$ *in the risky*

assets to maximize $\mathbb{E}[-e^{-\rho\widetilde{w}_{i2+}} \mid \mathcal{I}_{l1}]$*, taking the asset price and signal realizations, with no expectations that the market will be open up to international investors in Period 2, subject to budget constraint (1.2);*

(c) At the portfolio choice stage of Period 2*, investor* l $(l \in \{i, j, k\})$ *invests* $\boldsymbol{q}_{l2}$ *in the risky assets to maximize* $\mathbb{E}[-e^{-\rho\widetilde{w}_{l2+}} \mid \mathcal{I}_{l2}]$*, taking the asset price and signal realizations as given, subject to budget constraint (1.1), and foreign institutional investor* k *subject to constraint* $q_{k2}^{(\mathrm{II})} = 0$*;*

(d) The price of the risky asset equates demand and supply:

$$\int_i \boldsymbol{q}_{i1}\,\mathrm{d}i + \int_j \boldsymbol{q}_{j1}\,\mathrm{d}j = \bar{x}_1 + \widetilde{x}_1. \tag{1.4}$$

$$\int_i \boldsymbol{q}_{i2}\,\mathrm{d}i + \int_j \boldsymbol{q}_{j2}\,\mathrm{d}j + \int_k \boldsymbol{q}_{k2}\,\mathrm{d}k = \bar{x}_2 + \widetilde{x}_2. \tag{1.5}$$

Portfolio choices In Period 2, investor l faces a simple one-shot portfolio choice problem, which has an optimal solution[3]

$$\boldsymbol{q}_{l2} = \frac{1}{\rho}\hat{\boldsymbol{\Omega}}_{L2}\hat{\boldsymbol{\mu}}_{l2}, \tag{1.6}$$

where $\hat{\boldsymbol{\mu}}_{l2} \equiv \mathbb{E}[\widetilde{\boldsymbol{d}_2} + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r \mid \mathcal{I}_{l2}]$ is the conditional mean of Period-2 payoff and $\hat{\boldsymbol{\Omega}}_{L2} \equiv \mathbb{V}ar^{-1}(\widetilde{\boldsymbol{d}_2} + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r \mid \mathcal{I}_{l2})$ is the conditional precision of Period-2 payoff.

The portfolio choice in Period 1 can be written as

$$\boldsymbol{q}_{l1} = \frac{1}{\rho r}\hat{\boldsymbol{\Omega}}_{L1}(\hat{\boldsymbol{\mu}}_{l1} - \rho\boldsymbol{\mathcal{C}}_L), \tag{1.7}$$

where $\hat{\boldsymbol{\mu}}_{l1} \equiv \mathbb{E}[\widetilde{\boldsymbol{d}_1} + \boldsymbol{p}_2 - \boldsymbol{p}_1 r \mid \mathcal{I}_{l1}]$ is the conditional mean of Period-1 payoff, $\hat{\boldsymbol{\Omega}}_{L1} \equiv \mathbb{V}ar^{-1}(\widetilde{\boldsymbol{d}_1} + \boldsymbol{p}_2 - \boldsymbol{p}_1 r \mid \mathcal{I}_{l1})$ is the conditional variance of Period-1 payoff, and $\boldsymbol{\mathcal{C}}_L \equiv \mathbb{C}ov(\widetilde{\boldsymbol{d}_1} + \boldsymbol{p}_2^{\dagger} - \boldsymbol{p}_1 r, \boldsymbol{q}_{l2}^{\dagger}(\widetilde{\boldsymbol{d}_2} + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^{\dagger} r) \mid \mathcal{I}_{l1})$, with $\boldsymbol{q}_{l1}\boldsymbol{\mathcal{C}}_L$ being the conditional covariance of Period-1 and Period-2 portfolio payoffs. Notice the difference between $\boldsymbol{q}_{l2}$ and $\boldsymbol{q}_{l2}^{\dagger}$. While $\boldsymbol{q}_{l2}$ is the actual investment that

[3]For notatonal simplicity, I adopt the notational style that uses a hat to distinguish conditional moments at the portfolio choice stage (*interim*) with those at the data choice stage (*ex ante*). These conditional moments are derived in the appendix.

domestic investor l, $l = i, j$ makes at the portfolio choice stage of Period 2 when international institutional investors enter the market, $\boldsymbol{q}_{l2}^{\dagger}$ is what investor l predict herself to hold in Period 2 at the portfolio choice stage in Period 1 when she does not expect that the market would open to international investors in the next period.

Equilibrium price relation The signal-to-noise ratio

$$\boldsymbol{\zeta}_2 = -\frac{1}{\rho r}(r\boldsymbol{I} - \boldsymbol{G})\left[\left(\lambda_I \beta_I^2 \boldsymbol{M}_{I2} + \lambda_K \beta^2 \left(\boldsymbol{E}_{11} - \frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{12}\right)\boldsymbol{M}_{K2}\right)\boldsymbol{\Omega}_e + \lambda_1 \boldsymbol{\Omega}_{\epsilon}\right]. \quad (1.8)$$

An interesting observation is that the signal-to-noise ratio matrix is not symmetric. If foreign institutional investors choose to acquire data on type-II risky asset, the information they obtain after data processing is not aggregated into the equilibrium price of type-II risky asset since it is not open for international trade. However, this information will be partially reflected in the equilibrium price of type-I risky asset. This is because the two types of assets are correlated and thus information on type-II asset also affects their posterior belief on type-I asset. This creates information inefficiency.

The average posterior precision of dividend innovation is

$$\bar{\hat{\boldsymbol{\Omega}}}_{y2} = \boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + \boldsymbol{\Lambda}^{-1}(\lambda_I \beta_I^2 \boldsymbol{M}_{I2} + \lambda_K \beta^2 \boldsymbol{E}_{11}\boldsymbol{M}_{K2})\boldsymbol{\Omega}_e + \boldsymbol{\Lambda}^{-1}\lambda_1 \boldsymbol{\Omega}_{\epsilon} - \frac{\lambda_K \hat{\mathcal{C}}_{Ky2}^2}{\lambda_2 \hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{11}.$$

On the one hand, since foreign institutional investors do not possess soft information, their entry dilute the density of the soft information for type-II risky asset. On the other hand, they also add to aggregated information acquired from data. Notice that their contribution to aggregated data precision is corrected down by the last term, which comes from a revision of the portfolio position due to the correlation between assets and is present regardless of their data choice.

In equilibrium, the price coefficients are characterized by

$$\boldsymbol{A}_2 = \frac{1}{r}\boldsymbol{\mu}_v + (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - \rho r(r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\boldsymbol{\Lambda}^{-1}(r\boldsymbol{I} - \boldsymbol{G})^{-1}\bar{\boldsymbol{x}}_2, \tag{1.9}$$

$$\boldsymbol{B}_2 = (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}, \tag{1.10}$$

$$\boldsymbol{C}_2 = (r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left[\bar{\hat{\boldsymbol{\Omega}}}_{y2} - \boldsymbol{\Omega}_y + \frac{\lambda_K\hat{\mathcal{C}}^2_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{11} - \frac{\lambda_K\hat{\mathcal{C}}_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}(\boldsymbol{\Omega}_{p2} + \beta^2\boldsymbol{M}_{K2}\boldsymbol{\Omega}_e)\right] \tag{1.11}$$

$$\boldsymbol{D}_2 = \boldsymbol{C}_2\boldsymbol{\zeta}_2^{-1}. \tag{1.12}$$

where $\lambda_1 \equiv \lambda_I + \lambda_J$, $\lambda_2 \equiv \lambda_I + \lambda_J + \lambda_K$ and $\boldsymbol{\Lambda} = \left(\begin{smallmatrix}\lambda_2 & 0\\ 0 & \lambda_1\end{smallmatrix}\right)$.

$\boldsymbol{B}_2$ is the constant discounted dividend growth rate, and together $\boldsymbol{A}_2$ and $\boldsymbol{B}_2$ govern the level of the equilibrium price. $\boldsymbol{C}_2$ is the price sensitivity to dividend innovation, and $\boldsymbol{D}_2$ regulates the amount of demand shocks contained in the price. Purchased data amount $\boldsymbol{M}_{Lt}$, data processing capacity $\bar{M}$, the data processing technological β, β_I and the precision of soft information $\boldsymbol{\Omega}_\epsilon$ all enter the price coefficients via the signal-to-noise ratio and the average posterior precision of dividend innovation.

Data choices Domestic and foreign institutional investors face slightly data choices problems. I start by characterizing the optimal data choice of domestic institutional investor i. Her *ex ante* utility can be found in the Appendix. What is crucial is summarized in the following lemma.

LEMMA 1. *The Period-2 data choice of domestic institutional investor i only enter her Period-2 ex ante utility function via $\left|\hat{\boldsymbol{\Sigma}}_{i2}\right|$, where $\hat{\boldsymbol{\Sigma}}_{i2}$ is her posterior variance of Period-2 payoff.*

Since the data choice $(M^{(\mathrm{I})}_{i2}, M^{(\mathrm{II})}_{i2})$ only enter the *ex ante* utility function though the determinant of the posterior variance of payoff $\left|\hat{\boldsymbol{\Sigma}}_{i2}\right|$, the data choice problem reduces to allocating data points to maximize the determinant of the posterior precision $\left|\hat{\boldsymbol{\Omega}}_{i2}\right|$. The F.O.C. yields $\frac{\hat{\Omega}^{(\mathrm{I})}_{iy2}}{\beta_I^2\Omega_e^{(\mathrm{I})}} = \frac{\hat{\Omega}^{(\mathrm{II})}_{iy2}}{\beta_I^2\Omega_e^{(\mathrm{II})}}$. Intuitively, domestic institutional investor i allocates data points such that the equivalent data amount to the posterior precision of current dividend innovation are the same for each type of risky assets. Her optimal data allocation depends on the posterior precision equivalent data amount, which I formally define next.

Definition 2. *The equivalent data amount $m^{(\mathrm{I})}$ to the posterior precision of dividend innovation of type-I risky asset for domestic institutional investor i is the amount of data that would be needed if the posterior precision of dividend innovation were to be obtained by data alone. Formally,* $m^{(\mathrm{I})}(M_{i2}^{(\mathrm{I})}) = \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\beta_I^2 M_{i2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}}$.

Denote the optimal data allocation profile of domestic institutional investor i to be $(M_{i2}^*, \bar{M} - M_{i2}^*)$. Intuitively, if at the full data processing capacity the equivalent data amount $m^{(\mathrm{I})}(\bar{M})$ to the posterior precision of dividend innovation of type-II risky asset is still less than the equivalent amount of data when processing zero data for type-I risky asset, then domestic institutional investor i will opt to not learn about type-I asset and direct all her data processing resources to type-II asset, so $M_{i2}^* = 0$. In the opposite case, if the equivalent data amount $m^{(\mathrm{II})}(0)$ to the posterior precision of dividend innovation when processing zero data for type-II risky asset is larger than the equivalent amount of data at full data processing capacity for type-I risky asset, then domestic institutional investor i will only learn about type-I asset, so $M_{i2}^* = \bar{M}$. Therefore, her optimal data acquisition amount for type-I risky asset is

$$M_{i2}^* = \begin{cases} 0, & \text{if } m^{(\mathrm{I})}(0) \geq m^{(\mathrm{II})}(\bar{M}), \\ \bar{M}, & \text{if } m^{(\mathrm{II})}(0) \geq m^{(\mathrm{I})}(\bar{M}), \\ \frac{1}{2}\left[m^{(\mathrm{II})}(\bar{M}) - m^{(\mathrm{I})}(0)\right], & \text{otherwise.} \end{cases} \tag{1.13}$$

The conditions when investor i would tilt data processing capacity towards type-I risky asset is thus clear.

Proposition 1. *Domestic investor i would learn more about type-I than type-II risky asset if $m^{(\mathrm{I})}(0) < m^{(\mathrm{II})}(0)$. If further $m^{(\mathrm{I})}(\bar{M}) \leq m^{(\mathrm{II})}(0)$, then investor i would concentrate all data processing capacity on type-I risky asset.*

Foreign institutional investor k is only allowed to trade type-I risky asset but her data choice problem is not a priori trivial; it is possibly in her best interests to acquire and process data points for type-II risky asset. The way her data choices enter her *ex ante* utility is summarized in the following lemma.

LEMMA 2. *The Period-2 data choice of foreign institutional investor k only enter her Period-2 ex ante utility function via $\hat{\Sigma}^{(\mathrm{I})}_{k2}$, where $\hat{\Sigma}^{(\mathrm{I})}_{k2}$ denotes the first diagonal element of her posterior variance for Period-2 payoff and is the posterior variance for type-I risky asset.*

Her data choice problem reduces to maximize $\hat{\Omega}^{(\mathrm{I})}_{ky2} - \frac{\hat{\mathcal{C}}^2_{ky2}}{\hat{\Omega}^{(\mathrm{II})}_{ky2}}$, reflecting a revision of portfolio position resulted from the information on type-II risky asset used to update the posterior belief of type-I risky asset. The F.O.C. yields $\beta^2\Omega^{(\mathrm{I})}_e = \left(\frac{\hat{\mathcal{C}}_{ky2}}{\hat{\Omega}^{(\mathrm{II})}_{ky2}}\right)^2 \beta^2\Omega^{(\mathrm{II})}_e$. Intuitively, the marginal gain in posterior precision of dividend innovation for type-I risky asset directly from learning about type-I risky asset and indirectly from learning about type-II risky asset should be equal. To proceed, I define an important concept for information spillover.

DEFINITION 3. *The information spillover rate is defined as $\kappa(M^{(\mathrm{II})}_{k2}) \equiv \left(\frac{\hat{\mathcal{C}}_{ky2}}{\hat{\Omega}^{(\mathrm{II})}_{ky2}}\right)^2 \frac{\Omega^{(\mathrm{II})}_e}{\Omega^{(\mathrm{I})}_e}$, where $\hat{\mathcal{C}}_{ky2}$ and $\hat{\Omega}^{(\mathrm{II})}_{ky2}$ are the off-diagonal and the second diagonal element of investor k's posterior precision for dividend innovations, respectively.*

Her optimal data allocation profile depends on the information spillover rate. One unit of precision for type-II risky asset's dividend innovation converts to κ units of precision for type-I risky asset. If the information spillover rate when she does not learn about type-II risky asset is already less than or equal to 1, then the needed data amount to make the marginal gain in posterior precision of dividend innovation for type-I risky asset directly from learning about type-I equal that indirectly from learning is negative, so she will only learn about type-I asset, $M^*_{k2} = 0$. In the opposite case, if the information spillover rate at full data processing capacity is still larger than or equal to 1, then it is in her best interests to learn as much about type-II risky asset as possible, $M^*_{k2} = \bar{M}$. Then her data choice can be expressed as $(\bar{M} - M^*_{k2}, M^*_{k2})$, where

$$M^*_{k2} = \begin{cases} 0, & \text{if } \kappa(0) \leq 1, \\ \bar{M}, & \text{if } \kappa(\bar{M}) \geq 1, \\ \frac{\Omega^{(\mathrm{II})}_y + \hat{\Omega}^{(\mathrm{II})}_{p2}}{\beta^2\Omega^{(\mathrm{II})}_e} \left[\left|\sqrt{\kappa(0)}\right| - 1\right], & \text{otherwise.} \end{cases} \tag{1.14}$$

While foreign institutional investor k could optimally choose to learn about type-II risky asset, the generated information is not aggregated at the market level efficiently. Efficiency is achieved

when $M_k^{(\text{II})} = 0$. Intuitively, if the two types of risky assets are uncorrelated, learning about type-II risky asset would not help learn about type-I risky asset, foreign institutional investors would allocate all data processing capacity to type-I asset. This is summarized in the following lemma.

LEMMA 3. *If $\gamma = 0$, that is type-I and -II risky assets are not correlated, the information spillover rate $\kappa = 0$ and the amount of data acquired and processed by foreign institutional inventors for type-II risky asset $M_{k2}^* = 0$.*

When the two type of risky assets are correlated, the amount of data on type-II risky assets acquired and processed by institutional investors declines with the increase of their technology level, as shown in Figure 1.2.

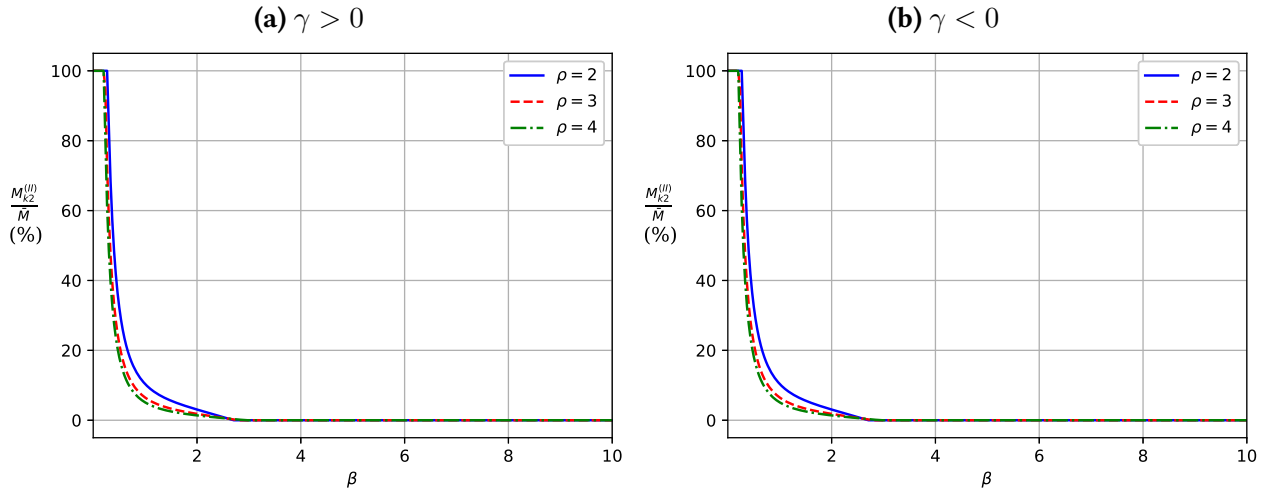


Figure 1.2. Data amount acquired for type-II asset by foreign institutional investors in response to the foreign data processing technology level

The left and right panels are plotted with positively and negatively correlated type-I and -II assets, respectively. As long as the two types of assets are correlated, the data acquired and processed by foreign institutional investors for type-II risky assets decrease with their data processing technology level.

It is ensured to happen if both the domestic and the domestic institutional investors are sufficiently technologically advanced or underdeveloped for processing the hard data for type-II risky assets, regardless of the difference of their technology levels per se.

Proposition 2. *When $\Omega_e^{(\text{II})}$ is sufficiently large or sufficiently close to zero, foreign institutional investors do not learn about type-II risky asset. Mathematically,*

$$\lim_{\Omega_e^{(\text{II})}\to\infty} M_{k2}^{(\text{II})} = 0 \text{ and } \lim_{\Omega_e^{(\text{II})}\to 0} M_{k2}^{(\text{II})} = 0.$$

When $\Omega_e^{(\text{II})}$ is sufficiently large, the signal-to-noise ratio for type-II risky asset will be very large too, which in turn makes the information spillover rate very small. When $\Omega_e^{(\text{II})}$ is sufficiently close to zero, very few information can be transmitted to type-I risky asset, and the information spillover rate will also be very small. When the information spillover rate falls below one, then learning directly from type-I risky asset is more efficient than learning indirectly from type-II risky asset. Thus in these cases foreign institutional investors do not learn about type-II risky asset. As we can see in Figure 1.4, the proportion of data processing capacity allocated to type-II assets by foreign institutional investors drops to 0 at the both end of the data precision axis.

There is a less strict condition for zero acquisition and processing in type-II data. Intuitively, When the unit precision of hard information for type-I asset directly from processing data on type-I asset exceeds that acquired through processing data on type-II asset via information spillover, foreign institutional inventors will not acquire data on type-II asset. This condition is formally summarized in the following proposition.

Proposition 3. *A sufficient condition for foreign institutional investors to exhaust data processing capacity $\bar{M}$ on type-I risky asset is $\Omega_e^{(\text{I})} > \frac{\Omega_e^{(\text{II})}|\mathcal{C}_y|}{\Omega_y^{(\text{II})}}$.*

If the above condition is satisfied, then the existence and accuracy of soft information does not affect foreign institutional investors' data acquisition. Otherwise, data acquisition for type-II asset declines with the abundancy of soft information, as is shown in Figure 1.5.

Foreign institutional investor's acquisition of hard information for shares not eligible for foreign investments decline with the abundancy of soft information, regardless of investability by foreign investors, available to domestic investors, as shown in Figure 1.5.

Figure 1.7 plots the reactions of signal-to-noise ratios to foreign data processing technology

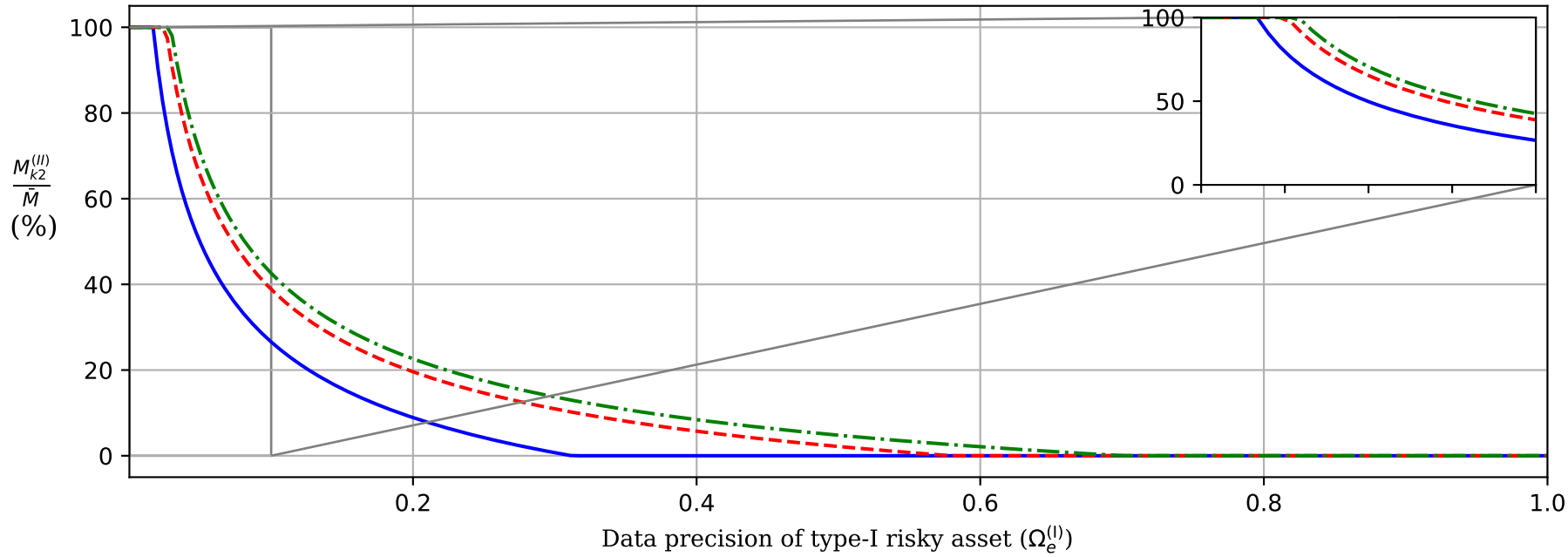


Figure 1.3. Data amount for type-II asset by foreign institutional investors in response to data precision for type-I asset

This figure plots the acquired and processed data amount for type-II risky asset (foreign ineligible) by foreign institutional investors in proportion to their data processing capacity against the data precision for type-I asset. The two types of risky assets are correlated with correlation coefficient $\gamma = 0.64$.

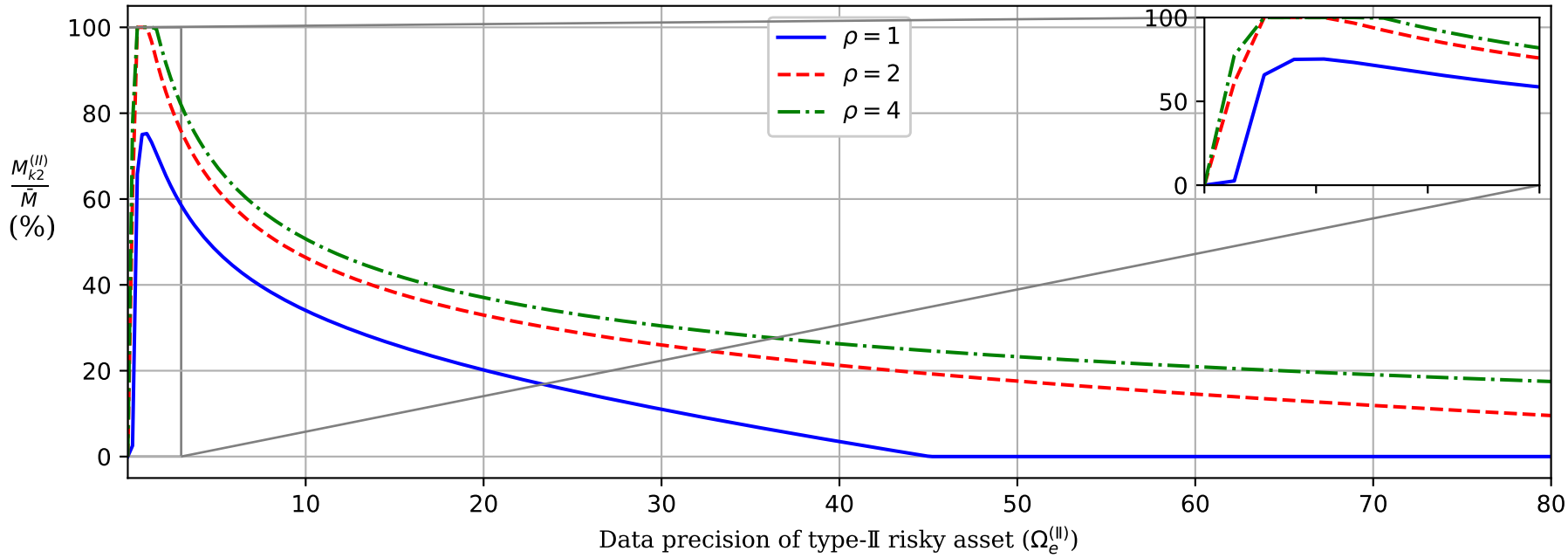


Figure 1.4. Data amount acquired for type-Ⅱ asset by foreign institutional investors in response to data precision of type-Ⅱ asset

This figure plots the acquired and processed data amount for type-Ⅱ risky asset (not foreign-investable) by foreign institutional investors in proportion to their data processing capacity against the data precision for that asset. The two types of risky assets are correlated with correlation coefficient $\gamma = 0.64$.

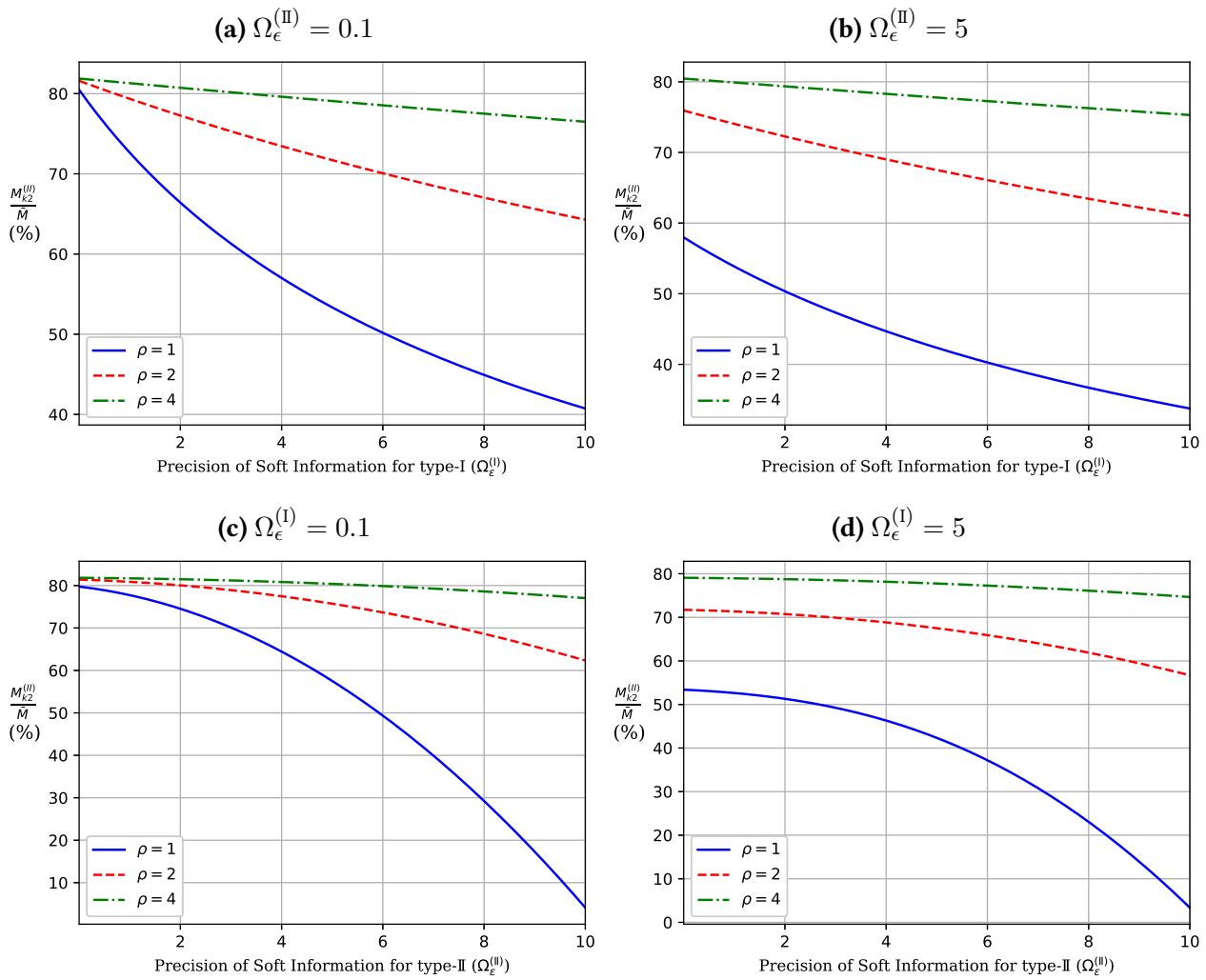


Figure 1.5. Data amount acquired for type-II asset by foreign institutional investors in response to soft information precision

This figure plots how foreign institutional investors' data acquisition on type-II risky asset change with the soft information held by domestic investors. Seen in all the four panels is that foreign institutional investor's acquisition on hard information for foreign-ineligible assets decline with the abundancy of soft information available to domestic investors, regardless of investability by foreign investors.

levels. As the technology level goes up, the price of type-I risky asset contains more information on its fundamental innovation. The effects come from two sources. On the one hand, the higher the foreign data processing technology level is, the more data foreign institutional investors would choose to acquire and process. On the other hand, the higher average data processing technology level also increases the average hard information in the market.

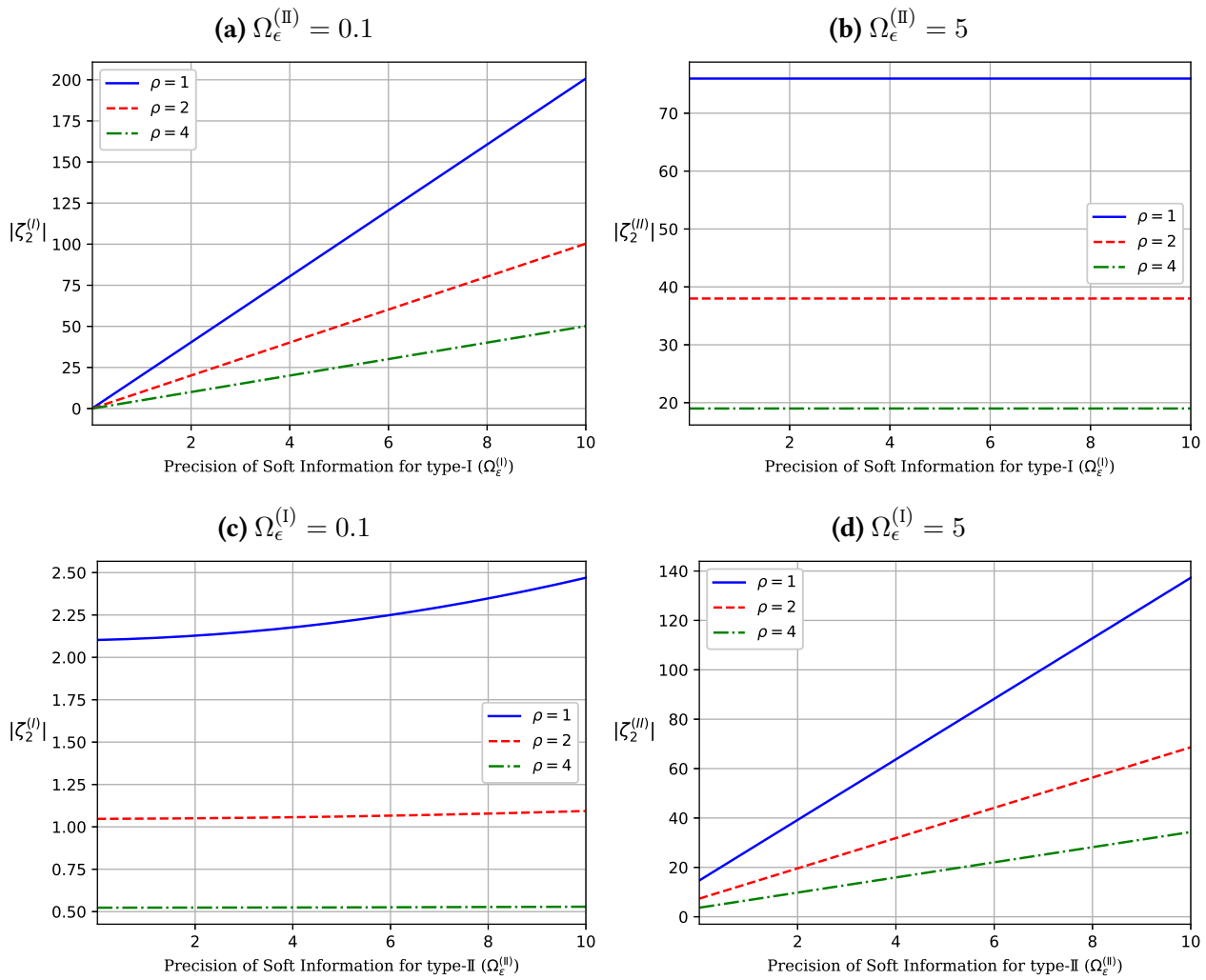


Figure 1.6. Reactions of the signal-to-noise ratios to soft information precision

This figure plots how the signal-to-noise ratio of one asset increases with the precision of soft information about that asset. With an increase in the soft information precision for the other type of asset, the change of the signal-to-noise ratio is small.

1.2.3 Impacts of partial market opening

Market opening influences the equilibrium price via three channels. Firstly, through the signal-to-noise ratio ζ_2. There is a spillover effect on type-Ⅱ risky asset via data acquisition. The entry of foreign institutional investors dilutes soft information but brings in more information from data. In response, domestic institutional investors could reallocate her data acquisition between the two types of assets. How the type-I risky asset is affected is more complicated. It

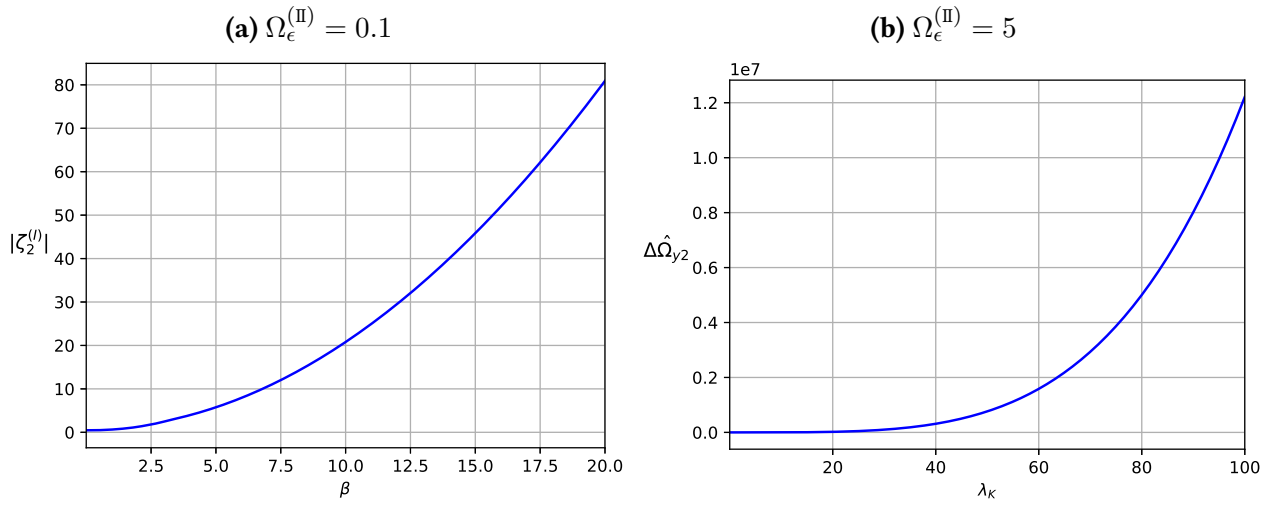


Figure 1.7. Reactions of the signal-to-noise ratios to foreign technology level

This figure plots how signal-to-noise ratios reacts to foreign data processing technology levels. As the technology level goes up, the price of type-I risky asset contains more information on its fundamental innovation. With the same amount of data acquired, the investors will be able to extract more hard information.

comprises of a possible change in domestic institutional investors' data choice, the addition of foreign institutional investors' data processing, and a dilution of soft information. Moreover, if foreign institutional investors choose to acquire and process some data for type-II risky asset, the resulted information will be inefficiently aggregated for type-I.

Secondly, through the weighted average of the posterior variance of the dividend innovation $\bar{\hat{\boldsymbol{\Sigma}}}_{y2}$.

Thirdly, through the unit demand level.

Change in data choices of domestic institutional investors The counterfactual where in Period 2 the domestic market remains closed is readily available from the partial opening model if we set $\lambda_K = 0$.

Definition 4. *The equilibrium equivalent data amount $m_e^{(\mathrm{I})}$ to the posterior precision of dividend innovation of type-I risky asset for domestic institutional investor i is the amount of data that would be needed if the posterior precision of dividend innovation were to be obtained by data alone in equilibrium. Formally, $m_e^{(\mathrm{I})}(M_{I2}^{(\mathrm{I})}) = \frac{\Omega_y^{(\mathrm{I})} + (\frac{r-g_1}{\rho r}(\lambda_1\Omega_\epsilon^{(\mathrm{I})} + \lambda_I\beta_I^2 M_{I2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})}))^2\Omega_x^{(\mathrm{I})} + \beta_I^2 M_{I2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}}$.*

Proposition 4. *In equilibrium,*

1) in the counterfactual world where the market remains closed in Period 2, if the equilibrium equivalent data amount with zero data processing for type-I risky asset is still larger than or equal to the equilibrium equivalent data amount at data processing capacity for type-II formally $m_e^{(\mathrm{I})}(0) \geq m_e^{(\mathrm{II})}(\bar{M})$*, then domestic institutional investors will not acquire and process data for type-I;*

2) if in addition the unit data precision for type-I is smaller than the unit precision converted from type-II data to type-I, i.e. $\Omega_e^{(\mathrm{I})} \leq z\Omega_e^{(\mathrm{II})}$*, then domestic institutional investors will not acquire and process data for type-I when the market is partially open.*

Signal-to-noise ratio Partial market opening have impacts on the signal-to-noise ratio for type-I and type-II risky assets. Apart from the common channel via the change of domestic institutional investors' data choices that affects both types of assets, the two types of risky assets are affected differently.

For type-I risky asset, the entry of foreign institutional investors could have direct and/or indirect impacts on its signal-to-noise ratio. If they choose to learn about type-I risky asset, the information extracted from acquire-and-processed data is aggregated into its equilibrium price, thus directly affecting the signal-to-noise ratio for type-I asset. The effects depend positively on the number of foreign institutional investors, data processing capacity, and data signal precision. If they choose to learn about type-II risky asset, the extracted information is not aggregated into the equilibrium price for type-II but aggregate into that for type-I via information spillover, which indirectly affects the signal-to-noise ratio of type-I risky asset.

For type-II risky asset, domestic institutional investors are the sole contributors to the signal-to-noise ratio. Since market opening has no effect on their endowed soft information, the direction in which the signal-to-noise ratio changes only depends on the data choice of domestic institutional investors: if they learn less about type-II risky asset, i.e. $M^{(\mathrm{II})} < M^{\dagger(\mathrm{II})}$, then the signal-to-noise ratio of type-II risky asset drops; if they learn more, then the signal-to-noise ratio increases.

To compare the signal-to-noise ratio after market opening $\boldsymbol{\zeta}_2$ with the counterfactual $\boldsymbol{\zeta}_2^\dagger$ where the market remain closed to international investors, let me take the difference of their negatives. To simplify notations, let me define $\omega_L^{(\mathrm{I})}(M_{L2}^{(\mathrm{I})}, \beta_L) = \lambda_L \beta_L^2 M_{L2}^{(\mathrm{I})} \Omega_e^{(\mathrm{I})}$ to be the aggregated data signal precision of type-I risky asset for investor type L in Period 2 and similarly $\omega_L^{(\mathrm{II})}(M_{L2}^{(\mathrm{II})}, \beta_L) = \lambda_L \beta_L^2 M_{L2}^{(\mathrm{II})} \Omega_e^{(\mathrm{II})}$ for type-II risky asset. Then $\Delta\left|\zeta_2^{(\mathrm{I})}\right| = \frac{r-g_1}{\rho r}[\omega_I^{(\mathrm{I})}(M_{I2}^{(\mathrm{I})} - M_{I2}^{\dagger(\mathrm{I})}, \beta_I) + \omega_K^{(\mathrm{I})}(M_{K2}^{(\mathrm{I})}, \beta)]$, $\Delta\left|\zeta_2^{(\mathrm{II})}\right| = \frac{r-g_2}{\rho r}\omega_I^{(\mathrm{I})}(M_{I2}^{(\mathrm{II})} - M_{I2}^{\dagger(\mathrm{II})}, \beta_I)$, and $\Delta\left|\zeta_2^{(c)}\right| = \frac{r-g_1}{\rho r}\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}\omega_L^{(\mathrm{II})}(M_{K2}^{(\mathrm{II})}, \beta)$. Let me call $\zeta_2^{(\mathrm{I})}$ and $\zeta_2^{(\mathrm{II})}$ the self-signal-to-noise ratio, which is the rate of its own dividend innovation signal to its noise contained in the equilibrium price; $\zeta_2^{(c)}$ the cross-signal-to-noise ratio, which is the rate of type-II dividend innovation signal to the noise of type-I risky asset contained in its equilibrium price. The signal-to-noise ratios depend on the risk-free rate, the dividend growth rates, the risk-aversion coefficient, numbers of investors, data processing technology β_I and β, and the data choices of institutional investors.

Notice that $\zeta_2^{(c)}$ is only present if foreign institutional investors choose to learn about type-II risky asset. In the extreme cases where foreign institutional investors only learn about type-II risky asset, this information does not aggregate into the either of the two self-signal-to-noise ratios. $\Delta\left|\zeta_2^{(\mathrm{I})}\right|$ and $\Delta\left|\zeta_2^{(\mathrm{II})}\right|$ thus have opposite signs.

PROPOSITION 5. *Given that* $\Omega_e^{(\mathrm{I})} > \frac{\Omega_e^{(\mathrm{II})}}{\Omega_y^{(\mathrm{II})}}|\mathcal{C}_y|$ *or* $\kappa(0) \leq 1$,

1) A sufficient condition for partial market opening to have positive effect on the self-signal-to-noise ratio for type-I risky asset is $\lambda_K > \lambda_I \frac{\beta_I^2}{\beta^2}$.

2) If the number of foreign institutional investors is smaller than $\frac{(M^{\dagger(\mathrm{I})} - M^{(\mathrm{I})})\beta_I^2}{\bar{M}\beta^2}\lambda_I$ *but their entry induces domestic institutional investors to learn less about type-I risky asset, then the self-signal-to-noise ratio for for type-I risky asset decreases after market opening.*

3) The cross-signal-to-noise is zero.

PROPOSITION 6. *Given that* $\kappa(\bar{M}) \geq 1$,

1) Partial market opening has positive impacts on the self-signal-to-noise ratio if domestic institutional investors learn more about that asset with the entry of foreign investors.

2) The changes in the self-signal-to-noise ratios for the two type of risky assets are additive inverses of each other.

Average posterior precision of dividend innovation Similarly, I define and compute the effect of market opening on average posterior precision of dividend innovation as $\Delta\bar{\hat{\Omega}}_{y2}^{(\mathrm{I})} = (\zeta_2^{(\mathrm{I})2} - \zeta_2^{\dagger(\mathrm{I})2})\Omega_x^{(\mathrm{I})} + (\frac{\lambda_I}{\lambda_2}M^{(\mathrm{I})}\beta_I^2 - \frac{\lambda_I}{\lambda_1}M^{\dagger(\mathrm{I})}\beta_I^2 + \frac{\lambda_K}{\lambda_2}\bar{M}\beta^2)\Omega_e^{(\mathrm{I})} + (\frac{\lambda_I}{\lambda_2} - \frac{\lambda_I}{\lambda_1})\Omega_\epsilon^{(\mathrm{I})}$ and $\Delta\bar{\hat{\Omega}}_{y2}^{(\mathrm{II})} = (\zeta_2^{(\mathrm{II})2} - \zeta_2^{\dagger(\mathrm{II})2})\Omega_x^{(\mathrm{II})} + (\frac{\lambda_I}{\lambda_2}M^{(\mathrm{II})} - \frac{\lambda_I}{\lambda_1}M^{\dagger(\mathrm{II})})\beta_I^2\Omega_e^{(\mathrm{II})} + (\frac{\lambda_I}{\lambda_2} - \frac{\lambda_I}{\lambda_1})\Omega_\epsilon^{(\mathrm{II})}$. The change of average posterior precision of dividend innovation can be decomposed into three parts. Firstly, a change in signal-to-noise ratio leads to a change in investors' average ability to extract information from the price for both types of risky assets and thus a change in the precision of price signals. As with the different effects of market opening on signal-to-noise ratios, here the effects on changes in the precision of price signals for type-I and type-II risky assets are also different. Secondly, a change in the amount of acquired and processed data per investor, or *data density*, leads to a change in the precision of private data signals. For both types of risky assets, the possible change of data choice by domestic institutional investors and the change of their weight after market opening leads to a change in data density. For type-I risky asset, foreign institutional investors bring in an extra $\lambda_K\bar{M}$ points of data in total. Thirdly, the integrated soft information in the market is fixed and thus entering of foreign investors lowers the density of soft information, bringing in a negative impact on the average posterior precision of dividend innovation.

Proposition 7. *For type-I risky asset,*

1) If the number of foreign institutional investors is smaller than $\frac{(M^{\dagger(\mathrm{I})} - M^{(\mathrm{I})})\beta_I^2}{\bar{M}\beta^2}\lambda_I$ *but their entry induces domestic institutional investors to learn less about type-I risky asset, then there is less information about type-I dividend innovation in the market, that is the average posterior precision of dividend innovation decreases.*

2) Given the number of foreign institutional investors is large enough to ensure increasing of signal-to-noise ratio, that is $\lambda_K > \lambda_I\frac{\beta_I^2}{\beta^2}$,

(i) if inclusion of new data and introduction of perfect data processing technology through for-

eign institutional investors outweighs the dilution of soft information held by domestic institutional investors, then the average precision of dividend innovation increases when domestic institutional investors learn more about it, or when they learn less, but the number of foreign institutional investors is large enough that

$$\lambda_K > \frac{(M^{\dagger(\mathrm{I})} - M^{(\mathrm{I})})\beta_I^2\Omega_e^{(\mathrm{I})}}{\bar{M}\beta^2\Omega_e^{(\mathrm{I})} - \frac{\lambda_I}{\lambda_1}(M^{\dagger(\mathrm{I})}\beta_I^2\Omega_e^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})})}\lambda_I;$$

(ii) if the dilution of soft information outweighs inclusion of new data and introduction of perfect data processing technology, then the average precision of dividend innovation increases only when the number of foreign institutional investors is small enough, that is

$$\lambda_K < \frac{(M^{\dagger(\mathrm{I})} - M^{(\mathrm{I})})\beta_I^2\Omega_e^{(\mathrm{I})}}{\bar{M}\beta^2\Omega_e^{(\mathrm{I})} - \frac{\lambda_I}{\lambda_1}(M^{\dagger(\mathrm{I})}\beta_I^2\Omega_e^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})})}\lambda_I,$$

such that domestic institutional investors learn more about type-I dividend innovation.

Proposition 8. *For type-II risky asset,*

1) If the entry of foreign institutional investors induces domestic institutional investors to learn less about it, i.e. $M^{(\mathrm{II})} < M^{\dagger(\mathrm{II})}$, then the average posterior precision of dividend innovation decreases.

2) If the number of foreign institutional investors is not too large, or specifically

$$\lambda_K < \frac{(M^{(\mathrm{II})} - M^{\dagger(\mathrm{II})})\Omega_e^{(\mathrm{II})}}{M^{\dagger(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}}(\lambda_I + \lambda_J)$$

and their entry induces domestic institutional investors to learn more about the type-II risky asset, i.e. $M^{(\mathrm{II})} > M^{\dagger(\mathrm{II})}$, then the average posterior precision of dividend innovation increases.

1.3 Shanghai-Hong Kong Stock Connect

This section turns to the Chinese capital market liberalization, for which I specifically calibrate the partial-opening model to the Shanghai-Hong Kong Stock Connect Program. The China case

is selected for its fitness to the soft information and data processing technology assumptions, as well as its huge impacts on world economy.

1.3.1 Background

The domestic Chinese stock market (the "A-share" market) started in 1990 with the establishment of two independently operated domestic stock exchanges: the Shanghai Stock Exchange (SSE hereinafter) and the Shenzhen Stock Exchange (SZSE hereinafter). However, the Chinese stock market is not well-connected with the international markets. Table 1.2 shows a selected list of world markets with their percentage of shares and market value eligible for foreign investment during their capital market liberalization process. The Shanghai-Hong Kong Stock Connect Program is an important step China takes to open its capital markets to the outside world. Launched on November 17, 2014, this program allows investors in Hong Kong, local and global, to trade shares in the SSE (north-bound trades) via Hong Kong Stock Exchange (HKSE hereinafter). See Figure A.2 in the appendix for the details of how international trade is carried out. The Shanghai-Hong Kong Connect mainly covers index component stocks with large market cap listed in the SSE and HKEX (for the details of trading eligibility refer to Table A.2 in the appendix).

1.3.2 Calibration

Investor composition The aggregate data on composition of shareholders are obtained from China Securities Regulatory Commission (CSRC). The largest holder of Chinese A-shares is industrial capital, including non-financial companies and general legal persons. Considering that industry capital and original shareholders' holdings cannot be freely circulated[4], they are not included in the discussions of this chapter. In the same vein, the measure of λ_L's is calculated by *negotiable market capitalization*, i.e. market capitalization for tradable common listed shares only.

[4]These are non-tradable shares (e.g. common shares held by state legal persons or orginal promotors.) in Chinese stock markets. They cannot be traded in the exchange, but can be traded by auction or transfer by agreement under the approval of the CSRC.

Markets	Date	Percentage of Shares	Percentage of Market Value
Argentina	Sep-89	58.3	-
Brazil	May-91	47.7	-
Chile	Oct-89	34.6	-
Colombia	Dec-91	25.0	-
India	Nov-92	62.9	-
Korea	Jan-92	91.7	-
Mexico	May-89	12.9	-
Pakistan	Feb-91	7.3	-
Taiwan (China)	Jan-91	100.0	100.0
Turkey	Aug-89	64.3	-
Venezuela	Jan-90	38.5	-
Shanghai (China)	Nov-14	59.0	82.8
Shenzhen (China)	Dec-16	53.1	67.0

Table 1.2: Percentage of foreign-investable shares in markets when liberalization was implemented

Percentages for Shanghai and Shenzhen are calculated with the end-of-month values. Eligibility is provided by HKSE, and market values by SSE and SZSE. Percentage of shares for other markets are calculated based on the raw data provided in Tables I and II by Chari and Henry (2004).

The second largest shareholders in the A-share stock market are individual investors. In 2013, the shareholding ratio of individual investors exceeded 33%. The ratio peaked in 2015 exceeding 39%, and then gradually started to decrease, below 30% in 2023. For the time window surrounding the launch of Shanghai-Hong Kong Stock Connect, I take the simple average and estimate $\lambda_J = 0.35$. The third largest shareholders in the A-share stock market are institutional investors. They hold around 15% of the A-shares and thus $\lambda_I = 0.15$. Foreign capital holds about 3% of the A-shares. Therefore $\lambda_K = 0.03$. Notice that λ's are not the proportion of each type of investors in the model since I consider only the negotiable market capitalization. The proportions of domestic individual investors, domestic institutional investors, and foreign institutional investors are 66%, 28%, and 6%, respectively.

Risk-free rate The risk-free rate $r = 1.042$ is taken as the mean value of yield to maturity on 10 year government bond in the year of 2014. The government bond yield curve is published by China Central Depository & Clearing Co., Ltd. (CCDC).

Risky assets Data on the daily prices of shares prices and eligibility details of the Stock Connect Programs are obtained from the China Stock Market & Accounting Research (CSMAR) database. Grouped by their eligibility for northbound trade, shares are aggregated by negotiable market capitalization. The variance-covariance structure of the two aggregate shares are $\begin{pmatrix} 4.1721 & 2.9818 \\ 2.9818 & 2.5778 \end{pmatrix}$.

Following the measurement methodology in the literature, I use unexpected earnings to proxy for dividend innovation $\widetilde{y}_t$. Empirically, unexpected earnings, or earnings surprises, are typically obtained by comparing analyst forecasts and earnings announcement, or using time-series model forecasts.

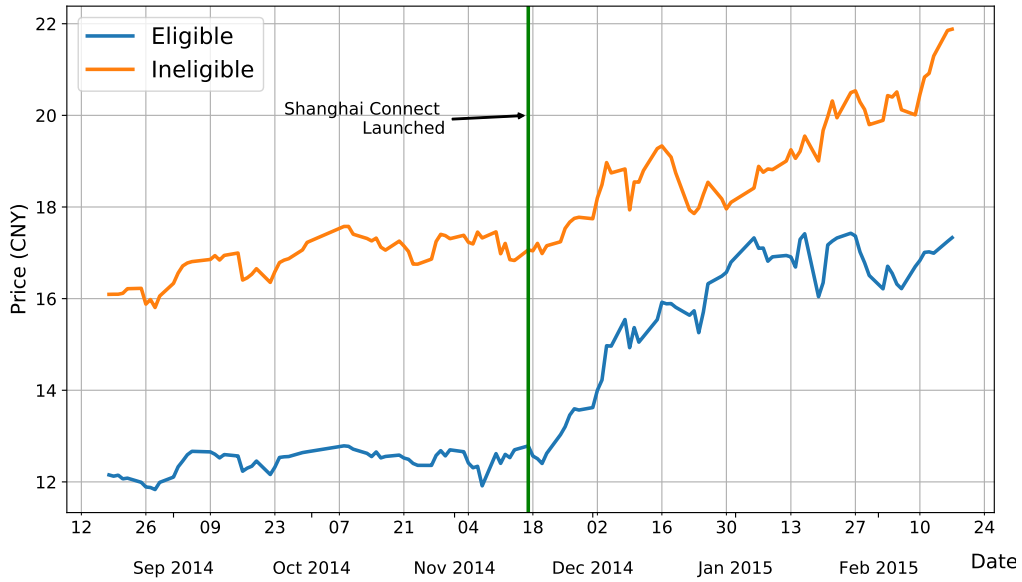
For measurement through analyst forecasts, I obtain all analyst forecasts on earnings per share and the corresponding annual eps announcement for the relevant companies. To ensure a forecast window of one year, I only keep forecasts that were reported in the previous fiscal year of the forecast end date. Within this set of earnings forecasts, I drop all but the latest forecasts from each analyst and then take the median across analysts for each firm and year. The unexpected earnings are constructed by subtracting the median forecasts from the announced earnings. It is normalized by multiplying the number of shares outstanding, and then divided by the total assets. The resulting measure is denoted by $ue_{i,t}$.

For measurement through time-series model forecasts, I obtain annual earnings announcement data for each firm from 2010 to 2018. It is then fitted to a VAR(1) model. The model residuals are taken to be earnings surprise.

Last period's dividend is measured by the EPS announcement for the previous year. It is also multiplied by the number of shares outstanding, and divided by total assets. The resulting measure is denoted by $ea_{i,t-1}$.

The internationally tradable shares are constituents of the SSE 180 index, which are selected based on total market capitalization, negotiable market capitalization and industry. As a result, shares that are eligible and ineligible for international trade through the Connect program are fundamentally different. To deal with this problem, I adopt the propensity score matching

(a) Aggregate price levels



(b) Annualized volatility

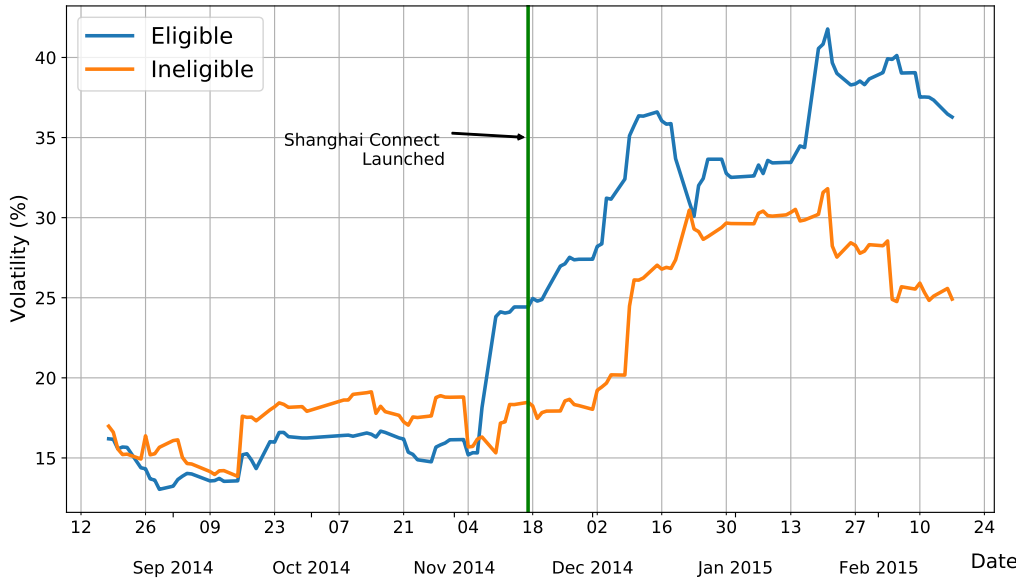


Figure 1.8. Prices and volatilities of A-shares around the launch of the Shanghai-Hong Kong Stock Connect Program

Panel (a) plots the price levels aggregated by negotiable market capitalization and Panel (b) plots the annualized rolling volatility, with a window of 30 trading days. Type-I security is eligible for northbound trade via the Connect while type-II is ineligible.

technique and calculate the propensity scores on total market capitalization, negotiable market capitalization and industry dummies. The kNN algorithm is used to match each type-I asset to

its closest neighboring type-Ⅱ asset.

Since previous earnings and unexpected earnings are both normalized by multiplying the number of shares outstanding and divided by total assets, the same operation is done to market prices, which is the market capitalization divided by total assets, denoted by $ma_{i,t}$.

Estimation To obtain estimates for price coefficients, I run the following cross-sectional regression for Period $t = 1, 2$,

$$
\begin{aligned}
ma_{i,t}^{(\mathrm{I})} &= \hat{A}_t^{(\mathrm{I})} + \hat{B}_t^{(\mathrm{I})} ea_{i,t-1}^{(\mathrm{I})} + \hat{C}_t^{(\mathrm{I})} ue_{i,t}^{(\mathrm{I})} + \hat{C}_t^{(c)} ue_{i,t}^{(\mathrm{II})} + \hat{K}_{i,t}^{(\mathrm{I})} Z_{i,t}^{(\mathrm{I})} + \hat{\varepsilon}_{i,t}^{(\mathrm{I})}, \\
ma_{i,t}^{(\mathrm{II})} &= \hat{A}_t^{(\mathrm{II})} + \hat{B}_t^{(\mathrm{II})} ea_{i,t-1}^{(\mathrm{II})} + \hat{C}_t^{(\mathrm{II})} ue_{i,t}^{(\mathrm{II})} + \hat{K}_{i,t}^{(\mathrm{II})} Z_{i,t}^{(\mathrm{II})} + \hat{\varepsilon}_{i,t}^{(\mathrm{II})},
\end{aligned}
$$

where $Z_{i,t}$ is a collection of dummies at the 2012 CSRC sub-industry level. The estimated coefficients are imperfect measures for the price coefficients $\boldsymbol{A}_t$, $\boldsymbol{B}_t$ and $\boldsymbol{C}_t$. The residuals are taken as an estimate for noise contained in the price, that is $\boldsymbol{D}_t \boldsymbol{\Sigma}_x \boldsymbol{D}_t'$.

There is a preliminary test for the model according to Lemma 4, which requires the estimated coefficient $\hat{C}_t^{(\mathrm{I})}$ to be positive. $\hat{C}_t^{(\mathrm{I})}$ is estimated to be 3.71, significant at the 5% level.

Lemma 4. *If* $C_2^{(\mathrm{I})} < 0$, *then* $\gamma < 0$.

Since the signal-to-noise ration $\boldsymbol{\zeta}_t$ and posterior variances all depend on data choices, I take a two-step estimation method. First, I specify data choices for foreign institutional investors $(\bar{M} - M_{K2}^{(\mathrm{II})}, M_{K2}^{(\mathrm{II})})$, then domestic institutional investors' data choices can be expressed in terms of $M_{K2}^{(\mathrm{II})}$ and model primitives. Second, use the GMM method for Equations (1.15-1.18) to estimate the parameters $\rho, g_1, g_2, \beta, \Omega_\epsilon^{(\mathrm{I})}, \Omega_\epsilon^{(\mathrm{II})}, \Omega_x^{(\mathrm{I})}, \Omega_x^{(\mathrm{II})}$ (denoted by $\boldsymbol{\theta}$ on the left hand of the equations for notational simplicity) and check if the condition for that specific set of data choices to be optimal is satisfied. In particular, the estimate of $C_2^{(c)}$ is non-zero with t-value -1.399, we estimate 6 particular sets of data choices, I specify 3 particular sets of data choices for foreign institutional investors: 1. $(0, \bar{M})$; 2. $(\bar{M}, 0)$; 3. $(\bar{M} - M_{K2}^c, M_{K2}^c)$.

$$g_{1,2}(\boldsymbol{\theta}, r, \boldsymbol{X}) = (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G} - \boldsymbol{B}_2, \tag{1.15}$$

$$g_{2,2}(\boldsymbol{\theta}, r, \boldsymbol{X}) = (r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left[\bar{\hat{\boldsymbol{\Omega}}}_{y2} - \boldsymbol{\Omega}_y + \frac{\lambda_K \hat{\mathcal{C}}_{Ky2}^2}{\lambda_2 \hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{11} - \frac{\lambda_K \hat{\mathcal{C}}_{Ky2}}{\lambda_2 \hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{12}(\boldsymbol{\Omega}_{p2} + \beta^2 \boldsymbol{M}_{K2}\boldsymbol{\Omega}_e)\right] - \boldsymbol{C}_2, \tag{1.16}$$

$$\begin{aligned} g_{3,2}(\boldsymbol{\theta}, r, \boldsymbol{X}) = {} & -\rho r (r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left[\bar{\hat{\boldsymbol{\Omega}}}_{y2} - \boldsymbol{\Omega}_y + \frac{\lambda_K \hat{\mathcal{C}}_{Ky2}^2}{\lambda_2 \hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{11} - \frac{\lambda_K \hat{\mathcal{C}}_{Ky2}}{\lambda_2 \hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{12}(\boldsymbol{\Omega}_{p2} + \beta^2 \boldsymbol{M}_{K2}\boldsymbol{\Omega}_e)\right] \\ & \left[\left(\lambda_I \beta_I^2 \boldsymbol{M}_{I2} + \lambda_K \beta^2 \left(\boldsymbol{E}_{11} - \frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{12}\right)\boldsymbol{M}_{K2}\right)\boldsymbol{\Omega}_e + \lambda_1 \boldsymbol{\Omega}_\epsilon\right]^{-1}(r\boldsymbol{I} - \boldsymbol{G})^{-1} - \boldsymbol{D}_2 \end{aligned} \tag{1.17}$$

$$\begin{aligned} g_{4,2}^{(\mathrm{I})}(\boldsymbol{\theta}, r, \boldsymbol{X}) = {} & \mathcal{C}_y - \left(\frac{r - g_1}{\rho r}\right)^2 \lambda_K \beta^2 z M_{K2}^{(\mathrm{II})} \Omega_e^{(\mathrm{II})} \left[(\lambda_I \beta_I^2 M_{I2}^{(\mathrm{I})} + \lambda_K \beta^2 M_{K2}^{(\mathrm{I})})\Omega_e^{(\mathrm{I})} + \lambda_1 \Omega_\epsilon^{(\mathrm{I})}\right]\Omega_x^{(\mathrm{I})} \\ & - z\Bigg\{\Omega_y^{(\mathrm{II})} + \beta^2 M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \left[\frac{r - g_1}{\rho r}\lambda_K \beta^2 z M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}\right]^2 \Omega_x^{(\mathrm{I})} \\ & + \left[\frac{r - g_2}{\rho r}\left[\lambda_I \beta_I^2 M_{I2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \lambda_1 \Omega_\epsilon^{(\mathrm{II})}\right]\right]^2 \Omega_x^{(\mathrm{II})}\Bigg\}. \end{aligned} \tag{1.18}$$

1.3.3 Estimated parameters

As reported in Table 1.3, the case of China is characterized by a very small proportion of foreign investors ($\lambda_K = 0.03$, with a proportion of 6%), a large proportion of domestic individual investors ($\lambda_J = 0.35$, with a proportion of 66%), a large disparity between domestic and foreign data processing technology ($\frac{\beta}{\beta_I} = 1.39$), and high precision of soft information ($\Omega_\epsilon^{(\mathrm{I})} = 3.34, \Omega_\epsilon^{(\mathrm{I})} = 4.71$).

ρ	r	g_1	g_2	β	$\Omega_\epsilon^{(\mathrm{I})}$	$\Omega_\epsilon^{(\mathrm{II})}$	$\Omega_x^{(\mathrm{I})}$	$\Omega_x^{(\mathrm{II})}$	λ_I	λ_J	λ_K
2.07	1.042	0.75	0.82	1.39	3.34	4.71	2.35	1.97	0.15	0.35	0.03

Table 1.3: Parameters calibrated for the partial-opening model

With these calibrated parameters, foreign institutional investors would allocate their data processing capacity in both type of shares, and the signal-to-noise ratio for foreign investable shares decreases. Domestic institutional investors are also induced to learn slightly less about

those shares, and thus the average posterior precision of dividend innovation decreases.

	Closed			Partial Opening			Full Opening		
	Model	Data	Difference	Model	Data	Difference	Model	Data	Difference
Type I	5.823	6.328	0.505	3.732	4.478	0.746	8.334	7.293	-1.041**
	(1.883)	(2.919)	(0.493)	(1.477)	(2.263)	(0.898)	(3.037)	(3.283)	(0.448)
Type II	2.457	2.134	-0.323	3.272	2.981	-0.291	4.790	4.478	-0.312*
	(0.542)	(0.896)	(0.475)	(1.235)	(0.933)	(0.407)	(2.817)	(2.349)	(0.213)

Table 1.4: Price informativeness generated by models and data

Both the model-derived and the data-calculated indicator for price informativeness of type-I asset first decreases and then increases. For closed and partial-opening cases, model-derived price informativeness is not significantly different from the data. In the full-opening case, they still show the same direction. Since China has not yet implemented full opening, I used data from 2019 to approximate.

1.4 Full Opening of the Market

Instead of a partial market opening, economies theoretically could also directly enter a full market opening. The baseline full-opening model in this section on the one hand is useful in its own regard, and on the other hand provides a benchmark and long-term vision for partially-open markets.

In a full-opening economy, all securities are available for international trade. In addition, λ_H foreign individual investors also enter the domestic market. Foreign institutional investors might allocate data processing to both types of risky assets. As a result, information will be less concentrated on a specific type of risky asset. Entry of foreign individual investors also lower the information density.

1.4.1 Equilibrium characterization

I use a ‡ to denote variables in this equilibrium. The optimal portfolio choices have the same form as in the partial opening model. The optimal data choice problem of domestic institutional investors does not change, but that of foreign institutional investors is no longer trivial since they are also allowed to invest in type-I risky asset. Now they are allocating $\bar{M}$ data processing

capacity to both type of risky assets as the domestic institutional investors do. The form of price coefficients slightly changes since average level of demand changes (see coefficients A.38, A.39, A.40 and A.41) with the type of investors allowed in each type of risky asset.

The signal-to-noise ratio and the average posterior precision of dividend innovation are $\zeta_2^{\ddagger(\mathrm{I})} = -\frac{r-g}{\rho r}[(\lambda_I M_I^{\ddagger(\mathrm{I})}\beta_I^2+\lambda_K M_K^{\ddagger(\mathrm{I})}\beta^2)\Omega_e^{(\mathrm{I})}+\lambda_1\Omega_\epsilon^{(\mathrm{I})}]$ and $\bar{\hat{\Omega}}_{y2}^{\ddagger(\ddagger)} = \Omega_y^{(\mathrm{I})}+\zeta_2^{\ddagger 2(\mathrm{I})}\Omega_x^{(\mathrm{I})}+\frac{\lambda_I M_I^{\ddagger(\mathrm{I})}\beta_I^2+\lambda_K M_K^{\ddagger(\mathrm{I})}\beta^2}{\lambda_3}\Omega_e^{(\mathrm{I})}+\frac{\lambda_1}{\lambda_3}\Omega_\epsilon^{(\mathrm{I})}$, respectively. Intuitively, both the signal-to-noise ratio and the average posterior precision of dividend innovation are more balanced and less concentrated in the full market opening equilibrium than in the partial market opening equilibrium.

1.4.2 Impacts of full market opening

Full market opening influences the signal-to-noise ratios in a slightly different way. What foreign institutional investors have learned about type-II risky asset, if any, is aggregated by the market. In light of that domestic institutional inventors may change their data acquisition accordingly. The results for type-I and II risky assets are summarized in Propositions 9 and 10.

PROPOSITION 9. *For type-I risky asset, the signal-to-noise ratio*

1) decreases if domestic institutional investors learn less about it;

2) increases if domestic institutional investors learn more about it and the number of foreign institutional investors $\lambda_K < \frac{(M_{I2}^{\ddagger(\mathrm{I})}-M_{I2}^{(\mathrm{I})})\beta_I^2}{(\bar{M}-M_{K2}^{(\mathrm{I})\ddagger})\beta^2}\lambda_I$.

PROPOSITION 10. *For type-II risky asset,*

1) the signal-to-noise ratio increases if domestic institutional investors learn more about it.

2) A necessary condition for the signal-to-noise to decrease is that domestic institutional investors learn less about it, or $\lambda_K < \frac{M_{K2}^{(\mathrm{II})\ddagger}\beta^2}{\bar{M}\beta_I^2}\lambda_I$.

Without the distortion created by partial market opening, the average posterior precision of dividend innovation is aligned with the signal-to-noise ratio.

PROPOSITION 11. *The average posterior precision of dividend innovation changes in the same direction as the signal-to-noise ratio for both types of risky assets.*

In recent years, as China gradually widens and deepens its capital market liberalization, some researchers have found new evidence that stock prices in China have become as informative about future profits as they are in the US (Carpenter et al., 2021). Figure 1.9 shows how signal-to-noise ratio changes (full market opening vs. closed market) reacts to the proportion of foreign institutional investors of China Stock Connect Programs. In this scenario, entry of the foreign institutional investors always increase the signal-to-noise ratio, compatible with the above empirical findings. As more and more foreign institutional investors enter the markets, the change increases at an increasing speed since more hard information is aggregated. However, as more and more foreign investors enter, the increased hard information is balanced out by the lost of soft information.

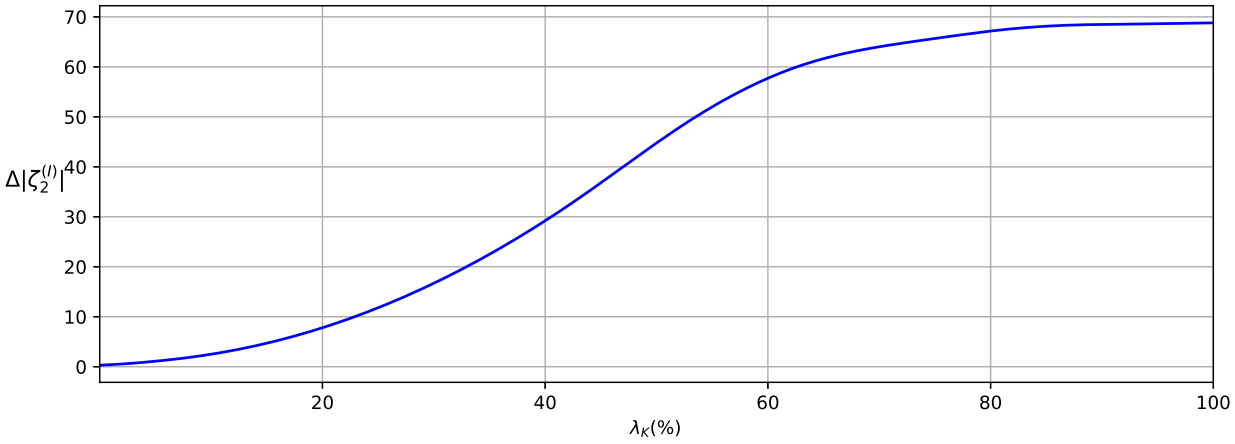


Figure 1.9. Reactions of signal-to-noise ratios to the size of foreign institutional investors

This figure plots how signal-to-noise ratio changes in the full market opening case compared with the closed-market case reacts to the proportion of foreign institutional investors of China Stock Connect Programs. Entry of the foreign institutional investors increase the signal-to-noise ratio. As more and more foreign institutional investors enter the markets, the change increases at an increasing speed since more hard information is aggregated. However, as more and more foreign investors enter, the increased hard information is balanced out by the lost of soft information.

1.5 Technological advancement

Data are not equal to information; they need to be converted into information via certain data processing technology. Investors are endowed with heterogeneous imperfect data processing technologies. Processing data into information might incur information loss or even induce

errors. Investors can launch technology innovation[5], which would incur an immediate fixed cost but would ensure a correct data processing process if successful. Or they can hire data analysts who have superior data processing technology with a certain probability. This resembles Dvořák (2005), domestic clients of global brokerages have higher profits than foreign clients of global brokerages, suggesting that the combination of local information and global expertise leads to higher profits.

Suppose that data processing technology can be improved upfront with a cost $\frac{1}{2}a\beta_{i2}^2$. In this case, domestic institutional investor i would be both choosing the data processing technology β_{i2} and allocating data points between risky assets. The objective of her data choice problem becomes $\ln\left|\hat{\Omega}_{iy2}\right| - \rho r a\beta_{i2}^2$. The F.O.C. with respect to the data choice $M_{i2}^{(\mathrm{I})}$ remains the same. The F.O.C. with respect to β_i gives $\rho r a = \frac{\bar{M}\Omega_e^{(\mathrm{I})}}{\hat{\Omega}_{iy2}^{(\mathrm{I})}}\frac{\hat{\Omega}_{iy2}^{(\mathrm{I})}\hat{\Omega}_{iy2}^{(\mathrm{II})}}{\left|\hat{\mathbf{\Omega}}_{iy2}\right|}$. The technology and data choices are summarized in the following proposition.

Proposition 12. *Suppose that data processing technology can be improved upfront with a cost $\frac{1}{2}a\beta_{i2}^2$. Then domestic investor i choose a technology level*

$$\beta_{i2}^* = \begin{cases} \beta_I, & \textit{if } \beta_{i2}^c \leq \beta_I \\ \beta, & \textit{if } \beta_{i2}^c \geq \beta \\ \beta_{i2}^c, & \textit{otherwise} \end{cases},$$

where

$$\beta_{i2}^c = \sqrt{\frac{1}{\bar{M}\Omega_e^{(\mathrm{I})}}[\hat{\Omega}_{iy2}^{(\mathrm{I})} - (\Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})})] + \frac{1}{\bar{M}\Omega_e^{(\mathrm{II})}}[\frac{\Omega_e^{\mathrm{II}}}{\Omega_e^{(\mathrm{I})}}\hat{\Omega}_{iy2}^{(\mathrm{I})} - (\Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})})]}.$$

Then M_{i2}^* is obtained by $M_{i2}^* = \frac{1}{\beta_{i2}^{*2}\Omega_e^{(\mathrm{I})}}[\hat{\Omega}_{iy2}^{(\mathrm{I})} - (\Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})})]$.

An investor's decision for technology innovation depends on how much data is available on the market and what her data processing capacity is. The effect of data amount on investors' choice of technological innovation is ambiguous. On the one hand, the more data is available

[5]Here the definition of innovation is similar to but slightly different from that in the learning-by-doing literature. For example, Jovanovic and MacDonald (1994) defined innovation as "putting what has been learned to work commercially".

for the investors, the more valuable technological innovation is if successful. On the other hand, when the data processing capacity is sufficiently high, investors are not incentivised to innovate data processing technology due to the immediate fixed cost. Intuitively, the farther an investor's data processing technology is to the perfect data processing technology, the more incentivised she is to innovate.

It does not distinguish technological deepening and technological opening, as proposed by Jovanovic and Nyarko (1996)[6]. For example, institutions developing a new and better text mining algorithm is technological opening, professional traders receiving investment trainings is technological deepening.

1.6 Conclusion

When foreign institutional investors enter the domestic market, they bring in better data processing technology for hard information but stand with the lack of soft information. This mismatch creates information inefficiency, and the two opposite forces also reconcile the mixed empirical evidence on whether domestic investors have advantage over foreign traders in the domestic stock markets when the domestic capital markets start to liberalize.

With partial market opening, foreign institutional investors allocate their data processing capacity to internationally untradable shares if the information spillover rate is larger than 1, resulting in information inefficiency on the market. The hard information obtained from processing data for internationally untradable shares is not aggregated into their equilibrium price but partially aggregated into the equilibrium price of foreign investable shares given that the two types of shares are correlated. Thus a distortion of price informativeness is created. This distortion is further strengthened when hard information of eligible companies comes with higher precision.

In the case of China, foreign institutional investors would allocate their data processing capac-

[6]Jovanovic and Nyarko (1996) proposed the ideas of technological deepening and technological opening. Technological deepening is learning more about a given grade of a technology. The more the agent uses a technology, the better he learns its parameters, and the more productive he gets. And technological opening is learning about a new technology. Once the productivity gains on a given technology are exhausted, further growth can occur only by switching to a better technology. An agent may be so skilled at a technology that he will refuse to switch to a better, but unfamiliar one and in the long run be overtaken by an agent who initially is less skilled.

ity in both type of shares, and the signal-to-noise ratio for foreign investable shares decreases. Domestic institutional investors are also induced to learn slightly less about those shares, and thus the average posterior precision of dividend innovation decreases. In the full opening counterfactual, the signal-to-noise raios for both type of shares increase.

This chapter abstracts from the managerial side. It could be interesting to zoom in on the firm side in further research, and see how firm investment, agent costs, voluntary disclosure change in capital market opening. Ruan et al. (2021) found that after the implementation of the Shanghai-Hong Kong Stock Connect program, the quality of internationally tradable companies' public information disclosure significantly improved. In the framework of this model, firm controls data release and decides the precision of data. In each period, there is a certain amount from the current dividend that can be diverted by the Manager. Manager would incur a personal cost for diverting. The diversion cannot be too large, otherwise Manager's value diversion would be detected and punished. On the one hand, managerial data choice directly affects the posterior standard deviation of current dividend innovation. On the other hand, the entering of foreign investors could also lower the safe bar. The choice of data precision balances the capital agency costs reduction and the increase of the information publication costs (Holland, 1998), given that the mandatory disclosure requirement is satisfied.

Chapter 2: Reading Tea Leaves: Benefits and Costs of Learning about Future Demand Shocks

2.1 Introduction

Noisy rational expectations equilibrium (NREE) models have been an established framework for information choice problems in financial markets. These models usually introduce exogenous noise trades, alternatively termed as shocks to the demand/supply of the risky assets, to prevent their prices from fully revealing private information observed by investors (Grossman and Stiglitz, 1980; Hellwig, 1980; Verrecchia, 1982). These shocks are orthogonal to dividend innovations and do not affect the fundamental value, or liquidation value, of the risky assets. It is now well-understood that noise trades can be micro-founded by hedging needs (Liu and Wang, 2016; Medrano and Vives, 2004), liquidity needs (Han et al., 2016), or sentiment (Shleifer and Summers, 1990). In this chapter, I abstract from micro-foundations of noise trades and hereinafter refer to them as demand shocks.

Traditionally, learning about future demand shocks affects the financial markets via the front-running channel, that is if an investor can learn about future shocks, then she has motives to buy (sell) shares now and sell (buy) them tomorrow if tomorrow's demand is predicted to be high (low) tomorrow. In this chapter, I propose a *future uncertainty channel*, through which future demand shocks affect the market regardless of whether the front-running motive is in place or not. If we shut down the front-running channel, then future demand shocks, if investors cannot learn about them today, *should* be irrelevant for investor behavior and the current asset price. The future uncertainty channel, in fact, makes the future relevant to the present when there is information acquisition on future demand shocks. The future uncertainty channel works through two forces to increase the investor's *ex ant*e utility: on the aggregate level, market-wise dividend

information acquisition decreases, which makes dividends riskier with higher returns; on the individual level, information acquisition on future demand shocks makes the conditional variance of future excess payoffs lower. Higher expected returns and lower conditional variance of future excess payoffs both increase investors' *ex ante* utility.

In the literature, investors are usually worse off if every one holds more private information. This is a manifestation of *fallacy of composition*, i.e. the error of assuming what is true of a member of a group is true for the group as a whole. In the context of financial markets, more private information is better for an individual investor. On aggregate, however, when all investors acquire more private information, assets become less risky with lower returns, as discussed by Kurlat and Veldkamp (2015). Learning about current demand shocks is not sufficient to break the fallacy; but learning about *future* demand shocks does. It creates heterogeneity between generations; it introduces risk-sharing across time and information spillover between generations; and it reduces market-wise dividend information acquisition and thus makes the asset riskier with a higher risk premium.

Acquiring a signal about future demand shocks also echoes the general idea of news-driven business cycles in the macroeconomic literature. The news shocks are mainly shocks on future total factor productivity, ignoring an important aspect of price formation: demand shocks can drive the prices of goods away from their fundamental values. Learning about future demand shocks is even more important when we turn our attention to the financial market, which is the focus of this chapter.

This chapter studies how learning about future demand shocks affects investor behavior, including choices of information and investment, and market outcomes, including price informativeness and volatility. Specifically, I propose a model of a competitive CARA-normal economy with overlapping-generation (OLG) investors who live for two periods. In each period, investors first decide, subject to a linear information capacity constraint, how to allocate precision to signals on current dividend innovations, current demand shocks and future demand shocks. Then they make subsequent investment decisions after observing signal realizations and the equilib-

rium asset price. To disentangle the front-running channel and the future uncertainty channel, I study two versions of the model, one with a long-lived asset ("dynamic model") and one with dividend strips ("static model").

The model features risk sharing across time and information spillover between generations. On the one hand, with signals of future demand shocks, future risks are incorporated into today's price. As a result, tomorrow's market risks are partially shared by today's investors. On the other hand, only the young generation optimally choose to learn about future demand shocks. They will be able to extract *more* dividend innovation from the equilibrium price in their second period, but *less* in the current period. The old generation know more about the current demand shocks from the previous period, and their private information is aggregated by and then spill over to the young generation through the equilibrium price.

This model also features two margins of information choices. On the intensive margin, learning about tomorrow's demand shocks today and tomorrow are imperfect substitutes, with the former also affecting the marginal value of learning tomorrow. Assuming learning about future demand shocks is equally costly today and tomorrow, learning about tomorrow's demand shock today can fully substitute for learning about it tomorrow and offer an extra benefit from a higher marginal value of learning. On the extensive margin, we are weighing the cost-adjusted marginal benefits of one unit of precision allocated to signals about dividend innovations and demand shocks. The extra benefit brought in by learning about future demand shocks makes learning about them more beneficial than learning about current dividend innovations. Therefore, with information acquisition equally costly (both across signals and for each unit of precision of the same signal), investors specialize in information acquisition and optimally choose to solely acquire signals about future demand shocks in the first period and signals about dividend innovations in the second period. On the aggregate level, less dividend information is learned and thus price informativeness decreases, while price volatility increases due to more heterogeneous private information.

To put the question into empirical contexts, I discuss a few examples where learning about

future demand shocks plays an important role. One notable example is index recomposition. As Han et al. (2016) points out, institutional traders, e.g. index funds or ETFs, need to rebalance portfolios for non-informational reasons around index recomposition. The existence of index funds and the requirement that they replicate indices create demand shocks that are non-fundamental to the underlying securities. On the announcement date of index recomposition, investors receive a signal of effective-date demand shocks. Sophisticated investors can even adjust portfolios before the index recomposition announcement by making informed predictions about how index funds will have to rebalance. Another set of examples are in a similar vein as the news shock, except that here the news that investors receive is not about future total factor productivity, but about future demand shocks. One case with far-reaching impact is the BlackRock-BGI acquisition. BlackRock, Inc. announced on June 11, 2009 that it would acquire Barclays Global Investors (BGI), including iShares, from Barclays Bank PLC. The deal closed on December 1, 2009. Consequently, BlackRock became the world's largest investment management company. Observing the news of the BGI transaction on June 11, sophisticated investors would expect the iShares asset under management (AUM) to expand in the following years. For the component stocks of iShares ETFs, these are future demand shocks that are not related to their firm fundamentals. Some other examples may include the future liquidity crisis, sentiment, or behavioral bias. In the financial markets, it is usually the high-frequency traders (HFTs) who are learning about future demand shocks.

Contribution to the literature This chapter contributes to the literature on rational expectations models with asymmetric information. Our understanding of endogenous information acquisition started from static descriptions about financial markets where signals are only on current fundamentals (Grossman and Stiglitz, 1980; Hellwig, 1980; Verrecchia, 1982): there is strategic substitution in information acquisition of informed investors. On the one hand, they can gather private information about the liquidation value or dividend to make better portfolio choices, collectively making price more informative. On the other hand, when facing a more

informative price, investors would rely less on private signals and more on price, which reduces their incentive to acquire private information. There are more mechanisms at force when we consider a more complicated information structure. For example, when different investors are informed of different fundamentals of the asset payoff, trading aggressively on information about one fundamental reduces uncertainty of other fundamentals other investors face, encouraging those investors to learn more about those fundamentals (Goldstein and Yang, 2015). Goldstein and Yang (2019) further specify dimensions of fundamental uncertainty to be one that is related to the aggregate economy and one that is firm-specific. Disclosing firm-specific information that the real decision maker already knows improves price informativeness and real efficiency. Learning about current demand shocks plays a slightly different role. In Medrano and Vives (2004), Ganguli and Yang (2009) and Manzano and Vives (2011), an investor receives an endowment shock, which can be viewed as her private information about the aggregate endowment shock and helps her make more precise inference about fundamentals. Marmora and Rytchkov (2018) show that learning about noise increases price informativeness, and the price can be the most informative when the majority of investors acquire non-fundamental information.

In dynamic settings, interpretations about fundamental and non-fundamental information can be different, and it becomes meaningful to distinguish current and future shocks. However, learning is still limited to future dividend innovations, e.g. Cai (2017) and Avdis (2016). In Cai (2017), learning about the future is achieved through learning about the persistent part of current and past dividends, which can then be thought of as a signal for future dividends. And in Avdis (2016), the equilibrium price contains both dividend and discount rates information, where discount rates shocks are modeled as demand shocks. Thus dividend information acquisition offers an extra benefit of revealing non-dividend information about conditional expected returns. This chapter approaches dynamic information acquisition in a different way where signals of i.i.d. future demand shocks are acquired today. I emphasize the future uncertainty channel where the asset becomes riskier because less dividend information is learned on the aggregate level and an atomic investor is more certain about asset payoff in the second period thanks to the signal about

second-period demand shocks they acquire in the first period.

This chapter also connects with the macroeconomic literature on news-driven business cycles. As discussed by Beaudry and Portier (2014), "if an agent can properly anticipate a future need, he can gain by trying to preempt the market and invest early as to make good readily available when the predicted needs eventually appear". The front-running channel in this chapter echoes with this general idea. However, the macroeconomic literature usually treats news shock as signals about future fundamentals. For example, Jaimovich and Rebelo (2009) shows how news about future total factor productivity affects current output, consumption and investment. This chapter complements this literature by bringing to the spotlight the importance of learning about future demand shocks.

This chapter also contributes to the literature on HFT[1]. In reality, HFTs actively manage their limit orders and reveal information through quote updates (Hasbrouck and Saar, 2013). Empirical evidence also shows that HFTs predict order flows (Hirschey, 2018) or price changes (Harris and Saad, 2014), which can be mapped into acquiring a signal for future demand shocks in this chapter. This chapter also finds that with access to information about future demand shocks, investors are better off, which is consistent with Jovanovic and Menkveld (2017) who also find a moderately positive welfare effect.

The rest of this chapter is organized as follows. Section 2.2 sets up the model and present model solutions, Section 2.3 presents the analytical and numerical results from the static model, Section 2.4 discusses the empirical applications of the dynamic model, and Section 2.5 concludes.

2.2 The Model

2.2.1 Setup

Assets There is an infinite stream of dividend payments $\{d_t\}_0^\infty$, $\tilde{d}_t = \mu + g d_{t-1} + \tilde{y}_t$, where μ and g are known parameters, and the dividend shock $\tilde{y}_t \sim i.i.d.\ N(0, \tau_y^{-1})$. The dividend $\tilde{d}_t$ is paid out at the end of each period t.

[1] I refer to Menkveld (2016) for a more thorough literature review on HFT.

There is a single tradable, risky asset, with supply $\bar{x}_t$. To disentangle the future uncertainty and the front-running channels, I introduce a parameter $\alpha \in \{0, 1\}$ as in Farboodi and Veldkamp (2020). When $\alpha = 1$, a time-t asset pays $p_{t+1} + \tilde{d}_t$, the future price of the long-lived asset plus its dividend. In this case, future demand shocks will affect future equilibrium price, and thus affect reselling profits and current choices. I call this the "dynamic" model. When $\alpha = 0$, the asset is not long-lived, which can be reinterpreted it as a series of dividend strips. That is, in each period, there is a *new* one-period risky asset that can be traded. Signals about future demand shocks can still transmit via the future uncertainty channel, while the front-running channel is shut down. Using dividend strips allows me to isolate the new future uncertainty channel. I call this the "static" model. If demand shocks can be insured or hedged, then the static model and the dynamic model are equivalent, only differ in some constant and multiplier.[2] Thus the static model can be interpreted as equivalent to a model with a long-lived risky asset where demand shocks can be fully-insured or fully-hedged. Studying dividend strips per se is also important since they provide us information about how stock prices are formed[3].

Signals Each investor receives an unbiased signal $\tilde{\eta}^y_{it} = \tilde{y}_t + \tilde{\epsilon}^y_{it}$ about current dividend innovation $\tilde{y}_t$, where $\tilde{\epsilon}^y_{it} \sim N(0, \tau^{-1}_{\epsilon yt})$, and an unbiased signal $\tilde{\eta}^x_{it} = \tilde{x}_t + \tilde{\epsilon}^x_{it}$ about current demand shocks $\tilde{x}_t$, where $\tilde{x}_t \sim iid.\ N(0, \tau_x^{-1})$ and $\tilde{\epsilon}^x_{it} \sim N(0, \tau^{-1}_{\epsilon xt})$. Investors can also predict demand shocks in the next period $\tilde{x}_{t+1}$ by an unbiased signal $\tilde{\xi}^x_{it} = \tilde{x}_{t+1} + \tilde{\nu}^x_{it}$, where $\tilde{\nu}^x_{it} \sim i.i.d.\ N(0, \tau^{-1}_{\nu xt})$. All shocks are orthogonal. Signal precision[4] $\tau_{\epsilon xt}, \tau_{\epsilon yt}$, and $\tau_{\nu xt}$ are choice variables of investor i; she optimally allocate signal precision with a limited information capacity constraint K.

[2]Consider an insurance company. $\mathcal{G}_t$-investor i pays an insurance fee at the beginning of their life, and receives $f_t \mid \mathcal{I}^+_t$ (measurable at the end of period t) at the end of period t and $f_{t+1} \mid \mathcal{I}^+_{t+1}$ at the end of period $t+1$, where $\tilde{f}_t$ is the fundamental value of this long-living asset, and $\tilde{f}_t \mid \mathcal{I}^+_t = \frac{r^2\mu}{(r-1)(r-g)} + \frac{rg}{r-g} d_{t-1} + \frac{r}{r-g} y_t$. Thus the budget constraint is $\tilde{W}_{i,t+1^+} = r[rW_{it-1^+} + q'_{i1}(\tilde{f}_t - p_t)] + q'_{it+1}(\tilde{f}_{t+1} - p_{t+1}r)$. For investor i, current and future demand shocks are insured. The aforementioned model is different to the static model only by some constants and multipliers.

[3]For example, van Binsbergen et al. (2012) empirically measure dividend strip prices to study the term structure of the equity premium.

[4]The subscript i is dropped because investors are *ex ante* identical.

Investors At the start of each period t, a continuum of measure-$\frac{1}{2}$ investors, denoted by $\mathcal{G}_t$, with initial investable wealth $W_{it-1^+} = W_0$ are born. They live for two periods, making both information and portfolio choices in each period.[5] Investor i only consumes at the end of her life, and her utility function over her terminal wealth is $-e^{-\rho W_{it+1^+}}$, where ρ is the risk aversion coefficient.

At the portfolio choice stage of period $t+1$, after receiving private signals and observing the equilibrium price, $\mathcal{G}_t$-investor i maximizes her expected utility conditional on information they have by choosing how much to invest in the risky asset, that is she chooses the optimal portfolio $\hat{q}_{i2}$ to maximize her interim expected utility $\mathbb{E}[-e^{-\rho \tilde{W}_{it+1^+}} \mid \hat{\mathcal{I}}_{it+1}]$ subject to the budget constraint[6] :

$$\tilde{W}_{it+1^+} = rW_{i,t^+} + \hat{q}_{i2}(\tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1}r). \tag{2.1}$$

At the beginning of period $t+1$, she chooses signal precision $\hat{\tau}_{\epsilon yit+1}$, $\hat{\tau}_{\epsilon xit+1}$, and $\hat{\tau}_{\nu xit+1}$ to maximize expected utility $\mathbb{E}[-e^{-\rho \tilde{W}_{it+1^+}} \mid \hat{\mathcal{I}}_{it}^{+}]$ subject to the information capacity constraint

$$\gamma_{\epsilon y}\hat{\tau}_{\epsilon yit+1} + \gamma_{\epsilon x}\hat{\tau}_{\epsilon xit+1} + \gamma_{\nu x}\hat{\tau}_{\nu xit+1} \leq K. \tag{2.2}$$

The information constraint is assumed to be linear in signal precision, that is the marginal cost of an extra unit of signal precision is the same at all signal precision levels.[7] This assumption helps to isolate the benefits of learning about future demand shocks, i.e. the future uncertainty channel, from a cost reduction channel. Another popular choice of information constraint in the literature is a convex one, so that the marginal cost of learning within one signal is increasing as investors allocate more precision to that signal. In that case, the possibility to learn about future demand shocks today and tomorrow makes learning about demand shocks effectively cheaper, because investors can smooth learning costs. This cost reduction channel works in the same direction

[5] Several papers also model two trading periods, e.g. Holden and Subrahmanyam (2002), Chordia and Subrahmanyam (2004) and Avdis (2016), among others.

[6] Throughout this chapter, the notation $\hat{}$ is used on variables that are related to the old generation.

[7] The linear information capacity constraint is common in the literature; see Kacperczyk et al. (2016) among others.

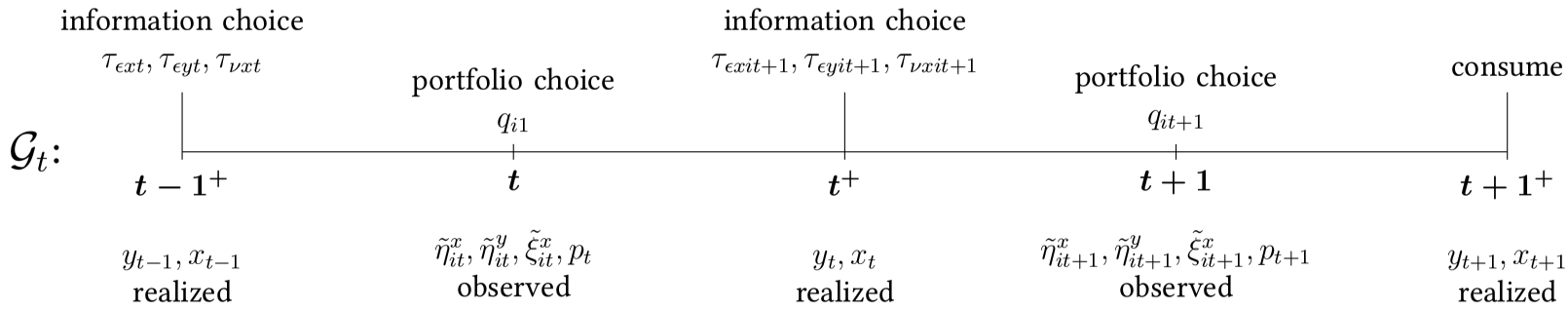


Figure 2.1. Life cycle of investors who are born at the beginning of period t.

Generation-t investors are born at the beginning of period t (denoted by $t-1^+$). They live for two periods: period t and period $t+1$. In period t, they first observe realizations of past dividend innovations and demand shocks and make information choices $\tau_{\epsilon xt}, \tau_{\epsilon yt}, \tau_{\nu xt}$. After receiving signals $\tilde{\eta}^x_{it}$, $\tilde{\eta}^y_{it}$, and observing the equilibrium asset price p_t, they make a portfolio choice q_{i1}. Then at the end of period t (denoted by t^+), period-t dividend d_t is realized and delivered. Then they repeat the two choices in period $t+1$. At the end of period $t+1$, investors consume all end-of-life wealth.

of the benefit increase through the future uncertainty channel, and will make the results even stronger. For the purpose of demonstrating the future uncertainty channel clearly, here I use a linear constraint.

$\mathcal{G}_t$-investor i is rational, which means that when she makes portfolio choices in period t, she also has to take into consideration her future choices in period $t+1$. Therefore, at the portfolio choice stage of period t, she chooses q_{i1} and q_{it+1} [8] to maximize the expected utility $\mathbb{E}[-e^{-\rho\tilde{W}_{it+1^+}} \mid \mathcal{I}_{it}]$, subject to the budget constraint

$$\tilde{W}_{it+1^+} = r^2 W_0 + r q_{i1}(\tilde{d}_t + \alpha p_{t+1} - p_t r) + \hat{q}_{i2}(\tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1} r). \tag{2.3}$$

Notice the difference between $\hat{q}_{i2}$ and q_{it+1}. While $\hat{q}_{i2}$ is the actual investment that $\mathcal{G}_t$-investor i makes at the portfolio choice stage of period $t+1$, based all previous dividend and signal realizations, $\hat{q}_{i2}$ is what $\mathcal{G}_t$-investor i predict herself to hold in period t based on only the information that is available to her at the portfolio choice stage in period t.

Finally, in the beginning of period t, $\mathcal{G}_t$-investor i makes period-t information choices $\tau_{\epsilon yt}$, $\tau_{\epsilon xt}$, and $\tau_{\nu xt}$ to maximize her *ex ante* utility $\mathbb{E}[-e^{-\rho\tilde{W}_{it+1^+}} \mid \mathcal{I}^+_{t-1}]$ subject to the information capacity constraint

$$\gamma_{\epsilon y}\tau_{\epsilon yt} + \gamma_{\epsilon x}\tau_{\epsilon xt} + \gamma_{\nu x}\tau_{\nu xt} \leq K. \tag{2.4}$$

Information structure The equilibrium price is a linear function of the signals, and since signals are unbiased, we can express the price as a linear function of shocks:

$$p_t = A_t + B_t d_{t-1} + C_t \tilde{y}_t + D_t \tilde{x}_t + E_t \tilde{x}_{t+1}. \tag{2.5}$$

[8] Technically, $\mathcal{G}_t$-investor i also needs to include three other choice variables, namely $\tau_{\epsilon xit+1}$, $\tau_{\epsilon yit+1}$ and $\tau_{\nu xit+1}$. They can be omitted, however, because of two reasons. Firstly, we only solve portfolio choices in a functional form of conditional expectations and variances of future payoffs. Although those moments depend on the optimal information choices in period $t+1$, that functional form does not. Secondly, solving for the optimal information choices in period $t+1$ is also useless since the investor will make better information choices after d_t is realized at the beginning of period $t+1$. The same reasoning goes for the information choice problem at the beginning of period t.

This is different from existing literature in that there is an extra term of the future demand shock, $\tilde{x}_{t+1}$, in the equilibrium price. For future demand shocks to enter the equilibrium price relation, we need two conditions. Firstly, there must be a signal about future demand shocks. Otherwise, as in Farboodi and Veldkamp (2020) which modeled a long-lived asset and signals about current dividend and demand shocks, but still the future demand shock $\tilde{x}_{t+1}$ does not enter the equilibrium price relation. Secondly, the asset has to be long-lived ($\alpha = 1$). Otherwise when $\alpha = 0$, $\tilde{d}_t$ and $\tilde{d}_{t+1}$ are different assets, and the demand shocks about $\tilde{d}_{t+1}$ does not affect the equilibrium price of $\tilde{d}_t$ directly, i.e. $\tilde{x}_{t+1}$ does not appear in the equilibrium price, but only indirectly though price coefficients due to information acquisition.

Define the price-to-current-noise ratio to be $\zeta_{1t} \equiv \frac{C_t}{D_t}$, and the price-to-future noise ratio to be $\zeta_{2t} \equiv \frac{C_t}{E_t}$, where C_t is the price sensitivity to dividend innovation, and D_t and E_t are the price sensitivity to current and future demand shocks, respectively, in the equilibrium price (2.5). A transformation of the equilibrium price p_t can generate an unbiased signal about current dividend shock. The two generations observe the same price but extract out of it different information about $\tilde{y}_t$. In period t, $\mathcal{G}_t$-investor i extracts a signal $\tilde{\eta}^p_{it} = \tilde{y}_t + \tilde{\nu}^p_{it}$, where $\tilde{\nu}^p_{it} = \frac{1}{\zeta_{1t}}\left(\tilde{x}_t - \mathbb{E}[\tilde{x}_t \mid \tilde{\eta}^x_{it}]\right) + \frac{1}{\zeta_{2t}}\left(\tilde{x}_{t+1} - \mathbb{E}[\tilde{x}_{t+1} \mid \tilde{\xi}^x_{it}]\right) \sim N\left(0, \tau_{pt}^{-1}\right)$, with the variance of the price signal noise being $\tau_{pt}^{-1} = \zeta_{1t}^{-2} V_{xt} + \zeta_{2t}^{-2} V^{\xi}_{xt}$.[9] And $\mathcal{G}_{t-1}$-investor i extracts a signal $\tilde{\hat{\eta}}^p_{it} = \tilde{y}_t + \hat{\tilde{\nu}}^p_{it}$, where $\hat{\tilde{\nu}}^p_{it} \equiv \frac{1}{\zeta_{1t}}\left(\tilde{x}_t - \mathbb{E}[\tilde{x}_t \mid \tilde{\eta}^x_{it}, \tilde{\xi}^x_{it-1}]\right) + \frac{1}{\zeta_{2t}}\left(\tilde{x}_{t+1} - \mathbb{E}[\tilde{x}_{t+1} \mid \tilde{\hat{\xi}}^x_{it}]\right) \sim N\left(0, \hat{\tau}_{pit}^{-1}\right)$, with the variance of the price signal noise being $\hat{\tau}_{pit}^{-1} = \zeta_{1t}^{-2} \hat{V}_{xit} + \zeta_{2t}^{-2} \hat{V}^{\xi}_{xit}$. I proceed by defining an important concept of *price informativeness*.

Definition 5. *Price informativeness for investor i is defined by precision of the price signal given the dividend realization and her information about demand shocks. That is, in period t, the price informativeness is $\mathbb{V}ar^{-1}\left(\tilde{\nu}^p_{it} \mid \tilde{\eta}^x_{it}, \tilde{\xi}^x_{it}\right)$ for $\mathcal{G}_t$ and $\mathbb{V}ar^{-1}\left(\hat{\tilde{\nu}}^p_{it} \mid \tilde{\eta}^x_{it}, \tilde{\xi}^x_{it-1}, \tilde{\hat{\xi}}^x_{it}\right)$ for $\mathcal{G}_{t-1}$.*

As we can see, price informativeness for an investor i equals τ_{pt} for $\mathcal{G}_t$-investor and $\hat{\tau}_{pit}$ for $\mathcal{G}_{t-1}$-investor, respectively.

[9] V_{xt}, V^{ξ}_{xt}, $\hat{V}_{xit}$ and $\hat{V}^{\xi}_{xit}$ are defined in the appendix. The subscript i is dropped from them and τ_{pt} because in period t, $\mathcal{G}_t$-investors are *ex ante* identical.

To summarize, each investor will synthesize all information and obtain two signals, the dividend signal $\tilde{\eta}_{it}^{y}$ and the price signal $\tilde{\eta}_{it}^{p}$ ($\tilde{\hat{\eta}}_{it}^{p}$), regarding the dividend innovation. Private knowledge about $\tilde{x}_t$ will affect their ability to extract information about the dividend innovation from the equilibrium price. In other words, it affects price informativeness. An investor can learn more from the price either when she learns more about $\tilde{x}_t$ so that she can better extract information from the price, or other investors have learned more about $\tilde{x}_t$ so that there is not much noise contained in the price. The acquired signal $\tilde{\xi}_{it}^{x}$ about future demand shocks will affect next period's learning abilities and the marginal value of learning.

Let $\mathcal{F}_{t-1}^{+}$ be the information set that contains all past realizations of dividend innovations, demand shocks and equilibrium prices. The two generations have different information sets at different choice stages, which can be summarized by Table 2.1.

Generation	Information Choice Stage	Portfolio Choice Stage
$\mathcal{G}_t$	$\mathcal{I}_{t-1}^{+} = \mathcal{F}_{t-1}^{+}$	$\mathcal{I}_{it} = \left\{\mathcal{F}_{t-1}^{+}, \eta_{it}^{y}, \eta_{it}^{x}, \xi_{it}^{x}, p_t\right\}$
$\mathcal{G}_{t-1}$	$\hat{\mathcal{I}}_{it-1}^{+} = \left\{\mathcal{F}_{t-1}^{+}, \xi_{it-1}^{x}\right\}$	$\hat{\mathcal{I}}_{it} = \left\{\mathcal{F}_{t-1}^{+}, \xi_{it-1}^{x}, \hat{\eta}_{it}^{y}, \hat{\eta}_{it}^{x}, \hat{\xi}_{it}^{x}, p_t\right\}$

$\mathcal{G}_t$- and $\mathcal{G}_{t-1}$-investors have different information sets, and the main difference is that what $\mathcal{G}_{t-1}$-investors have learnt about $\tilde{x}_t$ in period t carries on and affects their decision making in period t.

Table 2.1: Information sets of the two generations in Period t

$\mathcal{F}_{t-1}^{+}$ can be interpreted as what an outsider knows about the market at the beginning of period t. *Ex ante*, $\mathcal{G}_t$-investors know as much as an outsider. But at the portfolio choice stage, they will have heterogeneous information because they have received private signals. After the current dividend innovation is realized, signals about the current dividend innovation and demand shock no longer assist in future decision making. Only the signal realization of next period's demand shock carries on and affects future choices.

2.2.2 Equilibrium definition

Definition 6. *A rational expectations equilibrium in this economy is a sequence of information choices* $\{\tau_{\epsilon yt}, \tau_{\epsilon xt}, \tau_{\nu xt}, \hat{\tau}_{\epsilon yit+1}, \hat{\tau}_{\epsilon xit+1}, \hat{\tau}_{\nu xit+1}\}$*, prices of the risky asset* $\{p_t\}$*, and portfolio choices*

$\{q_{i1}, \hat{q}_{i2}\}, \forall t \in [0, \infty)$ *such that*

(a) At the beginning of period t*,* $\mathcal{G}_t$*-investor* i *chooses signal precision* $\tau_{\epsilon yt}$*,* $\tau_{\epsilon xt}$ *and* $\tau_{\nu xt}$ *to maximize ex ante utility* $\mathbb{E}[-e^{-\rho \tilde{W}_{it+1^+}} \mid \mathcal{I}^+_{t-1}]$ *subject to information capacity constraint (2.4);*

(b) At the portfolio choice stage of period t*,* $\mathcal{G}_t$*-investor* i *invests* q_{i1} *in the risky asset to maximize* $\mathbb{E}[-e^{-\rho \tilde{W}_{it+1^+}} \mid \mathcal{I}_{it}]$*, taking the asset price and signal realizations as given, subject to budget constraint (2.3);*

(c) At the beginning of period $t+1$*,* $\mathcal{G}_t$*-investor* i *chooses signal precision* $\hat{\tau}_{\epsilon yit+1}$*,* $\hat{\tau}_{\epsilon xit+1}$ *and* $\hat{\tau}_{\nu xit+1}$ *to maximize* $\mathbb{E}[-e^{-\rho \tilde{W}_{it+1^+}} \mid \hat{\mathcal{I}}^+_{it}]$*, taking private signals from period* t *as given, subject to information capacity constraint (2.2);*

(d) At the portfolio choice stage of period $t+1$*,* $\mathcal{G}_t$*-investor* i *invests* $\hat{q}_{i2}$ *in the risky asset to maximize* $\mathbb{E}[-e^{-\rho \tilde{W}_{it+1^+}} \mid \hat{\mathcal{I}}_{it+1}]$*, taking the asset price and signal realizations, including those from period* t*, as given, subject to budget constraint (2.1);*

(e) The price of the risky asset equates demand and supply:

$$\int_{i \in \mathcal{G}_t} q_{it} \, \mathrm{d}i + \int_{i \in \mathcal{G}_{t-1}} \hat{q}_{it} - q_{it-1} \, \mathrm{d}i = \bar{x}_t + \tilde{x}_t + \int_{i \in \mathcal{G}_{t-2}} \hat{q}_{it-1} \, \mathrm{d}i \,. \tag{2.6}$$

2.2.3 Model solutions

Portfolio choices

For notatonal simplicity, I define the expected period-t payoff and the expected period-$(t+1)$ payoff conditional on $\mathcal{G}_t$-investor i's interim information set in period t as $\mathcal{E}_{it} \equiv \mathbb{E}[\tilde{d}_t + \alpha p_{t+1} \mid \mathcal{I}_{it}] - p_t r$ and $\mathcal{E}_{2i} \equiv \mathbb{E}[\tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1} r \mid \mathcal{I}_{it}]$, respectively, and their variance as $\mathcal{V}_t \equiv \mathbb{V}ar(\tilde{d}_t + \alpha p_{t+1} \mid \mathcal{I}_{it})$ and $\mathcal{V}_2 \equiv \mathbb{V}ar(\tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1} r \mid \mathcal{I}_{it})$, respectively. The covariance of the period-t payoff and the period-$(t+1)$ payoff conditional on $\mathcal{I}_{it}$ is defined as $\mathcal{C} \equiv \mathbb{C}ov(\tilde{d}_t + \alpha p_{t+1}, \tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1} r \mid \mathcal{I}_{it})$. These conditional moments are derived in the appendix. Then

the portfolio choice of $\mathcal{G}_t$-investor i can be written as

$$q_{i1} = \frac{1}{\rho r}\left(\frac{\mathcal{V}_2}{\mathcal{V}_t\mathcal{V}_2 - \mathcal{C}^2}\mathcal{E}_{it} - \frac{\mathcal{C}}{\mathcal{V}_t\mathcal{V}_2 - \mathcal{C}^2}\mathcal{E}_{2i}\right). \tag{2.7}$$

If the conditional expectation about future excess payoffs ($\mathcal{E}_{2i}$) changes, the individual demand curve shifts parallelly. But when the conditional variance of future excess payoff becomes smaller or the conditional covariance of current and future excess payoffs becomes larger, the slope of the demand curve would become less steep.

For $\mathcal{G}_{t-1}$-investor i, she faces a simple one-shot portfolio choice problem, which has an optimal solution

$$\hat{q}_{it} = \frac{\hat{\mathcal{E}}_{it}}{\rho\hat{\mathcal{V}}_{it}}, \tag{2.8}$$

where $\hat{\mathcal{E}}_{it} \equiv \mathbb{E}[\tilde{d}_t + \alpha p_{t+1} \mid \hat{\mathcal{I}}_{it}] - p_t r$ and $\hat{\mathcal{V}}_{it} \equiv \mathbb{V}ar(\tilde{d}_t + \alpha p_{t+1} \mid \hat{\mathcal{I}}_{it})$.

Static case When $\alpha = 0$, the model is static: it does not involve portfolio rebalance and each period t is an identical snapshot from the timeline. One thing that is worth noticing is that the covariance term $\mathcal{C}_t = 0$ because conditional on $\mathcal{I}_{it}$, the random component in $\tilde{d}_t - p_t$ is $\tilde{y}_t \mid \tilde{\eta}^y_{it}, \tilde{\eta}^p_{it}$ and in $\tilde{d}_{t+1} - p_{t+1}$ are $\tilde{y}_{t+1}, \tilde{x}_{t+1} \mid \tilde{\xi}^x_{it}$ and $\tilde{x}_{t+2}$, respectively. They are orthogonal (see appendix for proof). This effectively separates an investor's two portfolio choices. Therefore, the optimal portfolio choices for $\mathcal{G}_t$- and $\mathcal{G}_{t-1}$-investors are $q_{i1} = \frac{\mathbb{E}[\tilde{d}_t|\mathcal{I}_{it}] - p_t r}{\rho r \mathbb{V}ar\left(\tilde{d}_t|\mathcal{I}_{it}\right)}$ and $\hat{q}_{it} = \frac{\mathbb{E}\left[\tilde{d}_t|\hat{\mathcal{I}}_{it}\right] - p_t r}{\rho \mathbb{V}ar\left(\tilde{d}_t|\hat{\mathcal{I}}_{it}\right)}$, respectively.

Equilibrium price relation

To simplify notations, let us define $\bar{\chi}_{yt} \equiv \frac{1}{r}\Omega_t V_{yt} + \int_{i\in\mathcal{G}_{t-1}} \frac{\hat{V}_{yit}}{\hat{\mathcal{V}}_{it}}\,\mathrm{d}i$ and $\bar{\chi}_{xt} \equiv \frac{1}{r}\Omega_t V_{xt} + \int_{i\in\mathcal{G}_{t-1}} \frac{\hat{V}_{xit}}{\hat{\mathcal{V}}_{it}}\,\mathrm{d}i$. Please refer to B.1.2 for details. In equilibrium, the period-t price coefficients

are recursively characterized by

$$A_t = \left(1 - \frac{\bar{\Sigma}_t}{2r}\frac{\mathcal{C}}{\mathcal{V}_t\mathcal{V}_2 - \mathcal{C}^2}\right)\frac{\mu}{r-\alpha g} - \frac{\bar{\Sigma}_t\rho}{r}\bar{x}_t + \frac{1}{r}\left(\alpha + \frac{\bar{\Sigma}_t}{2r}\frac{\mathcal{C}}{\mathcal{V}_t\mathcal{V}_2 - \mathcal{C}^2}\right)A_{t+1} - \frac{\bar{\Sigma}_t}{2r^2}\frac{\mathcal{C}}{\mathcal{V}_t\mathcal{V}_2 - \mathcal{C}^2}\alpha A_{t+2}, \tag{2.9}$$

$$B_t = \frac{g}{r-\alpha g}, \tag{2.10}$$

$$C_t = \frac{1}{r-\alpha g}\left(1 - \bar{\Sigma}_t\bar{\chi}_{yt}\tau_y\right) \tag{2.11}$$

$$D_t = -\frac{\rho\bar{\Sigma}_t}{r} + \frac{1}{r-\alpha g}\bar{\Sigma}_t\bar{\zeta}^{\dagger}_{1t}\tau_x,$$

$$E_t = \frac{1}{r}\left[\frac{\bar{\Sigma}_t}{2r}\frac{\mathcal{C}}{\mathcal{V}_t\mathcal{V}_2 - \mathcal{C}^2}V^{\xi}_{xt}\tau_{\nu xt} + \alpha\left(1 - \bar{\Sigma}_t\bar{\chi}_{xt}\tau_x\right)\right]D_{t+1} + \frac{1}{r-\alpha g}\bar{\Sigma}_t\bar{\zeta}^{\dagger}_{2t}\tau_x. \tag{2.12}$$

B_t is the constant dividend growth rate, and together A_t and B_t govern the level of the equilibrium price. C_t is the price sensitivity to dividend innovation, and D_t and E_t regulate the amount of current and future demand shocks contained in the price, respectively. $\bar{\zeta}^{\dagger}_{1t}$ and $\bar{\zeta}^{\dagger}_{2t}$ are the market-wide average future-uncertainty-adjusted dividend-to-total-uncertainty-weighted signal-to-current-noise and signal-to-future-noise ratios, defined by (B.43) and (B.44), respectively.

Static case When $\alpha = 0$, the equilibrium price relation greatly simplifies. Firstly, $E_t = 0$. This means that here $\tilde{x}_{t+1}$ does not enter the equilibrium price at t. This is because $\tilde{d}_t$ and $\tilde{d}_{t+1}$ are different assets, and the supply shock to $\tilde{d}_{t+1}$ does not affect the equilibrium price of $\tilde{d}_t$ directly, but only indirectly though price coefficients due to information acquisition. Secondly, $\mathcal{C}_t = 0$, that is the current and future excess payoffs are conditionally independent. Thirdly, $\mathcal{G}_{t-1}$-investors make homogeneous period-t information choices. Then price coefficients then become

$$A_t = \frac{1}{r}(\mu - \bar{\Sigma}_t\rho\bar{x}_t), \quad B_t = \frac{g}{r}, \quad C_t = \frac{1}{r}\left(1 - \frac{r+1}{2r}\bar{\Sigma}_t\tau_y\right), \quad D_t = \frac{1}{r}\left(-\rho\bar{\Sigma}_t + \frac{r+1}{2r}\bar{\Sigma}_t\zeta_t\tau_x\right). \tag{2.13}$$

Information choices

Ex ante expected utilities In period t, $\mathcal{G}_{t-1}$-investors are in the second period of their life. the objective function is $\mathbb{E}\left[-e^{-\rho\tilde{W}_{it^+}} \mid \hat{\mathcal{I}}^+_{it-1}\right]$ and the constraint of the information choice problem is very similar to (2.2), with only a change of subscript t. This yields an objective for investor $i \in \{\mathcal{G}_{t-1}\}$ that is increasing and concave in precision of the fundamental signal and the current demand signal:

$$\mathbb{E}\left[e^{-\rho r W_{it^+}} \mid \hat{\mathcal{I}}_{it-1^+}\right] = -e^{-\rho r W_{it-1^+}} \left|\frac{\hat{\Sigma}_{t-1^+}}{\hat{\mathcal{V}}_{it}}\right|^{-1/2} \exp\left\{-\frac{\mu^2_{\hat{\mathcal{E}}_t}}{2\hat{\Sigma}_{t-1^+}}\right\}. \tag{2.14}$$

where $\mu_{\hat{\mathcal{E}}_t}$ and $\hat{\Sigma}_{t-1^+}$ are defined by (B.18) and (B.17), respectively. Notice that in the static model, the investor optimally choose not to learn about future non-fundamental risks since she does not have a future to come and thus $\hat{\tau}_{\nu x i t+1}$ does not affect her ex-ante utility.

For $\mathcal{G}_t$-investors, the information choice problem is to maximize $\mathbb{E}\left[-e^{-\rho\tilde{W}_{it+1^+}} \mid \mathcal{I}^+_{t-1}\right]$ subject to (2.4). This yields an objective for $i \in \{\mathcal{G}_t\}$ that is increasing and concave in precision of the fundamental signal, the current demand signal and the future demand signal precision:

$$\mathbb{E}\left[-e^{-\rho W_{it+1^+}|\mathcal{I}^+_{t-1}}\right] = -e^{-\rho r^2 W_0} \times \sqrt{\frac{\hat{\mathcal{V}}_{it+1}}{\hat{\mathcal{V}}_{it+1} + \Sigma_{\hat{Z}_{t+1}}}} |I - 2\Sigma F|^{-\frac{1}{2}} e^{\frac{1}{2}G'(I-2\Sigma F)^{-1}\Sigma G + J}, \tag{2.15}$$

where $\hat{\mathcal{V}}_{it+1}$, $\Sigma_{\hat{Z}_{t+1}}$, Σ, F, G and J are defined in the appendix.

Information choices in the static model Information choices in the dynamic model are very complex. To develop intuitions, here I derive the analytical solution of information choices in the static model. To understand how she would allocate precision to signals about current fundamental and demand shocks, it is convenient to first define the concept of learning index.

Definition 7. *Learning index for a signal is defined as the increase in ex-ante marginal utility associated with an extra unit of precision assigned to this signal.*

For $\mathcal{G}_{t-1}$-investors, the learning index for the signals about current fundamental and demand shock are $\frac{\partial \widehat{EU}}{\partial \hat{\tau}_{\epsilon y i t}}$ and $\frac{\partial \widehat{EU}}{\partial \hat{\tau}_{\epsilon x i t}}$, respectively. The learning indices are functions of current precision choice $\hat{\tau}_{\epsilon y i t}$ and $\hat{\tau}_{\epsilon x i t}$, and also the signal precision $\tau_{\nu x t-1}$ and the realization of the signal about current demand shocks received in the previous period, ξ^x_{it-1}. Since all investors are ex-ante identical, the signal precision $\tau_{\nu x t-1}$ is the same across $\mathcal{G}_{t-1}$-investors. But signal realizations ξ^x_{it-1} are private information to investors, introducing heterogeneity in second-period learning indices. The effects of ξ^x_{it-1} on learning indices for signals about current dividend innovation and demand shocks are in such a proportional way that the weights of them do not depend on learning indices. This observation can be summarized by the following lemma.

Lemma 5 (Proportionality of learning indices). *The ratio of second-period learning indices for current dividend innovation and demand shocks only depend on the signal-to-noise ratio, not on ξ^x_{it-1}, the realization of the signal about current demand shocks received in the previous period, that is $\frac{\hat{L}_{\epsilon x i t}}{\hat{L}_{\epsilon y i t}} = \zeta_t^2$.*

To better characterize the benefit of learning about future demand shocks, for now I make the following assumption.

Assumption 1. *Acquiring signals about current dividend innovation, current and future demand shocks is equally costly, that is $\gamma_{\epsilon y} = \gamma_{\epsilon x} = \gamma_{\nu x}$.*

We can then summarize the information choices of $\mathcal{G}_{t-1}$-investors by the following proposition.

Proposition 13 (Second-period information choice). *In period t, $\mathcal{G}_{t-1}$-investors*

(a) optimally choose not to learn about future demand shocks, that is $\hat{\tau}_{\nu x i t} = 0$; and

(b) [under Assumption 1] only learn about current dividend innovation regardless of the realization of signal about $\tilde{x}_t$ they received in the previous period, that is $\hat{\tau}_{\epsilon y i t} = \frac{K}{\gamma_{\epsilon y}}$ and $\hat{\tau}_{\epsilon x i t} = 0$.

What $\mathcal{G}_{t-1}$-investors have learned in the past about today does not affect their information choice for today. But they do have an additional signal about today's asset. Their conditional

mean and variance about today's asset are different than those of $\mathcal{G}_t$-investors. Then the portfolio choice and thus payoffs are affected.

The proportionality of learning indices for signals about current non-fundamental and fundamental shocks preserves, as is summarized in Lemma 6

LEMMA 6 (Proportionality of learning indices). *The ratio of first-period learning indices for current dividend innovation and demand shocks only depend on the signal-to-noise ratio, that is* $\frac{L_{\epsilon xit}}{L_{\epsilon yit}} = \zeta_t^2$.

We can then summarize the first-period information choice by Proposition 14.

PROPOSITION 14 (First-period information choice). *Under Assumption 1,* $\mathcal{G}_t$*-investors optimally choose to only learn about future demand shocks, that is* $\tau_{\epsilon yt} = \tau_{\epsilon xt} = 0$ *and* $\tau_{\nu xt} = \frac{K}{\gamma_{\nu x}}$.

Because $\mathcal{G}_t$-investors are identical at the information choice stage of period t, they will make the same choice of precision, and thus will have the same learning indices. Compared with the benchmark, the learning indices have changed because the option to learn about future demand shocks affects $\bar{\Sigma}_t$, the across-generation average of posterior variance of $\tilde{y}_t$, in two ways. Firstly, investor $i \in \mathcal{G}_t$ optimally choose to only acquire signals about $\tilde{x}_{t+1}$ and not learn about any current shocks. Secondly, the learning indices of investor $i \in \mathcal{G}_{t-1}$ is affected by their observed private signal $\tilde{\hat{\xi}}_{it}^x$ and their portfolio choices change accordingly. Changes in learning indices of the old generation spill over to the young generation through $\bar{\Sigma}_t$.

Compared with learning about $\tilde{x}_{t+1}$ in period $t+1$, learning about it in period t has an extra bite: it decreases the variance of future excess payoffs conditional on today's information, which is directly affected by τ_y, τ_x and $\tau_{\nu xt}$.

2.3 Results: Static Model

With a long-lived risky asset, learning about future demand shocks affects the market via two mechanisms: the front-running channel and the future uncertainty channel. These two mechanisms work in the same direction. The new piece here is the future uncertainty channel. To see

how it works more clearly, in this section I study the static model and set $\alpha = 0$ to shut down the front-running channel.

2.3.1 Limiting cases

Benchmark I: no signal about future demand shocks ($\tau_{\nu xt} = 0$)

Proposition 15. *Under Assumption 1, in both periods, investors make identical information choices $\tau^B_{\epsilon yt} = \tau^B_{\epsilon yit+1} = \frac{K}{\gamma_{\epsilon y}}$ and $\tau^B_{\epsilon xt} = \tau^B_{\epsilon xit+1} = 0$ if $K < \frac{r}{r+1}\rho\sqrt{\gamma_{\epsilon y}\gamma_{\epsilon x}}$, and arbitrarily allocate precision to $\tau^B_{\epsilon yt}$ and $\tau^B_{\epsilon xt}$ if $K = \frac{r}{r+1}\rho\sqrt{\gamma_{\epsilon y}\gamma_{\epsilon x}}$. The signal-to-noise ratio and posterior average variance are $\zeta^B_t = -\frac{1}{\rho}\frac{r+1}{r}\frac{K}{\gamma_{\epsilon y}}$ and $\left(\bar{\Sigma}^B_t\right)^{-1} = \frac{r+1}{r}\left[\tau_y + \left(\zeta^B_t\right)^2\tau_x\right] - \rho\zeta^B_t$, respectively.*

Proposition 16. *Under Assumption 1, in the Benchmark I model without signals about future demand shocks:*

(a) The equilibrium exists when $K \leq \rho\frac{r}{r+1}\sqrt{\gamma_{\epsilon y}\gamma_{\epsilon x}}$.

(b) $C^B_t > 0, D^B_t < 0$.

(c) The equilibrium signal-to-noise ratio $|\zeta_t|$ is increasing in information processing capacity K, that is $\frac{\partial|\zeta^B_t|}{\partial K} > 0$, decreasing in learning costs associated with demand shocks $\gamma_{\epsilon x}$, that is $\frac{\partial|\zeta^B_t|}{\partial\gamma_{\epsilon x}} < 0$, and not affected by prior precision about fundamental shocks τ_y and demand shocks τ_x, that is $\frac{\partial|\zeta^B_t|}{\partial\tau_x} = \frac{\partial|\zeta^B_t|}{\partial\tau_y} = 0$.

(d) The posterior average precision $\left(\bar{\Sigma}^B_t\right)^{-1}$ is increasing in information processing capacity K, that is $\frac{\partial\left(\bar{\Sigma}^B_t\right)^{-1}}{\partial K} > 0$, and prior precision about fundamental shocks τ_y, that is $\frac{\partial\left(\bar{\Sigma}^B_t\right)^{-1}}{\partial\tau_y} > 0$, and demand shocks τ_x, that is $\frac{\partial\left(\bar{\Sigma}^B_t\right)^{-1}}{\partial\tau_x} > 0$, and decreasing in learning costs associated with demand shocks $\gamma_{\epsilon x}$, that is $\frac{\partial\left(\bar{\Sigma}^B_t\right)^{-1}}{\partial\gamma_{\epsilon x}} < 0$.

(d) As the information processing capacity K goes up, investor i will allocate higher precision to both fundamental and non-fundamental signals, that is $\frac{\mathrm{d}\tau^B_{\epsilon xt}}{\mathrm{d}K} > 0$ and $\frac{\mathrm{d}\tau^B_{\epsilon yt}}{\mathrm{d}K} > 0$. As the learning cost associated with demand shocks $\gamma_{\epsilon x}$ goes up, investor i will allocate higher precision to fundamental signals and lower precision to non-fundamental signals, that is $\frac{\mathrm{d}\tau^B_{\epsilon xt}}{\mathrm{d}\gamma_{\epsilon x}} < 0$ and

$\frac{\mathrm{d}\tau^B_{\epsilon y t}}{\mathrm{d}\gamma_{\epsilon x}} > 0$, *and information choices are not affected by prior uncertainty of fundamental or demand shocks, that is* $\frac{\mathrm{d}\tau^B_{\epsilon x t}}{\mathrm{d}\tau_x} = \frac{\mathrm{d}\tau^B_{\epsilon y t}}{\mathrm{d}\tau_x} = \frac{\mathrm{d}\tau^B_{\epsilon x t}}{\mathrm{d}\tau_y} = \frac{\mathrm{d}\tau^B_{\epsilon y t}}{\mathrm{d}\tau_y} = 0$.

When $\gamma_{\nu x} \to \infty$ in the main model, due to limited information capacity, learning about future demand shocks becomes too costly, and investors optimally choose not to learn about them. The model thus degenerates to Benchmark I.

Proposition 17. *When learning about future demand shocks is sufficiently costly, the main model is sufficiently close to Benchmark I, that is there exists a* $\bar{\gamma}_{\nu x}$ *such that* $\forall \gamma_{\nu x} > \bar{\gamma}_{\nu x}$*, Proposition 16 applies to the main model.*

Benchmark II: future demand shocks can be learned perfectly without cost

At t, $\mathcal{G}_{t-1}$-investors know $\tilde{x}_t$ and thus can infer $\tilde{y}_t$ from the price signal perfectly. If $d_t > p_t$, then they take an infinite long position in the risky asset, that is $\hat{q}_{it} = \infty$; if $d_t < p_t$, then they take an infinite short position, that is $\hat{q}_{it} = -\infty$. For the market of the risky asset to clear, we must have $p_t = d_t$, so price is fully revealing. All $\mathcal{G}_{t-1}$- and $\mathcal{G}_t$-investors will not acquire any information. Market collapse and no trade will take place.

Proposition 18 (Benchmark II). *If investors can freely observe the demand shock in the next period, then the equilibrium price is fully revealing. There will be no information acquisition or trades of assets.*

2.3.2 Price reveals less fundamental information

When we grant investors another source for non-fundamental information, the informational content of price will be changed accordingly. We now examine this issue by looking at the signal-to-noise ratio, the average posterior precision about fundamental shocks and price informativeness.

Understanding how the signal-to-noise ratio $|\zeta_t|$ changes with model parameters helps us figure out the information quality in the market. In Benchmark I and the second-period information

choice of the main mode, $|\zeta_t|$ summarizes all relevant informational contents in the market; together with parameters associated with learning costs ($\gamma_{\epsilon y}$, $\gamma_{\epsilon x}$, $\gamma_{\nu x}$ and K), the signal-to-noise ratio $|\zeta_t|$ determines how investors allocate signal precision in Benchmark I and in the second period of the main model.

Investors have a prior belief about the fundamental shocks. The average posterior precision about the dividend innovation $\bar{\Sigma}_t^{-1}$ is the sum of prior precision of the fundamental shocks, precision of dividend signal and precision of the price signal. is a function of the risk-free rate, the risk-aversion coefficient, prior precision of dividend innovation and demand shock, and the signal-to-noise ratio $|\zeta_t|$. $\bar{\Sigma}_t^{-1}$ and $|\zeta_t|$ often comove with changes in parameters, as is shown in Figure 2.2.

LEMMA 7. *The market-wide average posterior precision about the fundamental shock* ($\bar{\Sigma}_t^{-1}$) *comoves positively with the signal-to-noise ratio* ($|\zeta_t|$)*, if changes in* $|\zeta_t|$ *are due to changes in parameters* K, $\gamma_{\epsilon x}$, $\gamma_{\epsilon y}$, $\gamma_{\nu x}$ *or* τ_y*. Mathematically,* $\text{sign}\,\frac{\partial\bar{\Sigma}_t^{-1}}{\partial K} = \text{sign}\,\frac{\partial|\zeta_t|}{\partial K}$, $\text{sign}\,\frac{\partial\bar{\Sigma}_t^{-1}}{\partial\gamma_{\epsilon x}} = \text{sign}\,\frac{\partial|\zeta_t|}{\partial\gamma_{\epsilon x}}$, $\text{sign}\,\frac{\partial\bar{\Sigma}_t^{-1}}{\partial\gamma_{\epsilon y}} = \text{sign}\,\frac{\partial|\zeta_t|}{\partial\gamma_{\epsilon y}}$, $\text{sign}\,\frac{\partial\bar{\Sigma}_t^{-1}}{\partial\gamma_{\nu x}} = \text{sign}\,\frac{\partial|\zeta_t|}{\partial\gamma_{\nu x}}$ *and* $\text{sign}\,\frac{\partial\bar{\Sigma}_t^{-1}}{\partial\tau_y} = \text{sign}\,\frac{\partial|\zeta_t|}{\partial\tau_y}$.

PROPOSITION 19. *With access to information about future demand shock,*

(a) the signal-to-noise ratio $|\zeta_t|$ *and the market-wide average posterior precision about the fundamental shock* $\bar{\Sigma}_t^{-1}$ *increase with information processing capacity* K *and prior precision of demand shocks* τ_x*, i.e.* $\frac{\partial|\zeta_t|}{\partial K} > 0$, $\frac{\partial|\zeta_t|}{\partial\tau_x} > 0$, $\frac{\partial\bar{\Sigma}_t^{-1}}{\partial K} > 0$ *and* $\frac{\partial\bar{\Sigma}_t^{-1}}{\partial\tau_x} > 0$*;*

(b) Assumes that the learning cost associated with demand shocks are the same today and tomorrow, i.e. $\gamma_{\epsilon x} = \gamma_{\nu x} = \gamma_x$*. Then there exists* $\bar{K}$, γ_x *and* γ_x^* *such that when* $K > \bar{K}$ *and* $\gamma_x < \gamma_x$, $|\zeta_t|$ *and* $\bar{\Sigma}_t^{-1}$ *decrease with* γ_x *if* $\gamma_x < \gamma_x^*$ *and increase with* γ_x *if* $\gamma_x > \gamma_x^*$*. When* $K \leq \bar{K}$ *or* $\gamma_x \geq \gamma_x$, $|\zeta_t|$ *and* $\bar{\Sigma}_t^{-1}$ *increase with* γ_x*. Mathematically,*

$$\text{sign}\,\frac{\partial|\zeta_t|}{\partial\gamma_x} = \text{sign}\,\frac{\partial\bar{\Sigma}_t^{-1}}{\partial\gamma_x}\begin{cases} -, & \text{if } K > \bar{K}, \gamma_x < \gamma_x, \gamma_x < \gamma_x^*, \\ +, & \text{otherwise.} \end{cases}$$

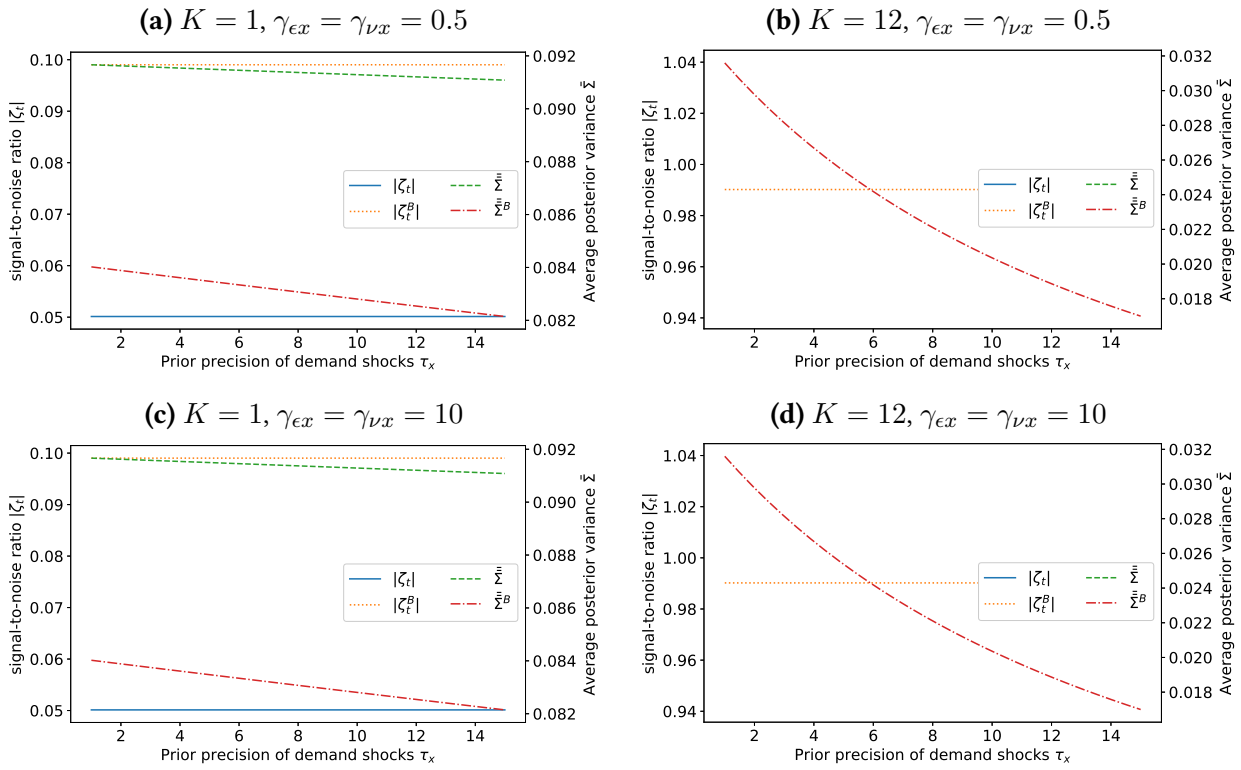


Figure 2.2. The signal-to-noise ratio and the average posterior precision

These figures plot how the signal-to-noise ratio $|\zeta_t|$ and the average posterior precision $\bar{\Sigma}_t$ change with K with different values of prior precision of demand shocks and the learning costs associated with learning about demand shocks. Compared with the benchmark model without information acquisition on future demand shocks, the signal-to-noise ratio is lower and the average posterior variance is higher. Both in the main model and the benchmark model, the signal-to-noise ratio increases and the average posterior variance decreases with information processing capacity K.

PROPOSITION 20. *Compared with Benchmark I, the signal-to-noise ratio* $|\zeta_t|$ *decreases when investors have an extra source of information on future demand shocks. The same results apply to average posterior precision* $\bar{\Sigma}_t^{-1}$. *That is* $\operatorname{sign} \Delta|\zeta_t| = \operatorname{sign} \Delta\bar{\Sigma}_t^{-1} = -$, *where* $\Delta|\zeta_t| = \zeta_t - \zeta_t^B$ *and* $\bar{\Sigma}_t^{-1} = \left(\bar{\Sigma}_t^{-1}\right)^B$.

THEOREM 1 (Price informativeness). *With access to information about future demand shocks, price informativeness decreases compared with Benchmark I .*

Figure 2.3a plots price informativeness against the information processing capacity. For the two periods in an investor's life, the equilibrium price is not equally informative for her. In the second period, price is more informative because she has required a signal about second-period

demand shocks in her first period. Although the price contains less fundamental information, her ability to extract other investors' private information out of the equilibrium price is higher, and the latter effect dominates. However, in the first period, her ability to extract information from price is not changed. Since price contains less fundamental information, it is less informative for her.

Figure 2.3b plots the average posterior variance of the dividend innovation against the information processing capacity. On the aggregate level, less fundamental information is acquired in the market. As we can see here, the average posterior variance of fundamental shocks is larger when investors can learn about future demand shocks. As a result, risk premium is higher.

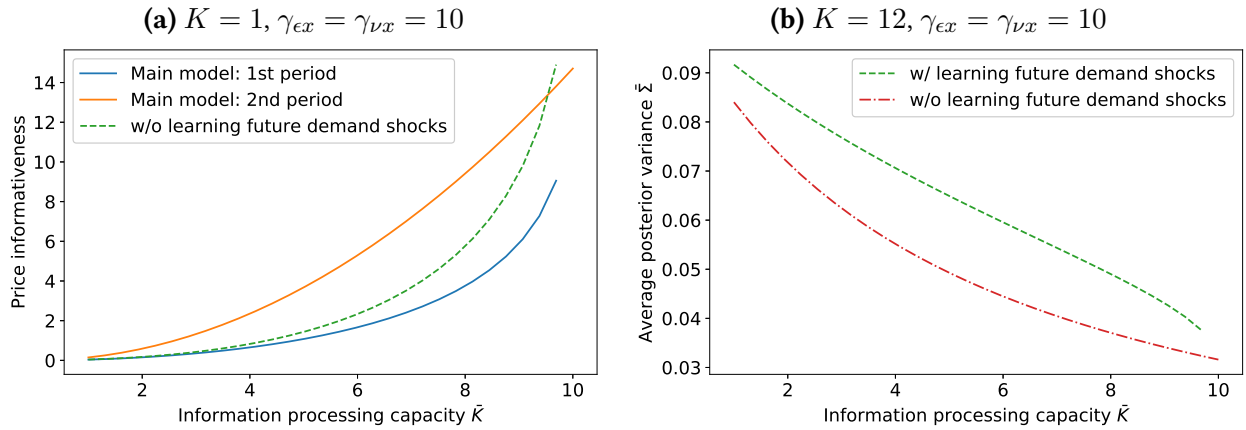


Figure 2.3. Impact of learning about future demand shocks on price informativeness and risk premium.

Panel (a) plots price informativeness against the information processing capacity. In general, the price informativeness is increasing in the information processing capacity. Investors in different stages of their life have different information sets and thus different abilities to extract dividend information from the equilibrium price. Panel (b) plots the average posterior variance of dividend innovation against the information processing capacity. Since the average posterior variance of dividend innovation is positively correlated with risk premium, the plot shows that the risk premium increases when investors are allowed to learn about future demand shocks.

2.3.3 Information choices

THEOREM 2. *When it is possible to learn from future demand shocks, investor i's information choices change in both stages of her life:*

*(a) [**fundamental vs. non-fundamental**] Agent i learns less about the fundamental shocks*

today compared with the benchmark;

(b) ***[current vs. future]*** *Agent i learns more about future demand shock than current demand shock.*

Since the marginal cost of allocating precision to one single signal is increasing, granting access of another signal about demand shocks to investors essentially makes demand shocks cheaper to learn. As a result, investor i will learn less about the fundamental shocks today. Precision of the signal of current demand shock and that of future demand shock contribute the conditional mean and variance of $\tilde{x}_t$ in a symmetric way. Starting from the benchmark equilibrium, every unit of $\tau_{\epsilon xt}$ investor i moves away, she can allocate more than one unit to $\tau_{\nu xt}$. Since investor i live for two periods, allocating precision to $\tau_{\nu xt}$ has an extra benefit of lowering the conditional variance of excess payoff from the perspective of today, which in turn affects tomorrow's learning index.

2.3.4 *Ex ante* utility

PROPOSITION 21. *With access to information about future demand shocks, the ex-ante utility for the young generation is improved, while the ex-ante utility of the old generation is improved when K, τ_x and $\gamma_{\epsilon x} = \gamma_{\nu x}$ are smaller than some boundaries.*

As showed in Figure 2.4, the ex-ante utility improves through the future uncertainty channel. On the aggregate level, less fundamental information acquisition leads to higher risk premium, while price is more informative and thus investors take less risk in the second period. In practice, high frequency traders are learning about future demand shocks. We tend to think of this as wasteful activity. But here I show that learning about future demand shocks indeed increase their ex-ante utility.

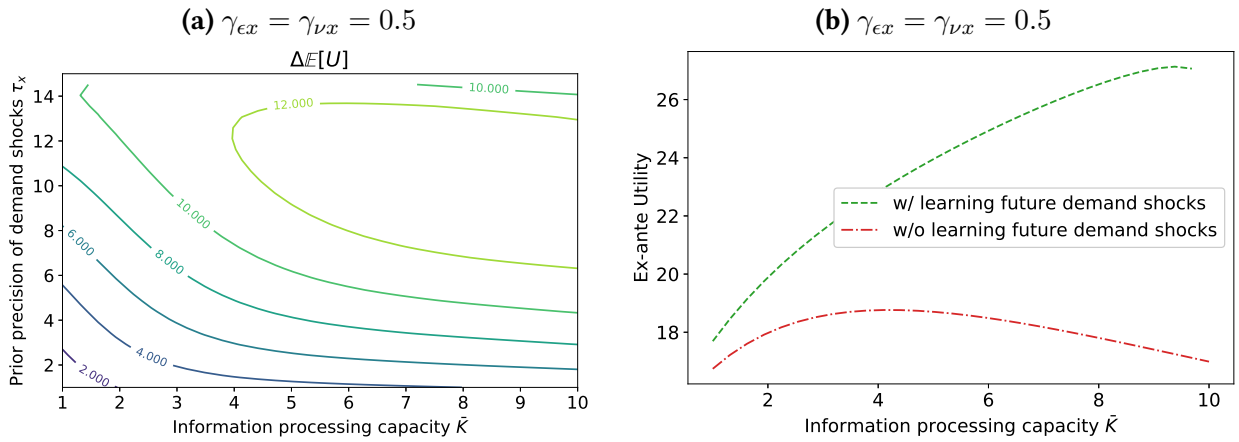


Figure 2.4. Information acquisition on future demand shocks improves ex-ante expected utility of investors

Panel (a) is the contour plot of $\Delta\mathbb{E}[U] = \mathbb{E}[U] - \mathbb{E}[U^B]$ under the linear information constraint. As the information processing capacity grows or the prior precision of demand shocks increases, information acquisition on future demand shocks brings a larger benefit in terms of beginning-of-life expected utility to investors. Panel (b) shows that compared with the benchmark model without information acquisition on future demand shocks, the expected utility is less concave in the information processing capacity.

2.3.5 Price volatility

From the perspective of $\mathcal{G}_{t-1}$-investor i, the price volatility at the beginning of period t is $\mathbb{V}ar\left(p_t \mid \hat{\mathcal{I}}^+_{it-1}\right) = C_t\tau_y^{-1}C_t + D_t\left(\tau_x + \tau_{\nu xt-1}\right)^{-1} D_t$. For an individual asset j, its conditional variance is $\mathbb{V}ar\left(p_t \mid \hat{\mathcal{I}}^+_{it-1}\right) = \frac{C_t^2}{\tau_y} + \frac{D_t^2}{\tau_x+\tau_{\nu xt-1}}$. Notice that each generation t' of investors only choose to learn about future demand shocks in period t', which affects market clearing in period $t'+1$. Recall that $\bar{\Sigma}_t^{-1} = \frac{1}{r}\left(\tau_y + \tau_{\epsilon yt} + \zeta_t\left(\tau_x + \tau_{\epsilon xt}\right)\zeta_t\right) + \tau_y + \hat{\tau}_{\epsilon yt} + \zeta_t\left(\tau_x + \hat{\tau}_{\epsilon xt} + \hat{\tau}_{\nu xt-1}\right)\zeta_t$. Comparing with the benchmark, due to acquisition of future demand signals at $t-1^+$ by generation $\mathcal{G}_t$, $\tau_{\epsilon xt}$ or $\tau_{\epsilon yt}$ is decreased. Due to acquisition of future demand signals by generation $\mathcal{G}_{t-1}$ at $t-1^+$, $\hat{\tau}_{\epsilon xt}$ or $\hat{\tau}_{\epsilon yt}$ is decreased, but their information choice at $t-2^+$ increases $\hat{\tau}_{\nu xt-1}$.

THEOREM 3 (Price volatility). *The price volatility $\mathbb{V}ar(p_t)$ changes with information processing capacity K. Given that $\frac{\partial|\zeta_t|}{\partial K} > 0$:*

(a) if $\tau_x \geq \bar{\Omega}_x$, then $\frac{\partial\mathbb{V}ar(p_t)}{\partial K} \geq 0$;

(b) if $\underline{\Omega}_x < \tau_x < \bar{\Omega}_x$, *then*

$$\frac{\partial \mathbb{V}ar(p_t)}{\partial K} \begin{cases} \geq 0, & \text{if } \zeta_t \leq \underline{\zeta}_t \text{ or } \bar{\zeta}_t \leq \zeta_t < 0, \\ < 0, & \text{if } \underline{\zeta}_t < \zeta_t < \bar{\zeta}_t; \end{cases}$$

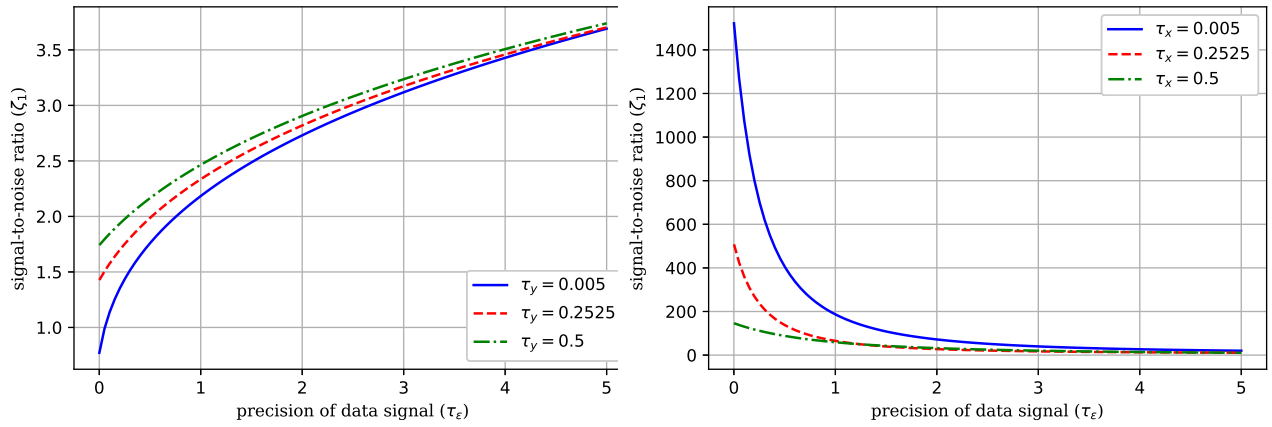
(c) If $0 < \tau_x \leq \underline{\Omega}_x$, *then*

$$\frac{\partial \mathbb{V}ar(p_t)}{\partial K} \begin{cases} \geq 0, & \text{if } \zeta_t \leq \underline{\zeta}_t, \\ < 0, & \text{if } \underline{\zeta}_t < \zeta_t < 0; \end{cases}$$

where $\bar{\Omega}_x \equiv \frac{5\rho^2}{4\left(\frac{r+1}{r}\right)^2 \tau_y}$, $\underline{\Omega}_x \equiv \frac{\rho^2}{\left(\frac{r+1}{r}\right)^2 \tau_y}$, $\bar{\zeta}_t \equiv \frac{r}{2(r+1)\tau_x}\left[-\rho + \sqrt{-4\left(\frac{r+1}{r}\right)^2 \tau_x \tau_y + 5\rho^2}\right]$, *and* $\underline{\zeta}_t \equiv \frac{r}{2(r+1)\tau_x}\left[-\rho - \sqrt{-4\left(\frac{r+1}{r}\right)^2 \tau_x \tau_y + 5\rho^2}\right]$.

From the perspective of the young generation $\mathcal{G}_t$, price volatility at the beginning of period t is not affected by whether learning about future demand shocks is possible.

(a) $\gamma_{\epsilon x} = \gamma_{\nu x} = 0.5$ **(b)** $\gamma_{\epsilon x} = \gamma_{\nu x} = 0.5$



f

Figure 2.5. Access to information about future demand shocks affects price volatility

Panel (a) is the contour plot of $\mathbb{V}ar\left(p_t r\right) - \mathbb{V}ar\left(p_t^B r\right)$. Panel (b) plots price volatility against the information processing capacity.

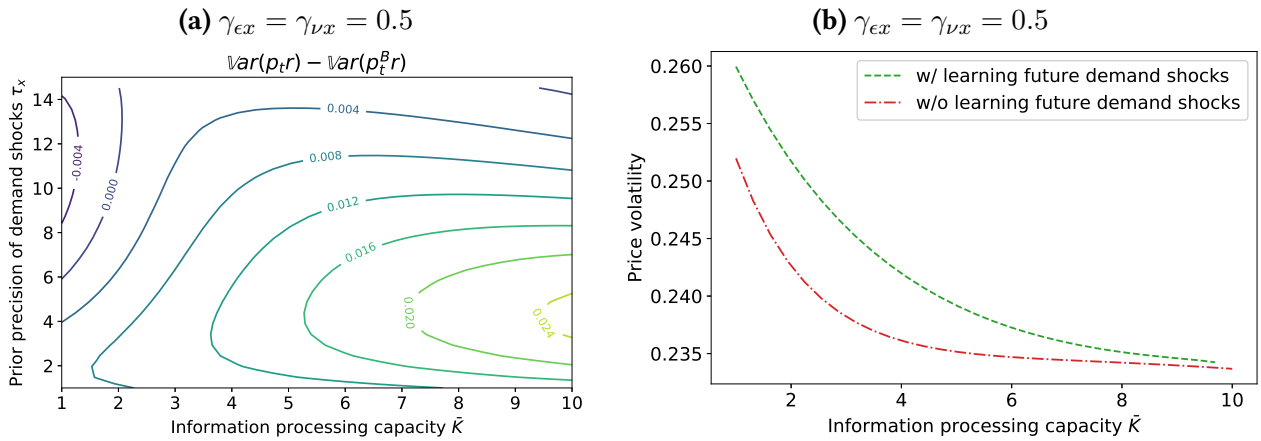


Figure 2.6. Access to information about future demand shocks affects price volatility

Panel (a) is the contour plot of $\mathbb{V}ar\left(p_t r\right) - \mathbb{V}ar\left(p_t^B r\right)$. Panel (b) plots price volatility against the information processing capacity.

Figure 2.6a shows that for given learning costs $\gamma_{\epsilon x} = \gamma_{\nu x} = 0.5$, compared with Benchmark I where learning about future demand shocks is impossible, price volatility increases when the price precision of demand shocks τ_x is low and/or the information processing capacity is high. Figure 2.6b plots price volatility against the information processing capacity. In each period t, there are two generations of investors. Without learning about future demand shocks, the two generations are identical. Learning about future demand shocks introduces more heterogeneity on private information on two margins. On the extensive margin, the older generation will not learn about future demand shocks because they do not have a future to come, while the younger generation do. The two generations make different information choices and receive different private information. On the intensive margin, the older generation have one more signal, thus their information set is more diverse. Heterogenous private information increases price volatility $\mathbb{V}ar(p_t r)$.

2.4 Empirical Applications: Dynamic Model

This model can be applied to a broad range of real-world scenarios where investors are able to learn about future demand shocks. Some examples include portfolio adjustment by index funds

upon index recomposition and the BlackRock-BGI acquisition. [10]

Index recomposition As Han et al. (2016) points out, institutional traders, e.g. index funds or ETFs, need to rebalance portfolios for non-informational reasons around index recomposition. The existence of index funds and the requirement that they replicate indices creates demand shocks that are non-fundamental to the underlying securities. On the announcement date of index recomposition, investors receive a signal of effective-date demand shocks. Sophisticated investors can even front run the market before the index recomposition announcement by making informed predictions about how index funds will have to rebalance.

The BGI transaction [11] BlackRock, Inc. announced on June 11, 2009 that it agreed to pay $13.5 billion to acquire Barclays Global Investors (BGI), including iShares, from Barclays Bank PLC.[12] When the deal closed on December 1, 2009, BlackRock paid approximately $6.65 billion in cash and issued capital stock valued at $8.53 billion to Barclays.[13] Consequently, BlackRock became the world's largest investment management company, managing, at the end of 2009, approximately $3.346 trillion worldwide.

While Barclays sold BGI to raise funds to strengthen its balance sheet, BlackRock struck the deal with Barclays, according to its Chairman and CEO Laurence Fink, to "broaden the firm's investment capability with passive and quantitative investment strategies, particularly exchange-traded funds, and to address client demand for outcome-oriented portfolio strategies and offer institutional and other investors a full spectrum of investment strategies, including passive and quantitative approaches". BGI's products complement BlackRock's fundamental active strategies that employ a bottom-up security selection process led by portfolio managers who use quantitative and fundamental analysis and their market experience to inform their judgment. BlackRock

[10] Zou (2019) used the BlackRock case to provide identification to the documented positive correlation between ETF flows and the price-to-fundamentals relation of underlying stocks.

[11] Calibration results would be included in the next version of the paper.

[12] Barclays PLC Annual Report (2009), https://home.barclays/content/dam/home-barclays/documents/investor-relations/annualreports/ar2009/2009-form-20-F.pdf (accessed December 23, 2019).

[13] BlackRock, Inc. Annual Report (2009), https://ir.blackrock.com/Cache/1001211956.PDF (accessed December 23, 2019).

aims at industry leadership in ETF globally. The BGI transaction will increase the scale benefits and introduce synergies between alpha and beta investments. Observing the news of the BGI transaction on June 11, sophisticated investors expect the iShares AUM to expand in the following years. For the component stocks of iShares ETFs, these are future demand shocks that are not related to their firm fundamentals.

2.5 Conclusion

In the literature of rational expectations equilibrium models with asymmetric information, research has been done to understand how information acquisition on current demand shocks, and current and future dividend innovation affects the market. This chapter tries to put together the missing piece and study how learning about future demand shocks would affect investors' information and investment choices and their ex-ante utility, as well as price informativeness and volatility.

Using a competitive CARA-normal NREE model with OLG investors who live for 2 periods, I show that when possible to learn about future demand shocks, investors choose to learn less about current dividend innovations and demand shocks and allocate precision to learn about future demand shocks. Learning about future demand shocks affects the financial market through the front-running channel and the future uncertainty channel. To isolate the future uncertainty channel, I derive a set of results for the static model where the asset is a series of dividend strips. If future demand shocks are independent from the demand shocks today, then future demand shocks per se do not affect today's asset price. It is *information acquisition* on future demand shocks that makes it relevant to current price through the *future uncertainty channel*. While the current equilibrium price still does not aggregate signals about future demand shocks, information acquisition on future demand shocks affects the risk premium and how the current equilibrium price aggregates signals about current shocks. The future uncertainty channel works through two forces: on the one hand, the aggregate level of dividend information acquisition decreases, which makes dividends more risky and leads to higher required returns; on the other

hand, an atomic investor is more certain about the asset payoff in the second period due to the signal about the second-period demand shocks she acquires in the first period. As a result, investors' ex-ante expected utility improves. Price volatility increases due to more heterogeneous private information, and learning about future demand shocks makes price more informative in the future but less informative now.

In the dynamic model, the front-running channel works in the same direction as the future uncertainty channel and will strengthen the results in the main model. I discuss a few examples on how this model can be applied to different empirical scenarios.

Learning about future demand shocks is important since it breaks the fallacy of composition and improves the investors' ex-ante expected utility. It also extends the concept of news shock and have implications on business cycle. The next version of this chapter will discuss it in more detail.

Chapter 3: From Persistent Disagreement to Agreement: Learning with Revisions of Perception Sets

3.1 Introduction

The world overflows with information that is visual, auditory, olfactory, gustatory, and etc. In order to make sense of all this *external* information, our body has developed a series of processes for sensing and perceiving the outside world. The first step is *sensation*, "the process of detecting external events with sense organs and turning those stimuli into neutral signals" (Krause et al., 2017). The sensory information is then relayed to the brain where the second step, *perception*, occurs. Perception is "the organization, identification, and interpretation of sensory information" (Schacter et al., 2014) in order to form an *internal* representation of the world. As an example, when investors are flipping through business news in the *Wall Street Journal*, the sensory receptors in their eyeballs are registering different patterns of light reflecting off the pages, converting them to neutral signals which are then integrated and processed by the brain into meaningful perception of words. However, humans are not machines; what is conveyed physically on the paper is likely not what readers understand.

According to Bruner and Postman (1949), when encountering new information, we are open to different cues. However, as we process more information, some cues become familiar and thus selective. We search for more cues that confirm the categorization of the target. We also actively ignore and even distort cues that violate our initial perceptions. Our perception becomes more selective and we finally re-create a consistent internal presentation of the new information. In other words, we interpret patterns of stimuli through a *perceptual set*—"a readiness to perceive a stimulus in a particular way" (Weiten, 2001).

An important factor that heavily influences our perception is our expectation (Bruner and

Minturn, 1955), which is characterized by the prior belief in this chapter. Perceptual sets are also affected by motivational forces—people have a tendency to see what they want to see (Balcetis and Dunning, 2006), and the selection of sensory input—people tend to perceive what they focus their attention on (Chun and Wolfe, 2005) and overlook unexpected items and events that do not receive their attention (Simons and Chabris, 1999). Another factor that affects perception sets is *confirmation bias*, that is people seek information that supports their prior beliefs and ignore information that contradicts them (Nickerson, 1998). Confirmation bias is also widely studied in the finance literature.

Motivational factors, selective perception and confirmation bias all vary from person to person; it is akin to "individual fixed effect" and thus summarized in this chapter by the blanket term *personal traits*. These personal traits make perception sets very persistent—it is hard for people to change how they view the world. Nevertheless, it is not completely impossible to achieve that, especially after disastrous events. Those events are so intense that people's attention is drawn to them even if it was placed elsewhere and that contradictions to prior beliefs is too significant to overlook. In other words, disastrous events post *perceptual shocks* that suddenly and greatly change perceptual sets.

Individuals process the same *sensory information* through different perceptual sets and thus form different *perceptual information*, that is different internal representation of the same external inputs, and then combine them with the prior belief to form posterior beliefs[1]. The processes of perception and belief updating are shown in Figure 3.1.

To model dynamic belief updating and jumps of perceptual sets, I introduce the tools that are frequently used in continuous-time finance. Models of utility in stochastic continuous-time settings typically assume that beliefs are represented by a probability measure. For example, Chen et al. (2024) models two competitive agents, one with the objective belief in the sense that her probability measure is consistent with the physical measure, and the other with a probability

[1]In psychology, belief is defined as the mental acceptance of something to be true (Weiten, 2001), which is consistent with its economic interpretation as a probability distribution over a set of possible states.

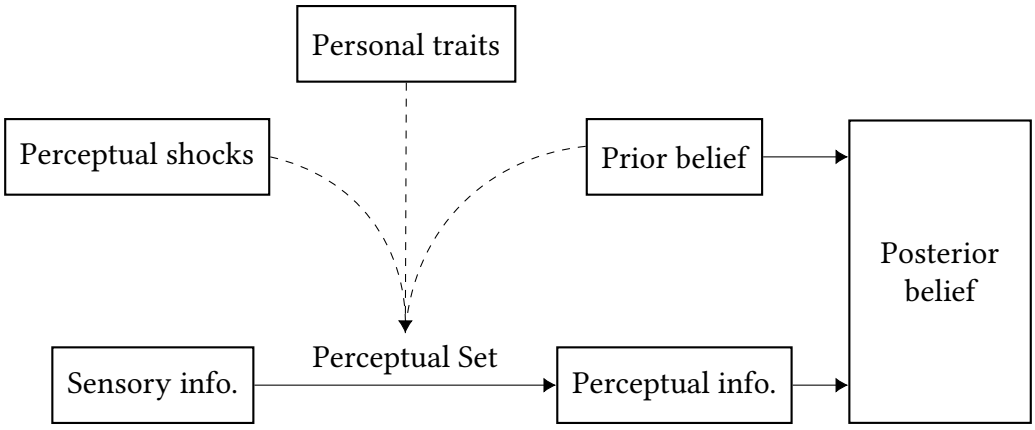


Figure 3.1. Sensory information, perception and belief

I abstract from sensation and directly take sensory information as raw input. It goes through the perception process, during which it is filtered by a subjective perceptual set and converted to *perceptual information*. It is then combined with the prior belief to form the posterior belief. The perceptual set is determined by personal traits (motivational factors, selective perception and confirmation bias), the prior belief and perceptual shocks. Perceptual shocks could make the perceptual set subject to sudden and big changes.

measure that is different from but equivalent to the physical measure. In this chapter, I model perceptual sets as probability measures, while beliefs about the unobservable process are obtained through Kalman filter. Individuals are not aware of the existence of perceptual sets; how the external information is converted to internal information is a blackbox to them. Starting from a common prior, individuals observe a random process $\{P_t\}$ and try to make inferences about a function of P_t and an unobservable random process $\{\theta_t\}$. The inference is de facto done under individuals' subjective probability measures. In the spirit of entrenchment for beliefs[2], I define a notion of *set entrenchment*. A perceptual set with low entrenchment is more subject to revision. Agents are endowed with an initial set entrenchment level. Arrivals of independent perceptual shocks and shocks from P_t will affect entrenchment. The perceptual set is revised when the entrenchment level hits the set revision threshold. Even with common prior and information, individuals have persistently different beliefs. In somes cases, we also observe side-switching.

Another interesting question to think about is to disentangle the effects of new information on the posterior belief versus changes in the posterior belief resulted from perceptual set revision.

[2]In belief contraction, the beliefs with the lowest entrenchment should be the ones that are most readily given up (Hansson, 2017).

The first one is the standard Bayesian updating, while the second one has the flavor of Collin-Dufresne et al. (2016), which implemented parameter learning for model selection. This chapter takes a different angle, in that independent perceptual shocks could lead to perceptual set revision, resulting a different posterior despite of the same old information. New information could be absent in this setting.

This general theoretical framework of subjective perceptual sets could be applied to different areas in economics and finance. We can introduce it to the financial markets to explain excess volatility and trading volume. It also explains why investors initially underreact to certain news but suddenly the market perspective is changed, and why tail events happen more often than common wisdom suggests. On the micro-level, if investors are aware of the existence of the perceptual set but do not know how it is biased, what can they do to improve their utility? And can a social planner devise any policy to increase investors' utility and stablize the market? The framework can also be used to explain belief polarization in political economy and bubbles in household expectations.

Literature There is a growing literature that tries to explain persistent disagreement among agents. The starting point of many studies is a well-known result by Savage (1954). He argued that repeated observation of signals will lead a Bayesian agent to assign probability 1 to a true event almost surely. What we observe, however, is frequently at odds with this standard theoretical framework; persistent disagreement among agents still exists despite of abundant information.

To explain this observation, some research has been done under the rational inattention framework. For example, in Nimark and Sundaresan (2019), agents can choose channels, which are noisy, costly, and have different precisions in different states of the world, to observe the reality. They allocate more precision to those states a priori more likely, which perpetuates differences in beliefs. Kominers et al. (2018) study Bayesian agents for whom computing posterior beliefs is costly; such agents face a tradeoffs between economizing on attention costs and having more accurate beliefs. Even small processing costs can lead to significant departures from the

standard costless processing model.

However, humans are not machines; we process information in different ways than a machine does. Psychologists and cognitive scientists have shown that individuals display biases in their perception of the world, upon which another strand of research in economics is based. In that literature, it is offen assumed that agents do not know certain state of the world *and* the quality of information regarding the state. For example, Cheng and Hsiaw (2022) drew on the *curse of knowledge*, that is individuals who have seen information often fail to ignore it when the context requires using only prior beliefs, to model *pre-screeners* who "double-dip" the data. Pre-screeners firstly form a first-stage belief about credibility based on newly-observed signals, and then use it instead of their prior to weight all signals. This proposed mistake predicts when individuals over- or underreact to new information, depending on the order in which they received previous signals, and explains why information that should objectively mitigate disagreement may amplify it instead. In a similar vein, Koçak (2018) showed that if a signal is informative on multiple dimensions, and if individuals can hold only a limited number of items in working memory, that is after the agent updates the first belief, she has to discard the prior of the now updated parameter, and use its posterior and the new information to update the second parameter, then the order with which beliefs are updated across different dimensions leads to different posteriors. This provides an underlying mechanism of why the curse of knowledge comes into effect.

In contrast, in this chapter individuals do not doubt the credibility of experts but process sensory information in an *unnoticeably* biased way. Individuals do not question the quality of information source, nor do they make active information acquisition. They are given a full set of information, but each individual process information under a different perceptual set, resulting persistent disagreement. The perceptual set is not just a result from experiences or prior information, but could also be biological, e.g. sensitivity to different types of sensory information. Thus individuals can be free of double-dipping the data but still updating information in a biased way.

In Chauvin (2023), a dissonance-reducing learner wishes to agree with his sources of information, and therefore distorts his subjective perception of the arguments before adding them to

the prior. Distortion is achieved through scaling the arguments by real-valued factors. Fryer et al. (2019) modeled "double updating" when agents are faced with ambiguous information. The agent first interpret each signal based on their current belief and then form a posterior on the sequence of interpreted signals. In this chapter, however, belief polarization happens even with unambiguous information. This chapter also extends the idea of subjective perception of parameters/information by introducing a general perceptual set.

3.2 Learning without perceptual bias

Although perceptual sets are generally shaped by personal traits, previous perceptual shocks and prior beliefs, they are very persistent. Once they are formed, the influence they receive from belief updating through combining new information is minimal. Revisions of perceptual sets occur mainly following huge perceptual shocks, which are unforeseeable rare events. In the absence of perceptual set revisions, the learning process can be modelled by optimal linear filtering. The agents are not aware of the perception process and bias, and the perceptual shocks are exogenous, unforeseeable and rare. Once perceptual shocks hit and the perceptual set is revised, the agent repeats the learning process and apply the optimal linear filtering under her new perceptual set. In this section, I introduce the general framework of optimal linear filtering, which is joined piecewise once the perceptual set revision is induced exogenously.

Let (V_t, P_t), $t \in [0, T]$ be a two-dimensional random process, with V_t being the unobservable state of the world that is to be learned, and P_t is observable. More specifically, P_t and V_t are related by

$$\mathrm{d}P_t = A_t V_t \,\mathrm{d}t + B_t(P_t)\,\mathrm{d}Z_t = \left[A_t V_t - \eta_t^i\right]\mathrm{d}t + B_t(P_t)\,\mathrm{d}Z_t^i, \tag{3.1}$$

where A_t is deterministic at time t, and $B_t(P_t)$ is $\mathcal{F}_t^P$-measurable. We are interested in the optimal[3] estimator for V_t. Assume that V_t can be represented by

$$\mathrm{d}V_t = a_t V_t \,\mathrm{d}t + b_t \,\mathrm{d}Z_t^f \,. \tag{3.2}$$

[3]It is optimal in the minimum mean-square error sense.

where H_t is a random process with $\mathbb{P}\left(\int_0^T |a_s V_t| \, \mathrm{d}t < \infty\right) = 1$ and $\int_0^T a_t^2 \mathbb{E}^i \left[V_t^2\right] \mathrm{d}t < \infty$, b_t is a random process with $\mathbb{P}\left(\int_0^T |b_t| \, \mathrm{d}t < \infty\right) = 1$, and Z_t^f is a standard Wiener process over $(\Omega, \mathcal{F}, \mathbb{P})$, independent from Z_t. Let $\mathcal{F}_t^P = \{P_t; 0 \leq s \leq t\}$ denote the information filtration generated by the history of signal realizations, which is publicly observable. Agents also have a common prior $V_0 \mid \mathcal{F}_0^P \sim N(m_0, \gamma_0)$. The conditional mean of V_t from agent i's perspective, which is summarized by the following proposition.

PROPOSITION 22. *Let* $m_t^i \equiv \mathbb{E}^i \left[V_t \mid \mathcal{F}_t^P\right]$. *Then*

$$\mathrm{d}m_t^i = a_t m_t^i \, \mathrm{d}t + A_t \gamma_t^i \, \mathrm{d}Z_t^{Pi}, \tag{3.3}$$

where $\mathrm{d}Z_t^{Pi} = \frac{1}{B_t(P_t)} \left[\mathrm{d}P_t - (A_t - \eta_t^i) m_t^i \, \mathrm{d}t\right]$.

See the appendix for proofs and the framework for a more general learning objective. By applying $h_t = V_t^2$ to Theorem 4 and Ito's Lemma to Proposition 28, we can get the tracking error (conditional variance) of V_t from agent i's perspective, which is summarized by the following proposition.

PROPOSITION 23. *Let* $\gamma_t^i \equiv \mathbb{E}^i \left[(V_t - m_t^i)^2 \mid \mathcal{F}_t^P\right]$. *Then (3.3) and*

$$m_t^i = m_0 + \int_0^t \left[2a_s(\gamma_s^i + (m_s^i)^2) + b_s^2\right] \mathrm{d}s + \int_0^t A_s \left[\left(m_s^i\right)^3 + 2m_s^i \gamma_s^i\right] \mathrm{d}Z_s^{Pi}, \tag{3.4}$$

$$\left(m_t^i\right)^2 = m_0^2 + \int_0^t \left[2a_s \left(m_s^i\right)^2 + A_s^2 \left(\gamma_s^i\right)^2\right] \mathrm{d}s + 2 \int_0^t A_s m_s^i \gamma_s^i \, \mathrm{d}Z_s^{Pi}, \tag{3.5}$$

$$\gamma_t^i = m_t^i - \left(m_t^i\right)^2 \tag{3.6}$$

jointly characterize γ_t^i.

3.3 Learning under perceptual bias

3.3.1 Perceptual sets

Consider the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Assumes that there is a continuum of measure one ex-ante identical behavioral agents who live over a finite time interval $[0, T]$. Agent i has a *subjective* perceptual set, which is implemented by probability measure $\mathbb{P}^i$, different from but equivalent to the *objective* perceptual set, that is the physical measure $\mathbb{P}$. The two measures are related by

$$\mathrm{d}Z_t^i = \mathrm{d}Z_t + \eta_t^i \, \mathrm{d}t, \tag{3.7}$$

where η_0^i is the initial perceptual bias. It only changes when set revision is triggered; otherwise it remains constant. As a benchmark, I refer to a comparable Bayesian agent who always has an objective perceptual set, that is she always process sensory information perfectly and thus not subject to set revision.

3.3.2 Revision of perceptual sets

Simplistic modelling of a rare disaster

At a fixed time t, agent i receive an exogenous perceptual shock, which changes their perceptual bias η_0^i by a multiplier q, $q \geq 0, q \neq 1$. When $q = 0$, agent i is no longer perceptually biased. When $0 < q < 1$, she becomes less perceptually biased. When $q > 1$, she becomes more perceptually biased.

This simplistic modelling of perceptual set revision is suitable for characterizing a rare disaster event. To see a clear-cut effect of a disaster, the smaller effects of non-perceptual shock factors that can affect perceptual set revision are normalized to zero. The large magnitude of the unpredicted disaster rationalizes the setting of a perceptual set revision.

Set entrenchment

More general scenarios allow perceptual shocks, personal traits and prior beliefs to jointly affect perceptual sets. To achieve that, I introduce set entrenchment level into the framework. At time $t = 0$, agent i has an initial set entrenchment level E_0^i. The observed observation of P_t not only take part in belief updating as the source of new information, but also affect the set entrenchment level. Normally, the latter effect is small and could be positive or negative. The change in set entrenchment level is predominantly affected by independent perceptual shocks, which always reduce the set entrenchment level, if occur. Mathematically, let E_t^i be the set entrenchment level for agent i at time t, follows: then

$$E_t^i = E_0^i + \alpha^i \int B_t(P_t)\,\mathrm{d}Z_t - \zeta^i \int k_t \beta_t\,\mathrm{d}t,$$

where ζ^i is the sensitivity to perceptual shocks, k_t is a Bernoulli random variable which takes the value 1 with probability p and the value 0 with probability $1 - p$, and β_t is an independent random variable that has a beta distribution. k_t governs whether or not a perceptual shock would occur, and β_t determines the intensity of the shock if it occurs. When the entrenchment level drops to zero, the agent revises her perceptual set towards the objective view by $\eta_t^i = (1 - \beta_t)\eta_{t^-}^i$. Perceptual shocks always reduce set entrenchment; regardless of the sign of the initial bias, perceptual shocks shift the bias towards zero.

3.3.3 Results

In this section, I first present the special case where the state of the world does not change, and then a more general case with a stochastic world state.

Special case: time-invariant state of the world

Consider the special case where the unknown state of the world V is time-invariant, and the observable P_t is simplified to $\mathrm{d}P_t = V_t\,\mathrm{d}t + \sigma\,\mathrm{d}Z_t$. Individuals observe $\{P_s\}, t = 1, \cdots, t$ to

make inference for V. Without the revision of perceptual shocks, the conditional expectation and variance of the state variable can be solved from equations 3.3-3.6, and the results are shown as follows.

PROPOSITION 24. *In the special case where $A_t = 1$, $B_t(P_t) = \sigma$, and the state of the world V is time-invariant, consider the Bayesian agent and the behavioral agent in the absence of perceptual set revision.*

1) The time-t conditional mean of V from the Bayesian agent's view is

$$m_t = \frac{m_0\sigma^2 + P_0\gamma_0}{\gamma_0 t + \sigma^2} + V\frac{\gamma_0 t}{\gamma_0 t + \sigma^2} + \frac{\gamma_0\sigma}{\gamma_0 t + \sigma^2}\int_0^t \mathrm{d}Z_t\,. \tag{3.8}$$

2) The behavioral agent's view differs by

$$m_t^i = m_t + \eta^i\frac{\gamma_0 t}{\gamma_0 t + \sigma^2}. \tag{3.9}$$

3) All agents agree on the conditional variance of the underlying state:

$$\gamma_t^i = \gamma_t = \frac{\gamma_0\sigma^2}{\gamma_t t + \sigma^2}. \tag{3.10}$$

4) The Bayesian agent eventually learns about the true state of the world V, while behavioral agent i eventually thinks the state of the world to be $V + \eta^i$.

In this special case, all agents, regardless of their perceptual sets, share the same conditional variance of the state variable, which is decreasing in t. The conditional expectations, however, vary with the perceptual bias. At a fixed time, perceptual shocks are imposed. The conditional moments can be calculated piecewise. After the perceptual shocks, agents still agree with the conditional variance of the state variable, the conditional expectation still varies by the remaining perceive bias.

If the intensity of or agents' sensitivity to perceptual shocks is large enough, revisions of perceptual sets occur for all agents. Further assume that the perceptual bias is completely corrected once perceptual set revision occurs, then the conditional expectation of the state variable viewed by all agents converge to the true value of the state variable. As an illustration, Figure 3.2 plots the empirical distribution of the conditional expectation of the state variable and the empirical distribution of the entrenchment level at time T. In this special case, the true value of V is set to be 1, and the default entrenchment level is normalized to 1. Tests do not reject the normal distribution hypothesis for either panel.

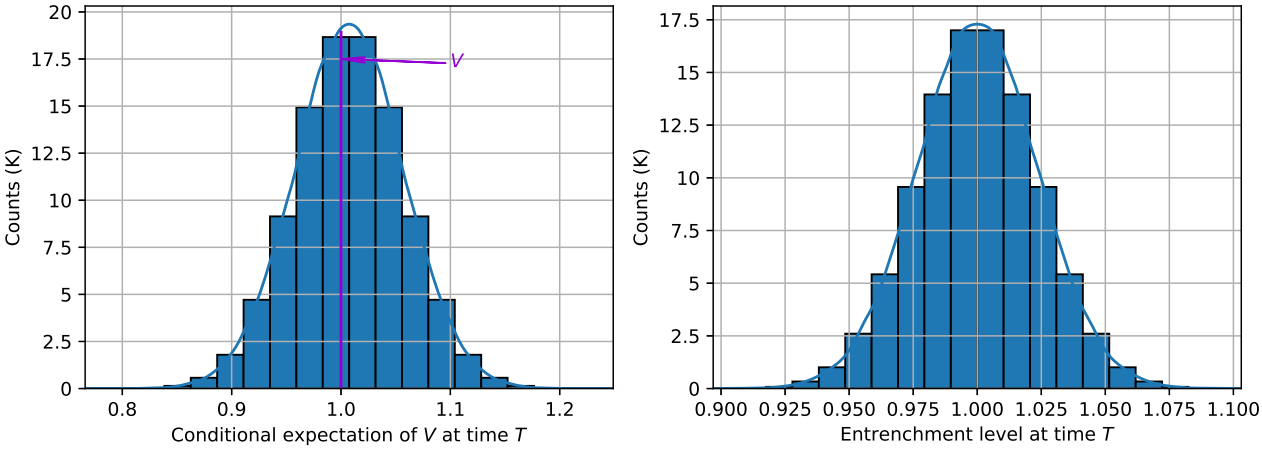


Figure 3.2. Distributions of conditional expectations of V and entrenchment levels

The left panel plots the empirical distribution of $m_T^i = \mathbb{E}^i[V \mid \mathcal{F}_T^\xi]$ with $V = 1$ and $\eta^i \sim N(0,1)$. The true value $V = 1$ is also noted on the plot. The right panel plots the empirical distribution of E_T^i with the default entrenchment level, which is the level to which entrenchment adjusts after a perceptual set revision, normalized to 1. Tests do not reject the normal distribution hypothesis for either panel.

Figure 3.3 plots the random paths of the conditional expectation of the underlying state. Absent of perceptual shocks, the conditional expectations change smoothly from updating with new information. When exogenous perceptual shocks are introduced at a fixed time ($t = 10$), perceptual sets are revised and thus conditional expectations experience jumps. The magnitude of the jumps reflects the effects of perceptual set revision, isolated from any incoming new information. Assume the sensitivity to perceptual shocks and the intensity of perceptual shocks are large enough, then the perceptual bias is immediately eliminated. The conditional expectations of the underlying state by all agents converge to its true value.

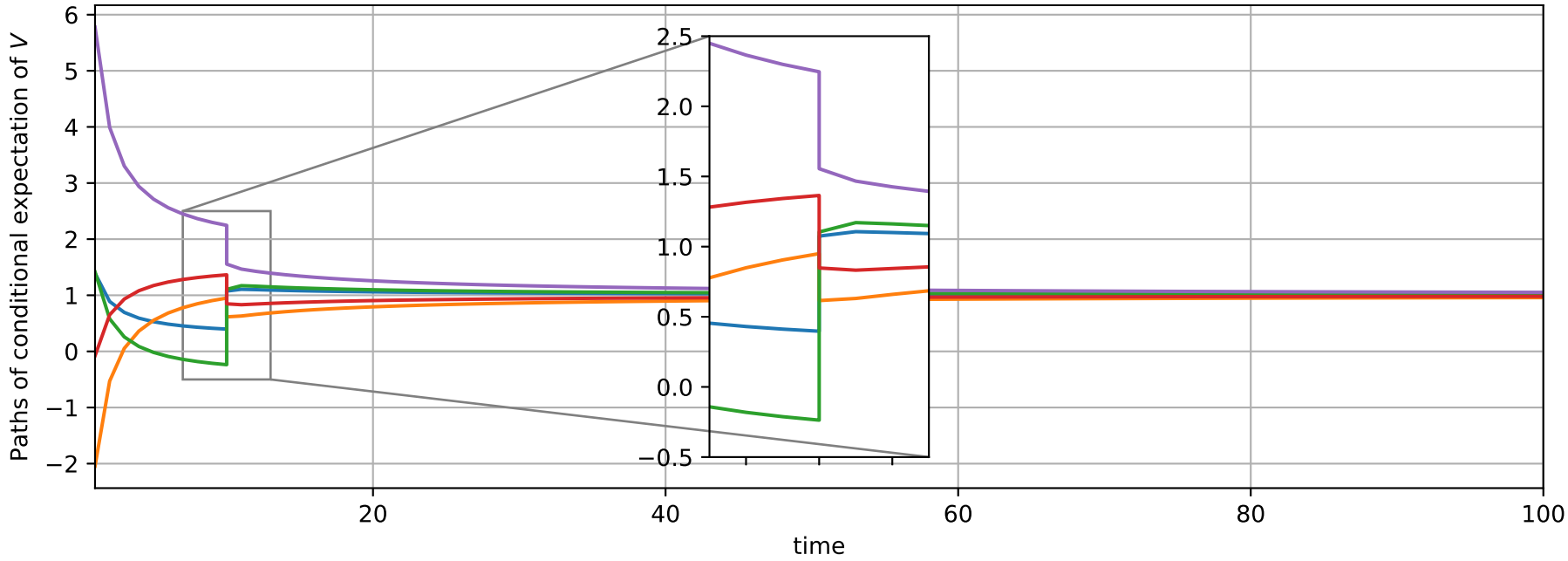


Figure 3.3. Paths of conditional expectations of θ with perceptual set revisions

Five random paths of m_t^i, the conditional expectation of the underlying state, are plotted. When exogenous perceptual shocks are introduced at time $t = 10$, perceptual sets are revised and thus conditional expectations experience jumps. The magnitude of the jumps reflects the effects of perceptual set revision, isolated from any incoming new information.

General case

Generally, the unknown state of the world V_t is a random process, as specified by (3.2). Individuals observe $\{P_s\}, s = 1, \cdots, t$ to make inference for V_t. In this part, I discuss a specification of $(V_t, P_t), t \in [0, T]$ with:

$$\mathrm{d}V_t = rV_t\,\mathrm{d}t + \sigma_v V_t\,\mathrm{d}Z_t^f, \tag{3.11}$$

$$\mathrm{d}P_t = A_t V_t\,\mathrm{d}t + B_t\sigma_p P_t\,\mathrm{d}Z_t, \tag{3.12}$$

where r and σ_v are unknown parameters, σ_p is a known function of P_t.

Let $\mathcal{F}_t^P = \{P_s; 0 \le s \le t\}$ denote the information filtration generated by the history of equilibrium price which are publicly observable by every investor. At time t, from investor i's perspective, the asset value V_t has conditional mean $m_t^i = \tilde{\mathbb{E}}^i\left[V_t \mid \mathcal{F}_t^P\right]$ and variance $\gamma_t^i = \tilde{\mathbb{E}}^i\left[\left(V_t - m_t^i\right)^2 \mid \mathcal{F}_t^P\right]$. The innovation process specifies to $\mathrm{d}Z_t^i = \frac{1}{\sigma}\left[P_t - (1-\eta_t^i)\,m_t^i\,\mathrm{d}t\right]$.

Figure 3.4 shows three random paths of the entrenchment level. Perceptual shocks are introduced at $t = 10$. Regardless of the entrenchment level at that time, it drops to 0 immediately and thus revision of perceptual sets occurs. The entrenchment level is then reset to the default value 1. When there are no perceptual shocks, the changes in the entrenchment level are very minor, insufficient to accumulate for other revisions of perceptual sets.

Figure 3.5 plots two sets of perceptual bias and entrenchment level paths. The initial (default) entrenchment level is normalized to 1. The perceptual bias is only revised when the entrenchment level drops to 0, after which the entrenchment level will reset to its default value 1. The entrenchment level changes slightly over time through belief updating but normally these small changes do not result in a revision of the perceptual set. In both panels, only at $t = 10$ when an exogenous perceptual shock is introduced does the perceptual set undergoes revision.

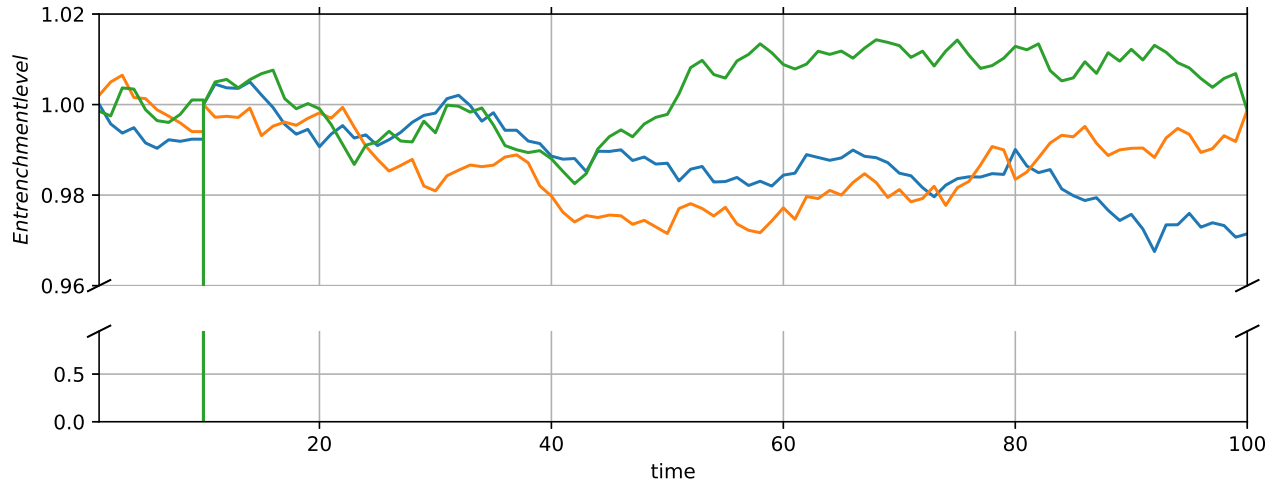


Figure 3.4. Paths of set entrenchment levels

Three random paths of the entrenchment level E_t^i are presented in this figure. Perceptual shocks are introduced at $t = 10$. Regardless of the entrenchment level at that time, it drops to 0 immediately and thus revision of perceptual sets occurs. The entrenchment level is then reset to the default value 1. When there are no perceptual shocks, the changes in the entrenchment level are very minor, insufficient to accumulate for other revisions of perceptual sets.

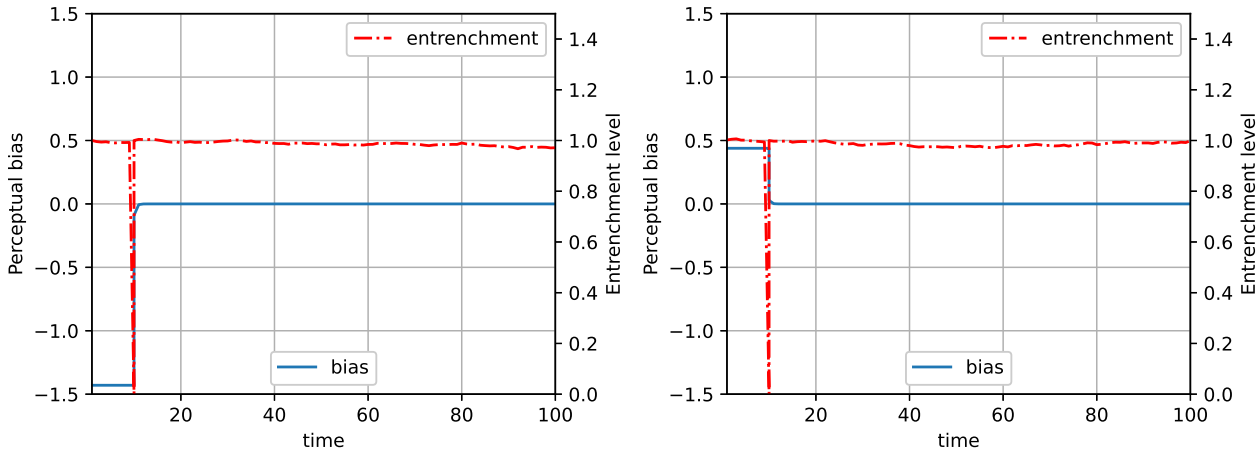


Figure 3.5. Examples of perceptual bias and entrenchment paths

The two panels plots two sets of perceptual bias and entrenchment level paths. The initial (default) entrenchment level is normalized to 1. The perceptual bias is only revised when the entrenchment level drops to 0, after which the entrenchment level will reset to its default value 1. The entrenchment level changes slightly over time through belief updating but normally these small changes do not result in a revision of the perceptual set. In both panels, only at $t = 10$ when an exogenous perceptual shock is introduced does the perceptual set undergoes revision.

Experience and set entrenchment

It is reasonable to think that experience affects entrenchment reduction. On the one hand, more experienced investors could initially hold their set more affirmatively. In other words, they will have higher initial set entrenchment level E_0^i. This is in a similar vein to Jovanovic and Nyarko (1996), which argued that an agent may be so skilled at a technology that he will refuse to switch to a better but unfamiliar one. On the other hand, more experienced investors could be better skilled at dealing with perceptual shocks, and thus they reduce set entrenchment at a higher rate, that is they have a larger ζ^i.

Overwhelming information and sensitivity to set shock

When the private signal about dividend growth rate is more precise, that is when σ_ξ is small, investors have more information to take in. The effects of having more information are two-fold. The good side is that more information helps investors make more informed portfolio choices. However, there is also a bad side. When new information becomes too overwhelming, investors become less sensitive to perceptual shocks, thus less likely to revise their sets.

3.4 Connecting with the financial markets

The modelling of learning with entrenchment sets can be connected with the financial markets. In Chen et al. (2024), investors "agree to disagree"; belief is modeled as a random process, while information and learning are not explicitly modeled. Chen and Epstein (2002) discussed risks and asset returns when an investor is using the wrong probability measure. With a continuous-time intertemporal version of *multiple-priors utility*, their model delivers restrictions on excess returns that admit interpretations reflecting a premium for risk and a separate premium for ambiguity. This chapter takes a different perspective in that I model information and learning explicitly, and distinguish beliefs and perceptual sets, where the latter are entrenched and hard to change.

3.4.1 Market setup

The perceptual bias model is readily re-interpretable as investors having perceptual biases towards the underlying value of the risky asset in the financial markets. Consider a continuous-time endowment economy over a finite time interval $[0, T]$. Investors have the same perceptual set structure as in the previous sections. In this section, I introduce assets into the model and specify the unknown state of the world and signal structures.

Assume that there is a riskless bond in the market with a constant risk-free rate $r > 0$. There is 1 share of aggregate stock, with underlying value V_t and price P_t, specified in the same way as (3.11) and (3.12). Z_t^f is the fundamental shock to the risky asset, which has a direct impact on the underlying value of the asset. Z_t represents the demand shock, which is orthogonal to the fundamental shock and only affects the asset price, not its value.

Investor i has initial wealth W_0 and can invest in both the risk-free asset and the stock. Her wealth dynamics is

$$\mathrm{d}W_t = \left(rW_t^i - C_t^i\right)\mathrm{d}t + \theta_t^i P_t(\mu - r)\,\mathrm{d}t + \theta_t^i P_t \sigma\,\mathrm{d}Z_t,$$

where the first term is the return earned on W_t^i if investor i invest solely in the risk-free market, and the second term is the total excess return earned from the θ_t shares of the stock. Equivalently,

$$\mathrm{d}W_t^i = \left[\left(r + \pi_t^i(\mu - r)\right)W_t^i - C_t^i\right]\mathrm{d}t + \pi_t^i \sigma W_t^i\,\mathrm{d}Z_t,$$

where $\pi_t^i = \frac{\theta_t^i P_t}{W_t^i}$ is the proportion of wealth invested in shares of stock.

Investor i has a time-additive separable utility, given by $\tilde{\mathbb{E}}_0\left[\int_0^T e^{-\rho s}u(C_s^i)\,\mathrm{d}s + e^{-\rho T}U(W_T^i)\right]$, where C is the consumption, ρ is investor's subjective discount rate, $u(C)$ is utility from the consumption flow, and $U(W_T^i)$ is the utility over terminal wealth. At time t, investor i's objective is to seek the optimal investment and consumption strategy to maximize the expected utility

given by

$$F\left(t, W_t\right) = \max_{\pi, C} \tilde{\mathbb{E}}^i \left[\int_t^T e^{-\rho(s-t)} u(C_s^i)\, \mathrm{d}s + e^{-\rho(T-t)} U(W_T^i) \mid \mathcal{F}_t^{\xi}, P_t \right].$$

Notice that when investor i makes decisions, she take expectations under the measure $\mathbb{P}^i$, not the risk-neutral measure $\mathbb{P}$, since she is unaware of the existence of perceptual bias.

The value function satisfies

$$\frac{\partial F}{\partial t} + \mu_W \frac{\partial F}{\partial t} + \frac{1}{2}\sigma_W^2 \frac{\partial^2 F}{\partial W^2} - \rho F + u(C^i) = 0 \tag{3.13}$$

3.4.2 Results

In the perceptual mechanism, there are two opposite forces that compete for overreaction and underreaction to sensory information. On the one hand, if the initial entrenchment level is high or the sensitivity to perceptual shocks is low, the entrenchment level would not drop to 0, which leads to an underreaction to shocks. On the other hand, overreaction could happen when perceptual sets are revised if investors have accumulated enough old sensory information. When investors revise their perceptual sets, previous sensory information would go through the new perceptual set to form revised perceptual information, which is then used to correct their posterior beliefs and previous underreaction to news. This is different from belief updating on new information. It is the effects of perceptual set revision with old information.

The trading volume can be obtained by aggregating the absolute value of $\mathrm{d}\theta_t^i$. When overreaction happens as a result of the perceptual set revision, the trading volume increases. This is because perceptual set revision creates a sudden jump in investors's expectations about the risky asset and thus investors need to adjust their portfolios on a bigger scale. Figure 3.6 plots the paths of the price and value of the risky asset. The upper panel plots an example of overreaction to perceptual shocks. Perceptual shocks are introduced at $t = 10$, indicated by a vertical dashed line. If the intensity of or the sensitivity to perceptual shocks are sufficiently large, perceptual

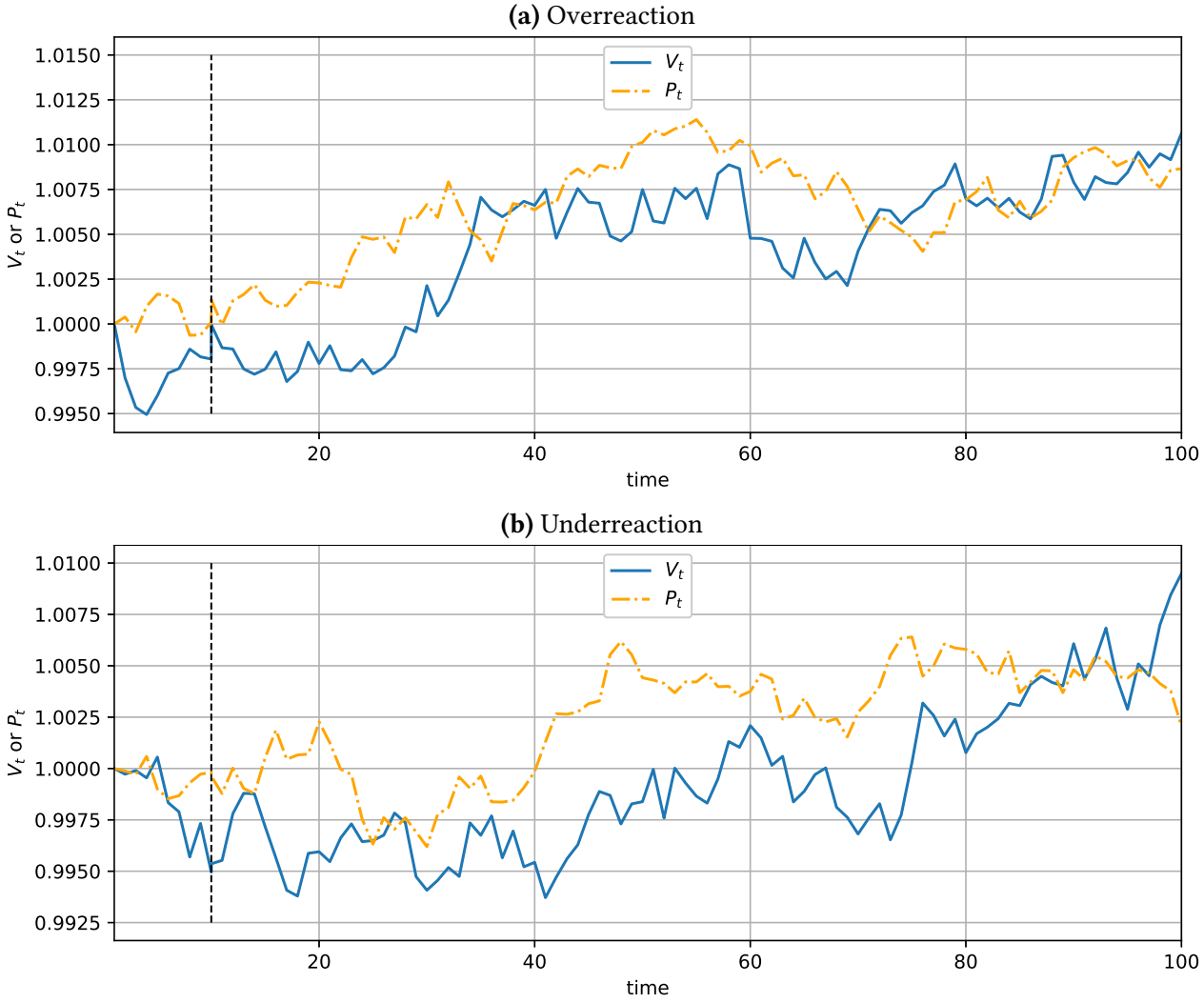


Figure 3.6. Evolutions of the price and value of the risky asset

This figure plots the paths of the price and value of the risky asset. The upper panel plots an example of overreaction to perceptual shocks. Perceptual shocks are introduced at $t = 10$, indicated by a vertical dashed line. If the intensity of or the sensitivity to perceptual shocks are sufficiently large, perceptual shocks induce a sudden jump in investors' conditional expectations of the risky asset. As a result, an overreaction to new information is reflected in the price. The lower panel plots an example of underreaction to perceptual shocks.

shocks induce a sudden jump in investors' conditional expectations of the risky asset. As a result, an overreaction to new information is reflected in the price. The lower panel plots an example of underreaction to perceptual shocks.

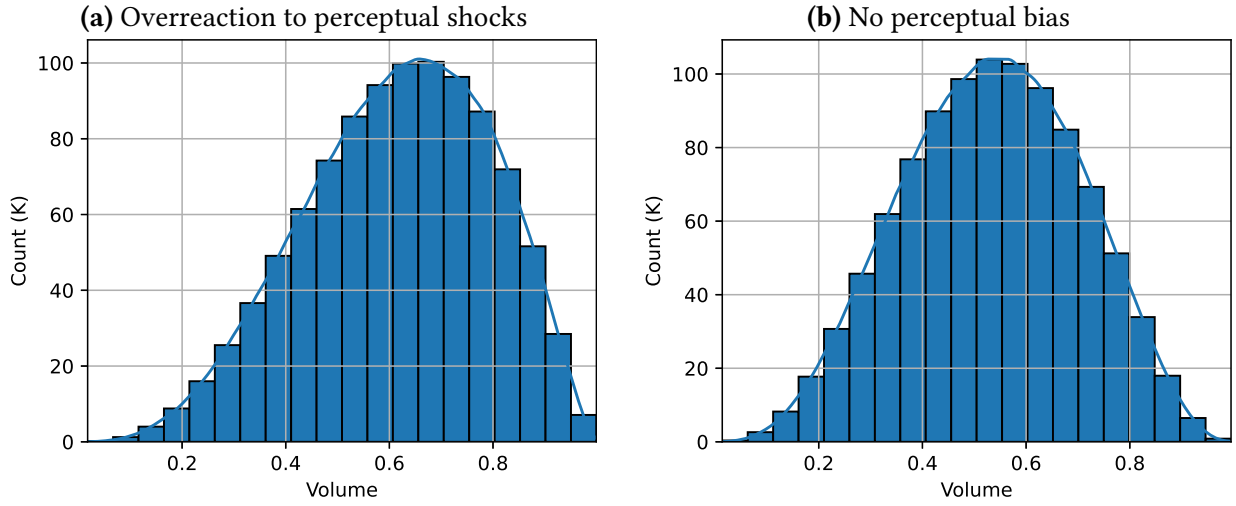


Figure 3.7. Distributions of trading volumes

This figure plots the histograms of the total trading volumes for the risky asset. In the left panel, perceptual shocks hit and result in an overreaction via perceptual set revisions. As investors adjust their perceptual sets, posterior beliefs are corrected under the new perceptual sets using all old and new information. The average total trading volume of the risky asset thus increases. The right panel plots the trading volumes in the case where there is no perceptual bias.

3.5 Conclusion

This chapter zooms in on the learning process. The sensation process transforms external inputs into sensory information, which is subsequently directed through the perceptual process to generate perceptual information. Perceptual information is then used to update prior beliefs to posterior beliefs. This chapter highlights the perception process and models it by perceptual sets. Agents have biased perceptual sets. Its entrenched nature is modelled by set entrenchment; once formed, the entrenchment level is hard to change. When perceptual shocks occur, however, the enacted perceptual sets would suddenly change and the entrenchment level would be reset to the default value, that is agents become entrenched with their new perceptual sets.

With perceptual bias, even after sufficient information is assimilated, there is still persistent

disagreement among agents. Revision of perceptual sets induced by perceptual shocks shifts disagreement to agreement. The magnitude and speed of shift are affected by agents' experience and sensitivity to perceptual shocks. The entrenched nature of perceptul sets and sudden set revision in the face of huge shocks allow room for both overreaction and underreaction to information. When overreaction happens, there is an increase in the trading volume of the risky asset.

The perceptual set mechanism can be applied to many other scenarios, for example, belief polarization in political economy. Traditionally, political polarization affects finance through two channels. Through the partisan utility channel, partisans obtain utility through from engaging with their fellow party members or from owning specific financial securities. My perceptual set mechanism is close to the partisan beliefs channel. Partisans may hold different beliefs regarding the state of the economy, expected risk-adjusted returns in the cross section of stocks, or the ability of in-group versus out-group partisans. The notion of a *partisan perceptual screen* was introduced in a seminal study by Campbell et al. (1960). Heterogeneity in beliefs could arise from differences in information environments (e.g., via partisan media), differences in attention, or differences in individuals' models of the world.

Bibliography

Efstathios Avdis. Information tradeoffs in dynamic financial markets. Journal of Financial Economics, 122(3):568–584, 2016.

Emily Balcetis and David Dunning. See what you want to see: motivational influences on visual perception. Journal of Personality and Social Psychology, 91(4):612, 2006.

Suleyman Basak. A general equilibrium model of portfolio insurance. The Review of Financial Studies, 8(4):1059–1090, 1995.

Paul Beaudry and Franck Portier. News-driven business cycles: Insights and challenges. Journal of Economic Literature, 52(4):993–1074, 2014.

Geert Bekaert and Campbell R Harvey. Foreign speculators and emerging equity markets. The Journal of Finance, 55(2):565–613, 2000.

Victor L Bernard and Jacob K Thomas. Post-earnings-announcement drift: delayed price response or risk premium? Journal of Accounting Research, 27:1–36, 1989.

Victor L Bernard and Jacob K Thomas. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. Journal of Accounting and Economics, 13(4): 305–340, 1990.

Jerome S Bruner and A Leigh Minturn. Perceptual identification and perceptual organization. The Journal of General Psychology, 53(1):21–28, 1955.

Jerome S Bruner and Leo Postman. On the perception of incongruity: A paradigm. Journal of Personality, 18(2):206–223, 1949.

Markus K Brunnermeier and Markus Konrad Brunnermeier. Asset pricing under asymmetric information: Bubbles, crashes, technical analysis, and herding. Oxford University Press on Demand, 2001.

Zhifeng Cai. Essays on macroeconomics and finance. PhD thesis, University of Minnesota, 2017.

A Campbell, Converse PE, Miller WE, and Stokes DE. The American Voter. Wiley, New York, 1960.

Jennifer N Carpenter, Fangzhou Lu, and Robert F Whitelaw. The real value of china' s stock market. Journal of Financial Economics, 139(3):679–696, 2021.

Marc K Chan and Simon Kwok. Risk-sharing, market imperfections, asset prices: Evidence from china' s stock market liberalization. Journal of Banking & Finance, 84:166–187, 2017.

Anusha Chari and Peter Blair Henry. Risk sharing and asset prices: evidence from a natural experiment. The Journal of Finance, 59(3):1295–1324, 2004.

Kyle Chauvin. Belief updating with dissonance reduction. 2023.

Hui Chen, Anton Petukhov, Jiang Wang, and Hao Xing. The dark side of circuit breakers. The Journal of Finance, 79(2):1405–1455, 2024.

Zengjing Chen and Larry Epstein. Ambiguity, risk, and asset returns in continuous time. Econometrica, 70(4):1403–1443, 2002.

Haw Cheng and Alice Hsiaw. Distrust in experts and the origins of disagreement. Journal of Economic Theory, 200:105401, 2022.

Hyuk Choe, Bong-Chan Kho, and René M Stulz. Do domestic investors have an edge? the trading experience of foreign investors in korea. The Review of Financial Studies, 18(3):795–829, 2005.

Tarun Chordia and Lakshmanan Shivakumar. Inflation illusion and post-earnings-announcement drift. Journal of Accounting Research, 43(4):521–556, 2005.

Tarun Chordia and Avanidhar Subrahmanyam. Order imbalance and individual stock returns: Theory and evidence. Journal of Financial Economics, 72(3):485–518, 2004.

Marvin M Chun and Jeremy M Wolfe. Visual attention. Blackwell handbook of sensation and perception, page 272–310, 2005.

Pierre Collin-Dufresne, Michael Johannes, and Lars A Lochstoer. Parameter learning in general equilibrium: The asset pricing implications. American Economic Review, 106(3):664–98, 2016.

John C Cox and Chi-fu Huang. Optimal consumption and portfolio policies when asset prices follow a diffusion process. Journal of economic theory, 49(1):33–83, 1989.

Douglas W Diamond and Robert E Verrecchia. Information aggregation in a noisy rational expectations economy. Journal of Financial Economics, 9(3):221–235, 1981.

Tomáš Dvořák. Do domestic investors have an information advantage? evidence from indonesia. The Journal of Finance, 60(2):817–839, 2005.

Eugene F Fama. Efficient capital markets: Ii. The Journal of Finance, 46(5):1575–1617, 1991.

Maryam Farboodi and Laura Veldkamp. Long-run growth of financial data technology. American Economic Review, 110(8), 2020.

Jr Fryer, Roland G, Philipp Harms, and Matthew O Jackson. Updating beliefs when evidence is open to interpretation: Implications for bias and polarization. Journal of the European Economic Association, 17(5):1470–1501, 10 2019. ISSN 1542-4766. doi: 10.1093/jeea/jvy025. URL https://doi.org/10.1093/jeea/jvy025.

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

Chao Gao and Wen Zeng. Near is dear: Remoteness, soft information, and stock returns. Soft Information, and Stock Returns (June 01, 2022), 2022.

Itay Goldstein and Liyan Yang. Information diversity and complementarities in trading and information acquisition. The Journal of Finance, 70(4):1723–1765, 2015.

Itay Goldstein and Liyan Yang. Good disclosure, bad disclosure. Journal of Financial Economics, 131(1):118–138, 2019.

John M Griffin, Patrick J Kelly, and Federico Nardari. Do market efficiency measures yield correct inferences? a comparison of developed and emerging markets. The Review of Financial Studies, 23(8):3225–3277, 2010.

Mark Grinblatt and Matti Keloharju. The investment behavior and performance of various investor types: a study of finland's unique data set. Journal of Financial Economics, 55(1):43–67, 2000.

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

Bing Han, Ya Tang, and Liyan Yang. Public information and uninformed trading: Implications for market liquidity and price efficiency. Journal of Economic Theory, 163:604–643, 2016.

John RM Hand. A test of the extended functional fixation hypothesis. Accounting Review, page 740–763, 1990.

Sven Hansson. Stanford encyclopedia of philosophy: Logic of belief revision. `https://plato.stanford.edu/entries/logic-belief-revision/`, 2017. [Online; accessed 5-November-2019].

Jeffrey H Harris and Mohsen Saad. The sound of silence. Financial Review, 49(2):203–230, 2014.

Joel Hasbrouck and Gideon Saar. Low-latency trading. Journal of Financial Markets, 16(4):646–679, 2013.

Harald Hau. Location matters: An examination of trading profits. The Journal of Finance, 56(5): 1959–1983, 2001.

Frederick A Hayek. The use of knowledge in society. In Knowledge management and organizational design, page 7–15. Routledge, 2009.

Martin F. Hellwig. On the aggregation of information in competitive markets. Journal of Economic Theory, 22(3):477–498, 1980.

Peter Blair Henry. Stock market liberalization, economic reform, and emerging market equity prices. The Journal of Finance, 55(2):529–564, 2000.

Peter Blair Henry. Capital-account liberalization, the cost of capital, and economic growth. American Economic Review, 93(2):91–96, 2003.

Nicholas Hirschey. Do high-frequency traders anticipate buying and selling pressure? Technical report, Available at SSRN 2238516, 2018.

Craig W Holden and Avanidhar Subrahmanyam. News events, information acquisition, and serial correlation. The Journal of Business, 75(1):1–32, 2002.

John Holland. Private voluntary disclosure, financial intermediation and market efficiency. Journal of Business Finance & Accounting, 25(1-2):29–68, 1998.

Nir Jaimovich and Sergio Rebelo. Can news about the future drive the business cycle? American Economic Review, 99(4):1097–1118, 2009.

Charles M Jones, Donghui Shi, Xiaoyan Zhang, and Xinran Zhang. Retail trading and return predictability in china. Journal of Financial and Quantitative Analysis, 60(1):68–104, 2025.

Boyan Jovanovic and Glenn M MacDonald. The life cycle of a competitive industry. Journal of Political Economy, 102(2):322–347, 1994.

Boyan Jovanovic and Albert J Menkveld. Dispersion and skewness of bid prices. Technical report, Available at SSRN 2463066, 2017.

Boyan Jovanovic and Yaw Nyarko. Learning by doing and the choice of technology. Econometrica, 64(6):1299–1310, 1996.

Marcin Kacperczyk, Stijn Van Nieuwerburgh, and Laura Veldkamp. A rational theory of mutual funds' attention allocation. Econometrica, 84(2):571–626, 2016.

Ioannis Karatzas, John P Lehoczky, Suresh P Sethi, and Steven E Shreve. Explicit solution of a general consumption/investment problem. Mathematics of Operations Research, 11(2):261–294, 1986.

E Han Kim and Vijay Singal. The fear of globalizing capital markets. Emerging Markets Review, 1(3):183–198, 2000.

Scott Duke Kominers, Xiaosheng Mu, and Alexander Peysakhovich. Paying (for) attention: The impact of information processing costs on bayesian inference. Available at SSRN 2857978, 2018.

Sri Prakash Kothari. Capital markets research in accounting. Journal of Accounting and Economics, 31(1-3):105–231, 2001.

Korhan Koçak. Sequential updating: A behavioral model of belief change. Technical report, Working Paper, 2018.

Mark Krause, Daniel Corts, Stephen Smith, and Dan Dolderman. Psychological Science. Pearson Cananda, Don Mills, second edition, 2017.

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

José María Liberti and Mitchell A Petersen. Information: Hard and soft. Review of Corporate Finance Studies, 8(1):1–41, 2019.

Robert S Liptser and Albert N Shiryaev. Statistics of random processes: I. General theory, volume 1. Springer Science & Business Media, 2001.

Hong Liu and Yajun Wang. Market making with asymmetric information and inventory risk. Journal of Economic Theory, 163:73–109, 2016.

Carolina Manzano and Xavier Vives. Public and private learning from prices, strategic substitutability and complementarity, and equilibrium multiplicity. Journal of Mathematical Economics, 47(3):346–369, 2011.

Paul Marmora and Oleg Rytchkov. Learning about noise. Journal of Banking & Finance, 89:209–224, 2018.

Luis Angel Medrano and Xavier Vives. Regulating insider trading when investment matters. Review of Finance, 8(2):199–277, 2004.

Albert J Menkveld. The economics of high-frequency trading: Taking stock. Annual Review of Financial Economics, 8:1–24, 2016.

Robert Merton. An intertemporal capital asset pricing model. Econometrica, 41(5):867–887, 1973.

Robert C Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. The review of Economics and Statistics, page 247–257, 1969.

Robert C Merton. Optimum consumption and portfolio rules in a continuous-time model. Journal of Economic Theory, 3(4):373–413, 1971.

Raymond S Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. Review of general psychology, 2(2):175–220, 1998.

Kristoffer Nimark and Savitar Sundaresan. Inattention and belief polarization. Journal of Economic Theory, 180:203–228, 2019.

Rui Ruan, Yuchen Sun, Yue Tang, and Huihua Nie. 资本市场开放能否提高企业信息披露质量?——基于“沪港通”和年报文本挖掘的分析 [can opening the capital market improve the quality o corporate information disclosure? an analysis based on the shanghai-hong kong stock connect and annual report texts]. 金融研究 [Journal of Financial Research], 488(2):188–206, 2021.

Leonard Savage. The Foundations of Statistics. Wiley, New York, 1954.

Daniel Schacter, Daniel Gilbert, Daniel Wegner, and Matthew Nock. Psychology, chapter 4. Worth, New York City, third edition, 2014.

Andrei Shleifer and Lawrence H Summers. The noise trader approach to finance. Journal of Economic perspectives, 4(2):19–33, 1990.

Daniel J Simons and Christopher F Chabris. Gorillas in our midst: Sustained inattentional blindness for dynamic events. perception, 28(9):1059–1074, 1999.

Cameron Truong. Post-earnings announcement abnormal return in the chinese equity market. Journal of International Financial Markets, Institutions and Money, 21(5):637–661, 2011.

Albert Tsang, Fei Xie, and Xiangang Xin. Foreign institutional investors and corporate voluntary disclosure around the world. The Accounting Review, 94(5):319–348, 2019.

Jules van Binsbergen, Michael Brandt, and Ralph Koijen. On the timing and pricing of dividends. American Economic Review, 102(4):1596–1618, 2012.

Stijn Van Nieuwerburgh and Laura Veldkamp. Information acquisition and under-diversification. The Review of Economic Studies, 77(2):779–805, 2010.

Laura Veldkamp. Information choice in macroeconomics and finance. Princeton University Press, 2011.

Robert E. Verrecchia. Information acquisition in a noisy rational expectations economy. Econometrica: Journal of the Econometric Society, page 1415–1430, 1982.

Wayne Weiten. Psychology: Themes and variations. Wadsworth/Thomson Learning, 2001.

Yuan Zou. Lost in the Rising Tide: Exchange-traded Fund Flows and Valuation. PhD thesis, Columbia University, 2019.

Appendix A: Appendix of Chapter 1

A.1 Partial-Opening Model: solutions and proofs

Since the two risky securities are uncorrelated and signal noises are independently distributed, solving the equilibrium for each type of firm separately is essentially the same as working with the two types together in the matrix form. To simplify notation, I will adopt the scalar form in this section of Appendix. To further avoid notational complexities of type-I and type-II firms, superscripts (I) and (II) are not used on model parameters and variables. Without extra notations, plain-vanilla scalars are used to refer to type-I risky asset, with the whole expression readily applicable to type-II risky asset.

A.1.1 Computing conditional moments

Recall that the dividend $\widetilde{\boldsymbol{d}}_t$ follows an AR(1) process, $\widetilde{\boldsymbol{d}}_t = \boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_{t-1} + \widetilde{\boldsymbol{y}}_t$. If we view the asset as a claim to an infinite series of dividends payments, then $\widetilde{\boldsymbol{v}}_t$, the ex-dividend value of the firm at the end of Period t can be interpreted as

$$\widetilde{\boldsymbol{v}}_t = \mathbb{E}\left[\sum_{s=1}^{\infty} \frac{\widetilde{\boldsymbol{d}}_{t+s}}{r^s} \mid \mathcal{I}_{l2^+}\right] = \boldsymbol{\mu}_v + (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\widetilde{\boldsymbol{d}}_t,$$

where $\boldsymbol{\mu}_v = (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu}\sum_{s=1}^{\infty}[(\frac{1}{r})^s I - (\frac{1}{r}\boldsymbol{G})^s] = \frac{r}{r-1}(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu}$.

The excess payoff that investors obtain is $\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^{\dagger} - \boldsymbol{p}_1 r$ in Period 1 and $\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^{\dagger} r$ in Period 2. They are normally distributed, and their conditional means and variances are crucial for the model. In this section, I calculate them conditional on different information sets. Throughout this chapter, I adopt the notational style that uses a hat to distinguish conditional moments at the portfolio choice stage (*interim*) with those at the data choice stage (*ex ante*).

Conditional moments of dividend innovations For both periods ($t = 1, 2$), the price signal is

$$\boldsymbol{s}_{pt} = \boldsymbol{C}_{\boldsymbol{t}}^{-1}(\boldsymbol{p}_t - \boldsymbol{A}_t - \boldsymbol{B}_t\boldsymbol{d}_{t-1}) = \widetilde{\boldsymbol{y}}_t + \boldsymbol{C}_{\boldsymbol{t}}^{-1}\boldsymbol{D}_t\widetilde{\boldsymbol{x}}_t, \tag{A.1}$$

with price noise $\widetilde{\boldsymbol{\nu}}_{pt} = \boldsymbol{C}_{\boldsymbol{t}}^{-1}\boldsymbol{D}_t\widetilde{\boldsymbol{x}}_t \sim N(0, \boldsymbol{\Sigma}_{pt})$, where $\boldsymbol{\Sigma}_{pt} = (\boldsymbol{C}_{\boldsymbol{t}}^{-1}\boldsymbol{D}_t)\boldsymbol{\Sigma}_x(\boldsymbol{C}_{\boldsymbol{t}}^{-1}\boldsymbol{D}_t)' = \boldsymbol{\Omega}_{pt}^{-1}$. Thus $\boldsymbol{\Omega}_{pt} = \boldsymbol{\zeta}_t'\boldsymbol{\Omega}_x\boldsymbol{\zeta}_t$.

Domestic individual investor

At the portfolio choice stage, domestic individual investor j is endowed with inalienable information $\widetilde{\boldsymbol{\xi}}_{jt}$, obtains the price signal $\boldsymbol{s}_{pt}$, but does not process data. Thus she updates the conditional expectation and variance about $\widetilde{\boldsymbol{y}}_t$ following

$$\begin{aligned}
\hat{\boldsymbol{\Sigma}}_{Jyt} &\equiv \mathbb{V}ar(\widetilde{\boldsymbol{y}}_t \mid \mathcal{I}_{jt}) = \mathbb{V}ar(\widetilde{\boldsymbol{y}}_t \mid \boldsymbol{p}_t) = (\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{pt} + \boldsymbol{\Omega}_\epsilon)^{-1}, \\
\hat{\boldsymbol{\mu}}_{Jyt} &\equiv \mathbb{E}[\widetilde{\boldsymbol{y}}_t \mid \mathcal{I}_{jt}] = \mathbb{E}[\widetilde{\boldsymbol{y}}_t \mid \boldsymbol{p}_t] = \hat{\boldsymbol{\Sigma}}_{Jyt}\boldsymbol{\Omega}_{pt}\boldsymbol{s}_{pt}.
\end{aligned}$$

Domestic institutional investor

Apart from information contained in the price, domestic institutional investor i is endowed with inalienable information and also exacts information from buyable data. The conditional moments are

$$\begin{aligned}
\hat{\boldsymbol{\Sigma}}_{Iyt} &\equiv \mathbb{V}ar(\widetilde{\boldsymbol{y}}_t \mid \mathcal{I}_{it}) = \mathbb{V}ar(\widetilde{\boldsymbol{y}}_t \mid \boldsymbol{p}_t, \boldsymbol{s}_{it}, \boldsymbol{\xi}_{it}) = (\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{pt} + \beta_I^2\boldsymbol{M}_{it}\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon)^{-1}, \\
\hat{\boldsymbol{\mu}}_{iyt} &\equiv \mathbb{E}[\widetilde{\boldsymbol{y}}_t \mid \mathcal{I}_{it}] = \mathbb{E}[\widetilde{\boldsymbol{y}}_t \mid \boldsymbol{p}_t, \boldsymbol{s}_{it}, \boldsymbol{\xi}_{it}] = \hat{\boldsymbol{\Sigma}}_{Iyt}(\boldsymbol{\Omega}_{pt}\boldsymbol{s}_{pt} + \beta_I^2\boldsymbol{\Omega}_e\boldsymbol{\iota}_{it}\boldsymbol{s}_{it} + \boldsymbol{\Omega}_\epsilon\boldsymbol{\xi}_{it}).
\end{aligned}$$

There are two things here worth special mentioning. One is that $\boldsymbol{\iota}_{it}\boldsymbol{s}_{it} = \begin{pmatrix} \sum_{m\in\{\text{(I)}\}} s_{it}^{(m)} \\ \sum_{m\in\{\text{(II)}\}} s_{it}^{(m)} \end{pmatrix}$ is a vector that denotes the sum of data signal acquired by investor i in Period t for each type of risky asset. The other is that due to the constraint of imperfect data processing technology, domestic institutional investors wrongly specifies $\beta_I < \beta$. This will in turn affect their portfolio choices.

International institutional investor

International institutional investors only enter the market in Period 2. Apart from the observed price $\boldsymbol{p}_2$, international institutional investor k also exacts information from data. Since her data choice problem is degenerate, she would allocate all $\bar{M}$ data points to type-I risky asset, that is $\boldsymbol{M}_{K2} = diag((\bar{M}, 0)')$, $\boldsymbol{\iota}_{K2}\boldsymbol{s}_{k2} = (\sum_{m=1}^{\bar{M}} s_{k2}^{(m)}, 0)'$. Then the conditional moments become

$$\hat{\boldsymbol{\Sigma}}_{Ky2} \equiv \mathbb{V}ar(\widetilde{\boldsymbol{y}}_2 \mid \mathcal{I}_{k2}) = \mathbb{V}ar(\widetilde{\boldsymbol{y}}_2 \mid \boldsymbol{p}_2, \boldsymbol{s}_{k2}) = (\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + \beta^2 \boldsymbol{M}_{K2}\boldsymbol{\Omega}_e)^{-1},$$
$$\hat{\boldsymbol{\mu}}_{ky2} \equiv \mathbb{E}[\widetilde{\boldsymbol{y}}_2 \mid \mathcal{I}_{k2}] = \mathbb{E}[\widetilde{\boldsymbol{y}}_2 \mid \boldsymbol{p}_2, \boldsymbol{s}_{k2}] = \hat{\boldsymbol{\Sigma}}_{Ky2}(\boldsymbol{\Omega}_{p2}\boldsymbol{s}_{p2} + \beta^2 \boldsymbol{\Omega}_e \boldsymbol{\iota}_{K2}\boldsymbol{s}_{k2}).$$

Conditional moments of payoffs Let us look at the portfolio choice stage of Period 1. $\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^\dagger - \boldsymbol{p}_1 r \mid \mathcal{I}_{l1}$ can be analyzed as

$$\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^\dagger - \boldsymbol{p}_1 r \mid \mathcal{I}_{l1} = \boldsymbol{A}_2^\dagger + \boldsymbol{B}_2^\dagger(\boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_0) + (\boldsymbol{I} + \boldsymbol{B}_2^\dagger)\widetilde{\boldsymbol{y}}_1 + \boldsymbol{C}_2^\dagger \widetilde{\boldsymbol{y}}_2 + \boldsymbol{D}_2^\dagger \widetilde{\boldsymbol{x}}_2 - \boldsymbol{p}_1 r \mid \mathcal{I}_{l1}.$$

Denote the variance and the expectation of the period-1 payoff $\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^\dagger - \boldsymbol{p}_1 r$ conditional on $\mathcal{I}_{l1}$ as $\hat{\boldsymbol{\Sigma}}_{L1}$ and $\hat{\boldsymbol{\mu}}_{l1}$, $l \in \{i, j\}$, $L \in \{I, J\}$. Then

$$\hat{\boldsymbol{\Sigma}}_{L1} = (\boldsymbol{I} + \boldsymbol{B}_2^\dagger)\hat{\boldsymbol{\Sigma}}_{Ly1}(\boldsymbol{I} + \boldsymbol{B}_2^\dagger)' + \boldsymbol{C}_2^\dagger \boldsymbol{\Sigma}_y \boldsymbol{C}_2'^\dagger + \boldsymbol{D}_2^\dagger \boldsymbol{\Sigma}_x \boldsymbol{D}_2'^\dagger, \tag{A.2}$$
$$\hat{\boldsymbol{\mu}}_{l1} = \boldsymbol{A}_2^\dagger + \boldsymbol{B}_2^\dagger(\boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_0) + (\boldsymbol{I} + \boldsymbol{B}_2^\dagger)\hat{\boldsymbol{\mu}}_{ly1} - \boldsymbol{p}_1 r, \tag{A.3}$$

respectively. At the data choice stage, $\boldsymbol{p}_1$ is not realized. Thus the period-1 excess payoff can be written as

$$\begin{aligned}
&\widetilde{\boldsymbol{d}}_1 + \widetilde{\boldsymbol{p}}_2^\dagger - \widetilde{\boldsymbol{p}}_1 r \mid \mathcal{I}_{l0} \\
&= \boldsymbol{A}_2^\dagger - r\boldsymbol{A}_1 + (\boldsymbol{I} + \boldsymbol{B}_2^\dagger)(\boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_0) - r\boldsymbol{B}_1\boldsymbol{d}_0 + \boldsymbol{C}_2^\dagger \widetilde{\boldsymbol{y}}_2 + (\boldsymbol{I} + \boldsymbol{B}_2^\dagger - r\boldsymbol{C}_1)\widetilde{\boldsymbol{y}}_1 + \boldsymbol{D}_2^\dagger \widetilde{\boldsymbol{x}}_2 - r\boldsymbol{D}_1 \widetilde{\boldsymbol{x}}_1.
\end{aligned}$$

Its conditional variance and expectation are

$$\boldsymbol{\Sigma}_{L1} = (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger} - r\boldsymbol{C}_1)^2\boldsymbol{\Sigma}_{Ly1} + \boldsymbol{C}_2^{2\dagger}\boldsymbol{\Sigma}_y + r^2\boldsymbol{D}_1^2\boldsymbol{\Sigma}_x + \boldsymbol{D}_2^{2\dagger}\boldsymbol{\Sigma}_x, \tag{A.4}$$

$$\boldsymbol{\mu}_{l1} = \boldsymbol{A}_2^{\dagger} - r\boldsymbol{A}_1 + (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})(\boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_0) - r\boldsymbol{B}_1\boldsymbol{d}_0 + (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger} - r\boldsymbol{C}_1)\boldsymbol{\mu}_{ly1}. \tag{A.5}$$

Domestic investors are endowed with inalienable information about $\widetilde{\boldsymbol{y}}_1$, thus $\boldsymbol{\Sigma}_{Ly1} = (\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_\epsilon)^{-1}$ and $\boldsymbol{\mu}_{ly1} = \boldsymbol{\Sigma}_{Ly1}\boldsymbol{\Omega}_\epsilon\boldsymbol{\xi}_{l1}$, $l = i, j$, $L = I, J$.

At the Period-2 portfolio choice stage, the Period-2 payoff $\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r$ can be analyzed as

$$\begin{aligned}\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r \mid \mathcal{I}_{l2} &= \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\widetilde{\boldsymbol{d}}_2 - \boldsymbol{p}_2 r \\ &= \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{d}_1 + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\widetilde{\boldsymbol{y}}_2 - \boldsymbol{p}_2 r.\end{aligned}$$

Its conditional variance and expectation are

$$\hat{\boldsymbol{\Sigma}}_{L2} = r^2(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\Sigma}}_{Ly2}(r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}, \tag{A.6}$$

$$\hat{\boldsymbol{\mu}}_{l2} = \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{d}_1 + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\mu}}_{ly2} - \boldsymbol{p}_2 r, \tag{A.7}$$

respectively. At the Period-2 data choice stage, the Period-2 excess payoff can be written as

$$\begin{aligned}&\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r \mid \mathcal{I}_{l2}- \\ &= \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - r\boldsymbol{A}_2 + [r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G} - r\boldsymbol{B}_2]\boldsymbol{d}_1 + [r(r\boldsymbol{I} - \boldsymbol{G})^{-1} - r\boldsymbol{C}_2]\widetilde{\boldsymbol{y}}_2 - r\boldsymbol{D}_2\widetilde{\boldsymbol{x}}_2 \mid \mathcal{I}_{l1}.\end{aligned}$$

Its conditional variance and expectation are

$$\boldsymbol{\Sigma}_{L2} = [r(r\boldsymbol{I} - \boldsymbol{G})^{-1} - r\boldsymbol{C}_2]\boldsymbol{\Sigma}_{Ly2}[r(r\boldsymbol{I} - \boldsymbol{G})^{-1} - r\boldsymbol{C}_2]' + r^2\boldsymbol{D}_2\boldsymbol{\Sigma}_x\boldsymbol{D}_2', \tag{A.8}$$

$$\boldsymbol{\mu}_{l2} = \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - r\boldsymbol{A}_2 + \left[r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G} - r\boldsymbol{B}_2\right]\boldsymbol{d}_1 + \left[r(r\boldsymbol{I} - \boldsymbol{G})^{-1} - r\boldsymbol{C}_2\right]\boldsymbol{\mu}_{ly2}, \tag{A.9}$$

respectively. Before foreign institutional investors make data choice, they have no information on Period-2 dividend innovation $\widetilde{\boldsymbol{y}}_2$, thus $\boldsymbol{\Sigma}_{Ky2} = \boldsymbol{\Sigma}_y$, $\boldsymbol{\mu}_{ky2} = 0$. Domestic investors have inalienable information $\boldsymbol{\xi}_{l2}$, so $\boldsymbol{\Sigma}_{Ly2} = (\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_\epsilon)^{-1}$ and $\boldsymbol{\mu}_{ly2} = \boldsymbol{\Sigma}_{Ly2}\boldsymbol{\Omega}_\epsilon\boldsymbol{\xi}_{l2}$, $l = i, j, L = I, J$.

A.1.2 Optimal portfolio choice

In period 2, the optimizing problems faced by all the three types of investors have the same form. For any domestic investor l, $l \in \{i, j\}$,

$$\max_{\{\boldsymbol{q}_{l2}\}} \quad \mathbb{E}[-e^{-\rho w_{l2}} \mid \mathcal{I}_{l2}] \tag{A.10}$$

$$\text{s.t.} \quad w_{l2} = rw_{l1} + \boldsymbol{q}_{l2}'(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r). \tag{A.11}$$

This is a one-shot problem. Substituting (A.11) into (A.10) and using the mean formula for log-normal variables[1], we have $\boldsymbol{q}_{l2}^* = \arg\max_{\boldsymbol{q}_{l2}} \boldsymbol{q}_{l2}'\hat{\boldsymbol{\mu}}_{l2} - \frac{\rho}{2}\boldsymbol{q}_{l2}\hat{\boldsymbol{\Sigma}}_{L2}\boldsymbol{q}_{l2}'$. Then the first-order condition gives

$$\boldsymbol{q}_{l2} = \frac{1}{\rho}\hat{\boldsymbol{\Omega}}_{L2}\hat{\boldsymbol{\mu}}_{l2}. \tag{A.12}$$

Note that since the payoffs of type-I and II risky assets are correlated, $\hat{\boldsymbol{\Omega}}_{L2}$ is not diagonal. So the posterior expectation, variance of one asset's payoff, and covariance of the payoffs with the other risky asset all affect the position of the other risky asset. Foreign institutional investor k faces a constrained problem with $q_{k2}^{(\mathrm{II})} = 0$. Thus

$$\boldsymbol{q}_{k2} = \begin{pmatrix} \frac{\hat{\mu}_{k2}^{(\mathrm{I})}}{\rho\hat{\sigma}_{K2}^{2(\mathrm{I})}} \\ 0 \end{pmatrix}.$$

In this case, the posterior variance and expectation of type-I risky asset alone determine the asset position $q_{k2}^{(\mathrm{I})}$. However, this does not mean that information about type-II risky asset has no use, as they could play a role through $\hat{\sigma}_{K2}^{2(\mathrm{I})}$ and $\hat{\mu}_{k2}^{(\mathrm{I})}$. To use a uniform notation, denote $\hat{\boldsymbol{\Omega}}_{K2} =$

[1] If $\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\mathbb{E}[e^{\boldsymbol{x}}] = e^{\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{\Sigma}}$.

$\begin{pmatrix} \hat{\Omega}_{K2}^{(\mathrm{I})} & \hat{c}_{K2} \\ \hat{c}_{K2} & \hat{\Omega}_{K2}^{(\mathrm{II})} \end{pmatrix}$. Then the first diagonal element $\hat{\sigma}_{K2}^{2(\mathrm{I})}$ of the posterior variance-covariance matrix $\hat{\boldsymbol{\Sigma}}_{K2}$ is $\frac{\hat{\Omega}_{K2}^{(\mathrm{II})}}{\hat{\Omega}_{K2}^{(\mathrm{I})}\hat{\Omega}_{K2}^{(\mathrm{II})}-\hat{c}_{K2}^2}$. Define $\hat{\boldsymbol{\Omega}}_{K2}^{\S} = \begin{pmatrix} \hat{\Omega}_{K2}^{(\mathrm{I})} - \frac{\hat{c}_{K2}^2}{\hat{\Omega}_{K2}^{(\mathrm{II})}} & 0 \\ 0 & 0 \end{pmatrix}$. With these notations, the portfolio choice of investor k can be written as

$$\boldsymbol{q}_{k2} = \frac{1}{\rho}\hat{\boldsymbol{\Omega}}_{K2}^{\S}\hat{\boldsymbol{\mu}}_{k2}.$$

Note that the first diagonal elements are different for $\hat{\tilde{\Omega}}_{K2}$ and $\hat{\tilde{\Omega}}_{K2}^{\S}$.

In Period 1, only domestic investors are in the market. Consider domestic investor $l, l \in \{i, j\}$. Since the market opening is an exogenous shock at the beginning of Period 2, when domestic investors make investment decisions in Period 1, a hypothetical Period 2 when market opening does not happen is considered. Let us denote the corresponding information set as $\mathcal{I}_{l2}^{\dagger}$. Then her optimal portfolio choice in the hypothetical Period 2 is $\boldsymbol{q}_{l2}^{\dagger}$, which has the same form as $\boldsymbol{q}_{l2}$ except for the substitution of $\mathcal{I}_{l2}$ by $\mathcal{I}_{l2}^{\dagger}$, in which the sole difference is that the Period-2 price is the would-be price if the market does not open to international investors. Therefore in Period 1 she faces the following optimization problem:

$$\max_{\{\boldsymbol{q}_{l1}\}} \quad \mathbb{E}[-e^{-\rho w_{l2}} \mid \mathcal{I}_{l1}] \tag{A.13}$$

$$\text{s.t.} \quad w_{l2} = r^2 w_{l0} + r\boldsymbol{q}_{l1}'(\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^{\dagger} - \boldsymbol{p}_1 r) + \boldsymbol{q}_{l2}'^{\dagger}(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^{\dagger} r). \tag{A.14}$$

By plugging $\boldsymbol{q}_{l2}'^{\dagger}$ and (A.14) into (A.13) and using properties of the log-normal random variable, one can reduce the portfolio choice problem to $\max_{\{\boldsymbol{q}_{l1}\}} r\boldsymbol{q}_{l1}'\hat{\boldsymbol{\mu}}_{l1} - \frac{\rho}{2}\mathbb{V}ar(\boldsymbol{q}_{l1}'(\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^{\dagger} - \boldsymbol{p}_1 r) + \boldsymbol{q}_{l2}'^{\dagger}(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^{\dagger} r) \mid \mathcal{I}_{l1})$.

Note that $\mathbb{V}ar(\boldsymbol{q}_{l1}'(\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^{\dagger} - \boldsymbol{p}_1 r) + \boldsymbol{q}_{l2}'^{\dagger}(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^{\dagger} r) \mid \mathcal{I}_{l1}) = \boldsymbol{q}_{l1}'\mathbb{V}ar(\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^{\dagger} - \boldsymbol{p}_1 r \mid \mathcal{I}_{l1})\boldsymbol{q}_{l1}' + \mathbb{V}ar(\widetilde{\boldsymbol{d}}_2 + \boldsymbol{q}_{l2}'^{\dagger}(\widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^{\dagger} r) \mid \mathcal{I}_{l1}) + 2\mathbb{C}ov(\boldsymbol{q}_{l1}'(\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^{\dagger} - \boldsymbol{p}_1 r), \boldsymbol{q}_{l2}'^{\dagger}(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^{\dagger} r) \mid \mathcal{I}_{l1})$, where the second item is irrelevant to the choice of $\boldsymbol{q}_{l1}$. Therefore the optimization problem ultimately

reduces to

$$\max_{\{\boldsymbol{q}_{l1}\}} \quad \boldsymbol{q}_{l1}'\hat{\boldsymbol{\mu}}_{l1} - \frac{\rho r}{2}\boldsymbol{q}_{l1}\hat{\boldsymbol{\Sigma}}_{L1}\boldsymbol{q}_{l1}' - \rho\boldsymbol{q}_{l1}\boldsymbol{\mathcal{C}}_L,$$

where the covariance term evaluates

$$\boldsymbol{\mathcal{C}}_L \equiv \mathbb{C}ov\left(\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^\dagger - \boldsymbol{p}_1 r, \boldsymbol{q}_{l2}'^\dagger\left(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^\dagger r\right) \mid \mathcal{I}_{l1}\right).$$

The F.O.C. leads to

$$\boldsymbol{q}_{l1} = \frac{1}{\rho r}\hat{\boldsymbol{\Omega}}_{L1}(\hat{\boldsymbol{\mu}}_{l1} - \rho\boldsymbol{\mathcal{C}}_L).$$

To compute $\boldsymbol{\mathcal{C}}_L$, first note that conditional on $\mathcal{I}_{l1}$, $\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^\dagger = \boldsymbol{A}_2^\dagger + (\boldsymbol{I} + \boldsymbol{B}_2^\dagger)(\boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_0 + \widetilde{\boldsymbol{y}}_1) + \boldsymbol{C}_2^\dagger\widetilde{\boldsymbol{y}}_2 + \boldsymbol{D}_2^\dagger\widetilde{\boldsymbol{x}}_2$. $\boldsymbol{A}_2^\dagger$, $\boldsymbol{B}_2^\dagger$, $\boldsymbol{C}_2^\dagger$ and $\boldsymbol{D}_2^\dagger$ are functions of model parameters, in particular, $\boldsymbol{B}_2^\dagger = (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}$. Thus

$$\widetilde{\boldsymbol{d}}_1 + \boldsymbol{p}_2^\dagger = \boldsymbol{A}_2^\dagger + (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}(\boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_0) + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\widetilde{\boldsymbol{y}}_1 + \boldsymbol{C}_2^\dagger\widetilde{\boldsymbol{y}}_2 + \boldsymbol{D}_2^\dagger\widetilde{\boldsymbol{x}}_2.$$

Denote $a_1 = \boldsymbol{A}_2^\dagger + (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}(\boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_0)$, $b_1 = \boldsymbol{C}_2^\dagger$ and $c_1 = \boldsymbol{D}_2^\dagger$.

Since all the shocks are independently normally distributed with zero mean, then we only need to calculate the constant and the coefficients of $\widetilde{\boldsymbol{y}}_1$, $\widetilde{\boldsymbol{y}}_2$, $\widetilde{\boldsymbol{x}}_2$, $\widetilde{\boldsymbol{y}}_1^2$, $\widetilde{\boldsymbol{y}}_2^2$ and $\widetilde{\boldsymbol{x}}_2^2$ terms of $\boldsymbol{q}_{l2}'^\dagger(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^\dagger r)$. On the one hand, $\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^\dagger r = \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}(\boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_1 + \widetilde{\boldsymbol{y}}_2) - r\boldsymbol{A}_2 - r\boldsymbol{B}_2\widetilde{\boldsymbol{d}}_1 - r\boldsymbol{C}_2\widetilde{\boldsymbol{y}}_2 - r\boldsymbol{D}_2\widetilde{\boldsymbol{x}}_2$. Since $\boldsymbol{B}_2^\dagger = (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}$, the $\widetilde{\boldsymbol{d}}_1$ terms cancel out and no $\widetilde{\boldsymbol{y}}_1$ shock is included, that is $\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^\dagger r = \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - r\boldsymbol{A}_2 + [r(r\boldsymbol{I} - \boldsymbol{G})^{-1} - r\boldsymbol{C}_2]\widetilde{\boldsymbol{y}}_2 - r\boldsymbol{D}_2\widetilde{\boldsymbol{x}}_2$. On the other hand, $\boldsymbol{q}_{l2}^\dagger = \frac{1}{\rho}\hat{\boldsymbol{\Omega}}_{L2}\hat{\boldsymbol{\mu}}_{l2} = \frac{1}{\rho}\hat{\boldsymbol{\Omega}}_{Ly2}(\boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\widetilde{\boldsymbol{d}}_1 + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\mu}}_{ly2} - r\boldsymbol{A}_2 - r\boldsymbol{B}_2\widetilde{\boldsymbol{d}}_1 - r\boldsymbol{C}_2\widetilde{\boldsymbol{y}}_2 - r\boldsymbol{D}_2\widetilde{\boldsymbol{x}}_2)$. Again, $\widetilde{\boldsymbol{d}}_1$ terms cancel out. Recall that $\hat{\boldsymbol{\Omega}}_{Iy2} = (\frac{r-g}{r})^2(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + M^{(\mathrm{I})}\beta_I^2\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon)$, $\hat{\boldsymbol{\Omega}}_{Jy2} = (\frac{r-g}{r})^2(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + \boldsymbol{\Omega}_\epsilon)$ and $\hat{\boldsymbol{\mu}}_{iy2} = \hat{\boldsymbol{\Sigma}}_{Iy2}(\boldsymbol{\Omega}_{p2}\boldsymbol{s}_{p2} + \beta_I^2\boldsymbol{\Omega}_e\sum_{m\in\{(\mathrm{I})\}}s_{i2}^{(m)} + \boldsymbol{\Omega}_\epsilon\boldsymbol{\xi}_{i2})$, $\hat{\boldsymbol{\mu}}_{jy2} = \hat{\boldsymbol{\Sigma}}_{Jy2}(\boldsymbol{\Omega}_{p2}\boldsymbol{s}_{p2} + \boldsymbol{\Omega}_\epsilon\boldsymbol{\xi}_{j2})$. So $\widetilde{\boldsymbol{y}}_1$ also does

not enter $\boldsymbol{q}_{l2}^{\dagger}$. Collecting terms that yield non-zero result after the covariance operation, we have $\hat{\boldsymbol{\mu}}_{iy2} \to \hat{\boldsymbol{\Sigma}}_{Iy2}(\boldsymbol{\Omega}_{p2} + M^{(\mathrm{I})}\beta_I^2\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon)\widetilde{\boldsymbol{y}}_2 + \hat{\boldsymbol{\Sigma}}_{Iy2}\boldsymbol{\Omega}_{p2}\frac{1}{\boldsymbol{\zeta}_2}\widetilde{\boldsymbol{x}}_2$. Thus

$$
\begin{aligned}
& \boldsymbol{q}_{i2}^{\dagger}(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2^{\dagger}r) \\
\to & \frac{(r-g)^2}{\rho r^2}(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + M^{(\mathrm{I})}\beta_I^2\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon) \\
& \times \left[\boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - r\boldsymbol{A}_2 + [r(r\boldsymbol{I} - \boldsymbol{G})^{-1} - r\boldsymbol{C}_2]\widetilde{\boldsymbol{y}}_2 - r\boldsymbol{D}_2\widetilde{\boldsymbol{x}}_2\right] \\
& \times \Big[\boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - r\boldsymbol{A}_2 + (r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\Sigma}}_{Iy2}(\boldsymbol{\Omega}_{p2} + M^{(\mathrm{I})}\beta_I^2\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon) - r\boldsymbol{C}_2)\widetilde{\boldsymbol{y}}_2 \\
& + (r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\Sigma}}_{Iy2}\boldsymbol{\Omega}_{p2}\frac{1}{\boldsymbol{\zeta}_2} - r\boldsymbol{D}_2)\widetilde{\boldsymbol{x}}_2\Big].
\end{aligned}
$$

Working out the constant and the coefficients of $\widetilde{\boldsymbol{y}}_2$, $\widetilde{\boldsymbol{x}}_2$, $\widetilde{\boldsymbol{y}}_2^2$ and $\widetilde{\boldsymbol{x}}_2^2$, we have $a_2 = (\boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - r\boldsymbol{A}_2)^2$, $b_2 = (\boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - r\boldsymbol{A}_2)[r(r\boldsymbol{I} - \boldsymbol{G})^{-1} - r\boldsymbol{C}_2 + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\Sigma}}_{Iy2}(\boldsymbol{\Omega}_{p2} + M^{(\mathrm{I})}\beta_I^2\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon) - r\boldsymbol{C}_2]$, $c_2 = (\boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - r\boldsymbol{A}_2)(r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\Sigma}}_{Iy2}\boldsymbol{\Omega}_{p2}\frac{1}{\boldsymbol{\zeta}_2} - 2r\boldsymbol{D}_2)$, $d_2 = [r(r\boldsymbol{I} - \boldsymbol{G})^{-1} - r\boldsymbol{C}_2](r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\Sigma}}_{Iy2}(\boldsymbol{\Omega}_{p2} + M^{(\mathrm{I})}\beta_I^2\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon) - r\boldsymbol{C}_2)$ and $e_2 = -r\boldsymbol{D}_2(r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\Sigma}}_{Iy2}\boldsymbol{\Omega}_{p2}\frac{1}{\boldsymbol{\zeta}_2} - r\boldsymbol{D}_2)$. Therefore

$$
\boldsymbol{\mathcal{C}}_I = \frac{(r-g)^2}{\rho r^2}(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + M^{(\mathrm{I})}\beta_I^2\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon)[a_1a_2 + (a_1d_2 + b_1b_2)\boldsymbol{\Sigma}_y + (a_1e_2 + c_1c_2)\boldsymbol{\Sigma}_x].
$$

Setting all $M^{(\mathrm{I})}$ to zero, we have $\boldsymbol{\mathcal{C}}_J$.

A.1.3 Equilibrium price coefficients

Period 2

In Period 2, all investors, domestic-institutional, domestic-individual, and foreign-institutional, participate in the domestic type-I stock market. For clarity, in the following subsection of deriving Period-2 price coefficients, I will only the (I) superscript for $M_2^{(\mathrm{I})}$, the data amount acquired by domestic institutional investors for type-I firm, and drop it for any other model parameter and variable. The net demand comes from all investors and equals firms' supply of shares. Thus

market clearing requires demand equals supply:

$$\int_i (\boldsymbol{q}_{i2} - \boldsymbol{q}_{i1})\,\mathrm{d}i + \int_j (\boldsymbol{q}_{j2} - \boldsymbol{q}_{j1})\,\mathrm{d}j + \int_k \boldsymbol{q}_{k2}\,\mathrm{d}k = \bar{\boldsymbol{x}}_2 + \widetilde{\boldsymbol{x}}_2 - (\bar{\boldsymbol{x}}_1 + \widetilde{\boldsymbol{x}}_1).$$

Subtracting the period-1 market clearing from it yields

$$\int_i \boldsymbol{q}_{i2}\,\mathrm{d}i + \int_j \boldsymbol{q}_{j2}\,\mathrm{d}j + \int_k \boldsymbol{q}_{k2}\,\mathrm{d}k = \bar{\boldsymbol{x}}_2 + \widetilde{\boldsymbol{x}}_2. \tag{A.15}$$

Notice the difference of these two equations. The first one equates demand and supply, which are both flow variables, while the second one is built on the fact that the aggregate position, which is a stock variable, held by all investors is equal to the total number of shares issued by the firm.

Let me start by aggregate investor i' posterior expectations on dividend innovation. The noises in data signals and in inalienable information both integrate to zero. We have

$$\int_i \hat{\boldsymbol{\mu}}_{iy2}\,\mathrm{d}i = \hat{\boldsymbol{\Sigma}}_{Iy2} \int_i \boldsymbol{\Omega}_{p2}\boldsymbol{s}_{p2} + \beta_I^2 \boldsymbol{\Omega}_e \boldsymbol{\iota}_{i2}\boldsymbol{s}_{i2} + \boldsymbol{\Omega}_\epsilon \boldsymbol{\xi}_{i2}\,\mathrm{d}i = \lambda_I \hat{\boldsymbol{\Sigma}}_{Iy2}[\boldsymbol{\Omega}_{p2}\boldsymbol{s}_{p2} + (\beta_I^2 \boldsymbol{M}_{I2}\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon)\widetilde{\boldsymbol{y}}_2].$$

On the other hand, the aggregate demand of all domestic-institutional investors evaluates to

$$\begin{aligned}
&\int_i \boldsymbol{q}_{i2}\,\mathrm{d}i = \frac{1}{\rho}\hat{\boldsymbol{\Omega}}_{I2}\int_i \hat{\boldsymbol{\mu}}_{i2}\,\mathrm{d}i \\
&= \frac{1}{\rho r^2}(r\boldsymbol{I} - \boldsymbol{G})\hat{\boldsymbol{\Omega}}_{Iy2}(r\boldsymbol{I} - \boldsymbol{G})' \\
&\quad \times \int_i \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{d}_1 + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\mu}}_{iy2} - \boldsymbol{p}_2 r\,\mathrm{d}i\,.
\end{aligned}$$

Plugging $\int_i \hat{\boldsymbol{\mu}}_{iy2}\,\mathrm{d}i$ we have

$$\begin{aligned}
\int_i \boldsymbol{q}_{i2}\,\mathrm{d}i = \frac{\lambda_I}{\rho r^2}(r\boldsymbol{I} - \boldsymbol{G})\hat{\boldsymbol{\Omega}}_{Iy2}(r\boldsymbol{I} - \boldsymbol{G})'[\boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{d}_1 - \boldsymbol{p}_2 r] \\
+\frac{\lambda_I}{\rho r}(r\boldsymbol{I} - \boldsymbol{G})[\boldsymbol{\Omega}_{p2}\boldsymbol{s}_{p2} + (\beta_I^2 \boldsymbol{M}_{I2}\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon)\widetilde{\boldsymbol{y}}_2].
\end{aligned} \tag{A.16}$$

The aggregate demand of all domestic individual investors evaluates to

$$\begin{aligned}\int_j \boldsymbol{q}_{j2}\,\mathrm{d}j = &\frac{\lambda_J}{\rho r^2}(r\boldsymbol{I}-\boldsymbol{G})\hat{\boldsymbol{\Omega}}_{Jy2}(r\boldsymbol{I}-\boldsymbol{G})'[\boldsymbol{\mu}_v + r(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{d}_1 - \boldsymbol{p}_2 r]\\ &+\frac{\lambda_J}{\rho r}(r\boldsymbol{I}-\boldsymbol{G})[\boldsymbol{\Omega}_{p2}\boldsymbol{s}_{p2} + \boldsymbol{\Omega}_{\epsilon}\widetilde{\boldsymbol{y}}_2]. \end{aligned} \tag{A.17}$$

Finally, the aggregate demand of all foreign institutional investors evaluates to $\int_k \boldsymbol{q}_{k2}\,\mathrm{d}k = \frac{1}{\rho}\hat{\boldsymbol{\Omega}}^{\S}_{K2}\int_k \hat{\boldsymbol{\mu}}_{k2}\,\mathrm{d}k$. To aggregate their demands with other investors in a uniform notation, I make a connection between $\hat{\boldsymbol{\Omega}}^{\S}_{K2}$ and $\hat{\boldsymbol{\Omega}}_{Ky2}$. First, let me write out $\hat{\boldsymbol{\Omega}}_{Ky2}$. Let $\hat{\boldsymbol{\Omega}}_{Ky2} = \begin{pmatrix} \hat{\Omega}^{(\mathrm{I})}_{Ky2} & \hat{\mathcal{C}}_{Ky2} \\ \hat{\mathcal{C}}_{Ky2} & \hat{\Omega}^{(\mathrm{II})}_{Ky2}\end{pmatrix}$. Then

$$\begin{pmatrix} \hat{\Omega}^{(\mathrm{I})}_{K2} & \hat{\mathcal{C}}_{K2} \\ \hat{\mathcal{C}}_{K2} & \hat{\Omega}^{(\mathrm{II})}_{K2}\end{pmatrix} \equiv \hat{\boldsymbol{\Omega}}_{K2} = \frac{1}{r^2}(r\boldsymbol{I}-\boldsymbol{G})\hat{\boldsymbol{\Omega}}_{Ky2}(r\boldsymbol{I}-\boldsymbol{G})' = \begin{pmatrix} \frac{(r-g_1)^2}{r^2}\hat{\Omega}^{(\mathrm{I})}_{Ky2} & \frac{(r-g_1)(r-g_2)}{r^2}\hat{\mathcal{C}}_{Ky2} \\ \frac{(r-g_1)(r-g_2)}{r^2}\hat{\mathcal{C}}_{Ky2} & \frac{(r-g_2)^2}{r^2}\hat{\Omega}^{(\mathrm{II})}_{Ky2}\end{pmatrix}.$$

Thus $\hat{\Omega}^{(\mathrm{I})}_{K2} = \frac{(r-g_1)^2}{r^2}\hat{\Omega}^{(\mathrm{I})}_{Ky1}$, $\hat{\Omega}^{(\mathrm{II})}_{K2} = \frac{(r-g_2)^2}{r^2}\hat{\Omega}^{(\mathrm{II})}_{Ky2}$ and $\hat{\mathcal{C}}_{K2} = \frac{(r-g_1)(r-g_2)}{r^2}\hat{\mathcal{C}}_{Ky2}$. Define $\hat{\boldsymbol{\Omega}}^{\S}_{Ky2} \equiv \begin{pmatrix} \hat{\Omega}^{(\mathrm{I})}_{Ky2} - \frac{\hat{\mathcal{C}}^2_{Ky2}}{\hat{\Omega}^{(\mathrm{II})}_{Ky2}} & 0 \\ 0 & 0\end{pmatrix} = \boldsymbol{E}_{11}\hat{\boldsymbol{\Omega}}_{Ky2} - \frac{\hat{\mathcal{C}}^2_{Ky2}}{\Omega^{(\hat{\mathrm{II}})}_{Ky2}}\boldsymbol{E}_{11}$, where $\boldsymbol{E}_{11} = \left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right), \boldsymbol{E}_{12} = \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)$ are matrix units with the $(1,1)$-th and $(1,2)$-th element being 1 and all other elements zero. Then

$$\frac{1}{r^2}(r\boldsymbol{I}-\boldsymbol{G})\hat{\boldsymbol{\Omega}}^{\S}_{Ky2}(r\boldsymbol{I}-\boldsymbol{G})' = \begin{pmatrix} \frac{(r-g_1)^2}{r^2}(\hat{\Omega}^{(\mathrm{I})}_{Ky2} - \frac{\hat{\mathcal{C}}^2_{Ky2}}{\hat{\Omega}^{(\mathrm{II})}_{Ky2}}) & 0 \\ 0 & 0\end{pmatrix} = \begin{pmatrix} \hat{\Omega}^{(\mathrm{I})}_{K2} - \frac{\hat{\mathcal{C}}^2_{K2}}{\hat{\Omega}^{(\mathrm{II})}_{K2}} & 0 \\ 0 & 0\end{pmatrix} = \hat{\boldsymbol{\Omega}}^{\S}_{K2}.$$

Therefore,

$$\begin{aligned}&\int_k \boldsymbol{q}_{k2}\,\mathrm{d}k = \frac{1}{\rho}\hat{\boldsymbol{\Omega}}^{\S}_{K2}\int_k \hat{\boldsymbol{\mu}}_{k2}\,\mathrm{d}k\\ =&\frac{1}{\rho r^2}(r\boldsymbol{I}-\boldsymbol{G})\hat{\boldsymbol{\Omega}}^{\S}_{Ky2}(r\boldsymbol{I}-\boldsymbol{G})'\\ &\times\int_k \boldsymbol{\mu}_v + r(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{d}_1 + r(r\boldsymbol{I}-\boldsymbol{G})^{-1}\hat{\boldsymbol{\mu}}_{ky2} - \boldsymbol{p}_2 r\,\mathrm{d}k\\ =&\frac{\lambda_K}{\rho r^2}(r\boldsymbol{I}-\boldsymbol{G})\hat{\boldsymbol{\Omega}}^{\S}_{Ky2}(r\boldsymbol{I}-\boldsymbol{G})'[\boldsymbol{\mu}_v + r(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{d}_1 - \boldsymbol{p}_2 r]\\ &+\frac{\lambda_K}{\rho r}(r\boldsymbol{I}-\boldsymbol{G})\hat{\boldsymbol{\Omega}}^{\S}_{Ky2}\hat{\boldsymbol{\Sigma}}_{Ky2}[\boldsymbol{\Omega}_{p2}\boldsymbol{s}_{p2} + \beta^2\boldsymbol{M}_{K2}\boldsymbol{\Omega}_e\widetilde{\boldsymbol{y}}_2]. \end{aligned} \tag{A.18}$$

Note that

$$\hat{\boldsymbol{\Omega}}^{\S}_{Ky2}\hat{\boldsymbol{\Sigma}}_{Ky2} = \begin{pmatrix} \hat{\Omega}^{(\mathrm{I})}_{Ky2} - \frac{\hat{\mathcal{C}}^2_{Ky2}}{\hat{\Omega}^{(\mathrm{II})}_{Ky2}} & 0 \\ 0 & 0 \end{pmatrix} \times \frac{1}{\hat{\Omega}^{(\mathrm{I})}_{Ky2}\hat{\Omega}^{(\mathrm{II})}_{Ky2} - \hat{\mathcal{C}}^2_{Ky2}} \begin{pmatrix} \hat{\Omega}^{(\mathrm{II})}_{Ky2} & -\hat{\mathcal{C}}_{Ky2} \\ -\hat{\mathcal{C}}_{Ky2} & \hat{\Omega}^{(\mathrm{I})}_{Ky2} \end{pmatrix} = \begin{pmatrix} 1 & -\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}^{(\mathrm{II})}_{Ky2}} \\ 0 & 0 \end{pmatrix}$$
$$= \boldsymbol{E}_{11} - \frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}.$$

Plug (A.1), (A.16), (A.17) and (A.18) back into (A.15) and we have

$$\begin{aligned}\rho(\bar{\boldsymbol{x}}_2 + \widetilde{\boldsymbol{x}}_2) = {} & \frac{1}{r^2}(r\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\Lambda}\bar{\hat{\boldsymbol{\Omega}}}_{y2}(r\boldsymbol{I} - \boldsymbol{G})'[\boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{d}_1 - \boldsymbol{p}_2 r] \\ & + \frac{1}{r}(r\boldsymbol{I} - \boldsymbol{G})\left(\boldsymbol{\Lambda} - \lambda_K \frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{\Omega}_{p2}\boldsymbol{C}_2^{-1}(\boldsymbol{p}_2 - \boldsymbol{A}_2 - \boldsymbol{B}_2\boldsymbol{d}_1) \\ & + \frac{1}{r}(r\boldsymbol{I} - \boldsymbol{G})[(\lambda_I\beta_I^2\boldsymbol{M}_{I2} + \lambda_K\beta^2\hat{\boldsymbol{\Omega}}^{\S}_{Ky2}\hat{\boldsymbol{\Sigma}}_{Ky2}\boldsymbol{M}_{K2})\boldsymbol{\Omega}_e + \lambda_1\boldsymbol{\Omega}_\epsilon]\widetilde{\boldsymbol{y}}_t,\end{aligned}$$

where $\lambda_1 = \lambda_I + \lambda_J$, $\lambda_2 = \lambda_I + \lambda_J + \lambda_K$, $\boldsymbol{\Lambda} = \left(\begin{smallmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{smallmatrix}\right)$, and $\bar{\hat{\boldsymbol{\Omega}}}_{y2} \equiv \boldsymbol{\Lambda}^{-1}(\lambda_I\hat{\boldsymbol{\Omega}}_{Iy2} + \lambda_J\hat{\boldsymbol{\Omega}}_{Jy2} + \lambda_K\hat{\boldsymbol{\Omega}}^{\S}_{Ky2})$. Moving all $\boldsymbol{p}_2$'s to the left-hand side and the rest to the right-hand side, one has

$$\begin{aligned} & \left[\boldsymbol{I} - (r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left(\boldsymbol{I} - \frac{\lambda_K\hat{\mathcal{C}}_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{\Omega}_{p2}\boldsymbol{C}_2^{-1}\right]\boldsymbol{p}_2 \\ = {} & \frac{1}{r}[\boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{d}_1] \\ & - (r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left(\boldsymbol{I} - \frac{\lambda_K\hat{\mathcal{C}}_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{\Omega}_{p2}\boldsymbol{C}_2^{-1}(\boldsymbol{A}_2 + \boldsymbol{B}_2\boldsymbol{d}_1) \\ & + (r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\boldsymbol{\Lambda}^{-1}[(\lambda_I\beta_I^2\boldsymbol{M}_{I2} + \lambda_K\beta^2\hat{\boldsymbol{\Omega}}^{\S}_{Ky2}\hat{\boldsymbol{\Sigma}}_{Ky2}\boldsymbol{M}_{K2})\boldsymbol{\Omega}_e + \lambda_1\boldsymbol{\Omega}_\epsilon]\widetilde{\boldsymbol{y}}_t \\ & - \rho r(r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\boldsymbol{\Lambda}^{-1}(r\boldsymbol{I} - \boldsymbol{G})^{-1}(\bar{\boldsymbol{x}}_2 + \widetilde{\boldsymbol{x}}_2),\end{aligned}$$

We then comparing with $\boldsymbol{p}_2 = \boldsymbol{A}_2 + \boldsymbol{B}_2\boldsymbol{d}_1 + \boldsymbol{C}_2\widetilde{\boldsymbol{y}}_2 + \boldsymbol{D}_2\widetilde{\boldsymbol{x}}_2$ to determine the coefficients.

$\underline{\boldsymbol{A}_2}$: Collect the constant terms we have

$$\begin{aligned}\boldsymbol{A}_2 = &\left[\boldsymbol{I} - (r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left(\boldsymbol{I} - \frac{\lambda_K\hat{\mathcal{C}}_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{\Omega}_{p2}\boldsymbol{C}_2^{-1}\right]^{-1} \times \\ &\left[\frac{1}{r}\boldsymbol{\mu}_v + (r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{\mu} - (r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left(\boldsymbol{I} - \frac{\lambda_K\hat{\mathcal{C}}_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{\Omega}_{p2}\boldsymbol{C}_2^{-1}\boldsymbol{A}_2\right. \\ &\left.-\rho r(r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\boldsymbol{\Lambda}^{-1}(r\boldsymbol{I}-\boldsymbol{G})^{-1}\bar{\boldsymbol{x}}_2\right].\end{aligned}$$

Denote $\boldsymbol{H} = (r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left(\boldsymbol{I} - \frac{\lambda_K\hat{\mathcal{C}}^2_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{\Omega}_{p2}\boldsymbol{C}_2^{-1}$ for readability. Moving all the $\boldsymbol{A}_2$ terms to the left-hand side, and , we have

$$\boldsymbol{A}_2 = \left[\boldsymbol{I} + (\boldsymbol{I}-\boldsymbol{H})^{-1}\boldsymbol{H}\right]^{-1}[\boldsymbol{I}-\boldsymbol{H}]^{-1}\left[\frac{1}{r}\boldsymbol{\mu}_v + (r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{\mu} - \frac{\rho r}{\lambda_2}(r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}(r\boldsymbol{I}-\boldsymbol{G})^{-1}\bar{\boldsymbol{x}}_2\right].$$

Note that

$$\begin{aligned}\left[\boldsymbol{I} + (\boldsymbol{I}-\boldsymbol{H})^{-1}\boldsymbol{H}\right]^{-1}[\boldsymbol{I}-\boldsymbol{H}]^{-1} = \boldsymbol{I} &\iff \left[\boldsymbol{I} + (\boldsymbol{I}-\boldsymbol{H})^{-1}\boldsymbol{H}\right]^{-1} = \boldsymbol{I}-\boldsymbol{H} \\ &\iff \boldsymbol{I} = \boldsymbol{I}-\boldsymbol{H}+\boldsymbol{H} = \boldsymbol{I}.\end{aligned}$$

Therefore,

$$\boldsymbol{A}_2 = \frac{1}{r}\boldsymbol{\mu}_v + (r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{\mu} - \rho r(r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\boldsymbol{\Lambda}^{-1}(r\boldsymbol{I}-\boldsymbol{G})^{-1}\bar{\boldsymbol{x}}_2.$$

$\underline{\boldsymbol{B}_2}$: Collect all the terms on $\boldsymbol{d}_1$, we have

$$\boldsymbol{B}_2 = [\boldsymbol{I}-\boldsymbol{H}]^{-1}\left[(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{G} - \boldsymbol{H}\boldsymbol{B}_2\right].$$

Moving all $\boldsymbol{B}_2$ to the left-hand side, we have

$$\begin{aligned}\boldsymbol{B}_2 &= \left[\boldsymbol{I} + (\boldsymbol{I}-\boldsymbol{H})^{-1}\boldsymbol{H}\right]^{-1}[\boldsymbol{I}-\boldsymbol{H}]^{-1}(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{G} \\ &= (r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{G}.\end{aligned}$$

$\underline{\boldsymbol{C}_2}$: The coefficient on $\widetilde{\boldsymbol{y}}_2$ is

$$\begin{aligned}\boldsymbol{C}_2 = &\left[\boldsymbol{I} - (r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left(\boldsymbol{I} - \frac{\lambda_K\hat{\mathcal{C}}_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{\Omega}_{p2}\boldsymbol{C}_2^{-1}\right]^{-1}\times \\ &(r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\boldsymbol{\Lambda}^{-1}\left[\left(\lambda_I\beta_I^2\boldsymbol{M}_{I2} + \lambda_K\beta^2\left(\boldsymbol{E}_{11} - \frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{M}_{K2}\right)\boldsymbol{\Omega}_e + \lambda_1\boldsymbol{\Omega}_\epsilon\right].\end{aligned}$$

Left multiplying $\left[\boldsymbol{I} - (r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left(\boldsymbol{I} - \frac{\lambda_K\hat{\mathcal{C}}_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{\Omega}_{p2}\boldsymbol{C}_2^{-1}\right]$ on both sides leads to

$$\begin{aligned}\boldsymbol{C}_2 = &(r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\times \\ &\boldsymbol{\Lambda}^{-1}\left[\left(\boldsymbol{\Lambda} - \lambda_K\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{\Omega}_{p2} + \left(\lambda_I\beta_I^2\boldsymbol{M}_{I2} + \lambda_K\beta^2\left(\boldsymbol{E}_{11} - \frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{12}\right)\boldsymbol{M}_{K2}\right)\boldsymbol{\Omega}_e + \lambda_1\boldsymbol{\Omega}_\epsilon\right].\end{aligned}$$

On the other hand,

$$\begin{aligned}\bar{\hat{\boldsymbol{\Omega}}}_{y2} &= \boldsymbol{\Lambda}^{-1}\left(\lambda_I\hat{\boldsymbol{\Omega}}_{Iy2} + \lambda_J\hat{\boldsymbol{\Omega}}_{Jy2} + \lambda_K\hat{\boldsymbol{\Omega}}^{\S}_{Ky2}\right) \\ &= \boldsymbol{\Lambda}^{-1}\Bigg[\lambda_I(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + \beta_I^2\boldsymbol{M}_{I2}\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon) + \lambda_J(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + \boldsymbol{\Omega}_\epsilon) \\ &\quad + \lambda_K\left(\boldsymbol{E}_{11}(\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + \beta^2\boldsymbol{M}_{K2}\boldsymbol{\Omega}_e) - \frac{\hat{\mathcal{C}}^2_{Ky2}}{\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{11}\right)\Bigg] \\ &= \boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + \boldsymbol{\Lambda}^{-1}(\lambda_I\beta_I^2\boldsymbol{M}_{I2} + \lambda_K\beta^2\boldsymbol{E}_{11}\boldsymbol{M}_{K2})\boldsymbol{\Omega}_e + \boldsymbol{\Lambda}^{-1}\lambda_1\boldsymbol{\Omega}_\epsilon - \frac{\lambda_K\hat{\mathcal{C}}^2_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{11}.\end{aligned}$$

Thus $\boldsymbol{\Omega}_{p2} + \boldsymbol{\Lambda}^{-1}(\lambda_I\beta_I^2\boldsymbol{M}_{I2} + \lambda_K\beta^2\boldsymbol{E}_{11}\boldsymbol{M}_{K2})\boldsymbol{\Omega}_e + \boldsymbol{\Lambda}^{-1}\lambda_1\boldsymbol{\Omega}_\epsilon = \bar{\hat{\boldsymbol{\Omega}}}_{y2} - \boldsymbol{\Omega}_y + \frac{\lambda_K\hat{\mathcal{C}}^2_{Ky2}}{\lambda_2\hat{\Omega}^{(\mathrm{II})}_{Ky2}}\boldsymbol{E}_{11}$. Plugging

it back, then

$$\boldsymbol{C}_2 = (r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left[\bar{\hat{\boldsymbol{\Omega}}}_{y2} - \boldsymbol{\Omega}_y + \frac{\lambda_K \hat{\mathcal{C}}_{Ky2}^2}{\lambda_2 \hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{11} - \frac{\lambda_K \hat{\mathcal{C}}_{Ky2}}{\lambda_2 \hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{12}(\boldsymbol{\Omega}_{p2} + \beta^2 \boldsymbol{M}_{K2}\boldsymbol{\Omega}_e)\right]$$

$\underline{\boldsymbol{D}_2}$: First note that the signal-to-noise ratio

$$\boldsymbol{\zeta}_2 \equiv \boldsymbol{D}_{\boldsymbol{2}}^{-1}\boldsymbol{C}_2 = -\frac{1}{\rho r}(r\boldsymbol{I} - \boldsymbol{G})\left[\left(\lambda_I \beta_I^2 \boldsymbol{M}_{I2} + \lambda_K \beta^2\left(\boldsymbol{E}_{11} - \frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{12}\right)\boldsymbol{M}_{K2}\right)\boldsymbol{\Omega}_e + \lambda_1 \boldsymbol{\Omega}_\epsilon\right]. \tag{A.19}$$

Thus $\boldsymbol{D}_2$ can be derived from $\boldsymbol{D}_2 = \boldsymbol{C}_2\boldsymbol{\zeta}_2^{-1}$.

In sum, the Period-2 price coefficients are

$$\boldsymbol{A}_2 = \frac{1}{r}\boldsymbol{\mu}_v + (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - \rho r(r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\boldsymbol{\Lambda}^{-1}(r\boldsymbol{I} - \boldsymbol{G})^{-1}\bar{\boldsymbol{x}}_2, \tag{A.20}$$

$$\boldsymbol{B}_2 = (r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}, \tag{A.21}$$

$$\boldsymbol{C}_2 = (r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}_{y2}\left[\bar{\hat{\boldsymbol{\Omega}}}_{y2} - \boldsymbol{\Omega}_y + \frac{\lambda_K \hat{\mathcal{C}}_{Ky2}^2}{\lambda_2 \hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{11} - \frac{\lambda_K \hat{\mathcal{C}}_{Ky2}}{\lambda_2 \hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{12}(\boldsymbol{\Omega}_{p2} + \beta^2 \boldsymbol{M}_{K2}\boldsymbol{\Omega}_e)\right] \tag{A.22}$$

$$\boldsymbol{D}_2 = \boldsymbol{C}_2\boldsymbol{\zeta}_2^{-1}. \tag{A.23}$$

Period 1

In Period 1, only domestic investors participate in the market. The price coefficients of type-I and type-II risky asset have the same form. Market clearing requires

$$\int_i \boldsymbol{q}_{i1}\,\mathrm{d}i + \int_j \boldsymbol{q}_{j1}\,\mathrm{d}j = \bar{\boldsymbol{x}}_1 + \widetilde{\boldsymbol{x}}_1. \tag{A.24}$$

We integrate domestic institutional investor i's portfolio choices:

$$\begin{aligned}\int_i \boldsymbol{q}_{i1}\,\mathrm{d}i &= \int_i \frac{1}{\rho r}\hat{\boldsymbol{\Omega}}_{I1}(\hat{\boldsymbol{\mu}}_{i1}-\rho\boldsymbol{\mathcal{C}}_I)\,\mathrm{d}i \\ &= \frac{1}{\rho r}\hat{\boldsymbol{\Omega}}_{I1}\left[(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)\int_i \hat{\boldsymbol{\mu}}_{iy1}\,\mathrm{d}i + \lambda_I(\boldsymbol{A}_2^\dagger+\boldsymbol{B}_2^\dagger(\boldsymbol{\mu}+\boldsymbol{G}\boldsymbol{d}_0)-\rho\boldsymbol{\mathcal{C}}_I-\boldsymbol{p}_1 r)\right]. \quad \text{(A.25)}\end{aligned}$$

Note that

$$\int_i \hat{\boldsymbol{\mu}}_{iy1}\,\mathrm{d}i = \hat{\boldsymbol{\Sigma}}_{Iy1}\int_i \boldsymbol{\Omega}_{p1}\boldsymbol{s}_{p1}+\beta_I^2\boldsymbol{\Omega}_e\boldsymbol{\iota}_{i1}\boldsymbol{s}_{i1}+\boldsymbol{\Omega}_\epsilon\boldsymbol{\xi}_{i1}\,\mathrm{d}i = \lambda_I\hat{\boldsymbol{\Sigma}}_{Iy1}[\boldsymbol{\Omega}_{p1}\boldsymbol{s}_{p1}+(\beta_I^2\boldsymbol{M}_{I1}\boldsymbol{\Omega}_e+\boldsymbol{\Omega}_\epsilon)\widetilde{\boldsymbol{y}}_1].$$

Thus

$$\begin{aligned}\int_i \boldsymbol{q}_{i1}\,\mathrm{d}i = &\frac{\lambda_I}{\rho r}\hat{\boldsymbol{\Omega}}_{I1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)\hat{\boldsymbol{\Sigma}}_{Iy1}[\boldsymbol{\Omega}_{p1}\boldsymbol{s}_{p1}+(\beta_I^2\boldsymbol{M}_{I1}\boldsymbol{\Omega}_e+\boldsymbol{\Omega}_\epsilon)\widetilde{\boldsymbol{y}}_1] \\ &+\frac{\lambda_I}{\rho r}\hat{\boldsymbol{\Omega}}_{I1}(\boldsymbol{A}_2^\dagger+\boldsymbol{B}_2^\dagger(\boldsymbol{\mu}+\boldsymbol{G}\boldsymbol{d}_0)-\rho\boldsymbol{\mathcal{C}}_I-\boldsymbol{p}_1 r).\end{aligned}$$

We can write $\hat{\tilde{\Sigma}}_{Iy1} = (\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1\prime}\hat{\boldsymbol{\Sigma}}_{I1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}-(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1\prime}(\boldsymbol{C}_2^\dagger\boldsymbol{\Sigma}_y\boldsymbol{C}_2^{\prime\dagger}+\boldsymbol{D}_2^\dagger\boldsymbol{\Sigma}_x\boldsymbol{D}_2^{\prime\dagger})(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}$. Then

$$\hat{\boldsymbol{\Omega}}_{I1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)\hat{\boldsymbol{\Sigma}}_{Iy1} = \left[\boldsymbol{I}-\hat{\boldsymbol{\Omega}}_{I1}\left(\boldsymbol{C}_2^\dagger\boldsymbol{\Sigma}_y\boldsymbol{C}_2^{\prime\dagger}+\boldsymbol{D}_2^\dagger\boldsymbol{\Sigma}_x\boldsymbol{D}_2^{\prime\dagger}\right)\right](\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}.$$

$\hat{\boldsymbol{\Omega}}_{I1}\left(\boldsymbol{C}_2^\dagger\boldsymbol{\Sigma}_y\boldsymbol{C}_2^{\prime\dagger}+\boldsymbol{D}_2^\dagger\boldsymbol{\Sigma}_x\boldsymbol{D}_2^{\prime\dagger}\right)$ is the proportion of the conditional variance of future noises in the total conditional variance of payoffs. Let me call it the *future uncertainty ratio* and denote it by $\hat{\boldsymbol{\Delta}}_{Lf1}$. Similarly, the *current uncertainty ratio* can be defined as the proportion of the conditional variance of current noises in the total conditional variance of payoffs and denoted by $\hat{\boldsymbol{\Delta}}_{Lc1}$. Then $\hat{\boldsymbol{\Delta}}_{Lc1} = \boldsymbol{I}-\hat{\boldsymbol{\Delta}}_{Lf1}$. Thus

$$\begin{aligned}\int_i \boldsymbol{q}_{i1}\,\mathrm{d}i = &\frac{\lambda_I}{\rho r}\hat{\boldsymbol{\Delta}}_{Ic1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}[\boldsymbol{\Omega}_{p1}\boldsymbol{s}_{p1}+(\beta_I^2\boldsymbol{M}_{I1}\boldsymbol{\Omega}_e+\boldsymbol{\Omega}_\epsilon)\widetilde{\boldsymbol{y}}_1] \\ &+\frac{\lambda_I}{\rho r}\hat{\boldsymbol{\Omega}}_{I1}(\boldsymbol{A}_2^\dagger+\boldsymbol{B}_2^\dagger(\boldsymbol{\mu}+\boldsymbol{G}\boldsymbol{d}_0)-\rho\boldsymbol{\mathcal{C}}_I-\boldsymbol{p}_1 r). \quad \text{(A.26)}\end{aligned}$$

The portfolio choices of domestic individual investors are aggregated to

$$\int_j \boldsymbol{q}_{j1}\,\mathrm{d}j = \frac{\lambda_J}{\rho r}\hat{\boldsymbol{\Delta}}_{Jc1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}(\boldsymbol{\Omega}_{p1}\boldsymbol{s}_{p1}+\boldsymbol{\Omega}_\epsilon\widetilde{\boldsymbol{y}}_1)+\frac{\lambda_J}{\rho r}\hat{\boldsymbol{\Omega}}_{J1}(\boldsymbol{A}_2^\dagger+\boldsymbol{B}_2^\dagger(\boldsymbol{\mu}+\boldsymbol{G}\boldsymbol{d}_0)-\rho\boldsymbol{\mathcal{C}}_J-\boldsymbol{p}_1 r). \tag{A.27}$$

The average current uncertainty ratio is $\bar{\hat{\boldsymbol{\Delta}}}_{c1}=\frac{1}{\lambda_1}(\lambda_I\hat{\boldsymbol{\Delta}}_{Ic1}+\lambda_J\hat{\boldsymbol{\Delta}}_{Jc1})$. Additionally, define the average precision of excess payoff to be $\bar{\hat{\boldsymbol{\Omega}}}_1=\frac{1}{\lambda_1}(\lambda_I\hat{\boldsymbol{\Omega}}_{I1}+\lambda_J\hat{\boldsymbol{\Omega}}_{J1})$. Plugging (A.26) and (A.27) into (A.24), one has

$$\begin{aligned}\bar{\boldsymbol{x}}_1+\widetilde{\boldsymbol{x}}_1=&\frac{1}{\rho r}\lambda_1\bar{\hat{\boldsymbol{\Delta}}}_{c1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}\boldsymbol{\Omega}_{p1}\boldsymbol{C}_1^{-1}(\boldsymbol{p}_1-\boldsymbol{A}_1-\boldsymbol{B}_1\boldsymbol{d}_0)\\&+\frac{1}{\rho r}\lambda_I\hat{\boldsymbol{\Delta}}_{Ic1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}\beta_I^2\boldsymbol{M}_{I1}\boldsymbol{\Omega}_\epsilon\widetilde{\boldsymbol{y}}_1+\frac{1}{\rho r}\lambda_1\bar{\hat{\boldsymbol{\Delta}}}_{c1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}\boldsymbol{\Omega}_\epsilon\widetilde{\boldsymbol{y}}_1\\&+\frac{1}{\rho r}\lambda_1\bar{\hat{\boldsymbol{\Omega}}}_1\left[\boldsymbol{A}_2^\dagger+\boldsymbol{B}_2^\dagger(\boldsymbol{\mu}+\boldsymbol{G}\boldsymbol{d}_0)-\boldsymbol{p}_1 r\right]-\frac{1}{r}\left(\lambda_I\hat{\boldsymbol{\Omega}}_{I1}\boldsymbol{\mathcal{C}}_I+\lambda_J\hat{\boldsymbol{\Omega}}_{J1}\boldsymbol{\mathcal{C}}_J\right).\end{aligned}$$

Moving all $\boldsymbol{p}_1$'s to the left-hand side and the rest to the right-hand side, one has

$$\begin{aligned}&\left[\boldsymbol{I}-\frac{1}{r}\bar{\hat{\boldsymbol{\Sigma}}}_1\bar{\hat{\boldsymbol{\Delta}}}_{c1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}\boldsymbol{\Omega}_{p1}\boldsymbol{C}_1^{-1}\right]\boldsymbol{p}_1\\=&-\frac{1}{r}\bar{\hat{\boldsymbol{\Sigma}}}_1\bar{\hat{\boldsymbol{\Delta}}}_{c1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}\boldsymbol{\Omega}_{p1}\boldsymbol{C}_1^{-1}(\boldsymbol{A}_1+\boldsymbol{B}_1\boldsymbol{d}_0)\\&+\frac{1}{r}\bar{\hat{\boldsymbol{\Sigma}}}_1\left[\hat{\boldsymbol{\Delta}}_{Ic1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}\frac{\lambda_I}{\lambda_1}\beta_I^2\boldsymbol{M}_{I1}\boldsymbol{\Omega}_e+\bar{\hat{\Delta}}_{c1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}\boldsymbol{\Omega}_\epsilon\right]\widetilde{\boldsymbol{y}}_1\\&+\frac{1}{r}\bar{\hat{\boldsymbol{\Sigma}}}_1\left[\bar{\hat{\boldsymbol{\Omega}}}_1(\boldsymbol{A}_2^\dagger+\boldsymbol{B}_2^\dagger(\boldsymbol{\mu}+\boldsymbol{G}\boldsymbol{d}_0))-\frac{\rho}{\lambda_1}(\lambda_I\hat{\boldsymbol{\Omega}}_{I1}\boldsymbol{\mathcal{C}}_I+\lambda_J\hat{\boldsymbol{\Omega}}_{J1}\boldsymbol{\mathcal{C}}_J)-\frac{\rho r}{\lambda_1}(\bar{\boldsymbol{x}}_1+\widetilde{\boldsymbol{x}}_1)\right].\end{aligned}$$

We then match the coefficients with $\boldsymbol{p}_1=\boldsymbol{A}_1+\boldsymbol{B}_1\boldsymbol{d}_0+\boldsymbol{C}_1\widetilde{\boldsymbol{y}}_1+\boldsymbol{D}_1\widetilde{\boldsymbol{x}}_1$.

$\underline{\boldsymbol{A}_1}$: Let $\boldsymbol{H}=\frac{1}{r}\bar{\hat{\boldsymbol{\Sigma}}}_1\bar{\hat{\boldsymbol{\Delta}}}_{c1}(\boldsymbol{I}+\boldsymbol{B}_2^\dagger)^{-1}\boldsymbol{\Omega}_{p1}\boldsymbol{C}_1^{-1}$. Then

$$\begin{aligned}\boldsymbol{A}_1=&\left[\boldsymbol{I}+(\boldsymbol{I}-\boldsymbol{H})^{-1}\boldsymbol{H}\right]^{-1}[\boldsymbol{I}-\boldsymbol{H}]^{-1}\\&\times\frac{1}{r}\bar{\hat{\boldsymbol{\Sigma}}}_1\left[\bar{\hat{\boldsymbol{\Omega}}}_1(\boldsymbol{A}_2^\dagger+\mu\boldsymbol{B}_2^\dagger)-\frac{\rho}{\lambda_1}(\lambda_I\hat{\boldsymbol{\Omega}}_{I1}\boldsymbol{\mathcal{C}}_I+\lambda_J\hat{\boldsymbol{\Omega}}_{J1}\boldsymbol{\mathcal{C}}_J)-\frac{\rho r}{\lambda_1}\bar{\boldsymbol{x}}_1\right]\\=&\frac{1}{r}\bar{\hat{\boldsymbol{\Sigma}}}_1\left[\bar{\hat{\boldsymbol{\Omega}}}_1(\boldsymbol{A}_2^\dagger+\mu\boldsymbol{B}_2^\dagger)-\frac{\rho}{\lambda_1}(\lambda_I\hat{\boldsymbol{\Omega}}_{I1}\boldsymbol{\mathcal{C}}_I+\lambda_J\hat{\boldsymbol{\Omega}}_{J1}\boldsymbol{\mathcal{C}}_J)-\frac{\rho r}{\lambda_1}\bar{\boldsymbol{x}}_1\right].\end{aligned}$$

$\underline{\boldsymbol{B}_1}$: Collect all the terms on $\boldsymbol{d}_0$, we have

$$\boldsymbol{B}_1 = [\boldsymbol{I} - \boldsymbol{H}]^{-1} \frac{1}{r} \bar{\hat{\boldsymbol{\Sigma}}}_1 \left(\bar{\hat{\boldsymbol{\Omega}}}_1 \boldsymbol{B}_2^{\dagger} \boldsymbol{G} - \bar{\hat{\boldsymbol{\Delta}}}_{c1} (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})^{-1} \boldsymbol{\Omega}_{p1} \boldsymbol{C}_1^{-1} \boldsymbol{B}_1 \right).$$

Moving all $\boldsymbol{B}_2$ terms to the left-hand side, we have

$$\boldsymbol{B}_1 = \left[\boldsymbol{I} + (\boldsymbol{I} - \boldsymbol{H})^{-1} \boldsymbol{H} \right]^{-1} [\boldsymbol{I} - \boldsymbol{H}]^{-1} \frac{1}{r} \bar{\hat{\boldsymbol{\Sigma}}}_1 \bar{\hat{\boldsymbol{\Omega}}}_1 \boldsymbol{B}_2^{\dagger} \boldsymbol{G} = \frac{1}{r} (r\boldsymbol{I} - \boldsymbol{G})^{-1} \boldsymbol{G}.$$

$\underline{\boldsymbol{C}_1}$: Collecting terms on $\widetilde{\boldsymbol{y}}_1$, we have

$$\begin{aligned} \boldsymbol{C}_1 = & \left[\boldsymbol{I} - \frac{1}{r} \bar{\hat{\boldsymbol{\Sigma}}}_1 \bar{\hat{\boldsymbol{\Delta}}}_{c1} (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})^{-1} \boldsymbol{\Omega}_{p1} \boldsymbol{C}_1^{-1} \right]^{-1} \\ & \times \frac{1}{r} \bar{\hat{\boldsymbol{\Sigma}}}_1 \left[\hat{\boldsymbol{\Delta}}_{Ic1} (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})^{-1} \frac{\lambda_I}{\lambda_1} \beta_I^2 \boldsymbol{M}_{I1} \boldsymbol{\Omega}_e + \bar{\hat{\Delta}}_{c1} (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})^{-1} \boldsymbol{\Omega}_{\epsilon} \right]. \end{aligned}$$

Multiplying $\left[\boldsymbol{I} - \frac{1}{r} \bar{\hat{\boldsymbol{\Sigma}}}_1 \bar{\hat{\boldsymbol{\Delta}}}_{c1} (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})^{-1} \boldsymbol{\Omega}_{p1} \boldsymbol{C}_1^{-1} \right]$ on both sides, we have

$$\boldsymbol{C}_1 = \frac{1}{r} \bar{\hat{\boldsymbol{\Sigma}}}_1 \left[\bar{\hat{\boldsymbol{\Delta}}}_{c1} (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})^{-1} \boldsymbol{\Omega}_{p1} + \hat{\boldsymbol{\Delta}}_{Ic1} (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})^{-1} \frac{\lambda_I}{\lambda_1} \beta_I^2 \boldsymbol{M}_{I1} \boldsymbol{\Omega}_e + \bar{\hat{\Delta}}_{c1} (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})^{-1} \boldsymbol{\Omega}_{\epsilon} \right].$$

The term in the square brackets can be interpreted as the market average posterior precision of payoff weighted by current uncertainty ratio that can be explained by current dividend innovation. Or alternatively,

$$\begin{aligned} \bar{\hat{\boldsymbol{\Omega}}}_1 &= \frac{1}{\lambda_1} (\lambda_I \hat{\boldsymbol{\Omega}}_{I1} + \lambda_J \hat{\boldsymbol{\Omega}}_{J1}) = \frac{1}{\lambda_1} \left[\lambda_I (\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p1} + \beta_I^2 \boldsymbol{M}_{I1} \boldsymbol{\Omega}_e + \boldsymbol{\Omega}_{\epsilon}) + \lambda_J (\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p1} + \boldsymbol{\Omega}_{\epsilon}) \right] \\ &= \boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p1} + \frac{\lambda_I}{\lambda_1} \beta_I^2 \boldsymbol{M}_{I1} \boldsymbol{\Omega}_e + \boldsymbol{\Omega}_{\epsilon}. \end{aligned}$$

Plugging it back, we have

$$\boldsymbol{C}_1 = \frac{1}{r} \bar{\hat{\boldsymbol{\Sigma}}}_1 \left[\bar{\hat{\boldsymbol{\Delta}}}_{c1} (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})^{-1} (\bar{\hat{\boldsymbol{\Omega}}}_1 - \boldsymbol{\Omega}_y) + (\bar{\hat{\boldsymbol{\Delta}}}_{Ic1} - \bar{\hat{\boldsymbol{\Delta}}}_{c1}) (\boldsymbol{I} + \boldsymbol{B}_2^{\dagger})^{-1} \frac{\lambda_I}{\lambda_1} \beta_I^2 \boldsymbol{M}_{I1} \boldsymbol{\Omega}_e \right].$$

$\underline{\boldsymbol{D}_1}$: From the price relation,

$$\boldsymbol{\zeta}_1 = \boldsymbol{D}_1^{-1}\boldsymbol{C}_1 = -\frac{1}{\rho r}\left[\hat{\boldsymbol{\Delta}}_{Ic1}(\boldsymbol{I}+\boldsymbol{B}_2^{\dagger})^{-1}\lambda_I\beta_I^2\boldsymbol{M}_{I1}\boldsymbol{\Omega}_e + \bar{\hat{\boldsymbol{\Delta}}}_{I1}(\boldsymbol{I}+\boldsymbol{B}_2^{\dagger})^{-1}\lambda_1\boldsymbol{\Omega}_\epsilon\right].$$

Then $\boldsymbol{D}_1$ can be derived from $\boldsymbol{D_1} = \boldsymbol{C_1}\boldsymbol{\zeta}_1^{-1}$.

To put the price coefficients together, we have

$$\boldsymbol{A}_1 = \frac{1}{r}\bar{\hat{\boldsymbol{\Sigma}}}_1\left[\bar{\hat{\boldsymbol{\Omega}}}_1(\boldsymbol{A}_2^{\dagger}+\mu\boldsymbol{B}_2^{\dagger}) - \frac{\rho}{\lambda_1}(\lambda_I\hat{\boldsymbol{\Omega}}_{I1}\boldsymbol{\mathcal{C}}_I + \lambda_J\hat{\boldsymbol{\Omega}}_{J1}\boldsymbol{\mathcal{C}}_J) - \frac{\rho r}{\lambda_1}\bar{\boldsymbol{x}}_1\right], \tag{A.28}$$

$$\boldsymbol{B}_1 = \frac{1}{r}(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{G}, \tag{A.29}$$

$$\boldsymbol{C}_1 = \frac{1}{r}\bar{\hat{\boldsymbol{\Sigma}}}_1\left[\bar{\hat{\boldsymbol{\Delta}}}_{c1}(\boldsymbol{I}+\boldsymbol{B}_2^{\dagger})^{-1}\boldsymbol{\Omega}_{p1} + \hat{\boldsymbol{\Delta}}_{Ic1}(\boldsymbol{I}+\boldsymbol{B}_2^{\dagger})^{-1}\frac{\lambda_I}{\lambda_1}\beta_I^2\boldsymbol{M}_{I1}\boldsymbol{\Omega}_e + \bar{\hat{\Delta}}_{c1}(\boldsymbol{I}+\boldsymbol{B}_2^{\dagger})^{-1}\boldsymbol{\Omega}_\epsilon\right], \tag{A.30}$$

$$\boldsymbol{D}_1 = \boldsymbol{C}_1\boldsymbol{\zeta}_1^{-1}. \tag{A.31}$$

A.1.4 Optimal data choice

In this section, I calculate the *ex ante* expected utility of the institutional investors at the beginning of each period, and then solve their data choice problems, that is to maximize the *ex ante* expected utility by allocating $\bar{M}$ buyable data points between internationally tradable and untradable firms. I proceed backwards and start from the Period-2 data choice problem.

Domestic institutional investors

Proof of Lemma 1.

$\underline{\text{Step 1: Calculate } \mathbb{E}\left[-e^{-\rho w_{i2}} \mid \mathcal{I}_{i2}\right]}$

The interim utility for institutional investor l is

$$\begin{aligned}\mathbb{E}\left[-e^{-\rho w_{i2}} \mid \mathcal{I}_{i2}\right] &= -e^{-\rho r w_{i1}}\mathbb{E}\left[e^{-(\mathbb{E}[\widetilde{\boldsymbol{d}}_2+\widetilde{\boldsymbol{v}}_2-\boldsymbol{p}_2 r|\mathcal{I}_{i2}])'\mathbb{V}ar^{-1}(\widetilde{\boldsymbol{d}}_2+\widetilde{\boldsymbol{v}}_2|\mathcal{I}_{i2})(\widetilde{\boldsymbol{d}}_2+\widetilde{\boldsymbol{v}}_2-\boldsymbol{p}_2 r)} \mid \mathcal{I}_{i2}\right] \\ &= -e^{-\rho r w_{i1}} \times e^{-\frac{1}{2}\hat{\boldsymbol{\mu}}_{i2}'\hat{\boldsymbol{\Sigma}}_{i2}^{-1}\hat{\boldsymbol{\mu}}_{i2}} \\ &= -e^{-\rho r^2 w_{i0}} \exp\left\{-\frac{1}{2}\hat{\boldsymbol{\mu}}_{i2}'\hat{\boldsymbol{\Sigma}}_{i2}^{-1}\hat{\boldsymbol{\mu}}_{i2} - \rho r \boldsymbol{q}_{i1}'(\boldsymbol{d}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1 r)\right\},\end{aligned}$$

where $\hat{\boldsymbol{\mu}}_{i2} \equiv \mathbb{E}[\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r \mid \mathcal{I}_{i2}] = \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}(\boldsymbol{\mu} + \boldsymbol{G}\boldsymbol{d}_1) + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\mu}}_{iy2} - \boldsymbol{p}_2 r$ and $\hat{\boldsymbol{\Sigma}}_{i2}$ is a function of known model parameters. Note that before making data choice in Period 2, $-\frac{1}{2}\hat{\boldsymbol{\mu}}_{i2}'\hat{\boldsymbol{\Sigma}}_{i2}^{-1}\hat{\boldsymbol{\mu}}_{i2}$ is a quadratic function of normal variables $\boldsymbol{p}_2$ (also $\boldsymbol{s}_{i2}$ and $\boldsymbol{\xi}_{i2}$ for domestic institutional investors), while $-\rho r \boldsymbol{q}_{i1}'(\boldsymbol{d}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1 r)$ is a linear function of normal variables. It is useful to write them in terms of the same zero-mean normal variable. Before that, we need an intermediate step to take the expectation over private signals $\boldsymbol{s}_{i2}$ and $\boldsymbol{\xi}_{i2}$.

Step 2: Calculate $\mathbb{E}\left[-e^{-\rho w_{i2}} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right]$

Conditional on $\mathcal{I}_{i2^-}$ and $\boldsymbol{p}_2$, w_{i1} is non-random, and $\hat{\boldsymbol{\mu}}_{i2}$ is random with $\boldsymbol{e}_{i2}$. (Note that the inalienable information $\boldsymbol{\xi}_{i2}$ is known to domestic institutional investors under $\mathcal{I}_{i2^-}$). By law of iterated expectations, one has

$$\mathbb{E}\left[-e^{-\rho w_{i2}} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right] = \mathbb{E}\left[\mathbb{E}\left[-e^{-\rho w_{i2}} \mid \mathcal{I}_{i2}\right] \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right] = -e^{-\rho r w_{i1}} \times \mathbb{E}\left[e^{-\frac{1}{2}\hat{\boldsymbol{\mu}}_{i2}'\hat{\boldsymbol{\Sigma}}_{i2}^{-1}\hat{\boldsymbol{\mu}}_{i2}} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right].$$

Let $\boldsymbol{\Sigma}_{i2}^p \equiv \mathbb{V}ar(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2)$. Conditional on $\mathcal{I}_{i2^-}$ and $\boldsymbol{p}_2$, $\hat{\boldsymbol{\mu}}_{i2}$ is still normally distributed. By law of total variance, $\boldsymbol{\Sigma}_{i2}^p = \mathbb{E}\left[\hat{\boldsymbol{\Sigma}}_{i2} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right] + \mathbb{V}ar\left(\hat{\boldsymbol{\mu}}_{i2} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right)$. Since $\hat{\boldsymbol{\Sigma}}_{i2}$ is $\mathcal{I}_{i2^-}$-measurable, one has that $\mathbb{V}ar(\hat{\boldsymbol{\mu}}_{i2} - \boldsymbol{\mu}_{i2}^p \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2) = \mathbb{V}ar\left(\hat{\boldsymbol{\mu}}_{i2} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right) = \boldsymbol{\Sigma}_{i2}^p - \hat{\boldsymbol{\Sigma}}_{i2}$. To work out $\mathbb{E}\left[e^{-\frac{1}{2}\hat{\boldsymbol{\mu}}_{i2}'\hat{\boldsymbol{\Sigma}}_{i2}^{-1}\hat{\boldsymbol{\mu}}_{i2}} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right]$, it is useful to write it in terms of the mean-zero random variable $\hat{\boldsymbol{\mu}}_{i2} - \boldsymbol{\mu}_{i2}^p$:

$$\mathbb{E}\left[e^{-\frac{1}{2}\hat{\boldsymbol{\mu}}_{i2}'\hat{\boldsymbol{\Sigma}}_{i2}^{-1}\hat{\boldsymbol{\mu}}_{i2}} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right] = \mathbb{E}\left[e^{-\frac{1}{2}\left(\hat{\boldsymbol{\mu}}_{i2}-\boldsymbol{\mu}_{i2}^p\right)'\hat{\boldsymbol{\Sigma}}_{i2}^{-1}\left(\hat{\boldsymbol{\mu}}_{i2}-\boldsymbol{\mu}_{i2}^p\right)-\boldsymbol{\mu}_{i2}^{p\prime}\hat{\boldsymbol{\Sigma}}_{i2}^{-1}\left(\hat{\boldsymbol{\mu}}_{i2}-\boldsymbol{\mu}_{i2}^p\right)-\frac{1}{2}\boldsymbol{\mu}_{i2}^{p\prime}\hat{\boldsymbol{\Sigma}}_{i2}^{-1}\boldsymbol{\mu}_{i2}^p} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right].$$

Therefore,

$$\mathbb{E}\left[e^{-\frac{1}{2}\hat{\boldsymbol{\mu}}'_{i2}\hat{\boldsymbol{\Sigma}}_{i2}\hat{\boldsymbol{\mu}}_{i2}} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right] = \left|\boldsymbol{\Sigma}^p_{i2}\hat{\boldsymbol{\Sigma}}^{-1}_{i2}\right|^{-1/2} \exp\left\{-\frac{1}{2}\boldsymbol{\mu}^{p\prime}_{i2}\boldsymbol{\Sigma}^{p2}_{i2}\boldsymbol{\mu}^p_{i2}\right\}.$$

Step 3: Calculate $\mathbb{E}\left[-e^{-\rho w_{i2}} \mid \mathcal{I}_{i2^-}\right]$

Notice that both w_{i1} and $\boldsymbol{\mu}^p_{i2}$ are random with respect to $\mathcal{I}^+_{i1}$ because of the random component $\boldsymbol{p}_2$. By law of iterated expectations, one has

$$\begin{aligned}\mathbb{E}\left[-e^{-\rho w_{i2}} \mid \mathcal{I}_{i2^-}\right] &= \mathbb{E}\left[\mathbb{E}\left[-e^{-\rho w_{i2}} \mid \mathcal{I}_{i2^-}, \boldsymbol{p}_2\right] \mid \mathcal{I}_{i2^-}\right] \\ &= -e^{-\rho r^2 w_{i0}} \\ &\quad \times\left|\boldsymbol{\Sigma}^p_{i2}\hat{\boldsymbol{\Sigma}}^{-1}_{i2}\right|^{-1/2}\mathbb{E}\left[\exp\left\{-\frac{1}{2}\boldsymbol{\mu}^{p\prime}_{i2}(\boldsymbol{\Sigma}^p_{i2})^{-1}\boldsymbol{\mu}^p_{i2} - \rho r\boldsymbol{q}'_{i1}(\boldsymbol{d}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1 r)\right\} \mid \mathcal{I}_{i2^-}\right].\end{aligned}$$

Let $\boldsymbol{\Sigma}_{i2} \equiv \mathbb{V}ar(\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r \mid \mathcal{I}_{i2^-})$. Conditional on $\mathcal{I}_{i2^-}$, $\boldsymbol{\mu}^p_{i2}$ is still normally distributed. By law of total variance, $\boldsymbol{\Sigma}_{i2} = \mathbb{E}\left[\boldsymbol{\Sigma}^p_{i2} \mid \mathcal{I}_{i2^-}\right] + \mathbb{V}ar\left(\boldsymbol{\mu}^p_{i2} \mid \mathcal{I}_{i2^-}\right)$. Since $\boldsymbol{\Sigma}^p_{i2}$ is $\mathcal{I}_{i2^-}$-measurable, one has that $\mathbb{V}ar\left(\boldsymbol{\mu}^p_{i2} \mid \mathcal{I}_{i2^-}\right) = \mathbb{V}ar\left(\boldsymbol{\mu}^p_{i2} \mid \mathcal{I}_{i2^-}\right) = \boldsymbol{\Sigma}_{i2} - \boldsymbol{\Sigma}^p_{i2}$, which is the variance of $\boldsymbol{\mu}^p_{i2} - \boldsymbol{\mu}_{i2}$ conditional on $\mathcal{I}_{i2^-}$. To work out $E\left[\exp\left\{-\frac{1}{2}\boldsymbol{\mu}^{p\prime}_{i2}(\boldsymbol{\Sigma}^p_{i2})^{-1}\boldsymbol{\mu}^p_{i2} - \rho r\boldsymbol{q}'_{i1}(\boldsymbol{d}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1 r)\right\} \mid \mathcal{I}_{i2^-}\right]$, it is useful to write it in terms of the mean-zero random variable $\boldsymbol{\mu}^p_{i2} - \boldsymbol{\mu}_{i2}$. Recall that $\boldsymbol{\mu}^p_{i2} = \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{d}_1 + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu}^p_{iy2} - \boldsymbol{p}_2 r$ and $\boldsymbol{\mu}_{i2} = \boldsymbol{\mu}_v + r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\mu} - r\boldsymbol{A}_2 + (r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G} - r\boldsymbol{B}_2)\boldsymbol{d}_1 + (r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G} - r\boldsymbol{C}_2)\boldsymbol{\mu}_{iy2}$. The constructed zero-mean normal variable is $\boldsymbol{\mu}^p_{i2} - \boldsymbol{\mu}_{i2} = r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{\Sigma}^p_{Ly2}\boldsymbol{\Omega}_{p2}\boldsymbol{C}_2^{-1}(\boldsymbol{p}_2 - \boldsymbol{A}_2 - \boldsymbol{B}_2\boldsymbol{d}_1) - \boldsymbol{p}_2 r + r\boldsymbol{A}_2 + r\boldsymbol{B}_2\boldsymbol{d}_1 - (r(r\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G} - r\boldsymbol{C}_2)\boldsymbol{\mu}_{iy2}$.

$$\boldsymbol{\mu}^p_{i2} - \boldsymbol{\mu}_{i2} = \mathcal{M}_1\boldsymbol{p}_2 - \mathcal{M}_1(\boldsymbol{A}_2 + \boldsymbol{B}_2\boldsymbol{d}_1),$$

where $\mathcal{M} = r(r\boldsymbol{I} - \boldsymbol{G})^{-1}(\boldsymbol{\Sigma}_y + \boldsymbol{\Sigma}_{p2})^{-1}\boldsymbol{\Sigma}_{p2}\boldsymbol{C}_2^{-1} - rI$. Therefore

$$
\begin{aligned}
&\mathbb{E}\left[e^{-\rho r w_{i1} - \frac{1}{2}\boldsymbol{\mu}_{i2}^{p\prime}(\boldsymbol{\Sigma}_{i2}^{p})^{-1}\boldsymbol{\mu}_{i2}^{p\prime}} \mid \mathcal{I}_{i2^-}\right] \\
&= \mathbb{E}\left[e^{-\frac{1}{2}\left(\boldsymbol{\mu}_{i2}^{p}-\boldsymbol{\mu}_{i2}\right)'(\boldsymbol{\Sigma}_{i2}^{p})^{-1}\left(\boldsymbol{\mu}_{i2}^{p}-\boldsymbol{\mu}_{i2}\right)-\boldsymbol{\mu}_{i2}'(\boldsymbol{\Sigma}_{i2}^{p})^{-1}\left(\boldsymbol{\mu}_{i2}^{p}-\boldsymbol{\mu}_{i2}\right)-\frac{1}{2}\boldsymbol{\mu}_{i2}'(\boldsymbol{\Sigma}_{i2}^{p})^{-1}\boldsymbol{\mu}_{i2}-\rho r(w_{i0}+\boldsymbol{q}_{i1}'(\boldsymbol{d}_1+\boldsymbol{p}_2-\boldsymbol{p}_1 r))} \mid \mathcal{I}_{i2^-}\right] \\
&= e^{-\rho r w_{i0}-\rho r\boldsymbol{q}_{i1}'(\boldsymbol{d}_1-\boldsymbol{p}_1 r)} \\
&\quad\times\mathbb{E}\left[e^{-\frac{1}{2}\left(\boldsymbol{\mu}_{i2}^{p}-\boldsymbol{\mu}_{i2}\right)'(\boldsymbol{\Sigma}_{i2}^{p})^{-1}\left(\boldsymbol{\mu}_{i2}^{p}-\boldsymbol{\mu}_{i2}\right)-\left(\boldsymbol{\mu}_{i2}'(\boldsymbol{\Sigma}_{i2}^{p})^{-1}+\rho r\boldsymbol{q}_{i1}'\mathcal{M}^{-1}\right)\left(\boldsymbol{\mu}_{i2}^{p}-\boldsymbol{\mu}_{i2}\right)-\frac{1}{2}\boldsymbol{\mu}_{i2}'(\boldsymbol{\Sigma}_{i2}^{p})^{-1}\boldsymbol{\mu}_{i2}} \mid \mathcal{I}_{i2^-}\right] \\
&= e^{-\rho r w_{i0}-\rho r\boldsymbol{q}_{i1}'(\boldsymbol{d}_1-\boldsymbol{p}_1 r)}\left|\boldsymbol{\Sigma}_{i2}\left(\boldsymbol{\Sigma}_{i2}^{p}\right)^{-1}\right|^{-1/2}\times \\
&\quad e^{-\frac{1}{2}\boldsymbol{\mu}_{i2}'\boldsymbol{\Sigma}_{i2}^{-1}\boldsymbol{\mu}_{i2}+\boldsymbol{\mu}_{i2}'(I-\boldsymbol{\Sigma}_{i2}^{-1}\boldsymbol{\Sigma}_{i2}^{p})\rho r(\mathcal{M}^{-1})'\boldsymbol{q}_{i1}+\frac{1}{2}\rho^2 r^2\boldsymbol{q}_{i1}'\mathcal{M}^{-1}(\boldsymbol{\Sigma}_{i2}^{p}-\boldsymbol{\Sigma}_{i2}^{p}\boldsymbol{\Sigma}_{i2}^{-1}\boldsymbol{\Sigma}_{i2}^{p})(\mathcal{M}^{-1})'\boldsymbol{q}_{i1}},
\end{aligned}
$$

and thus the *ex ante* utility is

$$
\begin{aligned}
\mathbb{E}\left[e^{-\rho r w_{i2}} \mid \mathcal{I}_{i2^-}\right] = -\left|\boldsymbol{\Sigma}_{i2}\hat{\boldsymbol{\Sigma}}_{i2}^{-1}\right|^{-1/2} & e^{-\rho r W_{i0}-\rho r\boldsymbol{q}_{i1}'(\boldsymbol{d}_1-\boldsymbol{p}_1 r)-\rho r\boldsymbol{q}_{i1}'(\boldsymbol{A}_2+\boldsymbol{B}_2\boldsymbol{d}_1)} \\
&\times e^{-\frac{1}{2}\boldsymbol{\mu}_{i2}'\boldsymbol{\Sigma}_{i2}\boldsymbol{\mu}_{i2}+\boldsymbol{\mu}_{i2}'(I-\boldsymbol{\Sigma}_{i2}\boldsymbol{\Sigma}_{i2}^{p})\rho r(\mathcal{M}^{-1})'\boldsymbol{q}_{i1}+\frac{1}{2}\rho^2 r^2\boldsymbol{q}_{i1}'\mathcal{M}^{-1}(\boldsymbol{\Sigma}_{i2}^{p}-\boldsymbol{\Sigma}_{i2}^{p}\boldsymbol{\Sigma}_{i2}\boldsymbol{\Sigma}_{i2}^{p})(\mathcal{M}^{-1})'\boldsymbol{q}_{i1}}.
\end{aligned}
$$

□

Since the data choice $(M_{i2}^{(\mathrm{I})}, M_{i2}^{(\mathrm{II})})$ only enter the *ex ante* utility function though the posterior variance of payoff ($\hat{\boldsymbol{\Sigma}}_{i2}^{-1}$), the data choice problem reduces to allocating data points to maximize $\left|\hat{\boldsymbol{\Omega}}_{iy2}\right|$. Since $\hat{\boldsymbol{\Omega}}_{iy2} = \boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + \beta_I^2\boldsymbol{M}_{i2}\boldsymbol{\Omega}_e + \boldsymbol{\Omega}_\epsilon$, the data choices only enter through the $\beta_I^2\boldsymbol{M}_{i2}\boldsymbol{\Omega}_e$ term. In fact, the objective function can be further reduced to the product of the diagonal elements of the posterior precision matrix for dividend innovation, that is $\hat{\Omega}_{iy2}^{(\mathrm{I})}\hat{\Omega}_{iy2}^{(\mathrm{II})}$. To see this, let me write out the elements of $\hat{\boldsymbol{\Omega}}_{iy2}$.

$$
\hat{\boldsymbol{\Omega}}_{iy2} = \begin{pmatrix} \Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + \beta_I^2 M_{i2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})} & \mathcal{C}_y + \hat{\mathcal{C}}_{p2} \\ \mathcal{C}_y + \hat{\mathcal{C}}_{p2} & \Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + \beta_I^2 M_{i2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})} \end{pmatrix}.
$$

Then $\left|\hat{\boldsymbol{\Omega}}_{iy2}\right|$ evaluates to $(\Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + \beta_I^2 M_{i2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})})(\Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + \beta_I^2 M_{i2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}) -$

$(\mathcal{C}_y + \hat{\mathcal{C}}_{p2})^2$. The second term is irrelevant to investor i's data choice, which finally reduces to:

$$\max_{\{M_{i2}^{(\mathrm{I})}, M_{i2}^{(\mathrm{II})}\}} \quad (\Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + \beta_I^2 M_{i2}^{(\mathrm{I})} \Omega_e^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})})(\Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + \beta_I^2 M_{i2}^{(\mathrm{II})} \Omega_e^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}) \tag{A.32}$$

$$\text{s.t.} \quad M_{i2}^{(\mathrm{I})} + M_{i2}^{(\mathrm{II})} \leq \bar{M}. \tag{A.33}$$

Since it is always to acquire and process more data points, data constraint (A.33) will be satisfied with equality. Substituting $M_{i2}^{(\mathrm{II})} = \bar{M} - M_{i2}^{(\mathrm{I})}$ into the objective function (A.32), the F.O.C. is

$$\frac{\beta_I^2 \Omega_e^{(\mathrm{I})}}{\hat{\Omega}_{iy2}^{(\mathrm{I})}} = \frac{\beta_I^2 \Omega_e^{(\mathrm{II})}}{\hat{\Omega}_{iy2}^{(\mathrm{II})}}.$$

The amount of data points allocated to type-I asset is solved to be

$$M_{i2}^c = \frac{1}{2}\left[\frac{\Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2 \Omega_e^{(\mathrm{II})}} + \bar{M} - \frac{\Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2 \Omega_e^{(\mathrm{I})}}\right].$$

We then compare it with the boundaries 0 and $\bar{M}$. We have

$$\begin{aligned} M_{i2}^c \leq 0 &\iff \frac{\Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2 \Omega_e^{(\mathrm{I})}} - \frac{\Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2 \Omega_e^{(\mathrm{II})}} \geq \bar{M}, \\ M_{i2}^c \geq \bar{M} &\iff \frac{\Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2 \Omega_e^{(\mathrm{II})}} - \frac{\Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2 \Omega_e^{(\mathrm{I})}} \geq \bar{M}. \end{aligned}$$

It is more interpretable to rearrange the conditions as

$$\begin{aligned} M_{i2}^c \leq 0 &\iff \frac{\Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2 \Omega_e^{(\mathrm{I})}} \geq \frac{\Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + \bar{M}\beta_I^2 \Omega_e^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2 \Omega_e^{(\mathrm{II})}}, \\ M_{i2}^c \geq \bar{M} &\iff \frac{\Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2 \Omega_e^{(\mathrm{II})}} \geq \frac{\Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + \bar{M}\beta_I^2 \Omega_e^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2 \Omega_e^{(\mathrm{I})}}. \end{aligned}$$

Intuitively, if the amount of data needed to achieve the entire posterior precision of dividend innovation for type-II risky asset is still less than the equivalent amount of data to recreate the posterior precision only observing the equilibrium price for type-I risky asset, then domestic

institutional investor i will opt to not learn about type-I asset and direct all her data processing resources to type-II asset, so $M_{i2}^* = 0$. Conversely, if the amount of data needed to achieve the entire posterior precision of dividend innovation for type-I risky asset is more than the equivalent amount of data to match the posterior precision only observing the equilibrium price for type-II risky asset, then domestic institutional investor i will optimally only learn about type-I asset, so $M_{i2}^* = \bar{M}$.

In sum, the solution is

$$M_{i2}^* = \begin{cases} 0, & \text{if } \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}} \geq \frac{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\bar{M}\beta_I^2\Omega_e^{(\mathrm{II})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}}, \\ \bar{M}, & \text{if } \frac{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}} \geq \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\bar{M}\beta_I^2\Omega_e^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}}, \\ \frac{1}{2}\left[\frac{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}} + \bar{M} - \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}}\right], & \text{otherwise.} \end{cases}$$

Foreign institutional investors

Foreign institutional investor k faces a slightly different data choice problem. She is only allowed to invest in type-I risky asset, and her optimal portfolio choice is determined by her posterior expectation and variance of the payoff, but this does not naturally suggest that information on type-II risky asset could not help with her decision since the two risky assets are correlated. Her interim utility

$$\begin{aligned} \mathbb{E}[-e^{-\rho w_{k2}} \mid \mathcal{I}_{k2}] &= -e^{-\rho r w_{k1}} \times \mathbb{E}\left[\exp\left\{-\frac{\hat{\mu}_{k2}^{(\mathrm{I})}}{\hat{\sigma}_{k2}^{2(\mathrm{I})}}\left(\widetilde{d}_2^{(\mathrm{I})} + \widetilde{v}_2^{(\mathrm{I})} - p_2^{(\mathrm{I})} r\right)\right\} \mid \mathcal{I}_{k2}\right] \\ &= -e^{-\rho r w_{k1}} \times \exp\left\{-\frac{\hat{\mu}_{k2}^{(\mathrm{I})}}{2\hat{\sigma}_{k2}^{2(\mathrm{I})}}\right\}. \end{aligned}$$

Recall that $\hat{\boldsymbol{\Omega}}_{K2} = r^2(r\boldsymbol{I} - \boldsymbol{G})^{-1}\hat{\boldsymbol{\Sigma}}_{Ky2}(r\boldsymbol{I} - \boldsymbol{G})^{-1\prime}$ and $\hat{\boldsymbol{\Sigma}}_{Ky2} = (\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2} + \beta^2 \boldsymbol{M}_{K2}\boldsymbol{\Omega}_e)^{-1}$; they are either data choices or known at the data choice stage. Thus

$$\mathbb{E}[-e^{-\rho w_{k2}} \mid \mathcal{I}_{k2^-}] = -e^{-\rho r w_{k1}} \times \mathbb{E}\left[e^{-\frac{\hat{\mu}_{k2}^{(\mathrm{I})}}{2\hat{\sigma}_{k2}^{2(\mathrm{I})}}} \mid \mathcal{I}_{k2^-}\right] = -e^{-\rho r w_{k1}} \times \left|\sigma_{K2}^{2(\mathrm{I})}/\hat{\sigma}_{k2}^{2(\mathrm{I})}\right|^{-\frac{1}{2}} \times e^{-\frac{\mu_{K2}^{2(\mathrm{I})}}{2\sigma_{K2}^{2(\mathrm{I})}}}.$$

So the objective of her data choice problem is reduced to $\min \hat{\sigma}_{k2}^{2(\text{I})}$, which further reduces to $\min \hat{\sigma}_{ky2}^{2(\text{I})} = \frac{1}{\hat{\Omega}_{ky2}^{(\text{I})}\hat{\Omega}_{ky2}^{(\text{II})} - \hat{\mathcal{C}}_{ky2}^2} \hat{\Omega}_{ky2}^{(\text{II})} \iff \max \hat{\Omega}_{ky2}^{(\text{I})} - \frac{\hat{\mathcal{C}}_{ky2}^2}{\hat{\Omega}_{ky2}^{(\text{II})}}$. Recall that

$$\hat{\mathbf{\Omega}}_{ky2} = \begin{pmatrix} \Omega_y^{(\text{I})} + \hat{\Omega}_{p2}^{(\text{I})} + \beta^2 M_{k2}^{(\text{I})} \Omega_e^{(\text{I})} & \mathcal{C}_y + \hat{\mathcal{C}}_{p2} \\ \mathcal{C}_y + \hat{\mathcal{C}}_{p2} & \Omega_y^{(\text{II})} + \hat{\Omega}_{p2}^{(\text{II})} + \beta^2 M_{k2}^{(\text{II})} \Omega_e^{(\text{II})} \end{pmatrix}.$$

Therefore the data choice problem finally reduces to:

$$\max_{\{M_{k2}^{(\text{I})}, M_{k2}^{(\text{II})}\}} \quad \beta^2 M_{k2}^{(\text{I})} \Omega_e^{(\text{I})} - \frac{(\mathcal{C}_y + \hat{\mathcal{C}}_{p2})^2}{\Omega_y^{(\text{II})} + \hat{\Omega}_{p2}^{(\text{II})} + \beta^2 M_{k2}^{(\text{II})} \Omega_e^{(\text{II})}} \tag{A.34}$$

$$\text{s.t.} \quad M_{k2}^{(\text{I})} + M_{k2}^{(\text{II})} \leq \bar{M}. \tag{A.35}$$

Again the data constraint is satisfied with equality. Substituting $M_{k2}^{(\text{I})} + M_{k2}^{(\text{II})} = \bar{M}$ into the objective function, we have $obj = \beta^2(\bar{M} - M_{k2}^{(\text{II})})\Omega_e^{(\text{I})} - \frac{(\mathcal{C}_y + \hat{\mathcal{C}}_{p2})^2}{\Omega_y^{(\text{II})} + \hat{\Omega}_{p2}^{(\text{II})} + \beta^2 M_{k2}^{(\text{II})} \Omega_e^{(\text{II})}}$. The F.O.C.

$$\frac{\partial obj}{\partial M_{k2}^{(\text{II})}} = -\beta^2 \Omega_e^{(\text{I})} + (\mathcal{C}_y + \hat{\mathcal{C}}_{p2})^2 \frac{\beta^2 \Omega_e^{(\text{II})}}{(\Omega_y^{(\text{II})} + \hat{\Omega}_{p2}^{(\text{II})} + \beta^2 M_{k2}^{(\text{II})} \Omega_e^{(\text{II})})^2} = 0.$$

Intuitively, the marginal gain in posterior precision of dividend innovation for type-I risky asset from directly learning about type-I risky asset equals the gain in posterior precision of dividend innovation for type-I risky asset from indirectly learning about type-Ⅱ risky asset. If the needed data amount is negative, or $M_{k2}^c \leq 0 \implies \frac{|\mathcal{C}_y + \hat{\mathcal{C}}_{p2}|}{\Omega_y^{(\text{II})} + \hat{\Omega}_{p2}^{(\text{II})}} \sqrt{\frac{\Omega_e^{(\text{II})}}{\Omega_e^{(\text{I})}}} \leq 1$, that is indirectly learning about type-Ⅱ risky asset is less efficient than directly learning about typeI risky asset, then foreign institutional investor k will only learn about type-I asset, $M_{k2}^* = 0$. If the needed data amount is larger than $\bar{M}$, or $M_{k2}^c \geq 0 \implies \frac{|\mathcal{C}_y + \hat{\mathcal{C}}_{p2}|}{\Omega_y^{(\text{II})} + \hat{\Omega}_{p2}^{(\text{II})} + \bar{M}\beta^2 \Omega_e^{(\text{II})}} \sqrt{\frac{\Omega_e^{(\text{II})}}{\Omega_e^{(\text{I})}}} \geq 1$, that is indirectly learning about type-Ⅱ risky asset is more efficient than directly learning about typeI risky asset, then foreign institutional investor k will only learn about type-Ⅱ asset, $M_{k2}^* = \bar{M}$. Define the information spillover efficiency rate as $\kappa \equiv \left(\frac{\hat{\mathcal{C}}_{ky2}}{\hat{\Omega}_{ky2}^{(\text{II})}}\right)^2 \frac{\Omega_e^{(\text{II})}}{\Omega_e^{(\text{I})}}$. It shows one unit of precision for type-Ⅱ risky asset's dividend innovation converts to κ units of precision for type-I risky asset.

In sum, the solution is

$$M_{k2}^{*}=\begin{cases}0, & \text{if } \frac{\left|\mathcal{C}_y+\hat{\mathcal{C}}_{p2}\right|}{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}}\sqrt{\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}}\leq 1,\\ \bar{M}, & \text{if } \frac{\left|\mathcal{C}_y+\hat{\mathcal{C}}_{p2}\right|}{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\bar{M}\beta^2\Omega_e^{(\mathrm{II})}}\sqrt{\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}}\geq 1,\\ \frac{1}{\beta^2\Omega_e^{(\mathrm{II})}}\left[\left|\mathcal{C}_y+\hat{\mathcal{C}}_{p2}\right|\sqrt{\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}}-\Omega_y^{(\mathrm{II})}-\hat{\Omega}_{p2}^{(\mathrm{II})}\right], & \text{otherwise.}\end{cases}$$

Proof of Proposition 2. I prove this result by checking the limits of $\kappa(0)$.

$$\begin{aligned}&\lim_{\Omega_e^{(\mathrm{II})}\to 0}\kappa(0)=\lim_{\Omega_e^{(\mathrm{II})}\to 0}\frac{(\mathcal{C}_y+\hat{\mathcal{C}}_{p2})^2}{(\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})})^2}\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}\\ &=\lim_{\Omega_e^{(\mathrm{II})}\to 0}\frac{\left(\mathcal{C}_y-\left(\frac{r-g_1}{\rho r}\right)^2\lambda_K\beta^2\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}\left[(\lambda_I\beta_I^2M_{I2}^{(\mathrm{I})}+\lambda_K\beta^2M_{K2}^{(\mathrm{I})})\Omega_e^{(\mathrm{I})}+\lambda_1\Omega_\epsilon^{(\mathrm{I})}\right]\Omega_x^{(\mathrm{I})}\right)^2}{\left(\Omega_y^{(\mathrm{II})}+\left[\frac{r-g_1}{\rho r}\lambda_K\beta^2\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}\right]^2\Omega_x^{(\mathrm{I})}+\left[\frac{r-g_2}{\rho r}\left[\lambda_I\beta_I^2M_{I2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}+\lambda_1\Omega_\epsilon^{(\mathrm{II})}\right]\right]^2\Omega_x^{(\mathrm{II})}\right)^2}\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}\\ &=0,\end{aligned}$$

where the last equality is achieved by directly plugging in $\Omega_e^{(\mathrm{II})}=0$.

The other limit is calculated by simultaneously dividing the term $(\Omega_e^{(\mathrm{II})})^3$ in the numerator and the denominator,

$$\begin{aligned}&\lim_{\Omega_e^{(\mathrm{II})}\to\infty}\frac{(\mathcal{C}_y+\hat{\mathcal{C}}_{p2})^2}{(\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})})^2}\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}\\ &=\lim_{\Omega_e^{(\mathrm{II})}\to\infty}\frac{\left(-\left(\frac{r-g_1}{\rho r}\right)^2\lambda_K\beta^2\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}M_{K2}^{(\mathrm{II})}\left[(\lambda_I\beta_I^2M_{I2}^{(\mathrm{I})}+\lambda_K\beta^2M_{K2}^{(\mathrm{I})})\Omega_e^{(\mathrm{I})}+\lambda_1\Omega_\epsilon^{(\mathrm{I})}\right]\Omega_x^{(\mathrm{I})}\right)^2}{\left(\left[\frac{r-g_1}{\rho r}\lambda_K\beta^2\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}M_{K2}^{(\mathrm{II})}\right]^2\Omega_x^{(\mathrm{I})}+\left[\frac{r-g_2}{\rho r}\left[\lambda_I\beta_I^2M_{I2}^{(\mathrm{II})}\right]\right]^2\Omega_x^{(\mathrm{II})}\right)^2}\frac{1}{\Omega_e^{(\mathrm{I})}\Omega_e^{(\mathrm{II})}}\\ &=0.\end{aligned}$$

Given that $\lim_{\Omega_e^{(\mathrm{II})}\to\infty}\kappa(0)=0$ and $\lim_{\Omega_e^{(\mathrm{II})}\to 0}\kappa(0)=0$, it is direct to see $M_{k2}^{*}=0$ when $\Omega_e^{(\mathrm{II})}\to\infty$ or $\Omega_e^{(\mathrm{II})}\to 0$. □

Proof of Proposition 3. The interior solution of $M_{k2}^{(\mathrm{II})}$ is $\frac{1}{\beta^2\Omega_e^{(\mathrm{II})}}\left[\left|\mathcal{C}_y+\hat{\mathcal{C}}_{p2}\right|\sqrt{\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}}-\Omega_y^{(\mathrm{II})}-\hat{\Omega}_{p2}^{(\mathrm{II})}\right]$. From

(A.19), we can directly write out the elements of $\boldsymbol{\zeta}_2$:

$$\boldsymbol{\zeta}_2 = -\begin{pmatrix} \frac{r-g_1}{\rho r}\left[(\lambda_I\beta_I^2 M_{I2}^{(\mathrm{I})} + \lambda_K\beta^2 M_{K2}^{(\mathrm{I})})\Omega_e^{(\mathrm{I})} + \lambda_1\Omega_\epsilon^{(\mathrm{I})}\right] & -\frac{r-g_1}{\rho r}\lambda_K\beta^2\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} \\ 0 & \frac{r-g_2}{\rho r}\left[\lambda_I\beta_I^2 M_{I2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \lambda_1\Omega_\epsilon^{(\mathrm{II})}\right] \end{pmatrix}.$$

Thus $\hat{\Omega}_{p2}^{(\mathrm{II})} = \left[\frac{r-g_1}{\rho r}\lambda_K\beta^2\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}M_{K2}^{(\mathrm{II})}\Omega_\epsilon^{(\mathrm{II})}\right]^2\Omega_x^{(\mathrm{I})} + \left[\frac{r-g_2}{\rho r}\left[\lambda_I\beta_I^2 M_{I2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \lambda_1\Omega_\epsilon^{(\mathrm{II})}\right]\right]^2\Omega_x^{(\mathrm{II})}$ and $\hat{\mathcal{C}}_{p2} = -\left(\frac{r-g_1}{\rho r}\right)^2\lambda_K\beta^2\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}\left[(\lambda_I\beta_I^2 M_{I2}^{(\mathrm{I})} + \lambda_K\beta^2 M_{K2}^{(\mathrm{I})})\Omega_e^{(\mathrm{I})} + \lambda_1\Omega_\epsilon^{(\mathrm{I})}\right]\Omega_x^{(\mathrm{I})}$. Recall that $\hat{\mathcal{C}}_{p2} = \hat{\mathcal{C}}_{Ky2} - \mathcal{C}_y$. We can then solve for $\hat{\mathcal{C}}_{Ky2}$ and $\hat{\mathcal{C}}_{p2}$.

$$\hat{\mathcal{C}}_{Ky2} - \mathcal{C}_y = -\left(\frac{r-g_1}{\rho r}\right)^2\lambda_K\beta^2\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}\left[(\lambda_I\beta_I^2 M_{I2}^{(\mathrm{I})} + \lambda_K\beta^2 M_{K2}^{(\mathrm{I})})\Omega_e^{(\mathrm{I})} + \lambda_1\Omega_\epsilon^{(\mathrm{I})}\right]\Omega_x^{(\mathrm{I})}$$

$$\implies \hat{\mathcal{C}}_{Ky2} = \frac{\mathcal{C}_y}{1 + \left(\frac{r-g_1}{\rho r}\right)^2\lambda_K\beta^2\frac{1}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}\left[(\lambda_I\beta_I^2 M_{I2}^{(\mathrm{I})} + \lambda_K\beta^2 M_{K2}^{(\mathrm{I})})\Omega_e^{(\mathrm{I})} + \lambda_1\Omega_\epsilon^{(\mathrm{I})}\right]\Omega_x^{(\mathrm{I})}}.$$

Thus $\left|\hat{\mathcal{C}}_{Ky2}\right| < |\mathcal{C}_y|$. As a result,

$$\left|\mathcal{C}_y + \hat{\mathcal{C}}_{p2}\right|\sqrt{\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}} - \Omega_y^{(\mathrm{II})} - \hat{\Omega}_{p2}^{(\mathrm{II})} < |\mathcal{C}_y|\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}} - \Omega_y^{(\mathrm{II})}.$$

Therefore, a sufficient condition for the LHS to be negative is fro the RHS to be negative, i.e. $\Omega_e^{(\mathrm{I})} > \frac{\Omega_e^{(\mathrm{II})}|\mathcal{C}_y|}{\Omega_y^{(\mathrm{II})}}$. □

In equilibrium, the data choices for domestic and foreign institutional investors are deter-

mined simultaneously. They solve the following system of non-linear equations:

$$M_{I2}^{(\mathrm{I})} = \begin{cases} 0, & \text{if } \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}} \geq \frac{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\bar{M}\beta_I^2\Omega_e^{(\mathrm{II})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}}, \\ \bar{M}, & \text{if } \frac{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}} \geq \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\bar{M}\beta_I^2\Omega_e^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}}, \\ \frac{1}{2}\left[\frac{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}} + \bar{M} - \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}}\right], & \text{otherwise.} \end{cases}$$

$$M_{K2}^{(\mathrm{II})} = \begin{cases} 0, & \text{if } \frac{\left|\mathcal{C}_y+\hat{\mathcal{C}}_{p2}\right|}{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}}\sqrt{\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}} \leq 1, \\ \bar{M}, & \text{if } \frac{\left|\mathcal{C}_y+\hat{\mathcal{C}}_{p2}\right|}{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\bar{M}\beta^2\Omega_e^{(\mathrm{II})}}\sqrt{\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}} \geq 1, \\ \frac{1}{\beta^2\Omega_e^{(\mathrm{II})}}\left[\left|\mathcal{C}_y+\hat{\mathcal{C}}_{p2}\right|\sqrt{\frac{\Omega_e^{(\mathrm{II})}}{\Omega_e^{(\mathrm{I})}}} - \Omega_y^{(\mathrm{II})} - \hat{\Omega}_{p2}^{(\mathrm{II})}\right], & \text{otherwise.} \end{cases}$$

where

$$\begin{aligned}
\hat{\boldsymbol{\Omega}}_{p2} &= \boldsymbol{\zeta}_2'\boldsymbol{\Omega}_x\boldsymbol{\zeta}_2, \\
\boldsymbol{\zeta}_2 &= -\frac{1}{\rho r}(r\boldsymbol{I}-\boldsymbol{G})\left[\left(\lambda_I\beta_I^2\boldsymbol{M}_{I2} + \lambda_K\beta^2\left(\boldsymbol{E}_{11} - \frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}\boldsymbol{E}_{12}\right)\boldsymbol{M}_{K2}\right)\boldsymbol{\Omega}_e + \lambda_1\boldsymbol{\Omega}_\epsilon\right], \\
\hat{\Omega}_{Ky2}^{(\mathrm{II})} &= \Omega_y^{(\mathrm{II})} + \beta^2 M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \left[\frac{r-g_1}{\rho r}\lambda_K\beta^2 z(M_{I2}^{(\mathrm{I})}, M_{K2}^{(\mathrm{II})})M_{K2}^{(\mathrm{II})}\Omega_\epsilon^{(\mathrm{II})}\right]^2\Omega_x^{(\mathrm{I})} \\
&\quad + \left[\frac{r-g_2}{\rho r}\left[\lambda_I\beta_I^2 M_{I2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \lambda_1\Omega_\epsilon^{(\mathrm{II})}\right]\right]^2\Omega_x^{(\mathrm{II})}, \\
\hat{\mathcal{C}}_{Ky2} &= \mathcal{C}_y - \left(\frac{r-g_1}{\rho r}\right)^2\lambda_K\beta^2 z(M_{I2}^{(\mathrm{I})}, M_{K2}^{(\mathrm{II})})M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}\left[(\lambda_I\beta_I^2 M_{I2}^{(\mathrm{I})} + \lambda_K\beta^2 M_{K2}^{(\mathrm{I})})\Omega_e^{(\mathrm{I})} + \lambda_1\Omega_\epsilon^{(\mathrm{I})}\right]\Omega_x^{(\mathrm{I})}.
\end{aligned}$$

First of all, $\hat{\Omega}_{Ky2}^{(\mathrm{II})}$ and $\hat{\mathcal{C}}_{Ky2}$ can be solved in terms of model parameters and data choices. Let $z = \frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}$. Recall that

$$\begin{aligned}
\hat{\Omega}_{Ky2}^{(\mathrm{II})} &= \Omega_y^{(\mathrm{II})} + \beta^2 M_{K2}^{c}\Omega_e^{(\mathrm{II})} + \left[\frac{r-g_1}{\rho r}\lambda_K\beta^2 z M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}\right]^2\Omega_x^{(\mathrm{I})} + \left[\frac{r-g_2}{\rho r}\left[\lambda_I\beta_I^2 M_{I2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \lambda_1\Omega_\epsilon^{(\mathrm{II})}\right]\right]^2\Omega_x^{(\mathrm{II})}, \\
\hat{\mathcal{C}}_{Ky2} &= \mathcal{C}_y - \left(\frac{r-g_1}{\rho r}\right)^2\lambda_K\beta^2 z M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}\left[(\lambda_I\beta_I^2 M_{I2}^{(\mathrm{I})} + \lambda_K\beta^2 M_{K2}^{(\mathrm{I})})\Omega_e^{(\mathrm{I})} + \lambda_1\Omega_\epsilon^{(\mathrm{I})}\right]\Omega_x^{(\mathrm{I})}.
\end{aligned}$$

From here we can derive a cubic equation for z:

$$\begin{aligned}
&\mathcal{C}_y - \left(\frac{r-g_1}{\rho r}\right)^2 \lambda_K \beta^2 z M_{K2}^{(\text{II})} \Omega_e^{(\text{II})} \left[(\lambda_I \beta_I^2 M_{I2}^{(\text{I})} + \lambda_K \beta^2 M_{K2}^{(\text{I})})\Omega_e^{(\text{I})} + \lambda_1 \Omega_\epsilon^{(\text{I})}\right] \Omega_x^{(\text{I})} \\
&= z \left\{ \Omega_y^{(\text{II})} + \beta^2 M_{K2}^{(\text{II})} \Omega_e^{(\text{II})} + \left[\frac{r-g_1}{\rho r} \lambda_K \beta^2 z M_{K2}^{(\text{II})} \Omega_e^{(\text{II})}\right]^2 \Omega_x^{(\text{I})} + \left[\frac{r-g_2}{\rho r}\left[\lambda_I \beta_I^2 M_{I2}^{(\text{II})} \Omega_e^{(\text{II})} + \lambda_1 \Omega_\epsilon^{(\text{II})}\right]\right]^2 \Omega_x^{(\text{II})} \right\}.
\end{aligned}$$

If foreign institutional investors do not learn about type-II risky asset, i.e. $M_{K2}^{(\text{II})} = 0$, then the above equation reduces to a linear equation of z, with

$$z = \frac{\mathcal{C}_y}{\Omega_y^{(\text{II})} + \left[\frac{r-g_2}{\rho r}\left[\lambda_I \beta_I^2 M_{I2}^{(\text{II})} \Omega_e^{(\text{II})} + \lambda_1 \Omega_\epsilon^{(\text{II})}\right]\right]^2 \Omega_x^{(\text{II})}}.$$

When $M_{K2}^{(\text{II})} > 0$, to write it in the standard form $z^3 + pz + q = 0$, we have

$$\begin{aligned}
p = &\left\{ \Omega_y^{(\text{II})} + \beta^2 M_{K2}^{c} \Omega_e^{(\text{II})} + \left[\frac{r-g_2}{\rho r}\left[\lambda_I \beta_I^2 M_{I2}^{(\text{II})} \Omega_e^{(\text{II})} + \lambda_1 \Omega_\epsilon^{(\text{II})}\right]\right]^2 \Omega_x^{(\text{II})} \right. \\
&\left. + \left(\frac{r-g_1}{\rho r}\right)^2 \lambda_K \beta^2 M_{K2}^{(\text{II})} \Omega_e^{(\text{II})} \left[(\lambda_I \beta_I^2 M_{I2}^{(\text{I})} + \lambda_K \beta^2 M_{K2}^{(\text{I})})\Omega_e^{(\text{I})} + \lambda_1 \Omega_\epsilon^{(\text{I})}\right] \Omega_x^{(\text{I})} \right\} \\
&/\left\{ \left[\frac{r-g_1}{\rho r} \lambda_K \beta^2 M_{K2}^{(\text{II})} \Omega_e^{(\text{II})}\right]^2 \Omega_x^{(\text{I})} \right\} > 0, \\
q = &-\mathcal{C}_y / \left\{ \left[\frac{r-g_1}{\rho r} \lambda_K \beta^2 M_{K2}^{(\text{II})} \Omega_e^{(\text{II})}\right]^2 \Omega_x^{(\text{I})} \right\},
\end{aligned}$$

which are only functions of model parameters and data choices. Using Cardano's results, we know that $\frac{q^2}{4} + \frac{p^3}{27} > 0$, which implies that this cubic equation only has one real root:

$$\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\text{II})}} = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

For notational simplicity, let me denote the solution by $z(M_{I2}^{(\text{I})}, M_{K2}^{(\text{II})})$.

$$M_{I2}^{(\mathrm{I})} = \begin{cases} 0, & \text{if } \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}} \geq \frac{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\bar{M}\beta_I^2\Omega_e^{(\mathrm{II})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}}, \\ \bar{M}, & \text{if } \frac{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}} \geq \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\bar{M}\beta_I^2\Omega_e^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}}, \\ \frac{1}{2}\left[\frac{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}} + \bar{M} - \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}}\right], & \text{otherwise.} \end{cases}$$

$\hat{\Omega}_{p2}^{(\mathrm{II})} = \left[\frac{r-g_1}{\rho r}\lambda_K\beta^2\frac{\hat{\mathcal{C}}_{Ky2}}{\hat{\Omega}_{Ky2}^{(\mathrm{II})}}M_{K2}^{(\mathrm{II})}\Omega_\epsilon^{(\mathrm{II})}\right]^2\Omega_x^{(\mathrm{I})} + \left[\frac{r-g_2}{\rho r}\left[\lambda_I\beta_I^2M_{I2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}+\lambda_1\Omega_\epsilon^{(\mathrm{II})}\right]\right]^2\Omega_x^{(\mathrm{II})}$ and $\hat{\Omega}_{p2}^{(\mathrm{I})} = [\frac{r-g_1}{\rho r}[(\lambda_I\beta_I^2M_{I2}^{(\mathrm{I})}+\lambda_K\beta^2M_{K2}^{(\mathrm{I})})\Omega_e^{(\mathrm{I})}+\lambda_1\Omega_\epsilon^{(\mathrm{I})}]]^2\Omega_x^{(\mathrm{I})}$.

$$\begin{aligned} 2M_{I2}^{(\mathrm{I})} &= \frac{\Omega_y^{(\mathrm{II})}+\hat{\Omega}_{p2}^{(\mathrm{II})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}} + \bar{M} - \frac{\Omega_y^{(\mathrm{I})}+\hat{\Omega}_{p2}^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}} \\ &= \frac{\Omega_y^{(\mathrm{II})}+(\frac{r-g_2}{\rho r})^2[\lambda_I\beta_I^2(\bar{M}-M_{I2}^{(\mathrm{I})})\Omega_e^{(\mathrm{II})}+\lambda_1\Omega_\epsilon^{(\mathrm{II})}]^2\Omega_x^{(\mathrm{II})}+[\frac{r-g_1}{\rho r}\lambda_K\beta^2zM_{K2}^{(\mathrm{II})}\Omega_\epsilon^{(\mathrm{II})}]^2\Omega_x^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{II})}}{\beta_I^2\Omega_e^{(\mathrm{II})}} + \bar{M} \\ &\quad - \frac{\Omega_y^{(\mathrm{I})}+(\frac{r-g_1}{\rho r})^2[\lambda_I\beta_I^2M_{I2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})}+\lambda_K\beta^2M_{K2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})}+\lambda_1\Omega_\epsilon^{(\mathrm{I})}]^2\Omega_x^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}}. \end{aligned}$$

Moving one $M_{I2}^{(\mathrm{I})}$ to RHS and $\frac{\Omega_y^{(\mathrm{I})}+(\frac{r-g_1}{\rho r})^2[\lambda_I\beta_I^2M_{I2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})}+\lambda_K\beta^2M_{K2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})}+\lambda_1\Omega_\epsilon^{(\mathrm{I})}]^2\Omega_x^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}}$ to LHS, we have

$$\begin{aligned} &\frac{\Omega_y^{(\mathrm{I})}+(\frac{r-g_1}{\rho r})^2[\lambda_I\beta_I^2M_{I2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})}+\lambda_K\beta^2M_{K2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})}+\lambda_1\Omega_\epsilon^{(\mathrm{I})}]^2\Omega_x^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}+M_{I2}^{(\mathrm{I})}\beta_I^2\Omega_e^{(\mathrm{I})}}{\beta_I^2\Omega_e^{(\mathrm{I})}} \\ &= [\Omega_y^{(\mathrm{II})}+(\frac{r-g_2}{\rho r})^2[\lambda_I\beta_I^2(\bar{M}-M_{I2}^{(\mathrm{I})})\Omega_e^{(\mathrm{II})}+\lambda_1\Omega_\epsilon^{(\mathrm{II})}]^2\Omega_x^{(\mathrm{II})} \\ &\quad +\Big[\frac{r-g_1}{\rho r}\lambda_K\beta^2zM_{K2}^{(\mathrm{II})}\Omega_\epsilon^{(\mathrm{II})}]^2\Omega_x^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{II})}+(\bar{M}-M_{I2}^{(\mathrm{I})})\beta_I^2\Omega_e^{(\mathrm{II})}\Big]/(\beta_I^2\Omega_e^{(\mathrm{II})}) \end{aligned}$$

This is a quadratic equation (or reduced to a linear equation) of $M_{I2}^{(\mathrm{I})}$. To write it in the standard notation,

$$aM_{I2}^{(\mathrm{I})2}+bM_{I2}^{(\mathrm{I})}+c=0,$$

where

$$
\begin{aligned}
a &= \lambda_I^2\beta_I^4[(\frac{r-g_1}{\rho r})^2\Omega_e^{(\mathrm{I})2}\Omega_e^{(\mathrm{II})}\Omega_x^{(\mathrm{I})} - (\frac{r-g_2}{\rho r})^2\Omega_e^{(\mathrm{I})}\Omega_e^{(\mathrm{II})2}\Omega_x^{(\mathrm{II})}],\\
b &= 2[(\frac{r-g_1}{\rho r})^2\lambda_I\beta_I^2\Omega_e^{(\mathrm{I})}\Omega_e^{(\mathrm{II})}(\lambda_1\Omega_\epsilon^{(\mathrm{I})} + \lambda_K\beta^2 M_{K2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})})\Omega_x^{(\mathrm{I})}\\
&\quad +(\frac{r-g_2}{\rho r})^2\lambda_I\beta_I^2\Omega_e^{(\mathrm{I})}\Omega_e^{(\mathrm{II})}(\lambda_1\Omega_\epsilon^{(\mathrm{II})} + \lambda_I\beta_I^2\bar{M}\Omega_e^{(\mathrm{II})})\Omega_x^{(\mathrm{II})} + \beta_I^2\Omega_e^{(\mathrm{I})}\Omega_e^{(\mathrm{II})}],\\
c &= [\Omega_y^{(\mathrm{I})} + (\frac{r-g_1}{\rho r})^2(\lambda_1\Omega_\epsilon^{(\mathrm{I})} + \lambda_K\beta^2 M_{K2}^{(\mathrm{I})}\Omega_e^{(\mathrm{I})})^2\Omega_x^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})}]\Omega_e^{(\mathrm{II})}\\
&\quad -[\Omega_y^{(\mathrm{II})} + (\frac{r-g_2}{\rho r})^2(\lambda_1\Omega_\epsilon^{(\mathrm{II})} + \lambda_I\beta_I^2\bar{M})^2\Omega_x^{(\mathrm{II})} + [\frac{r-g_1}{\rho r}\lambda_K\beta^2 z M_{K2}^{(\mathrm{II})}\Omega_e^{(\mathrm{II})}]^2\Omega_x^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{II})} + \beta_I^2\bar{M}\Omega_e^{(\mathrm{II})}]\Omega_e^{(\mathrm{I})}.
\end{aligned}
$$

In the case where $a = 0$, that is $(r-g_1)^2\Omega_e^{(\mathrm{I})}\Omega_x^{(\mathrm{I})} = (r-g_2)^2\Omega_e^{(\mathrm{II})}\Omega_x^{(\mathrm{II})}$, the equation is reduced to a linear form, and $M_{I2}^c = -\frac{c}{b}$. If in addition $c \geq 0$, then $M_{I2}^* = 0$. If $-\frac{c}{b} \geq \bar{M}$, then $M_{I2}^* = \bar{M}$.

In the case where $a \neq 0$, it is a quadratic equation, and real solution exists if $b^2 - 4ac \geq 0$. Under that condition, the smaller solution is always negative. If $c \geq 0$, then both solutions are negative, $M_{I2}^* = 0$.

$$
M_{I2}^* = \begin{cases} 0, & \text{if } c \geq 0, \\ \bar{M}, & \text{if } a = 0, -\frac{c}{b} \geq \bar{M} \text{ or } a \neq 0, c < 0, \frac{-b+\sqrt{b^2-4ac}}{2a} \geq \bar{M}, \\ -\frac{c}{b}, & \text{if } a = 0, 0 < -\frac{c}{b} < \bar{M} \\ \frac{-b+\sqrt{b^2-4ac}}{2a}, & \text{if } a \neq 0, c < 0, 0 < \frac{-b+\sqrt{b^2-4ac}}{2a} < \bar{M}. \end{cases}
$$

In the counterfactual where the market remains closed to international investors, the solution has the same form but with a different set of b and c:

$$
\begin{aligned}
b^\dagger &= 2\Big[(\frac{r-g_1}{\rho r})^2\lambda_I\beta_I^2\Omega_e^{(\mathrm{I})}\Omega_e^{(\mathrm{II})}\lambda_1\Omega_\epsilon^{(\mathrm{I})}\Omega_x^{(\mathrm{I})} + (\frac{r-g_2}{\rho r})^2\lambda_I\beta_I^2\Omega_e^{(\mathrm{I})}\Omega_e^{(\mathrm{II})}(\lambda_1\Omega_\epsilon^{(\mathrm{II})}\\
&\quad +\lambda_I\beta_I^2\bar{M}\Omega_e^{(\mathrm{II})})\Omega_x^{(\mathrm{II})} + \beta_I^2\Omega_e^{(\mathrm{I})}\Omega_e^{(\mathrm{II})}\Big],\\
c^\dagger &= [\Omega_y^{(\mathrm{I})} + (\frac{r-g_1}{\rho r})^2\lambda_1^2\Omega_\epsilon^{(\mathrm{I})2}\Omega_x^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})}]\Omega_e^{(\mathrm{II})}\\
&\quad -[\Omega_y^{(\mathrm{II})} + (\frac{r-g_2}{\rho r})^2(\lambda_1\Omega_\epsilon^{(\mathrm{II})} + \lambda_I\beta_I^2\bar{M})^2\Omega_x^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})} + \beta_I^2\bar{M}\Omega_e^{(\mathrm{II})}]\Omega_e^{(\mathrm{I})}.
\end{aligned}
$$

Notice that in this case, $M_{I2}^{\dagger(\mathrm{I})}$ is directly expressed in terms of model primitives.

Proof of Proposition 4. 1) $m_e^{(\mathrm{I})}(0) \geq m_e^{(\mathrm{II})}(\bar{M}) \iff c^\dagger \geq 0 \implies M_{I2}^* = 0$.

2) Consider $c - c^\dagger$. Cancelling out common terms and rearrange, we have

$$
\begin{aligned}
& c - c^\dagger > 0 \\
\iff & \lambda_K \beta^2 (\Omega_e^{(\mathrm{I})} - z\Omega_e^{(\mathrm{II})}) M_{K2}^{(\mathrm{II})2} - 2(\lambda_K \beta^2 \bar{M} \Omega_e^{(\mathrm{I})} + \lambda_1 \Omega_e^{(\mathrm{I})}) M_{K2}^{(\mathrm{II})} + \lambda_K \beta^2 \bar{M}^2 \Omega_e^{(\mathrm{I})} + 2\bar{M}\lambda_1 \Omega_\epsilon^{(\mathrm{I})} > 0.
\end{aligned}
$$

The above quadratic expression of $M_{K2}^{(\mathrm{II})}$ is positive when $M_{K2}^{(\mathrm{II})} = 0$. The symmetry axis is $M_{K2}^{(\mathrm{II})} = \frac{\lambda_K \beta^2 \bar{M} \Omega_e^{(\mathrm{I})} + \lambda_1 \Omega_e^{(\mathrm{I})}}{\lambda_K \beta^2 (\Omega_e^{(\mathrm{I})} - z\Omega_e^{(\mathrm{II})})}$. It lies to the left of the y-axis when $\Omega_e^{(\mathrm{I})} - z\Omega_e^{(\mathrm{II})} \leq 0$. In that case, $M_{K2}^{(\mathrm{II})}$ is also increasing to the right of the symmetry axis, thus the value of the quadratic expression when $M_{K2}^{(\mathrm{II})}$ is larger than its value at $M_{K2}^{(\mathrm{II})} = 0$, which is positive. Therefore $\Omega_e^{(\mathrm{I})} - z\Omega_e^{(\mathrm{II})} \leq 0$ ensures $c - c^\dagger > 0$. Any positive value of c will lead to $M_{I2}^* = 0$. □

A.1.5 Proofs for partial market opening impacts

Proof of proposition 5. Given $\Omega_e^{(\mathrm{I})} > \frac{\Omega_e^{(\mathrm{II})}}{\Omega_y^{(\mathrm{II})}} |\mathcal{C}_y|$, foreign institutional investors will only learn about type-II risky asset, $M_{K2}^{(\mathrm{I})} = \bar{M}, M_{K2}^{(\mathrm{II})} = 0$. The change of the signal-to-noise ratio for type-I risky asset brought in by partial market opening is

$$
\Delta \left| \zeta_2^{(\mathrm{I})} \right| \equiv \left| \zeta_2^{(\mathrm{I})} \right| - \left| \zeta_2^{\dagger(\mathrm{I})} \right| = \frac{r - g_1}{\rho r} [\lambda_I (M_{I2}^{(\mathrm{I})} - M_{I2}^{\dagger(\mathrm{I})}) \beta_I^2 + \lambda_K \bar{M} \beta^2] \Omega_e^{(\mathrm{I})}.
$$

Notice that $\frac{r - g_1}{\rho r} > 0$, $\Omega_e^{(\mathrm{I})} > 0$, $\lambda_I > 0$ and $\lambda_K \bar{M} \beta^2 \geq 0$. Since data choices are limited by data processing capacity $\bar{M}$, a sufficient condition for $\Delta \left| \zeta_2^{(\mathrm{I})} \right| > 0$ is $-\lambda_I \bar{M} \beta_I^2 + \lambda_K \bar{M} \beta^2 > 0$, which gives $\lambda_K > \lambda_I \frac{\beta_I^2}{\beta^2}$.

For $\Delta \left| \zeta_2^{(\mathrm{I})} \right| < 0$, we need $\lambda_I (M^{(\mathrm{I})} - M^{\dagger(\mathrm{I})}) \beta_I^2 + \lambda_K \bar{M} \beta^2 < 0$, which necessarily requires that domestic institutional investors reduces learning for type-I risky asset, i.e. $M^{(\mathrm{I})} < M^{\dagger(\mathrm{I})}$. Then the number of foreign institutional investors $\lambda_K < \frac{(M^{\dagger(\mathrm{I})} - M^{(\mathrm{I})}) \beta_I^2}{\bar{M} \beta^2} \lambda_I$. □

Proof of Proposition 7. After market opening, the average posterior precision of dividend innova-

tion for type-I risky asset changes by

$$
\begin{aligned}
&\bar{\hat{\Omega}}^{(\mathrm{I})}_{y2} - \bar{\hat{\Omega}}^{\dagger(\mathrm{I})}_{y2} \\
&= (\tau_y^{(\mathrm{I})} + \zeta_2^{2(\mathrm{I})}\tau_x^{(\mathrm{I})} + \frac{\lambda_I M^{(\mathrm{I})}\beta_I^2 + \lambda_K \bar{M}\beta^2}{\lambda_2}\tau_e^{(\mathrm{I})} + \frac{\lambda_1}{\lambda_2}\tau_\epsilon^{(\mathrm{I})}) - (\tau_y^{(\mathrm{I})} + \boldsymbol{\zeta}_2^{\dagger 2(\mathrm{I})}\tau_x^{(\mathrm{I})} + \frac{\lambda_I M^{\dagger(\mathrm{I})}\beta_I^2}{\lambda_1}\tau_e^{(\mathrm{I})} + \tau_\epsilon^{(\mathrm{I})}) \\
&= (\zeta_2^{2(\mathrm{I})} - \zeta_2^{\dagger 2(\mathrm{I})})\tau_x^{(\mathrm{I})} + (\frac{\lambda_I}{\lambda_2}M^{(\mathrm{I})}\beta_I^2 - \frac{\lambda_I}{\lambda_1}M^{\dagger(\mathrm{I})}\beta_I^2 + \frac{\lambda_K}{\lambda_2}\bar{M}\beta^2)\Omega_e^{(\mathrm{I})} + (\frac{\lambda_1}{\lambda_2} - 1)\Omega_\epsilon^{(\mathrm{I})}.
\end{aligned}
$$

For $\Delta\bar{\hat{\Omega}}^{(\mathrm{I})}_{y2} < 0$, notice that the first term $\zeta_2^{2(\mathrm{I})} - \zeta_2^{\dagger 2(\mathrm{I})} = (\left|\zeta_2^{(\mathrm{I})}\right| - \left|\zeta_2^{\dagger(\mathrm{I})}\right|)(\left|\zeta_2^{(\mathrm{I})}\right| + \left|\zeta_2^{\dagger(\mathrm{I})}\right|)$. For it to be negative, we need the change in signal-to-noise ratio $\Delta\left|\zeta_2^{(\mathrm{I})}\right|$ to be negative, by Proposition 5, it is equivalent to $M^{(\mathrm{I})} < M^{\dagger(\mathrm{I})}$ and $\lambda_K < \frac{(M^{\dagger(\mathrm{I})} - M^{(\mathrm{I})})\beta_I^2}{\bar{M}\beta^2}\lambda_I$. Notice also that $\lambda_2 > \lambda_1$, so the third term $(\frac{\lambda_1}{\lambda_2} - 1)\Omega_\epsilon^{(\mathrm{I})}$ is always negative. What remains is to confirm that the second term is also negative under the current conditions. Since $\lambda_1 < \lambda_2$, we have $\frac{\lambda_I}{\lambda_2}M^{(\mathrm{I})}\beta_I^2 - \frac{\lambda_I}{\lambda_1}M^{\dagger(\mathrm{I})}\beta_I^2 + \frac{\lambda_K}{\lambda_2}\bar{M}\beta^2 < \frac{1}{\lambda_2}(\lambda_I M^{(\mathrm{I})}\beta_I^2 - \lambda_I M^{\dagger(\mathrm{I})}\beta_I^2 + \lambda_K\bar{M}\beta^2) < 0$.

For $\Delta\bar{\hat{\Omega}}^{(\mathrm{I})}_{y2} > 0$, we try to find a sufficient condition. The first term is positive when $\Delta\left|\zeta_2^{(\mathrm{I})}\right|$ is positive, which is satisfied if $\lambda_K > \lambda_I\frac{\beta_I^2}{\bar{\beta}^2}$. For the sum of the second and the third terms to be positive, it is equivalent to have

$$
\lambda_K\left[\bar{M}\beta^2\Omega_e^{(\mathrm{I})} - \frac{\lambda_I}{\lambda_1}M^{\dagger(\mathrm{I})}\beta_I^2\Omega_e^{(\mathrm{I})} - \Omega_\epsilon^{(\mathrm{I})}\right] > \lambda_I(M^{\dagger(\mathrm{I})} - M^{(\mathrm{I})})\beta_I^2\Omega_e^{(\mathrm{I})}. \tag{A.36}
$$

Case 1: $\lambda_1\Omega_\epsilon^{(\mathrm{I})} \leq (\lambda_1\bar{M}\beta^2 - \lambda_I\bar{M}\beta_I^2)\Omega_e^{(\mathrm{I})}$.

In this case, inclusion of new data and introduction of perfect data processing technology through foreign institutional investors outweighs the dilution of inalienable information held by domestic institutional investors. If market opening induces domestic institutional investors to learn more about type-I risky asset, i.e. $M^{\dagger(\mathrm{I})} \leq M^{(\mathrm{I})}$, then (A.36) is always true. If $M^{\dagger(\mathrm{I})} > M^{(\mathrm{I})}$, then we need a further condition for it to hold: $\lambda_K > \frac{(M^{\dagger(\mathrm{I})} - M^{(\mathrm{I})})\beta_I^2\Omega_e^{(\mathrm{I})}}{\bar{M}\beta^2\Omega_e^{(\mathrm{I})} - \frac{\lambda_I}{\lambda_1}M^{\dagger(\mathrm{I})}\beta_I^2\Omega_e^{(\mathrm{I})} - \Omega_\epsilon^{(\mathrm{I})}}\lambda_I$.

Case 2: $\lambda_1\Omega_\epsilon^{(\mathrm{I})} > (\lambda_1\bar{M}\beta^2 - \lambda_I\bar{M}\beta_I^2)\Omega_e^{(\mathrm{I})}$.

In this case, the dilution of inalienable information held by domestic institutional investors outweighs inclusion of new data and introduction of perfect data processing technology through

foreign institutional investors. If market opening induces domestic institutional investors to learn more about type-I risky asset, i.e. $M^{\dagger(\mathrm{I})} \leq M^{(\mathrm{I})}$, then we need $\lambda_K < \frac{(M^{(\mathrm{I})}-M^{\dagger(\mathrm{I})})\beta_I^2\Omega_e^{(\mathrm{I})}}{\frac{\lambda_I}{\lambda_1}M^{\dagger(\mathrm{I})}\beta_I^2\Omega_e^{(\mathrm{I})}+\Omega_\epsilon^{(\mathrm{I})}-\bar{M}\beta^2\Omega_e^{(\mathrm{I})}}\lambda_I$ for (A.36) to hold. If $M^{\dagger(\mathrm{I})} > M^{(\mathrm{I})}$, then (A.36) does not hold. □

Proof of Proposition 8. After market opening, the average posterior precision of dividend innovation for type-Ⅱ risky asset changes by

$$\Delta\bar{\hat{\Omega}}_{y2}^{(\mathrm{II})} = \bar{\hat{\Omega}}_{y2}^{(\mathrm{II})} - \bar{\hat{\Omega}}_{y2}^{\dagger(\mathrm{II})} = (\zeta_2^{2(\mathrm{II})} - \zeta_2^{\dagger 2(\mathrm{II})})\tau_x^{(\mathrm{II})} + (\frac{\lambda_I}{\lambda_2}M^{(\mathrm{II})} - \frac{\lambda_I}{\lambda_1}M^{\dagger(\mathrm{II})})\Omega_e^{(\mathrm{II})} + (\frac{\lambda_I}{\lambda_2} - \frac{\lambda_I}{\lambda_1})\Omega_\epsilon^{(\mathrm{II})} \tag{A.37}$$

For $\Delta\bar{\hat{\Omega}}_{y2}^{(\mathrm{II})} < 0$, it is sufficient to have $M^{(\mathrm{II})} < M^{\dagger(\mathrm{II})}$. To see this, firstly recall that $\Delta\left|\zeta_2^{(\mathrm{II})}\right| = \frac{r-g}{\rho r}\lambda_I(M^{(\mathrm{II})} - M^{\dagger(\mathrm{II})})\Omega_e^{(\mathrm{I})}$. So $M^{(\mathrm{II})} < M^{\dagger(\mathrm{II})} \iff \Delta\left|\zeta_2^{(\mathrm{II})}\right| < 0$. Then the first term $(\zeta_2^{2(\mathrm{II})} - \zeta_2^{\dagger 2(\mathrm{II})})\tau_x^{(\mathrm{II})} = \Delta\left|\zeta_2^{(\mathrm{II})}\right|(\left|\zeta_2^{(\mathrm{II})}\right| + \left|\zeta_2^{\dagger(\mathrm{II})}\right|)\tau_x^{(\mathrm{II})} < 0$. Then I pin down the sign of the sum of the second and third terms:

$$\begin{aligned}
&(\frac{\lambda_I}{\lambda_2}M^{(\mathrm{II})} - \frac{\lambda_I}{\lambda_1}M^{\dagger(\mathrm{II})})\Omega_e^{(\mathrm{II})} + (\frac{\lambda_1}{\lambda_2} - 1)\Omega_\epsilon^{(\mathrm{II})} \\
=& \frac{\lambda_I}{\lambda_2}(M^{(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}) - \frac{\lambda_I}{\lambda_1}(M^{\dagger(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}) + (\frac{\lambda_J}{\lambda_2} - \frac{\lambda_J}{\lambda_1})\Omega_\epsilon^{(\mathrm{II})} \\
=& \frac{\lambda_I}{\lambda_1\lambda_2}[\lambda_1(M^{(\mathrm{II})} - M^{\dagger(\mathrm{II})})\Omega_e^{(\mathrm{II})} - \lambda_K(M^{\dagger(\mathrm{II})}\Omega_e^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}) - \frac{\lambda_J}{\lambda_2}\lambda_K\Omega_\epsilon^{(\mathrm{II})}] < 0.
\end{aligned}$$

For $\Delta\bar{\hat{\Omega}}_{y2}^{(\mathrm{II})} > 0$, $M^{(\mathrm{II})} > M^{\dagger(\mathrm{II})}$ is equivalent to the first term of (A.37) being positive. For the sum of the second and the third term to be positive, we need an extra condition: $\lambda_K < \frac{(M^{(\mathrm{II})}-M^{\dagger(\mathrm{II})})\Omega_e^{(\mathrm{II})}}{M^{\dagger(\mathrm{II})}\Omega_e^{(\mathrm{II})}+\frac{\lambda_1}{\lambda_I}\Omega_\epsilon^{(\mathrm{II})}}(\lambda_I + \lambda_J)$. □

Proof. The volatility of type-I risky asset is $\mathbb{V}ar(\boldsymbol{p}_2^{(\mathrm{I})} \mid \mathcal{I}_{l1^+}) = \boldsymbol{C}_2^2\boldsymbol{\Sigma}_y + \boldsymbol{D}_2^2\boldsymbol{\Sigma}_x$.

$$\begin{aligned}
\Delta\mathbb{V}ar(\boldsymbol{p}_2^{(\mathrm{I})} \mid \mathcal{I}_{l1^+}) &= (\boldsymbol{C}_2^2 - \boldsymbol{C}_2^{2\dagger})\boldsymbol{\Sigma}_y + (\boldsymbol{D}_2^2 - \boldsymbol{D}_2^{2\dagger})\boldsymbol{\Sigma}_x \\
&= (\boldsymbol{C}_2 - \boldsymbol{C}_2^\dagger)(\boldsymbol{C}_2 + \boldsymbol{C}_2^\dagger)\boldsymbol{\Sigma}_y + (\boldsymbol{D}_2 - \boldsymbol{D}_2^\dagger)(\boldsymbol{D}_2 + \boldsymbol{D}_2^\dagger)\boldsymbol{\Sigma}_x.
\end{aligned}$$

A sufficient condition for $\Delta\mathbb{V}ar(\boldsymbol{p}_2^{(\mathrm{I})} \mid \mathcal{I}_{l1^+})$ to be positive is the two components both being positive. For the first one, □

Proof. Recall that $C_2^{(\mathrm{I})} = \frac{1}{r-g}\bar{\tilde{\sigma}}_{y2}^{2(\mathrm{I})}\left(\bar{\hat{\Omega}}_{y2}^{(\mathrm{I})} - \tau_y^{(\mathrm{I})}\right) > 0$, so the change in dividend innovation coefficient $\Delta C_2^{(\mathrm{I})} = C_2^{(\mathrm{I})} - C_2^{\dagger(\mathrm{I})} = \frac{\Omega_y^{(\mathrm{I})}}{r-g}(\bar{\tilde{\sigma}}_{y2}^{2\dagger(\mathrm{I})} - \bar{\tilde{\sigma}}_{y2}^{2(\mathrm{I})})$ are in the same direction as $\Delta\bar{\hat{\Omega}}_{y2}^{(\mathrm{I})}$. This line of reasoning directly applies to type-II risky asset as well. □

Proof. Recall that $D_2^{(\mathrm{I})} = \frac{1}{r-g}\frac{1}{\zeta_2^{(\mathrm{I})}}(1 - \frac{\Omega_y^{(\mathrm{I})}}{\bar{\hat{\Omega}}_{y2}^{(\mathrm{I})}}) < 0$, so $\Delta\left|D_2^{(\mathrm{I})}\right| = \frac{1}{r-g}\frac{1}{\left|\zeta_2^{(\mathrm{I})}\right|\left|\zeta^{\dagger(\mathrm{I})}\right|}(\left|\zeta_2^{\dagger(\mathrm{I})}\right| - \left|\zeta_2^{\dagger(\mathrm{I})}\right|\frac{\Omega_y^{(\mathrm{I})}}{\bar{\hat{\Omega}}_{y2}^{(\mathrm{I})}} - \left|\zeta_2^{(\mathrm{I})}\right| + \left|\zeta_2^{(\mathrm{I})}\right|\frac{\Omega_y^{(\mathrm{I})}}{\bar{\tilde{\Omega}}_{y2}^{\dagger(\mathrm{I})}})$. Substitute in $\left|\zeta_2^{\dagger(\mathrm{I})}\right| = \left|\zeta_2^{(\mathrm{I})}\right| - \Delta\left|\zeta_2^{(\mathrm{I})}\right|$, and we have $\Delta\left|D_2^{(\mathrm{I})}\right| = -\Delta\left|\zeta_2^{(\mathrm{I})}\right|(1 - \frac{\Omega_y^{(\mathrm{I})}}{\bar{\hat{\Omega}}_{y2}^{(\mathrm{I})}}) + \frac{\left|\zeta_2^{(\mathrm{I})}\right|\Omega_y^{(\mathrm{I})}}{\hat{\Omega}_{y2}^{(\mathrm{I})}\hat{\Omega}_{y2}^{\dagger(\mathrm{I})}}\Delta\bar{\hat{\Omega}}_{y2}^{(\mathrm{I})}$. □

A.2 Extension to full Opening: solutions and proofs

A.2.1 Equilibrium

In this extension, a measure of λ_H foreign individual investors also enter the domestic market, and both types of risky assets are available for foreign investment. At the portfolio choice stage, foreign individual investor h only has the price signal $\boldsymbol{s}_{pt}$, thus updating the conditional expectation and variance about $\widetilde{\boldsymbol{y}}_2$ following

$$\hat{\boldsymbol{\Sigma}}_{Hy2}^{\ddagger} \equiv \mathbb{V}ar(\widetilde{\boldsymbol{y}}_2 \mid \mathcal{I}_{h2}^{\ddagger}) = \mathbb{V}ar(\widetilde{\boldsymbol{y}}_2 \mid \boldsymbol{p}_2^{\ddagger}) = (\boldsymbol{\Omega}_y + \boldsymbol{\Omega}_{p2}^{\ddagger})^{-1},$$
$$\hat{\boldsymbol{\mu}}_{Hy2}^{\ddagger} \equiv \mathbb{E}[\widetilde{\boldsymbol{y}}_2 \mid \mathcal{I}_{h2}^{\ddagger}] = \mathbb{E}[\widetilde{\boldsymbol{y}}_2 \mid \boldsymbol{p}_2^{\ddagger}] = \hat{\boldsymbol{\Sigma}}_{Hy2}^{\ddagger}\boldsymbol{\Omega}_{p2}^{\ddagger}\boldsymbol{s}_{p2}^{\ddagger}.$$

The signal-to-noise ratio becomes

$$\boldsymbol{\zeta}_2^{\ddagger} = -\frac{1}{\rho r}(r\boldsymbol{I} - \boldsymbol{G})[(\lambda_I\beta_I^2\boldsymbol{M}_I^{\ddagger} + \lambda_K\beta^2\boldsymbol{M}_K^{\ddagger})\boldsymbol{\Omega}_e + \lambda_1\boldsymbol{\Omega}_\epsilon],$$

and the average posterior precision of dividend innovation is

$$\begin{aligned}\bar{\hat{\boldsymbol{\Omega}}}^{\ddagger}_{y2} &= \frac{1}{\lambda_3}\left(\lambda_I\hat{\boldsymbol{\Omega}}^{\ddagger}_{Iy2}+\lambda_J\hat{\boldsymbol{\Omega}}^{\ddagger}_{Jy2}+\lambda_K\hat{\boldsymbol{\Omega}}^{\ddagger}_{Ky2}+\lambda_H\hat{\boldsymbol{\Omega}}^{\ddagger}_{Hy2}\right)\\ &= \frac{1}{\lambda_3}\left[\lambda_I(\boldsymbol{\Omega}_y+\boldsymbol{\Omega}^{\ddagger}_{p2}+\beta_I^2\boldsymbol{M}_I^{\ddagger}\boldsymbol{\Omega}_e+\boldsymbol{\Omega}_\epsilon)+\lambda_J(\boldsymbol{\Omega}_y+\boldsymbol{\Omega}^{\ddagger}_{p2}+\boldsymbol{\Omega}_\epsilon)\right]\\ &\quad+\frac{1}{\lambda_3}\left[\lambda_K(\boldsymbol{\Omega}_y+\boldsymbol{\Omega}^{\ddagger}_{p2}+\beta^2\boldsymbol{M}_K^{\ddagger}\boldsymbol{\Omega}_e)+\lambda_H(\boldsymbol{\Omega}_y+\boldsymbol{\Omega}^{\ddagger}_{p2})\right]\\ &= \boldsymbol{\Omega}_y+\boldsymbol{\Omega}^{\ddagger}_{p2}+\frac{\lambda_I\beta_I^2\boldsymbol{M}_I^{\ddagger}+\lambda_K\beta^2\boldsymbol{M}_K^{\ddagger}}{\lambda_3}\boldsymbol{\Omega}_e+\frac{\lambda_1}{\lambda_3}\boldsymbol{\Omega}_\epsilon,\end{aligned}$$

where $\lambda_3=\lambda_I+\lambda_J+\lambda_K+\lambda_H$. The price coefficients update to

$$\boldsymbol{A}_2^{\ddagger}=\frac{1}{r}\boldsymbol{\mu}_v+(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{\mu}-\rho r(r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}\bar{\hat{\boldsymbol{\Sigma}}}^{\ddagger}_{y2}(r\boldsymbol{I}-\boldsymbol{G})^{-1}\frac{1}{\lambda_3}\bar{\boldsymbol{x}}_2, \tag{A.38}$$

$$\boldsymbol{B}_2^{\ddagger}=(r\boldsymbol{I}-\boldsymbol{G})^{-1}\boldsymbol{G}, \tag{A.39}$$

$$\boldsymbol{C}_2^{\ddagger}=(r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}(\boldsymbol{I}-\bar{\hat{\boldsymbol{\Sigma}}}^{\ddagger}_{y2}\boldsymbol{\Omega}_y^{\ddagger}), \tag{A.40}$$

$$\boldsymbol{D}_2^{\ddagger}=(r\boldsymbol{I}-\boldsymbol{G})^{-1\prime}(\boldsymbol{I}-\bar{\hat{\boldsymbol{\Sigma}}}^{\ddagger}_{y2}\boldsymbol{\Omega}_y^{\ddagger})\boldsymbol{\zeta}_2^{-1\ddagger}. \tag{A.41}$$

While the data choice problem of foreign institutional investors is distorted in the partial market opening model, when the market fully opens, they face a similar optimization problem to that of the domestic institutional investors. I first give their *ex ante* expected utility.

LEMMA 8. *At the beginning of Period* 2*, the ex ante utility of foreign institutional investor* k *is*

$$\mathbb{E}\left[e^{-\rho r w_{k2}}\mid\mathcal{I}_{k2^-}\right]=-e^{\rho r w_{k1}}\times\left|\boldsymbol{\Sigma}_{k2}\hat{\boldsymbol{\Sigma}}_{k2}^{-1}\right|^{-1/2}\times e^{-\frac{1}{2}\boldsymbol{\mu}_{k2}'\boldsymbol{\Sigma}_{k2}\boldsymbol{\mu}_{k2}}, \tag{A.42}$$

where $\boldsymbol{\Sigma}_{k2}$ *and* $\boldsymbol{\mu}_{k2}$*, the variance and expectation of excess payoff, are defined by (A.2) and (A.7), respectively.*

Proof.

The interim utility at the portfolio choice stage for institutional investor k is

$$\begin{aligned}\mathbb{E}\left[-e^{-\rho w_{k2}} \mid \mathcal{I}_{k2}\right] &= -e^{-\rho r w_{k1}}\mathbb{E}\left[e^{-(\mathbb{E}[\widetilde{\boldsymbol{d}}_2+\widetilde{\boldsymbol{v}}_2-\boldsymbol{p}_2 r|\mathcal{I}_{k2}])'\mathbb{V}ar^{-1}(\widetilde{\boldsymbol{d}}_2+\widetilde{\boldsymbol{v}}_2|\mathcal{I}_{k2})(\widetilde{\boldsymbol{d}}_2+\widetilde{\boldsymbol{v}}_2-\boldsymbol{p}_2 r)} \mid \mathcal{I}_{k2}\right] \\ &= -e^{-\rho r w_{k1}} \times e^{-\frac{1}{2}\hat{\boldsymbol{\mu}}_{k2}'\hat{\boldsymbol{\Sigma}}_{k2}^{-1}\hat{\boldsymbol{\mu}}_{k2}},\end{aligned}$$

where $\hat{\boldsymbol{\mu}}_{k2} \equiv \mathbb{E}[\widetilde{\boldsymbol{d}}_2 + \widetilde{\boldsymbol{v}}_2 - \boldsymbol{p}_2 r \mid \mathcal{I}_{k2}] = \boldsymbol{\mu}_v + r(r\boldsymbol{I}-\boldsymbol{G})^{-1}(\boldsymbol{\mu}+\boldsymbol{G}\boldsymbol{d}_1) + r(r\boldsymbol{I}-\boldsymbol{G})^{-1}\hat{\boldsymbol{\mu}}_{ky2} - \boldsymbol{p}_2 r$ and $\hat{\boldsymbol{\Sigma}}_{k2}$ is a function of known model parameters and data choice $(M_{K2}^{(\mathrm{I})}, M_{K2}^{(\mathrm{II})})$. Note that before making data choice in Period 2, $-\frac{1}{2}\hat{\boldsymbol{\mu}}_{k2}'\hat{\boldsymbol{\Sigma}}_{k2}^{-1}\hat{\boldsymbol{\mu}}_{k2}$ is a quadratic function of normal variables. It is useful to write them in terms of zero-mean normal variable $\hat{\boldsymbol{\mu}}_{k2} - \boldsymbol{\mu}_{k2}$. One has[2]

$$\begin{aligned}&\exp\left\{-\frac{1}{2}\hat{\boldsymbol{\mu}}_{k2}'\hat{\boldsymbol{\Sigma}}_{k2}^{-1}\hat{\boldsymbol{\mu}}_{k2}\right\} \\ =&\exp\left\{-\frac{1}{2}(\hat{\boldsymbol{\mu}}_{k2}-\boldsymbol{\mu}_{k2})'\hat{\boldsymbol{\Sigma}}_{k2}^{-1}(\hat{\boldsymbol{\mu}}_{k2}-\boldsymbol{\mu}_{k2}) + \boldsymbol{\mu}_{k2}\hat{\boldsymbol{\Sigma}}_{k2}^{-1}(\hat{\boldsymbol{\mu}}_{k2}-\boldsymbol{\mu}_{k2}) - \frac{1}{2}\boldsymbol{\mu}_{k2}'\hat{\boldsymbol{\Sigma}}_{k2}^{-1}\boldsymbol{\mu}_{k2}\right\} \\ =&\left|\boldsymbol{\Sigma}_{l2}\hat{\boldsymbol{\Sigma}}_{k2}^{-1}\right|^{-1/2}\exp\left\{-\frac{1}{2}\boldsymbol{\mu}_{k2}'\boldsymbol{\Sigma}_{k2}^{-1}\boldsymbol{\mu}_{k2}\right\}.\end{aligned}$$

□

Since $\boldsymbol{\mu}_{k2}^{-1}$ and $\boldsymbol{\Sigma}_{k2}^{-1}$ are unconditional moments that are only functions of known model parameters, the data choice $(M_{K2}^{(\mathrm{I})}, M_{K2}^{(\mathrm{II})})$ only enter the ex ante utility function though the posterior variance of excess payoff $\hat{\boldsymbol{\Sigma}}_{k2}^{-1}$. Then the data choice problem reduces to allocate data points to maximize the product of the posterior precision of the two types of firms, that is

$$\max_{\{M_{K2}^{(\mathrm{I})}, M_{K2}^{(\mathrm{II})}\}} \quad (\Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + M_{k2}^{(\mathrm{I})}\beta^2\Omega_e^{(\mathrm{I})})(\Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + M_{k2}^{(\mathrm{II})}\beta^2\Omega_e^{(\mathrm{II})}) \tag{A.43}$$

$$\text{s.t.} \quad M_{K2}^{(\mathrm{I})} + M_{K2}^{(\mathrm{II})} \leq M. \tag{A.44}$$

Intuitively, investors would

[2]See Veldkamp (2011). The general formula for the mean of the exponential of a multivariate quadratic form of normal variables is the following: If $z \sim N(0,\Sigma)$, F, G and H are non-random, then $\mathbb{E}\left[e^{z'Fz+G'z+H}\right] = |I-2\Sigma F|^{-1/2}\exp\left\{\frac{1}{2}G'(I-2\Sigma F)^{-1}\Sigma G + H\right\}$.

A.2.2 Proofs for full market opening impacts

Proof of proposition 9. The change of the signal-to-noise ratio for type-I risky asset brought in by full market opening is

$$\Delta\left|\boldsymbol{\zeta}_2^{\ddagger(\mathrm{I})}\right| \equiv \left|\boldsymbol{\zeta}_2^{\ddagger(\mathrm{I})}\right| - \left|\boldsymbol{\zeta}_2^{(\mathrm{I})}\right| = \frac{r-g}{\rho r}[\lambda_I(M_{I2}^{\ddagger(\mathrm{I})} - M_{I2}^{(\mathrm{I})})\beta_I^2 + \lambda_K(M_{K2}^{(\mathrm{I})\ddagger} - \bar{M})\beta^2]\Omega_e^{(\mathrm{I})}.$$

Notice that $M_{K2}^{(\mathrm{I})\ddagger} - \bar{M} < 0$. A sufficient condition for $\Delta\left|\boldsymbol{\zeta}_2^{\ddagger(\mathrm{I})}\right| < 0$ is $M_{I2}^{\ddagger(\mathrm{I})} < M_{I2}^{(\mathrm{I})}$, that is domestic institutional investors learn less about type-I risky asset if the market full opens instead of partially opening.

For $\Delta\left|\zeta_2^{(\mathrm{I})}\right| > 0$, we need $\lambda_I(M_{I2}^{\ddagger(\mathrm{I})} - M_{I2}^{(\mathrm{I})})\beta_I^2 + \lambda_K(M_{K2}^{(\mathrm{I})\ddagger} - \bar{M})\beta^2 > 0$, which necessarily requires that domestic institutional investors increases learning for type-I risky asset, i.e. $M_{I2}^{\ddagger(\mathrm{I})} > M_{I2}^{(\mathrm{I})}$. Then the number of foreign institutional investors $\lambda_K < \frac{(M_{I2}^{\ddagger(\mathrm{I})} - M_{I2}^{(\mathrm{I})})\beta_I^2}{(\bar{M} - M_{K2}^{(\mathrm{I})\ddagger})\beta^2}\lambda_I$. □

Proof of proposition 10. The change of the signal-to-noise ratio for type-II risky asset brought in by full market opening is

$$\Delta\left|\boldsymbol{\zeta}_2^{\ddagger(\mathrm{II})}\right| \equiv \left|\boldsymbol{\zeta}_2^{\ddagger(\mathrm{II})}\right| - \left|\boldsymbol{\zeta}_2^{(\mathrm{II})}\right| = \frac{r-g}{\rho r}[\lambda_I(M_{I2}^{\ddagger(\mathrm{II})} - M_{I2}^{(\mathrm{II})})\beta_I^2 + \lambda_K M_{K2}^{(\mathrm{II})\ddagger}\beta^2]\Omega_e^{(\mathrm{II})}.$$

Since $\lambda_K M_{K2}^{(\mathrm{II})\ddagger}\beta^2 > 0$, a sufficient condition for $\Delta\left|\boldsymbol{\zeta}_2^{\ddagger(\mathrm{II})}\right| > 0$ is $M_{I2}^{\ddagger(\mathrm{II})} > M_{I2}^{(\mathrm{I})}$, that is domestic institutional investors learn more about type-II risky asset if the market full opens instead of partially opening.

For $\Delta\left|\zeta_2^{(\mathrm{I})}\right| < 0$, we need $\lambda_I(M_{I2}^{\ddagger(\mathrm{II})} - M_{I2}^{(\mathrm{II})})\beta_I^2 + \lambda_K M_{K2}^{(\mathrm{II})\ddagger}\beta^2 < 0$, which necessarily requires that domestic institutional investors decreases learning for type-II risky asset, i.e. $M_{I2}^{\ddagger(\mathrm{II})} < M_{I2}^{(\mathrm{II})}$. A necessary condition is $-\lambda_I\bar{M}\beta_I^2 + \lambda_K M_{K2}^{(\mathrm{II})\ddagger}\beta^2 < 0$, i.e. $\lambda_K < \frac{M_{K2}^{(\mathrm{II})\ddagger}\beta^2}{\bar{M}\beta_I^2}\lambda_I$. □

Proof of Proposition 11. The change of the average posterior precision of dividend innovation for

type-I risky asset brought in by full market opening is

$$\begin{aligned}\Delta\bar{\hat{\mathbf{\Omega}}}_{y2}^{\ddagger(\mathrm{I})} \equiv \bar{\hat{\mathbf{\Omega}}}_{y2}^{\ddagger(\mathrm{I})} - \bar{\hat{\Omega}}_{y2}^{(\mathrm{I})} &= (\tau_y^{(\mathrm{I})} + \boldsymbol{\zeta}_2^{\ddagger 2(\mathrm{I})}\tau_x^{(\mathrm{I})} + \frac{\lambda_I M_{I2}^{\ddagger(\mathrm{I})}\beta_I^2 + \lambda_K M_{K2}^{\ddagger(I)}\beta^2}{\lambda_3}\tau_e^{(\mathrm{I})} + \frac{\lambda_1}{\lambda_3}\tau_\epsilon^{(\mathrm{I})}) \\ &\quad -(\tau_y^{(\mathrm{I})} + \zeta_2^{2(\mathrm{I})}\tau_x^{(\mathrm{I})} + \frac{\lambda_I M^{(\mathrm{I})}\beta_I^2 + \lambda_K \bar{M}\beta^2}{\lambda_3}\tau_e^{(\mathrm{I})} + \frac{\lambda_1}{\lambda_3}\tau_\epsilon^{(\mathrm{I})}) \\ &= (\boldsymbol{\zeta}_2^{\ddagger 2(\mathrm{I})} - \boldsymbol{\zeta}_2^{2(\mathrm{I})})\tau_x^{(\mathrm{I})} + \frac{1}{\lambda_3}[\lambda_I(M_{I2}^{\ddagger(\mathrm{I})} - M_{I2}^{(\mathrm{I})})\beta_I^2 + \lambda_K(M_{K2}^{\ddagger(\mathrm{I})} - \bar{M})\beta^2]\Omega_e^{(\mathrm{I})} \\ &= \Delta\left|\zeta^{\ddagger(\mathrm{I})}\right|\left[(\left|\zeta_2^{\ddagger(\mathrm{I})}\right| + \left|\zeta_2^{(\mathrm{I})}\right|)\tau_x^{(\mathrm{I})} + \frac{1}{\lambda_3}\frac{\rho r}{r-g}\right].\end{aligned}$$

Therefore, $\Delta\bar{\hat{\mathbf{\Omega}}}_{y2}^{\ddagger(\mathrm{I})}$ and $\Delta\left|\zeta^{\ddagger(\mathrm{I})}\right|$ have the same sign. The result applies to type-II risky asset as well. □

A.3 Technological advance

The *ex ante* utility $\mathbb{E}[-e^{-\rho r w_{i2}}] \propto -\left|\hat{\mathbf{\Omega}}_{i2}\right|^{-1/2} e^{\frac{1}{2}\rho r a \beta_{i2}^2}$. Thus the objective function of the data choice problem is equivalent to $\ln\left|\hat{\mathbf{\Omega}}_{I2}\right| - \rho r a \beta_{i2}^2$. Recall that $\hat{\mathbf{\Omega}}_{i2} = \frac{1}{r^2}(r\boldsymbol{I} - \boldsymbol{G})'\hat{\mathbf{\Omega}}_{Iy2}(r\boldsymbol{I} - \boldsymbol{G})$. Then the optimizing problem further reduces to

$$\max_{\{\beta_{i2}, M_{i2}^{(\mathrm{I})}, M_{i2}^{(\mathrm{II})}\}} \quad \ln(\hat{\Omega}_{iy2}^{(\mathrm{I})}\hat{\Omega}_{iy2}^{(\mathrm{II})} - \mathcal{C}_{Iy2}^2) - \rho r a \beta_{i2}^2 \tag{A.45}$$

$$\text{s.t.} \quad \beta_I \le \beta_{i2} \le \beta, \tag{A.46}$$

$$M_{i2}^{(\mathrm{I})} + M_{i2}^{(\mathrm{II})} \le \bar{M}, \tag{A.47}$$

$$M_{i2}^{(\mathrm{I})}, M_{i2}^{(\mathrm{II})} \ge 0. \tag{A.48}$$

Recall that $\hat{\Omega}_{iy2}^{(\mathrm{I})} = \Omega_y^{(\mathrm{I})} + \hat{\Omega}_{p2}^{(\mathrm{I})} + M_{i2}^{(\mathrm{I})}\beta_{i2}^2\Omega_e^{(\mathrm{I})} + \Omega_\epsilon^{(\mathrm{I})}$ and $\hat{\Omega}_{iy2}^{(\mathrm{II})} = \Omega_y^{(\mathrm{II})} + \hat{\Omega}_{p2}^{(\mathrm{II})} + M_{i2}^{(\mathrm{II})}\beta_{i2}^2\Omega_e^{(\mathrm{II})} + \Omega_\epsilon^{(\mathrm{II})}$. The data processing capacity will be implemented with equality. Substituting $M_{i2}^{(\mathrm{II})} = \bar{M} - M_{i2}^{(\mathrm{I})}$ into the objective function, the F.O.C. with respective to $M_{i2}^{(\mathrm{I})}$ becomes

$$\frac{1}{\left|\hat{\mathbf{\Omega}}_{iy2}\right|}[\hat{\Omega}_{iy2}^{(\mathrm{II})}\beta_{i2}^2\Omega_e^{(\mathrm{I})} - \hat{\Omega}_{iy2}^{(\mathrm{I})}\beta_{i2}^2\Omega_e^{(\mathrm{II})}] = 0 \implies \frac{\Omega_e^{(\mathrm{I})}}{\hat{\Omega}_{iy2}^{(\mathrm{I})}} = \frac{\Omega_e^{(\mathrm{II})}}{\hat{\Omega}_{iy2}^{(\mathrm{II})}}. \tag{A.49}$$

The F.O.C. with respect to β_{i2} is

$$\frac{1}{\left|\hat{\mathbf{\Omega}}_{iy2}\right|}[2\hat{\Omega}^{(\text{II})}_{iy2}M^{(\text{I})}_{i2}\Omega^{(\text{I})}_{e}\beta_{i2} + 2\hat{\Omega}^{(\text{I})}_{iy2}(\bar{M} - M^{(\text{I})}_{i2})\Omega^{(\text{II})}_{e}\beta_{i2}] - 2\rho ra\beta_{i2} = 0$$

Plug in (A.49), rearrange, and we have

$$\bar{M}\hat{\Omega}^{(\text{I})}_{iy2}\Omega^{(\text{II})}_{e} = \rho ra\left|\hat{\mathbf{\Omega}}_{iy2}\right|,$$

which further reduces to a quadratic function of $\hat{\Omega}^{(\text{I})}_{iy2}$:

$$\rho ra\frac{\Omega^{(\text{II})}_{e}}{\Omega^{(\text{I})}_{e}}(\hat{\Omega}^{(\text{I})}_{iy2})^2 - \bar{M}\Omega^{(\text{II})}_{e}\hat{\Omega}^{(\text{I})}_{iy2} - \rho ra\hat{\mathcal{C}}^2_{Iy2} = 0.$$

Its determinant $\bar{M}^2(\Omega^{(\text{II})}_{e})^2 + 4(\rho ra)^2\hat{\mathcal{C}}^2_{Iy2}\frac{\Omega^{(\text{II})}_{e}}{\Omega^{(\text{I})}_{e}}$ is always positive. The unique positive solution is

$$\hat{\Omega}^{(\text{I})}_{iy2} = \frac{\bar{M}\Omega^{(\text{II})}_{e} + \sqrt{\bar{M}^2(\Omega^{(\text{II})}_{e})^2 + 4(\rho ra)^2\hat{\mathcal{C}}^2_{Iy2}\frac{\Omega^{(\text{II})}_{e}}{\Omega^{(\text{I})}_{e}}}}{2\rho ra\frac{\Omega^{(\text{II})}_{e}}{\Omega^{(\text{I})}_{e}}},$$

with the negative solution discarded. This interior solution has boundaries

$$\frac{\bar{M}\Omega^{(\text{I})}_{e}}{\rho ra} \leq \hat{\Omega}^{(\text{I})}_{iy2} \leq \sqrt{\left(\frac{\bar{M}\Omega^{(\text{I})}_{e}}{\rho ra}\right)^2 + 4\hat{\mathcal{C}}^2_{Iy2}\frac{\Omega^{(\text{I})}_{e}}{\Omega^{(\text{II})}_{e}}}.$$

Notice that $\hat{\Omega}^{(\text{I})}_{iy2} = \Omega^{(\text{I})}_{y} + \hat{\Omega}^{(\text{I})}_{p2} + \beta_i^2 M^{(\text{I})}_{i2}\Omega^{(\text{I})}_{e} + \Omega^{(\text{I})}_{\epsilon}$, so it is also bounded by $[\Omega^{(\text{I})}_{y} + \hat{\Omega}^{(\text{I})}_{p2} + \Omega^{(\text{I})}_{\epsilon}, \Omega^{(\text{I})}_{y} + \hat{\Omega}^{(\text{I})}_{p2} + \beta^2\bar{M}\Omega^{(\text{I})}_{e} + \Omega^{(\text{I})}_{\epsilon}]$. Therefore, if $\frac{\bar{M}\Omega^{(\text{I})}_{e}}{\rho ra} \geq \Omega^{(\text{I})}_{y} + \hat{\Omega}^{(\text{I})}_{p2} + \beta^2\bar{M}\Omega^{(\text{I})}_{e} + \Omega^{(\text{I})}_{\epsilon}$, then investor i would choose $\beta_{i2} = \beta$ and $M^{(\text{I})}_{i2} = \bar{M}$ to achieve the upper bound. If $\sqrt{\left(\frac{\bar{M}\Omega^{(\text{I})}_{e}}{\rho ra}\right)^2 + 4\hat{\mathcal{C}}^2_{Iy2}\frac{\Omega^{(\text{I})}_{e}}{\Omega^{(\text{II})}_{e}}} \leq \Omega^{(\text{I})}_{y} + \hat{\Omega}^{(\text{I})}_{p2} + \Omega^{(\text{I})}_{\epsilon}$, then investor i would choose $M^{(\text{I})}_{i2} = 0$ to achieve the lower bound. In this case,

β_{i2} is pinned down by $\hat{\Omega}^{(\text{II})}_{iy2}$. By F.O.C. $\hat{\Omega}^{(\text{II})}_{iy2} = \frac{\Omega^{\text{II}}_e}{\Omega^{(\text{I})}_e}\hat{\Omega}^{(\text{I})}_{iy2}$. Notice that

$$M^{(\text{I})}_{i2}\beta^2_{i2} = \frac{1}{\Omega^{(\text{I})}_e}[\hat{\Omega}^{(\text{I})}_{iy2} - (\Omega^{(\text{I})}_y + \hat{\Omega}^{(\text{I})}_{p2} + \Omega^{(\text{I})}_\epsilon)] \geq 0,$$
$$(\bar{M} - M^{(\text{I})}_{i2})\beta^2_{i2} = \frac{1}{\Omega^{(\text{II})}_e}[\hat{\Omega}^{(\text{II})}_{iy2} - (\Omega^{(\text{II})}_y + \hat{\Omega}^{(\text{II})}_{p2} + \Omega^{(\text{II})}_\epsilon)] \geq 0.$$

Adding them up, we have

$$\beta^c_{i2} = \sqrt{\frac{1}{\bar{M}\Omega^{(\text{I})}_e}[\hat{\Omega}^{(\text{I})}_{iy2} - (\Omega^{(\text{I})}_y + \hat{\Omega}^{(\text{I})}_{p2} + \Omega^{(\text{I})}_\epsilon)] + \frac{1}{\bar{M}\Omega^{(\text{II})}_e}[\frac{\Omega^{\text{II}}_e}{\Omega^{(\text{I})}_e}\hat{\Omega}^{(\text{I})}_{iy2} - (\Omega^{(\text{II})}_y + \hat{\Omega}^{(\text{II})}_{p2} + \Omega^{(\text{II})}_\epsilon)]}.$$

Thus

$$\beta^*_{i2} = \begin{cases} \beta_I, & \text{if } \beta^c_{i2} \leq \beta_I \\ \beta, & \text{if } \beta^c_{i2} \geq \beta \\ \beta^c_{i2}, & \text{otherwise} \end{cases}.$$

Then M^*_{i2} is obtained by $M^*_{i2} = \frac{1}{\beta^{*2}_{i2}\Omega^{(\text{I})}_e}[\hat{\Omega}^{(\text{I})}_{iy2} - (\Omega^{(\text{I})}_y + \hat{\Omega}^{(\text{I})}_{p2} + \Omega^{(\text{I})}_\epsilon)]$.

A.4 Institutional Background

A.4.1 Market

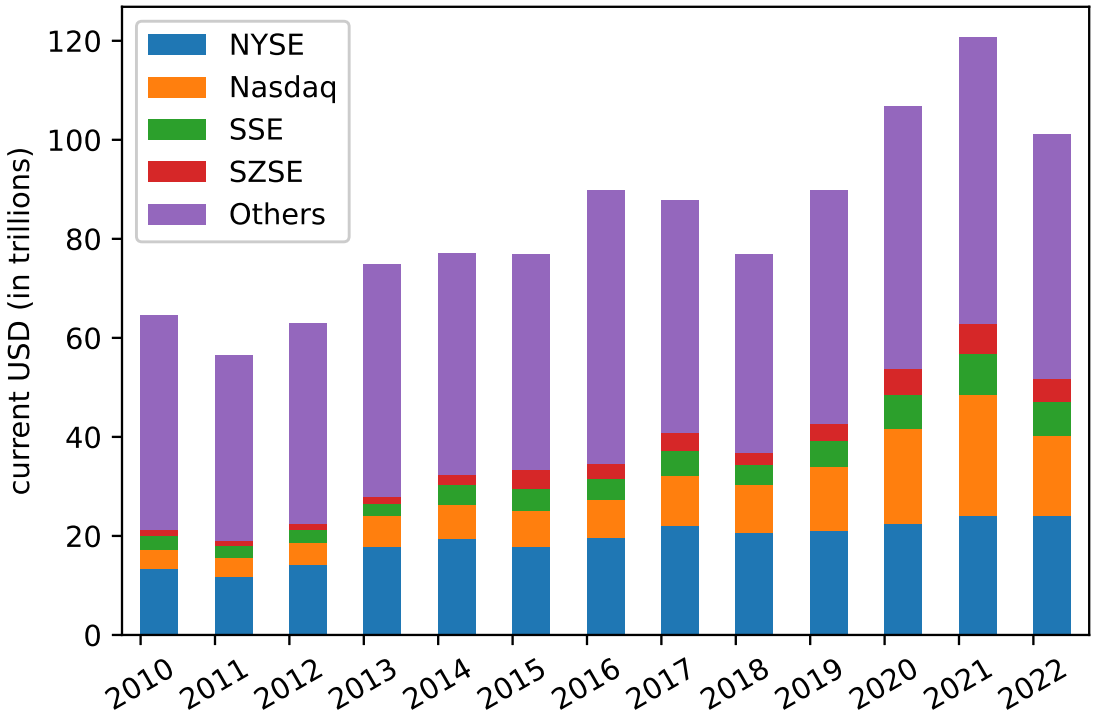


Figure A.1. Total equity market - market capitalization in current $trillion

A.4.2 China's Capital Market Opening Timeline

Following the Shanghai-Hong Kong Stock Connect, more programs are being launched by China to open up its capital market to the rest of the world, including the Shenzhen-Hong Kong Stock Connect (December 5, 2016) and the Shanghai-London Stock Connect (June 17, 2019). A large scale capital reallocation will be involved as China furthers its opening-up. As of November 22, a total of ¥273 billion capital net flow has been directed to China within one year. Furthermore, on Nov 7, 2019, the State Council of China announced a new policy that will "*eliminate all restrictions on the scope of business for foreign banks, securities companies and fund management*

Date	Event
Nov 17, 2014	The Shanghai-Hong Kong Stock Connect was launched.
Dec 5, 2016	The Shenzhen-Hong Kong Stock Connect was launched.
2018	China would relax the limit on foreign investment in joint-venture securities, fund management and futures companies to 51 percent.
June 17, 2019	The Shanghai-London Stock Connect was launched.
July 20, 2019	The State Council of China announced that the plan of lifting the foreign ownership limits on securities companies, fund management companies and futures companies would be advanced by one year to 2020.
Jan 1, 2020	The foreign ownership limit on futures companies was lifted. This limit was 51% for eligible foreign investors.
April 1, 2020	The foreign ownership limit on securities companies was lifted. The new rule is in accordance to implement Sino-US Phase 1 trade deal.

[a] The information was collected from the official website of various Chinese government agencies, including the State Council (http://english.www.gov.cn/statecouncil/).

Table A.1: Timeline of the Opening-Up of China's Capital Market[a]

companies, and remove requirements on total assets for the establishment of foreign-funded banks" (State Council of China, No. 23, 2019). Specifically in 2020, foreign investors will be allowed to hold more than 51% of shares in securities companies, fund management companies, futures companies, and life insurance companies. My model provides a new perspective from informational frictions result from imperfect data processing technologies and helps predict the impact of China's new policy of the full extension of business scope of foreign financial institutions.

A.4.3 Details of the Stock Connect Programs

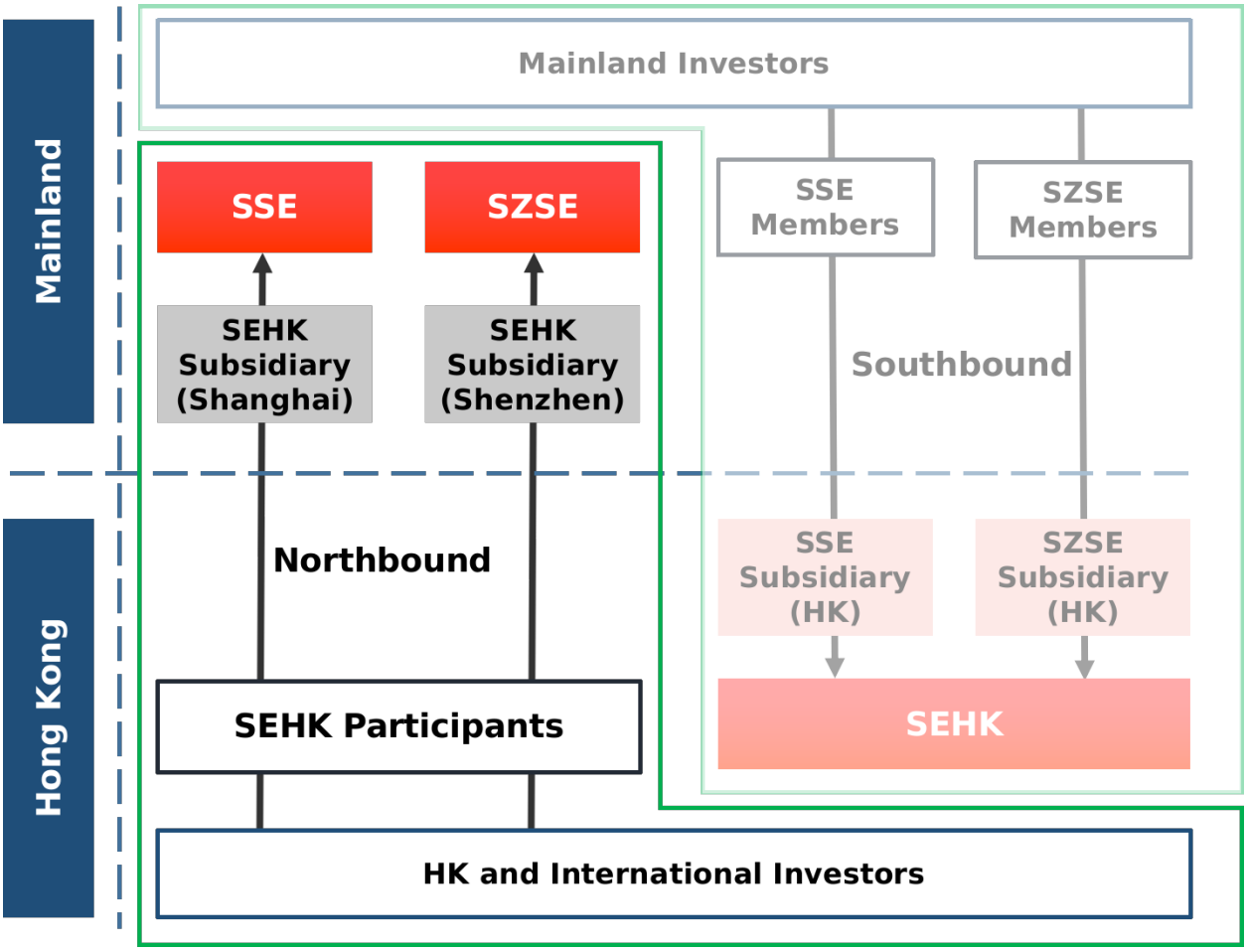


Figure A.2. Illustration of participants of the Shanghai and Shenzhen Connect programs

Table A.2: Details of the Stock Connect Programs

Panel A: The Shanghai-Hong Kong Stock Connect

Category	Restrictions
Northbound	
Eligible Shares	Constituents of the SSE 180 Index and SSE 380 Index, and shares of all companies listed on both SSE and SEHK.
Quotas[a]	Aggregate quota of RMB300 billion and a daily quota of RMB13 billion
Eligible Investors	No restrictions.
Southbound	
Eligible Shares	Constituents of the Hang Seng Composite LargeCap Index and Hang Seng Composite MidCap Index, and shares of all companies listed on both SSE and SEHK.
Quotas[a]	Aggregate quota of RMB250 billion and a daily quota of RMB10.5 billion.
Eligible Investors	Institutional investors, individual investors with an aggregate balance of not less than RMB 500,000 in their securities and cash accounts.

Panel B: The Shenzhen-Hong Kong Stock Connect

Category	Restrictions
Northbound	

(continued on next page)

Details of the Stock Connect Programs - continued

Eligible Shares	Constituents of the SZSE Component Index and SZSE Small/Mid Cap Innovation Index which has a market capitalisation of RMB6 billion or above, and all SZSE-listed shares of companies which have issued both A shares and H shares.
Quotas	Daily quota of RMB13 billion.
Eligible Investors	Investors eligible to trade shares that are listed on the ChiNext Board of SZSE under the Northbound Shenzhen Trading Link will be limited to institutional professional investors as defined in the relevant Hong Kong rules and regulations.

Southbound

Eligible Shares	Constituent stocks of the Hang Seng Composite LargeCap Index and Hang Seng Composite MidCap Index, any constituent stock of the Hang Seng Composite SmallCap Index which has a market capitalisation of HK$5 billion or above, and all SEHK-listed shares of companies which have issued both A shares and H shares.
Quotas	Daily quota of RMB10.5 billion.
Eligible Investors	Institutional investors, individual investors with an aggregate balance of not less than RMB 500,000 in their securities and cash accounts.

Panel C: The Shanghai-London Stock Connect

Category	Restrictions
Eastbound	

(continued on next page)

Details of the Stock Connect Programs - continued

Eligible Shares	Eligible companies listed on the London Stock Exchange will be able to issue Chinese Depository Receipts (CDRs) to Chinese investors and apply for them to be listed on the Main Board of the Shanghai Stock Exchange.
Quotas	Aggregate quota of RMB 250 billion.
Eligible Investors	Only qualified securities institutions in each of the two markets may conduct cross-border conversion business in relation to CDRs and GDRs.
Westbound	
Eligible Shares	Eligible companies listed on the Shanghai Stock Exchange will be able to issue Global Depository Receipts (GDRs) to UK and global investors and apply for them to be listed on the London Stock Exchange's Main Market.
Quotas	Aggregate quota of RMB 300 billion.
Eligible Investors	Only qualified securities institutions in each of the two markets may conduct cross-border conversion business in relation to CDRs and GDRs.

[a] The aggregate quota for the Shanghai-Hong Kong Stock Connect was abolished on Aug 16, 2016.

A-shares are issued to domestic investors and qualified foreign institutional investors (QFII). Launched on November 17, 2014, this program allows investors in Shanghai to trade stocks listed in the Hong Kong Exchange (HKEX, south-bound trades), and allows investors in Hong Kong to trade those in the SSE (north-bound trades). The Shanghai-Hong Kong Connect mainly covers index component stocks with large market cap listed in the SSE and HKEX. All Hong Kong and overseas investors will be allowed to trade eligible shares listed in Shanghai. However, only Mainland institutional investors and individual investors who have ¥500,000 in their investment and cash accounts are eligible to trade Hong Kong-listed shares. Since its inception, the total

trading volume through the Connect has amounted to ¥7.8 trillion, of which ¥4.3 trillion is for north-bound trades, and the rest is the trading volume of south-bound trades.

Appendix B: Appendix of Chapter 2

B.1 Model Solutions

B.1.1 Investors

This section works out investors' optimal information and portfolio choices, taking price coefficients and market aggregates as given. Without loss of generality, the model is solved in period t. Note that although A_{t+1} depends on $\tilde{x}_t$ (B.46), which is not $\mathcal{I}_{t-1}^{+}$- or $\mathcal{I}_{it}$-measurable, the solution to the optimization problems faced by investors are not affected.

Bayesian updating

At the beginning of period t, all $\mathcal{G}_t$-investors are identical and thus hold the same belief about period-t payoffs. Their information choices, $\tau_{\epsilon xt}$, $\tau_{\epsilon yt}$ and $\tau_{\nu xt}$, are also identical (we drop the i subscript for simplicity). Since the signal realizations $\mathcal{G}_t$-investors receive in period t are different, at the t-period portfolio choice stage, they hold heterogeneous beliefs about t-period payoffs. Some useful conditional variances that we can use to simplify notation is the conditional variance of dividend, current and future demand shocks in current period.

$$V_{yt} \equiv \mathbb{V}ar(\tilde{d}_t \mid \mathcal{I}_{it}) = \frac{1}{\tau_y + \tau_{\epsilon yt} + \tau_{pt}}, \tag{B.1}$$

$$V_{xt} \equiv \mathbb{V}ar(\tilde{x}_t \mid \mathcal{I}_{it}) = \frac{1}{\tau_x + \tau_{\epsilon xt}}, \tag{B.2}$$

$$V_{xt}^{\xi} \equiv \mathbb{V}ar(\tilde{x}_{t+1} \mid \mathcal{I}_{it}) = \frac{1}{\tau_x + \tau_{\nu xt}}. \tag{B.3}$$

On the aggregate level, the period-$(t+1)$ price coefficients are $\mathcal{I}_{it}$-measurable since there is a continuum of investors and thus the price coefficients depend on the distributions of shocks, not the realizations. Therefore, for $\mathcal{G}_t$-investor i, at period-t portfolio choice stage, the uncertainty

in the period-t excess payoff $\tilde{d}_t + \alpha p_{t+1} - p_t r$ comes from $\tilde{y}_t$, $\tilde{y}_{t+1}$, $\tilde{x}_t$, $\tilde{x}_{t+1}$ and $\tilde{x}_{t+2}$, and the uncertainty in the period-$(t+1)$ excess payoff $\tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1} r$ comes from $\tilde{y}_t$, $\tilde{y}_{t+1}$, $\tilde{y}_{t+2}$, $\tilde{x}_t$, $\tilde{x}_{t+1}$, $\tilde{x}_{t+2}$ and $\tilde{x}_{t+3}$. Thus the corresponding conditional expectation and variance of the period-t excess payoff are

$$\begin{aligned}
\mathcal{V}_t &\equiv \mathbb{V}ar(\tilde{d}_t + \alpha p_{t+1} - p_t r \mid \mathcal{I}_{it}) \\
&= (1 + \alpha B_{t+1})^2 V_{yt} + \alpha C_{t+1}^2 \tau_y^{-1} + \alpha D_{t+1}^2 V_{xt}^{\xi} + \alpha E_{t+1}^2 \tau_x^{-1}, \qquad \text{(B.4)} \\
\mathcal{E}_{it} &\equiv \mathbb{E}[\tilde{d}_t + \alpha p_{t+1} - p_t r \mid \mathcal{I}_{it}] \\
&= (1 + \alpha B_{t+1}) \left[\mu + g d_{t-1} + V_{yt} \left(\tau_{\epsilon yt} \tilde{\eta}_{it}^{y} + \tau_{pt} \tilde{\eta}_{it}^{p}\right)\right] + \alpha A_{t+1} + \alpha D_{t+1} V_{xt}^{\xi} \tau_{\nu xt} \tilde{\xi}_{it}^{x} - p_t r. \qquad \text{(B.5)}
\end{aligned}$$

For $\mathcal{G}_{t-1}$-investors, however, the signal realizations they receive in period $t-1$ are different, which adds another layer of heterogeneity to their information set. For $\mathcal{G}_{t-1}$-investor i at the period-t portfolio choice stage, conditional variances of the current dividend innovation, the current demand and the future demand shock are

$$\begin{aligned}
\hat{V}_{yit} &\equiv \mathbb{V}ar(\tilde{d}_t \mid \hat{\mathcal{I}}_{it}) = \frac{1}{\tau_y + \hat{\tau}_{\epsilon yit} + \hat{\tau}_{pit}}, \qquad \text{(B.6)} \\
\hat{V}_{xit} &\equiv \mathbb{V}ar(\tilde{x}_t \mid \hat{\mathcal{I}}_{it}) = \frac{1}{\tau_x + \hat{\tau}_{\epsilon xit} + \tau_{\nu xt-1}}, \qquad \text{(B.7)} \\
\hat{V}_{xit}^{\xi} &\equiv \mathbb{V}ar(\tilde{x}_{t+1} \mid \hat{\mathcal{I}}_{it}) = \frac{1}{\tau_x + \hat{\tau}_{\nu xit}}, \qquad \text{(B.8)}
\end{aligned}$$

For $\mathcal{G}_{t-1}$-investor i at period-t portfolio choice stage, the uncertainty in the period-t excess payoff $\tilde{d}_t + \alpha p_{t+1} - p_t r$ comes from $\tilde{y}_t$, $\tilde{y}_{t+1}$, $\tilde{x}_t$, $\tilde{x}_{t+1}$ and $\tilde{x}_{t+2}$, and the corresponding conditional expectation and variance of are

$$\begin{aligned}
\hat{\mathcal{V}}_{it} &\equiv \mathbb{V}ar(\tilde{d}_t + \alpha p_{t+1} - p_t r \mid \hat{\mathcal{I}}_{it}) = (1 + \alpha B_{t+1})^2 \hat{V}_{yit} + \alpha C_{t+1}^2 \tau_y^{-1} + \alpha D_{t+1}^2 \hat{V}_{xit}^{\xi} + \alpha E_{t+1}^2 \tau_x^{-1}, \qquad \text{(B.9)} \\
\hat{\mathcal{E}}_{it} &\equiv \mathbb{E}[\tilde{d}_t + \alpha p_{t+1} - p_t r \mid \hat{\mathcal{I}}_{it}] \\
&= (1 + \alpha B_{t+1}) \left[\mu + g d_{t-1} + \hat{V}_{yit} \left(\hat{\tau}_{\epsilon yit} \tilde{\hat{\eta}}_{it}^{y} + \hat{\tau}_{pit} \tilde{\hat{\eta}}_{it}^{p}\right)\right] + \alpha A_{t+1} + \alpha D_{t+1} \hat{V}_{xit}^{\xi} \hat{\tau}_{\nu xit} \tilde{\hat{\xi}}_{it}^{x} - p_t r. \qquad \text{(B.10)}
\end{aligned}$$

To obtain the expectation and variance of period-$(t+1)$ excess payoffs conditional on $\hat{\mathcal{I}}_{it+1}$, $\hat{V}_{it+1}$ and $\hat{E}_{it+1}$, we only need to change the time subscript t in (B.9) and (B.10) to $t+1$. Notice that $\hat{V}_i$ is $\mathcal{I}_{t-1}^+$-measurable, but $\hat{V}_{it+1}$ is only $\mathcal{I}_{it}$-measurable, not $\mathcal{I}_{t-1}^+$-measurable. This is left to verify after we have derived the ex-ante expected utility.

Optimal portfolio choice

In this section, I will solve for the optimal portfolio choice of $\mathcal{G}_t$-investor i in both period t and $t+1$. The solution will apply to any period t.

In period $t+1$, $\mathcal{G}_t$-investor i faces the standard one-shot portfolio choice problem, and the solution is

$$\hat{q}_{it+1} = \frac{\hat{\mathcal{E}}_{it+1}}{\rho \hat{\mathcal{V}}_{it+1}}. \tag{B.11}$$

This solution is taken as given in period t. But note that because the uncertainty in $\hat{\mathcal{E}}_{it+1}$ is not fully resolved at the portfolio choice stage of period t, the realization of $\hat{q}_{it+1}$ is not known, while its distribution is.

To solve the portfolio choice problem for $\mathcal{G}_t$-investor i in period t, plug in (2.3) to the expected utility $\mathbb{E}\left[-e^{-\rho \tilde{W}_{it+1^+}} \mid \mathcal{I}_{it}\right]$, which is maximized by choosing q_{i1}. The expected utility can be simplified to

$$\mathbb{E}\left[-e^{-\rho \tilde{W}_{it+1^+}} \mid \mathcal{I}_{it}\right] = -e^{-\rho r^2 W_0} \mathbb{E}\left[e^{-\rho r q_{i1}(\tilde{d}_t + \alpha p_{t+1} - p_t r) - \rho \hat{q}_{i2}(\tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1} r)} \mid \mathcal{I}_{it}\right].$$

To proceed, we first summarize the properties of the covariance term in the following lemma.

LEMMA 9. *At the t-period portfolio choice stage, the covariance of the t-period payoff and the $(t+1)$-period payoff of the risky asset for $\mathcal{G}_t$-investor i is*

$$\begin{aligned}
\mathcal{C}_t &\equiv \mathbb{C}ov\left(\hat{\mathcal{E}}_{it+1}\left(\tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1} r\right), \tilde{d}_t + \alpha p_{t+1} - p_t r \mid \mathcal{I}_{it}\right) \\
&= (2\gamma_1 \delta_2 \gamma_2^2 + \delta_1 \gamma_2 \gamma_2^\dagger) V_{yt} + \left[\gamma_1 \delta_3 (\gamma_3 + \gamma_3^\dagger) + \delta_1 \gamma_3 \gamma_3^\dagger\right] \tau_y \\
&\quad + \left[\gamma_1 \delta_4 (\gamma_4 + \gamma_4^\dagger) + \delta_1 \gamma_4 \gamma_4^\dagger\right] V_{xt}^\xi + \left[\gamma_1 \delta_5 (\gamma_5 + \gamma_5^\dagger) + \delta_1 \gamma_5 \gamma_5^\dagger\right] \tau_x.
\end{aligned} \tag{B.12}$$

For the static model, $\mathcal{C}_t = 0$.

Proof. For $\mathcal{G}_t$-investor i at her portfolio choice stage, the t-period payoff of the risky asset is

$$\tilde{d}_t + \alpha p_{t+1} \mid \mathcal{I}_{it} = \delta_1 + \delta_2 \tilde{y}_t \mid \tilde{\eta}_{it}^y, \tilde{\eta}_{it}^p + \delta_3 \tilde{y}_{t+1} + \delta_4 \tilde{x}_{t+1} \mid \tilde{\xi}_{it}^x + \delta_5 \tilde{x}_{t+2},$$

where $\delta_1 = (1 + \alpha B_{t+1})(\mu + g d_{t-1}) + \alpha A_{t+1}$, $\delta_2 = 1 + \alpha B_{t+1}$, $\delta_3 = \delta_3 = \alpha C_{t+1}$, $\delta_4 = \alpha D_{t+1}$ and $\delta_5 = \alpha E_{t+1}$.

The $(t+1)$-period excess payoff is

$$\begin{aligned}\tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1} r \mid \mathcal{I}_{it} = \gamma_1 + \gamma_2 \tilde{y}_t \mid \tilde{\eta}_{it}^y, \tilde{\eta}_{it}^p + \gamma_3^\dagger \tilde{y}_{t+1} + \gamma_4^\dagger \tilde{x}_{t+1} \mid \tilde{\xi}_{it}^x + \gamma_5^\dagger \tilde{x}_{t+2} \\ + \alpha \left(C_{t+2} \tilde{y}_{t+2} + E_{t+2} \tilde{x}_{t+3} \right),\end{aligned}$$

where $\gamma_1 = (1+\alpha B_{t+2})(\mu + g\mu + g^2 d_{t-1}) - r B_{t+1}(\mu + d_{t-1}) - r A_{t+1} + \alpha A_{t+2}$, $\gamma_2 = g(1 + \alpha B_{t+2}) - r B_{t+1}$, $\gamma_3^\dagger = \gamma_3^\dagger = 1 + \alpha B_{t+2} - r C_{t+1}$, $\gamma_4^\dagger = -r D_{t+1}$ and $\gamma_5^\dagger = \alpha D_{t+2} - r E_{t+1}$.

The expected period-$(t+1)$ excess payoff conditional on $\hat{\mathcal{I}}_{it+1}$ is a random variable at the portfolio choice stage of period t, denoted by $\hat{\mathcal{E}}_{it+1} \mid \mathcal{I}_{it}$. From (B.10), we have that

$$\begin{aligned}\hat{\mathcal{E}}_{it+1} \equiv (1 + \alpha B_{t+2}) \left[\mu + g \tilde{d}_t + \hat{V}_{yit+1} \left(\hat{\tau}_{\epsilon y it+1} \tilde{\hat{\eta}}_{it+1}^y + \hat{\tau}_{pit+1} \tilde{\hat{\eta}}_{it+1}^p \right) \right] \\ + \alpha A_{t+2} + \alpha D_{t+2} \hat{V}_{xit+1}^{\xi} \hat{\tau}_{\nu x it+1} \tilde{\hat{\xi}}_{it+1}^x - p_{t+1} r.\end{aligned}$$

Recall that the period-$t+1$ price signal for $\mathcal{G}_t$-investor i can be written as

$$\begin{aligned}\tilde{\hat{\eta}}_{it+1}^p &= \tilde{y}_{t+1} + \frac{1}{\zeta_{1t+1}} \left(\tilde{x}_{t+1} - \mathbb{E}\left[\tilde{x}_{t+1} \mid \tilde{\eta}_{it+1}^x, \tilde{\hat{\xi}}_{it}^x \right] \right) + \frac{1}{\zeta_{2t+1}} \left(\tilde{x}_{t+2} - \mathbb{E}\left[\tilde{x}_{t+2} \mid \tilde{\hat{\xi}}_{it+1}^x \right] \right) \\ &= \tilde{y}_{t+1} + \frac{1}{\zeta_{1t+1}} \left(\tilde{x}_{t+1} - \hat{V}_{xit+1} \left(\hat{\tau}_{\epsilon x it+1} \tilde{\eta}_{it+1}^x + \tau_{\nu x t} \tilde{\hat{\xi}}_{it}^x \right) \right) + \frac{1}{\zeta_{2t+1}} \left(\tilde{x}_{t+2} - \hat{V}_{xit+1}^{\xi} \hat{\tau}_{\nu x it+1} \tilde{\hat{\xi}}_{it+1}^x \right).\end{aligned}$$

From the above two equations, we know

$$\hat{\mathcal{E}}_{it+1} \mid \mathcal{I}_{it} = \gamma_1 + \gamma_2 \tilde{y}_t \mid \tilde{\eta}_{it}^y, \tilde{\eta}_{it}^p + \gamma_3 \tilde{y}_{t+1} + \gamma_4 \tilde{x}_{t+1} \mid \tilde{\xi}_{it}^x + \gamma_5 \tilde{x}_{t+2},$$

where $\gamma_3 = (1+\alpha B_{t+2})\hat{V}_{yit+1}(\hat{\tau}_{\epsilon yit+1}+\hat{\tau}_{pit+1}) - rC_{t+1}$, $\gamma_4 = \alpha D_{t+2}\hat{V}^{\xi}_{xit+1}\hat{\tau}_{\nu xit+1} - rD_{t+1} + (1+\alpha B_{t+2})\hat{V}_{yit+1}\hat{\tau}_{pit+1}\frac{1}{\zeta_{1t+1}}\hat{V}_{xit+1}\tau_x$ and $\gamma_5 = (1+\alpha B_{t+2})\hat{V}_{yit+1}\hat{\tau}_{pit+1}\frac{1}{\zeta_{2t+1}}\hat{V}^{\xi}_{xit+1}\tau_x - rE_{t+1}$.

Using the assumption that $\tilde{y}_t$, $\tilde{y}_{t+1}$, $\tilde{x}_{t+1}$ and $\tilde{x}_{t+2}$ are zero-mean and mutually independent, it follows that all third moments and cross moments are zero. Thus

$$\begin{aligned}\mathcal{C}_t &\equiv \mathbb{C}ov\left(\hat{\mathcal{E}}_{it+1}\left(\tilde{d}_{t+1}+\alpha p_{t+2}-p_{t+1}r\right), \tilde{d}_t+\alpha p_{t+1}-p_t r \mid \mathcal{I}_{it}\right)\\ &= (2\gamma_1\delta_2\gamma_2+\delta_1\gamma_2\gamma_2^{\dagger})V_{yt} + \left[\gamma_1\delta_3(\gamma_3+\gamma_3^{\dagger})+\delta_1\gamma_3\gamma_3^{\dagger}\right]\tau_y\\ &\quad + \left[\gamma_1\delta_4(\gamma_4+\gamma_4^{\dagger})+\delta_1\gamma_4\gamma_4^{\dagger}\right]V^{\xi}_{xt} + \left[\gamma_1\delta_5(\gamma_5+\gamma_5^{\dagger})+\delta_1\gamma_5\gamma_5^{\dagger}\right]\tau_x.\end{aligned}$$

The stationary solution of B_t gives that $\gamma_2 = (1+\alpha B_{t+2})g - rB_{t+1} = 0$. In the static model where $\alpha = 0$, we have $\mathcal{C}_t = 0$. □

On the other hand, the *ex ante* choice of precision of signals in the next period is irrelevant absent of realizations of those signals. In other words, $\tau_{\epsilon xit+1}$, $\tau_{\epsilon yit+1}$, $\tau_{\nu xit+1}$ do not enter the objective function and thus do not affect q_{i1} and q_{it+1} if investors are still at the portfolio choice stage of period t. We summarize it in the following lemma.

LEMMA 10. *At the portfolio choice stage of period t, $\mathcal{G}_t$-investor i does not need to consider information choice of next period.*

Therefore, the optimization problem can be simplified to

$$\max_{q_{i1}} rq_{it}\mathcal{E}_{it} - \frac{\rho}{2}r^2q_{i1}^2\mathcal{V}_t - r\mathcal{C}_t\frac{1}{\hat{\mathcal{V}}_{it+1}}q_{i1},$$

which implies

$$q_{i1} = \frac{\mathcal{E}_{it}}{\rho r\mathcal{V}_t} - \frac{\mathcal{C}_t}{\rho r\mathcal{V}_t\hat{\mathcal{V}}_{it+1}}. \tag{B.13}$$

In the static case ($\alpha = 0$), $\mathcal{C}_t = 0$, which greatly simplifies the portfolio choices to

$$q_{i1} = \frac{\mathbb{E}[\tilde{d}_t \mid \mathcal{I}_{it}] - p_t r}{\rho r\mathbb{V}ar(\tilde{d}_t \mid \mathcal{I}_{it})}, \quad \hat{q}_{it} = \frac{\mathbb{E}[\tilde{d}_t \mid \hat{\mathcal{I}}_{it}] - p_t r}{\rho\mathbb{V}ar(\tilde{d}_t \mid \hat{\mathcal{I}}_{it})}. \tag{B.14}$$

Optimal Information Choices

In this section, I calculate the *ex ante* expected utility of $\mathcal{G}_t$- and $\mathcal{G}_{t-1}$-investors at the beginning of period t when they have not made any period-t information or portfolio choice, and then solve their information choice problems, that is to maximize the *ex ante* expected utility subject to the information capacity constraint. The *ex ante* expected utility of $\mathcal{G}_{t-1}$-investor i is summarized by the following proposition.

PROPOSITION 25. *At the beginning of period t, the ex ante utility of $\mathcal{G}_{t-1}$ is*

$$\mathbb{E}\left[e^{-\rho W_{it^+}} \mid \hat{\mathcal{I}}_{it-1^+}\right] = -e^{-\rho r W_{it-1^+}} \left|\frac{\hat{\Sigma}_{t-1^+}}{\hat{\mathcal{V}}_{it}}\right|^{-1/2} \exp\left\{-\frac{\mu_{\hat{\mathcal{E}}_t}^2}{2\hat{\Sigma}_{t-1^+}}\right\}. \tag{B.15}$$

where $\mu_{\hat{\mathcal{E}}_t}$ and $\hat{\Sigma}_{t-1^+}$ are defined by (B.18) and (B.17), respectively.

Proof.

Step 1: Calculate $\mathbb{E}\left[-e^{-\rho W_{it^+}} \mid \hat{\mathcal{I}}_{it}\right]$

Conditional on $\hat{\mathcal{I}}_{it}$, the end-of-period-t wealth W_{it^+} is log-normally distributed, and thus $\mathbb{E}\left[-e^{-\rho W_{it^+}} \mid \hat{\mathcal{I}}_{it}\right] = -e^{-\rho\mathbb{E}\left[W_{it^+}|\hat{\mathcal{I}}_{it}\right]+\frac{\rho^2}{2}\mathbb{V}ar\left(W_{it^+}|\hat{\mathcal{I}}_{it}\right)}$. Combining the budget constraint $W_{it^+} = rW_{it-1^+}+\hat{q}_{it}\left(\tilde{d}_t + \alpha p_{t+1} - p_t r\right)$ with the portfolio choice of $\mathcal{G}_{t-1}$-investors (2.8), we have $\mathbb{E}\left[W_{it^+} \mid \hat{\mathcal{I}}_{it}\right] = rW_{it-1^+} + \frac{\hat{\mathcal{E}}_{it}^2}{\rho\hat{\mathcal{V}}_{it}}$ and $\mathbb{V}ar\left(W_{it^+} \mid \hat{\mathcal{I}}_{it}\right) = \frac{\hat{\mathcal{E}}_{it}^2}{\rho^2\hat{\mathcal{V}}_{it}}$. Therefore,

$$\mathbb{E}\left[-e^{-\rho W_{it^+}} \mid \hat{\mathcal{I}}_{it}\right] = -e^{-\rho r W_{it-1^+}} \times e^{-\frac{\hat{\mathcal{E}}_{it}^2}{2\hat{\mathcal{V}}_{it}}}. \tag{B.16}$$

Step 2: Calculate $\mathbb{E}\left[-e^{-\rho W_{it^+}} \mid \hat{\mathcal{I}}_{it-1^+}\right]$

By law of iterated expectations,

$$\mathbb{E}\left[-e^{-\rho W_{it^+}} \mid \hat{\mathcal{I}}_{it-1^+}\right] = \mathbb{E}\left[\mathbb{E}\left[-e^{-\rho W_{it^+}} \mid \hat{\mathcal{I}}_{it}\right] \mid \hat{\mathcal{I}}_{it-1^+}\right] = -e^{-\rho r W_{it-1^+}} \times \mathbb{E}\left[e^{-\frac{\hat{\mathcal{E}}_{it}^2}{2\hat{\mathcal{V}}_{it}}} \mid \hat{\mathcal{I}}_{it-1^+}\right].$$

Let $\mu_{\hat{\mathcal{E}}_t} \equiv \mathbb{E}\left[\hat{\mathcal{E}}_{it} \mid \hat{\mathcal{I}}^+_{it-1}\right]$ and $\hat{\Sigma}_{t-1^+} \equiv \mathbb{V}ar\left(\tilde{d}_t + \alpha p_{t+1} - p_t r \mid \hat{\mathcal{I}}^+_{it-1}\right)$. Conditional on $\hat{\mathcal{I}}_{it-1^+}$, $\hat{\mathcal{E}}_{it}$ is still normally distributed. By law of total variance, $\hat{\Sigma}_{t-1^+} = \mathbb{E}[\hat{\mathcal{V}}_{it} \mid \mathcal{I}^+_{it-1}] + \mathbb{V}ar(\hat{\mathcal{E}}_{it} \mid \mathcal{I}^+_{it-1})$. Since $\hat{V}_{it}$ is $\mathcal{I}^+_{it-1}$-measurable, we have $\mathbb{V}ar(\hat{\mathcal{E}}_{it} \mid \mathcal{I}^+_{it-1}) = \hat{\Sigma}_{t-1^+} - \hat{\mathcal{V}}_{it}$. And the total variance

$$\begin{aligned}
\hat{\Sigma}_{t-1^+} &\equiv \mathbb{V}ar\left(\tilde{d}_t + \alpha p_{t+1} - p_t r \mid \hat{\mathcal{I}}_{it-1^+}\right) \\
&= \left[(1 + \alpha B_{t+1} - rC_t)^2 + \alpha C^2_{t+1}\right]\tau_y + r^2 D_t^2 (\tau_x + \tau_{\nu x t-1})^{-1} \\
&\quad + \left[(\alpha D_{t+1} - rE_t)^2 + \alpha E^2_{t+1}\right]\tau_x. \qquad \text{(B.17)}
\end{aligned}$$

To compute the mean of $\hat{\mathcal{E}}_{it}$, by law of iterated expectations,

$$\begin{aligned}
\mu_{\hat{\mathcal{E}}_t} &\equiv \mathbb{E}\left[\hat{\mathcal{E}}_{it} \mid \hat{\mathcal{I}}_{it-1^+}\right] = \mathbb{E}\left[\tilde{d}_t + \alpha p_{t+1} - p_t r \mid \hat{\mathcal{I}}_{it-1^+}\right] \\
&= \frac{r}{r - \alpha g}\mu + \alpha A_{t+1} - rA_t - rD_t V^{\xi}_{xt-1}\tau_{\nu x t-1}\xi^x_{it-1}. \qquad \text{(B.18)}
\end{aligned}$$

To work out $\mathbb{E}\left[\exp\left\{-\frac{\hat{\mathcal{E}}^2_{it}}{2\hat{\mathcal{V}}_{it}}\right\} \mid \hat{\mathcal{I}}_{it-1^+}\right]$, it is useful to write it in terms of the mean-zero variable $\hat{\mathcal{E}}_{it} - \mu_{\hat{\mathcal{E}}_t}$[1]:

$$\mathbb{E}\left[e^{-\frac{\hat{\mathcal{E}}^2_{it}}{2\hat{\mathcal{V}}_{it}}} \mid \hat{\mathcal{I}}_{it-1^+}\right] = \mathbb{E}\left[\exp\left\{-\frac{\left(\hat{\mathcal{E}}_{it} - \mu_{\hat{\mathcal{E}}_t}\right)^2}{2\hat{\mathcal{V}}_{it}} - \frac{\left(\hat{\mathcal{E}}_{it} - \mu_{\hat{\mathcal{E}}_t}\right)\mu_{\hat{\mathcal{E}}_t}}{\hat{\mathcal{V}}_{it}} - \frac{\mu^2_{\hat{\mathcal{E}}_t}}{2\hat{\mathcal{V}}_{it}}\right\} \mid \hat{\mathcal{I}}_{it-1^+}\right].$$

Evaluating this expectation yields

$$\mathbb{E}\left[e^{-\frac{\hat{\mathcal{E}}^2_{it}}{2\hat{\mathcal{V}}_{it}}} \mid \hat{\mathcal{I}}_{it-1^+}\right] = \left|\frac{\hat{\Sigma}_{t-1^+}}{\hat{\mathcal{V}}_{it}}\right|^{-1/2} \exp\left\{-\frac{\mu^2_{\hat{\mathcal{E}}_t}}{2\hat{\Sigma}_{t-1^+}}\right\}.$$

□

The key feature of (B.15) is that period-t information choices only enter through $\hat{\mathcal{V}}_{it}$, which greatly simplifies the optimization problem, as stated in the following lemma.

[1] See Veldkamp (2011). The general formula for the mean of the exponential of a multivariate quadratic form of normal variables is the following: If $z \sim N(0, \Sigma)$, F, G and H are non-random, then $\mathbb{E}\left[e^{z'Fz+G'z+H}\right] = |I - 2\Sigma F|^{-1/2} \exp\left\{\frac{1}{2}G'(I - 2\Sigma F)^{-1}\Sigma G + H\right\}$.

LEMMA 11. *At the beginning of period t, the information choice facing $\mathcal{G}_{t-1}$-investor i is equivalent to*

$$\max_{\{\hat{\tau}_{\epsilon yit},\hat{\tau}_{\epsilon xit},\hat{\tau}_{\nu xit}\}} -\hat{\mathcal{V}}_{it}^{\dagger} \equiv -\frac{r^2}{(r-\alpha g)^2}\left[\tau_y+\hat{\tau}_{\epsilon yit}+\left(\frac{1}{\zeta_{1t}^2\left(\tau_x+\hat{\tau}_{\epsilon xit}+\tau_{\nu xt-1}\right)}+\frac{1}{\zeta_{2t}^2\left(\tau_x+\hat{\tau}_{\nu xit}\right)}\right)^{-1}\right]^{-1} -\alpha D_{t+1}^2\left(\tau_x+\hat{\tau}_{\nu xit}\right)^{-1}. \tag{B.19}$$

subject to information capacity constraint.

Proof. From (B.18) and (B.17), it is clear that $\mu_{\hat{\mathcal{E}}_t}$ and $\hat{\Sigma}_{t-1^+}$ are functions of model primitives, price coefficients, precision choice and signal realization of last period's future demand signal, but have nothing to do with this period's information choices, $\hat{\tau}_{\epsilon yit}$, $\hat{\tau}_{\epsilon xit}$ and $\hat{\tau}_{\nu xit}$. The information choices enter the *ex ante* utility (B.15) only through $\hat{\mathcal{V}}_{it}$. Therefore, maximizing the *ex ante* utility amounts to minimizing $\hat{\mathcal{V}}_{it}$, or maximizing $-\hat{\mathcal{V}}_{it}$. Using (B.9), and disposing of terms that only contains model primitives and price coefficients, one gets the equivalent objective function $\hat{\mathcal{V}}_{it}^{\dagger}$. □

PROPOSITION 26. *The optimal period-t information choice for $\mathcal{G}_{t-1}$-investor i exists and is unique. Furthermore, it is $\mathcal{I}_{t-1}^{+}$-measurable and thus homogeneous for all $\mathcal{G}_{t-1}$-investors.*

Proof. To begin with, let me calculate the marginal utility with respect to $\hat{\tau}_{\epsilon yit}$, $\hat{\tau}_{\epsilon xit}$ and $\hat{\tau}_{\nu xit}$.

$$\hat{L}_{\epsilon yit} \equiv -\frac{\partial\hat{\mathcal{V}}_{it}^{\dagger}}{\partial\hat{\tau}_{\epsilon yit}} = \frac{r^2}{(r-\alpha g)^2}\left[\tau_y+\hat{\tau}_{\epsilon yit}+\left(\frac{1}{\zeta_{1t}^2\left(\tau_x+\hat{\tau}_{\epsilon xit}+\tau_{\nu xt-1}\right)}+\frac{1}{\zeta_{2t}^2\left(\tau_x+\hat{\tau}_{\nu xit}\right)}\right)^{-1}\right]^{-2},$$

$$\hat{L}_{\epsilon xit} \equiv -\frac{\partial\hat{\mathcal{V}}_{it}^{\dagger}}{\partial\hat{\tau}_{\epsilon xit}} = \frac{r^2}{(r-\alpha g)^2}\left[\tau_y+\hat{\tau}_{\epsilon yit}+\left(\frac{1}{\zeta_{1t}^2\left(\tau_x+\hat{\tau}_{\epsilon xit}+\tau_{\nu xt-1}\right)}+\frac{1}{\zeta_{2t}^2\left(\tau_x+\hat{\tau}_{\nu xit}\right)}\right)^{-1}\right]^{-2}\times \left(\frac{1}{\zeta_{1t}^2\left(\tau_x+\hat{\tau}_{\epsilon xit}+\tau_{\nu xt-1}\right)}+\frac{1}{\zeta_{2t}^2\left(\tau_x+\hat{\tau}_{\nu xit}\right)}\right)^{-2}\zeta_{1t}^{-2}(\tau_x+\hat{\tau}_{\epsilon xit}+\tau_{\nu xt-1})^{-2},$$

$$\hat{L}_{\nu xit} \equiv -\frac{\partial\hat{\mathcal{V}}_{it}^{\dagger}}{\partial\hat{\tau}_{\nu xit}} = \frac{r^2}{(r-\alpha g)^2}\left[\tau_y+\hat{\tau}_{\epsilon yit}+\left(\frac{1}{\zeta_{1t}^2\left(\tau_x+\hat{\tau}_{\epsilon xit}+\tau_{\nu xt-1}\right)}+\frac{1}{\zeta_{2t}^2\left(\tau_x+\hat{\tau}_{\nu xit}\right)}\right)^{-1}\right]^{-2}\times \left(\frac{1}{\zeta_{1t}^2\left(\tau_x+\hat{\tau}_{\epsilon xit}+\tau_{\nu xt-1}\right)}+\frac{1}{\zeta_{2t}^2\left(\tau_x+\hat{\tau}_{\nu xit}\right)}\right)^{-2}\zeta_{1t}^{-2}(\tau_x+\hat{\tau}_{\nu xit})^{-2} +\alpha D_{t+1}^2(\tau_x+\hat{\tau}_{\nu xit})^{-2}.$$

It is easy to see that $\frac{\partial \hat{\mathcal{V}}_{it}^{\dagger}}{\partial \hat{\tau}_{\epsilon y i t}}, \frac{\partial \hat{\mathcal{V}}_{it}^{\dagger}}{\partial \hat{\tau}_{\epsilon x i t}}, \frac{\partial \hat{\mathcal{V}}_{it}^{\dagger}}{\partial \hat{\tau}_{\nu x i t}} > 0$. Taking second derivatives, one can also show that $\frac{\partial^2 \hat{\mathcal{V}}_{it}^{\dagger}}{\partial \hat{\tau}_{\epsilon y i t}^2}, \frac{\partial^2 \hat{\mathcal{V}}_{it}^{\dagger}}{\partial \hat{\tau}_{\epsilon x i t}^2}, \frac{\partial^2 \hat{\mathcal{V}}_{it}^{\dagger}}{\partial \hat{\tau}_{\nu x i t}^2} < 0$. Therefore, $\hat{\mathcal{V}}_{it}^{\dagger}$, thus the *ex ante* utility function (B.15), is strictly increasing and concave in information choices $\hat{\tau}_{\epsilon y i t}$, $\hat{\tau}_{\epsilon x i t}$ and $\hat{\tau}_{\nu x i t}$. Since the information capacity constraint is linear, the solution of the period-t information choice problem for $\mathcal{G}_{t-1}$ exists and is unique.

The F.O.C.'s for the constrained optimization problem is

$$\hat{L}_{\epsilon y i t} = \lambda \gamma_{\epsilon y}, \tag{B.20}$$

$$\hat{L}_{\epsilon x i t} = \lambda \gamma_{\epsilon x}, \tag{B.21}$$

$$\hat{L}_{\nu x i t} = \lambda \gamma_{\nu x}, \tag{B.22}$$

$$\gamma_{\epsilon y} \hat{\tau}_{\epsilon y i t} + \gamma_{\epsilon x} \hat{\tau}_{\epsilon x i t} + \gamma_{\nu x} \hat{\tau}_{\nu x i t} = K. \tag{B.23}$$

Dividing (B.20) by (B.21), one can express $\hat{\tau}_{\nu x i t}$ in terms of $\hat{\tau}_{\epsilon x i t}$, and dividing (B.20) by (B.22), one can express $\hat{\tau}_{\epsilon x i t}$ in terms of $\hat{\tau}_{\nu x i t}$ and $\hat{\tau}_{\epsilon x i t}$. Then plugging them back into (B.23) pins down $\hat{\tau}_{\epsilon x i t}$. Because the F.O.C.'s depend only on model primitives, price coefficients and $\tau_{\nu x t-1}$, and $\tau_{\nu x t-1}$ is by definition $\mathcal{I}_{t-1}^{+}$-measurable, period-t information choices of $\mathcal{G}_{t-1}$-investor i is also $\mathcal{I}_{t-1}^{+}$-measurable. □

This means that despite the fact that $\mathcal{I}_{it-1}^{+}$, the information set $\mathcal{G}_{t-1}$-investors hold at the beginning of period t is heterogeneous, the information that can be used for information choice is still homogeneous. Information choices for both periods are determined in the very beginning, and investors within the same generation all make the same information choices for both periods.

Lemma 12. *The uncertainty in the variance of excess payoff resolves earlier than the realization of corresponding shocks. More specifically,*

(1) $\mathcal{V}_t$, $\mathcal{V}_2$ and $\mathcal{C}$ are $\mathcal{I}_{t-1}^{+}$-measurable;

(2) $\hat{\mathcal{V}}_{it+1}$ is $\mathcal{I}_{t-1}^{+}$-measurable.

PROPOSITION 27. *The ex ante utility of $\mathcal{G}_t$ is*

$$\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}_{t-1}^+\right] = -e^{-\rho r^2 W_{it-1^+} + \frac{5\mathcal{C}_t^2}{8\mathcal{V}_t\hat{\mathcal{V}}_{it+1}^2} - \frac{\mathcal{C}_t^2}{8\Sigma_{t-1^+}^{\xi}\hat{\mathcal{V}}_{it+1}^2}} \left|\frac{\hat{\Sigma}_{t^+}}{\hat{\mathcal{V}}_{it+1}}\right|^{-1/2} \times \left|1 - 2\left(\tau_x^{-1} + \tau_{\nu xt}^{-1}\right)F\right|^{-1/2} e^{-\frac{1}{2}\frac{G^2}{\left(\tau_x^{-1}+\tau_{\nu xt}^{-1}\right)^{-1}-2F} + H}. \quad (B.24)$$

where F, G, and H are defined by (B.26), (B.27), and (B.28), respectively.

Proof. Plugging in the optimal portfolio choices into the *ex ante* expected utility, we have

$$\begin{aligned}&\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}_{t-1}^+\right] \\ &= -e^{-\rho r^2 W_0}\mathbb{E}\left[\exp\left\{-\left(\frac{\mathcal{E}_{it}}{\mathcal{V}_t} - \frac{\mathcal{C}_t}{\mathcal{V}_t\hat{\mathcal{V}}_{it+1}}\right)\left(\tilde{d}_t + \alpha p_{t+1} - p_t r\right) - \frac{\hat{\mathcal{E}}_{it+1}}{\hat{\mathcal{V}}_{it+1}}\left(\tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1} r\right)\right\}\right].\end{aligned}$$

According to Lemma 27, $\hat{\mathcal{V}}_{it+1}$ is $\mathcal{I}_{it}$-measurable. Thus the exponent term in the above expression is a quadratic form of normal variables under $\mathcal{I}_{it}$. Then we derive the *ex ante* expected utility in two steps.

Step 1: Calculate $\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \hat{\mathcal{I}}_{it+1}\right]$

Conditional on $\hat{\mathcal{I}}_{it+1}$, W_{it+1^+} is a normal random variable. Thus

$$\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \hat{\mathcal{I}}_{it+1}\right] = -e^{-\rho\mathbb{E}\left[W_{it+1^+}|\hat{\mathcal{I}}_{it+1}\right] + \frac{\rho^2}{2}\mathbb{V}ar\left(W_{it+1^+}|\hat{\mathcal{I}}_{it+1}\right)}. \quad (B.25)$$

Recall that $W_{it+1^+} = rW_{it^+} + \hat{q}_{i2}\left(\tilde{d}_{t+1} + \alpha p_{t+2} - p_{t+1}r\right)$. Plugging the optimal portfolio, we have $\mathbb{E}\left[W_{it+1^+} \mid \hat{\mathcal{I}}_{it+1}\right] = rW_{it^+} + \frac{\hat{\mathcal{E}}_{it+1}^2}{\rho\hat{\mathcal{V}}_{it+1}}$ and $\mathbb{V}ar\left(W_{it+1^+} \mid \hat{\mathcal{I}}_{it+1}\right) = \frac{\hat{\mathcal{E}}_{it+1}^2}{\rho^2\hat{\mathcal{V}}_{it+1}}$. Plugging them back into (B.25), we have

$$\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \hat{\mathcal{I}}_{it+1}\right] = -e^{-\rho r W_{it^+}} \times e^{-\frac{\hat{\mathcal{E}}_{it+1}^2}{2\hat{\mathcal{V}}_{it+1}}}.$$

Step 2: Calculating $\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}_{it}\right]$.

Applying (25) to period $t+1$, one has

$$\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}_{it}^+\right] = -e^{-\rho r W_{it^+}} \times \left|\frac{\hat{\Sigma}_{t^+}}{\hat{\mathcal{V}}_{it+1}}\right|^{-1/2} \exp\left\{-\frac{\mu^2_{\hat{\mathcal{E}}_{t+1}}}{2\hat{\Sigma}_{t^+}}\right\}.$$

On the one hand, note that $\hat{\Sigma}_{t^+}$ is $\mathcal{I}_{t-1}^+$-measurable, and that $\mu_{\hat{\mathcal{E}}_{t+1}}$ is $\mathcal{I}_{it}$-measurable. On the other hand, $W_{it^+} = rW_{it-1^+} + q_{i1}\left(\tilde{d}_t + \alpha p_{t+1} - p_t r\right)$, and $q_{i1} = \frac{\mathcal{E}_{it}}{\rho r \mathcal{V}_t} - \frac{\mathcal{C}_t}{\rho r \mathcal{V}_t \hat{\mathcal{V}}_{it+1}}$. Thus

$$\mathbb{E}\left[e^{-\rho r W_{it^+}} \mid \mathcal{I}_{it}\right] = e^{-\rho r^2 W_{it-1^+}} \times \mathbb{E}\left[e^{-\left(\frac{\mathcal{E}_{it}}{\mathcal{V}_t} - \frac{\mathcal{C}_t}{\mathcal{V}_t \hat{\mathcal{V}}_{it+1}}\right)\left(\tilde{d}_t + \alpha p_{t+1} - p_t r\right)} \mid \mathcal{I}_{it}\right].$$

The rightmost expectation is the expectation of the exponential of a linear combination of random variables, thus $\mathbb{E}\left[e^{-\left(\frac{\mathcal{E}_{it}}{\mathcal{V}_t} - \frac{\mathcal{C}_t}{\mathcal{V}_t \hat{\mathcal{V}}_{it+1}}\right)\left(\tilde{d}_t + \alpha p_{t+1} - p_t r\right)} \mid \mathcal{I}_{it}\right] = e^{-\frac{\mathcal{E}_{it}^2}{2\mathcal{V}_t} + \frac{\mathcal{C}_t \mathcal{E}_{it}}{2\mathcal{V}_t \hat{\mathcal{V}}_{it+1}} + \frac{\mathcal{C}_t^2}{2\mathcal{V}_t \hat{\mathcal{V}}_{it+1}^2}}$. Therefore, by law of iterated expectations,

$$\begin{aligned}
\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}_{it}\right] &= \mathbb{E}\left[\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}_{it}^+\right] \mid \mathcal{I}_{it}\right] \\
&= -e^{-\rho r^2 W_{it-1^+} + \frac{\mathcal{C}_t^2}{2\mathcal{V}_t \hat{\mathcal{V}}_{it+1}^2}} \left|\frac{\hat{\Sigma}_{t^+}}{\hat{\mathcal{V}}_{it+1}}\right|^{-1/2} \exp\left\{-\frac{\mathcal{E}_{it}^2}{2\mathcal{V}_t} + \frac{\mathcal{C}_t \mathcal{E}_{it}}{2\mathcal{V}_t \hat{\mathcal{V}}_{it+1}} - \frac{\mu^2_{\hat{\mathcal{E}}_{t+1}}}{2\hat{\Sigma}_{t^+}}\right\}.
\end{aligned}$$

Step 3: Calculate $\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}_{t-1}^+, \xi_{it}^x\right]$.

Conditional on $\mathcal{I}_{t-1}^+$ and ξ_t^x, randomness in the above equation only comes from $\mathcal{E}_{it}$. Let $\mu_{\mathcal{E}_t}^{\xi} \equiv \mathbb{E}\left[\mathcal{E}_{it} \mid \mathcal{I}_{t-1}^+, \xi_{it}^x\right]$ and $\Sigma_{t-1^+}^{\xi} \equiv \mathbb{V}ar\left(\tilde{d}_t + \alpha p_{t+1} - p_t r \mid \mathcal{I}_{t-1}^+, \xi_{it}^x\right)$.

$$\begin{aligned}
\mu_{\mathcal{E}_t}^{\xi} \equiv \mathbb{E}\left[\mathcal{E}_{it} \mid \mathcal{I}_{t-1}^+, \xi_t^x\right] &= \mathbb{E}\left[\tilde{d}_t + \alpha p_{t+1} - p_t r \mid \mathcal{I}_{t-1}^+, \xi_t^x\right] \\
&= \frac{r}{r - \alpha g}\mu + \alpha A_{t+1} - rA_t + \left(\alpha D_{t+1} - rE_t\right) V_{xt}^{\xi} \tau_{\nu x t} \xi_{it}^x.
\end{aligned}$$

And the total variance

$$\begin{aligned}
\Sigma_{t-1^+}^{\xi} &\equiv \mathbb{V}ar\left(\tilde{d}_t + \alpha p_{t+1} - p_t r \mid \hat{\mathcal{I}}_{it-1^+}\right) \\
&= \left[(1 + \alpha B_{t+1} - rC_t)^2 + \alpha C_{t+1}^2\right]\tau_y + \left(\alpha D_{t+1} - rE_t\right)^2 (\tau_x + \tau_{\nu x t})^{-1} + \left[r^2 D_t^2 + \alpha E_{t+1}^2\right]\tau_x.
\end{aligned}$$

It is useful to write $\mathcal{E}_{it}$ in terms of mean-zero variable $\mathcal{E}_{it} - \mu^{\xi}_{\mathcal{E}_t}$, that is $-\frac{\mathcal{E}_{it}^2}{2\mathcal{V}_t} = -\frac{\left(\mathcal{E}_{it}-\mu^{\xi}_{\mathcal{E}_t}\right)^2}{2\mathcal{V}_t} - \frac{\left(\mathcal{E}_{it}-\mu^{\xi}_{\mathcal{E}_t}\right)\mu^{\xi}_{\mathcal{E}_t}}{\mathcal{V}_t} - \frac{\left(\mu^{\xi}_{\mathcal{E}_t}\right)^2}{2\mathcal{V}_t}$. Therefore,

$$
\begin{aligned}
&\mathbb{E}\left[e^{-\frac{\mathcal{E}_{it}^2}{2\mathcal{V}_t}+\frac{\mathcal{C}_t\mathcal{E}_{it}}{2\mathcal{V}_t\hat{\mathcal{V}}_{it+1}}} \mid \mathcal{I}^+_{t-1}\right] \\
&= \mathbb{E}\left[\exp\left\{-\frac{\left(\mathcal{E}_{it}-\mu^{\xi}_{\mathcal{E}_t}\right)^2}{2\mathcal{V}_t} + \left[\frac{\mathcal{C}_t}{2\mathcal{V}_t\hat{\mathcal{V}}_{it+1}} - \frac{\mu^{\xi}_{\mathcal{E}_t}}{\mathcal{V}_t}\right]\left(\mathcal{E}_{it}-\mu^{\xi}_{\mathcal{E}_t}\right) - \frac{\left(\mu^{\xi}_{\mathcal{E}_t}\right)^2}{2\mathcal{V}_t} + \frac{\mathcal{C}_t\mu^{\xi}_{\mathcal{E}_t}}{2\mathcal{V}_t\hat{\mathcal{V}}_{it+1}}\right\} \mid \mathcal{I}^+_{t-1}\right] \\
&= \exp\left\{\frac{\mathcal{C}_t^2}{8\hat{\mathcal{V}}^2_{it+1}}\left(\frac{1}{\mathcal{V}_t} - \frac{1}{\Sigma^{\xi}_{t-1^+}}\right)\right\} \times \exp\left\{-\frac{\left(\mu^{\xi}_{\mathcal{E}_t}\right)^2}{2\Sigma^{\xi}_{t-1^+}} - \frac{\mathcal{C}_t\mu^{\xi}_{\mathcal{E}_t}}{2\Sigma^{\xi}_{t-1^+}\hat{\mathcal{V}}_{it+1}}\right\}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}^+_{t-1}, \xi^x_{it}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}_{it}\right] \mid \mathcal{I}^+_{t-1}, \xi^x_{it}\right] \\
&= -e^{-\rho r^2 W_{it-1^+} + \frac{5\mathcal{C}_t^2}{8\mathcal{V}_t\hat{\mathcal{V}}^2_{it+1}} - \frac{\mathcal{C}_t^2}{8\Sigma^{\xi}_{t-1^+}\hat{\mathcal{V}}^2_{it+1}}}\left|\frac{\hat{\Sigma}_{t^+}}{\hat{\mathcal{V}}_{it+1}}\right|^{-1/2} \exp\left\{-\frac{\mu^2_{\hat{\mathcal{E}}_{t+1}}}{2\hat{\Sigma}_{t^+}} - \frac{\left(\mu^{\xi}_{\mathcal{E}_t}\right)^2}{2\Sigma^{\xi}_{t-1^+}} - \frac{\mathcal{C}_t\mu^{\xi}_{\mathcal{E}_t}}{2\Sigma^{\xi}_{t-1^+}\hat{\mathcal{V}}_{it+1}}\right\}.
\end{aligned}
$$

Step 4: Calculate $\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}^+_{t-1}\right]$.

Conditional on $\mathcal{I}^+_{t-1}$, the only randomness in $-\frac{\mu^2_{\hat{\mathcal{E}}_{t+1}}}{2\hat{\Sigma}_{t^+}} - \frac{\left(\mu^{\xi}_{\mathcal{E}_t}\right)^2}{2\Sigma^{\xi}_{t-1^+}} - \frac{\mathcal{C}_t\mu^{\xi}_{\mathcal{E}_t}}{2\Sigma^{\xi}_{t-1^+}\hat{\mathcal{V}}_{it+1}}$ comes from $\mu_{\hat{\mathcal{E}}_{t+11}}$ and $\mu^{\xi}_{\mathcal{E}_t}$, both of which are functions of ξ^x_{it}. Let $M = \frac{r\mu}{r-\alpha g} + \alpha A_{t+1} - rA_t$. Then

$$
-\frac{\mu^2_{\hat{\mathcal{E}}_{t+1}}}{2\hat{\Sigma}_{t^+}} - \frac{\left(\mu^{\xi}_{\mathcal{E}_t}\right)^2}{2\Sigma^{\xi}_{t-1^+}} - \frac{\mathcal{C}_t\mu^{\xi}_{\mathcal{E}_t}}{2\Sigma^{\xi}_{t-1^+}\hat{\mathcal{V}}_{it+1}} = F^2\xi^x_{it} + G\xi_{it} + H,
$$

where

$$F = -\frac{1}{2}\left(V_{xt}^{\xi}\tau_{\nu xt}\right)^2\left[\hat{\Omega}_{t^+}r^2D_{t+1}^2 + \Omega_{t-1^+}^{\xi}\left(\alpha D_{t+1} - rE_t\right)^2\right], \tag{B.26}$$

$$G = V_{xt}^{\xi}\tau_{\nu xt}\left[\hat{\Omega}_{t^+}MrD_{t+1} - \Omega_{t-1^+}^{\xi}(\alpha D_{t+1} - rE_t)\left(M - \frac{\mathcal{C}_t}{2\hat{\mathcal{V}}_{it+1}}\right)\right], \tag{B.27}$$

$$H = -\frac{M^2}{2\hat{\Sigma}_{t^+}} - \frac{M^2}{2\Sigma_{t-1^+}^{\xi}} - \frac{\mathcal{C}_t M}{2\Sigma_{t-1}^{\xi}\hat{\mathcal{V}}_{it+1}}. \tag{B.28}$$

The *ex ante* utility for $\mathcal{G}_t$ investors is

$$\begin{aligned}
&\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}_{t-1}^{+}\right] = \mathbb{E}\left[\mathbb{E}\left[-e^{-\rho W_{it+1^+}} \mid \mathcal{I}_{t-1}^{+}, \xi_{it}^{x}\right] \mid \mathcal{I}_{t-1}^{+}\right] \\
&= -e^{-\rho r^2 W_{it-1^+} + \frac{5\mathcal{C}_t^2}{8\mathcal{V}_t\hat{\mathcal{V}}_{it+1}^2} - \frac{\mathcal{C}_t^2}{8\Sigma_{t-1^+}^{\xi}\hat{\mathcal{V}}_{it+1}^2}}\left|\frac{\hat{\Sigma}_{t^+}}{\hat{\mathcal{V}}_{it+1}}\right|^{-1/2}\left|1 - 2\left(\tau_x^{-1} + \tau_{\nu xt}^{-1}\right)F\right|^{-1/2} e^{-\frac{1}{2}\frac{G^2}{\left(\tau_x^{-1}+\tau_{\nu xt}^{-1}\right)^{-1} - 2F} + H}.
\end{aligned}$$

□

One can do a strictly increasing transformation of the *ex ante* utility:

$$V \overset{\text{def}}{=} -2\ln\left[\frac{-\mathbb{E}\left[-e^{-\rho W_{it+1^+}|\mathcal{I}_{t-1}^{+}}\right]}{e^{-\rho r^2 W_0}}\right] = \ln H_1 + \ln H_2 + H_3 + H_4, \tag{B.29}$$

where

$$H_1 = \frac{\hat{\Sigma}_{t^+}}{\hat{\mathcal{V}}_{it+1}}, \tag{B.30}$$

$$H_2 = 1 + V_{xt}^{\xi}\frac{\tau_{\nu xt}}{\tau_x}\left[\hat{\Omega}_{t^+}r^2D_{t+1}^2 + \Omega_{t-1^+}^{\xi}\left(\alpha D_{t+1} - rE_t\right)^2\right], \tag{B.31}$$

$$H_3 = -\frac{5\mathcal{C}_t^2}{4\mathcal{V}_t\hat{\mathcal{V}}_{it+1}^2} + \frac{\mathcal{C}_t^2}{4\Sigma_{t-1^+}^{\xi}\hat{\mathcal{V}}_{it+1}^2}, \tag{B.32}$$

$$H_4 = \frac{G^2}{\left(\tau_x^{-1} + \tau_{\nu xt}^{-1}\right)^{-1} - 2F} - H. \tag{B.33}$$

B.1.2 Market Clearing

In period t, $\mathcal{G}_t$-investors make the same information choices, while $\mathcal{G}_{t-1}$-investors choose precision heterogeneously, dependent on what they already learned in period $t-1$ about the period-t demand shock. This effectively makes the aggregation of signals and expected payoffs for the two generations different.

When aggregated, the noise in signals cancels out. The price signals of $\mathcal{G}_t$-investors aggregate to

$$\begin{aligned}\int_{i\in\mathcal{G}_t}\tilde{\eta}_{it}^{p}\,\mathrm{d}i &= \int_{i\in\mathcal{G}_t}\frac{1}{C_t}\left(p_t - A_t - B_t d_{t-1} - D_t\mathbb{E}\left[\tilde{x}_t \mid \tilde{\eta}_{it}^{x}\right] - E_t\mathbb{E}\left[\tilde{x}_{t+1} \mid \tilde{\xi}_{it}^{x}\right]\right)\mathrm{d}i \\ &= \frac{1}{2}\left[\frac{1}{C_t}(p_t - A_t - B_t d_{t-1}) - \frac{1}{\zeta_{1t}}V_{xt}\tau_{\epsilon xt}\tilde{x}_t - \frac{1}{\zeta_{2t}}V_{xt}^{\xi}\tau_{\nu xt}\tilde{x}_{t+1}\right], \end{aligned} \tag{B.34}$$

Then the expected period-t excess payoffs for $\mathcal{G}_t$-investors aggregate to

$$\begin{aligned}&\int_{i\in\mathcal{G}_t}\mathcal{E}_{it}\,\mathrm{d}i \\ &= \frac{1}{2}\Bigg\{(1+\alpha B_{t+1})(\mu + g d_{t-1}) + \alpha A_{t+1} + (1+\alpha B_{t+1})V_{yt}\tau_{\epsilon yt}\tilde{y}_t + \alpha D_{t+1}\left(1 - V_{xt}^{\xi}\tau_x\right)\tilde{x}_{t+1} - p_t r \\ &\quad + (1+\alpha B_{t+1})V_{yt}\tau_{pt}\left[\frac{1}{C_t}(p_t - A_t - B_t d_{t-1}) - \frac{1}{\zeta_{1t}}V_{xt}\tau_{\epsilon xt}\tilde{x}_t - \frac{1}{\zeta_{2t}}V_{xt}^{\xi}\tau_{\nu xt}\tilde{x}_{t+1}\right]\Bigg\}. \end{aligned} \tag{B.35}$$

For $\mathcal{G}_{t-1}$-investors, heterogeneity in conditional variances makes things more complicated:

$$\begin{aligned}&\int_{i\in\mathcal{G}_{t-1}}\frac{\hat{V}_{yit}\hat{\tau}_{pit}}{\hat{\mathcal{V}}_{it}}\tilde{\hat{\eta}}_{it}^{p}\,\mathrm{d}i \\ &= \int_{i\in\mathcal{G}_{t-1}}\frac{\hat{V}_{yit}\hat{\tau}_{pit}}{\hat{\mathcal{V}}_{it}}\frac{1}{C_t}\left(p_t - A_t - B_t d_{t-1} - D_t\mathbb{E}\left[\tilde{x}_t \mid \tilde{\hat{\eta}}_{it}^{x}, \tilde{\xi}_{it-1}^{x}\right] - E_t\mathbb{E}\left[\tilde{x}_{t+1} \mid \tilde{\hat{\xi}}_{it}^{x}\right]\right)\mathrm{d}i \\ &= \frac{1}{2}\left[\hat{\omega}_{pt}\frac{1}{C_t}\left(p_t - A_t - B_t d_{t-1}\right) - \frac{1}{\zeta_{1t}}\hat{\mathcal{Z}}_{1t}\tilde{x}_t - \frac{1}{\zeta_{2t}}\hat{\mathcal{Z}}_{2t}\tilde{x}_{t+1}\right], \end{aligned} \tag{B.36}$$

where $\hat{\mathcal{Z}}_{1t} \equiv 2\int_{i\in\mathcal{G}_{t-1}}\frac{\hat{V}_{yit}\hat{\tau}_{pit}}{\hat{\mathcal{V}}_{it}}\left(1 - \hat{V}_{xit}\tau_x\right)\mathrm{d}i$, $\hat{\mathcal{Z}}_{2t} \equiv 2\int_{i\in\mathcal{G}_{t-1}}\frac{\hat{V}_{yit}\hat{\tau}_{pit}}{\hat{\mathcal{V}}_{it}}\left(1 - \hat{V}_{xit}^{\xi}\tau_x\right)\mathrm{d}i$ and $\hat{\omega}_{pt} \equiv 2\int_{i\in\mathcal{G}_{t-1}}\frac{\hat{V}_{yit}\hat{\tau}_{pit}}{\hat{\mathcal{V}}_{it}}\,\mathrm{d}i$ are known by the market since the distributions of $\hat{\tau}_{\epsilon xit}$ and $\hat{\tau}_{\nu xit}$ are known.

Define the average posterior precision of period-t excess payoff to be $\hat{\Omega}_t \equiv 2\int_{i\in\mathcal{G}_{t-1}} \frac{1}{\hat{\mathcal{V}}_i}\,\mathrm{d}i$, the average ratio of the posterior variance of current demand shock to the posterior variance of period-t excess payoff to be $\hat{\chi}^{\xi}_{xt} \equiv 2\int_{i\in\mathcal{G}_{t-1}} \frac{\hat{V}^{\xi}_{xit}}{\hat{\mathcal{V}}_{it}}\,\mathrm{d}i$, and $\hat{\omega}_{yt} \equiv 2\int_{i\in\mathcal{G}_{t-1}} \frac{\hat{V}_{yit}\hat{\tau}_{\epsilon yit}}{\hat{\mathcal{V}}_{it}}\,\mathrm{d}i$. Then the risk premium of $\mathcal{G}_{t-1}$-investors aggregate to

$$
\begin{aligned}
&\int_{i\in\mathcal{G}_{t-1}} \frac{\hat{\mathcal{E}}_{it}}{\hat{\mathcal{V}}_{it}}\,\mathrm{d}i \\
&= \frac{1}{2}\Bigg\{ \hat{\Omega}_t\left[(1+\alpha B_{t+1})(\mu + g d_{t-1}) + \alpha A_{t+1} - p_t r\right] \\
&\quad +(1+\alpha B_{t+1})\hat{\omega}_{yt}\tilde{y}_t + \alpha D_{t+1}\left(\hat{\Omega}_t - \hat{\chi}_{xt}\tau_x\right)\tilde{x}_{t+1} \\
&\quad +(1+\alpha B_{t+1})\left[\hat{\omega}_{pt}\frac{1}{C_t}\Big[p_t - A_t - B_t d_{t-1}\Big] - \frac{1}{\zeta_{1t}}\hat{\mathcal{Z}}_{1t}\tilde{x}_t - \frac{1}{\zeta_{2t}}\hat{\mathcal{Z}}_{2t}\tilde{x}_{t+1}\right]\Bigg\}. \qquad \text{(B.37)}
\end{aligned}
$$

Recall that the market clearing condition is

$$
\bar{x}_t + \tilde{x}_t + \int_{i\in\mathcal{G}_{t-2}} \hat{q}_{it-1}\,\mathrm{d}i = \int_{i\in\mathcal{G}_t} q_{i1}\,\mathrm{d}i + \int_{i\in\mathcal{G}_{t-1}} (\hat{q}_{it} - q_{it-1})\,\mathrm{d}i\,.
$$

Plugging in portfolio choices, multiplying ρ on both sides and moving $\frac{\mathcal{C}_t}{r\mathcal{V}_t\hat{\mathcal{V}}_{it+1}}$ to LHS, we have

$$
\begin{aligned}
&\rho(\bar{x}_t + \tilde{x}_t) + \int_{i\in\mathcal{G}_t} \frac{\mathcal{C}_t}{r\mathcal{V}_t\hat{\mathcal{V}}_{it+1}}\,\mathrm{d}i + \int_{i\in\mathcal{G}_{t-1}} \frac{\mathcal{E}_{it-1}}{r\mathcal{V}_{t-1}} - \frac{\mathcal{C}_{t-1}}{r\mathcal{V}_{t-1}\hat{\mathcal{V}}_{it}}\,\mathrm{d}i + \int_{i\in\mathcal{G}_{t-2}} \frac{\hat{\mathcal{E}}_{it-1}}{\hat{\mathcal{V}}_{it-1}}\,\mathrm{d}i \\
&= \int_{i\in\mathcal{G}_t} \frac{\mathcal{E}_{it}}{r\mathcal{V}_t}\,\mathrm{d}i + \int_{i\in\mathcal{G}_{t-1}} \frac{\hat{\mathcal{E}}_{it}}{\hat{\mathcal{V}}_{it}}\,\mathrm{d}i\,.
\end{aligned}
$$

Let's work out RHS and LHS separately. RHS contains constants, p_t, $\tilde{y}_t$, $\tilde{x}_t$ and $\tilde{x}_{t+1}$:

$$
\begin{aligned}
&RHS \\
&= \frac{1}{2r}\Omega_t \Bigg\{ (1+\alpha B_{t+1})(\mu + g d_{t-1}) + \alpha A_{t+1} + (1+\alpha B_{t+1}) V_{yt}\tau_{\epsilon yt}\tilde{y}_t + \alpha D_{t+1}\left(1 - V_{xt}^{\xi}\tau_x\right)\tilde{x}_{t+1} \\
&\quad - p_t r + (1+\alpha B_{t+1}) V_{yt}\tau_{pt}\left[\frac{1}{C_t}(p_t - A_t - B_t d_{t-1}) - \frac{1}{\zeta_{1t}} V_{xt}\tau_{\epsilon xt}\tilde{x}_t - \frac{1}{\zeta_{2t}}\left(1 - V_{xt}^{\xi}\tau_x\right)\tilde{x}_{t+1}\right] \Bigg\} \\
&\quad + \frac{1}{2}\Bigg\{ \hat{\Omega}_t\left[(1+\alpha B_{t+1})(\mu + g d_{t-1}) + \alpha A_{t+1} - p_t r\right] + (1+\alpha B_{t+1})\hat{\omega}_{yt}\tilde{y}_t \\
&\quad + \alpha D_{t+1}\left(\hat{\Omega}_t - \hat{\chi}_{xt}\tau_x\right)\tilde{x}_{t+1} + (1+\alpha B_{t+1})\Big[\hat{\omega}_{pt}\frac{1}{C_t}\Big[p_t - A_t - B_t d_{t-1}\Big] \\
&\quad - \frac{1}{\zeta_{1t}}\hat{\mathcal{Z}}_{1t}\tilde{x}_t - \frac{1}{\zeta_{2t}}\hat{\mathcal{Z}}_{2t}\tilde{x}_{t+1}\Big]\Bigg\}.
\end{aligned}
$$

Collecting terms for $\tilde{y}_t$, $\tilde{x}_t$ and $\tilde{x}_{t+1}$, we have

$$
\begin{aligned}
&RHS \\
&= \frac{1}{2r}\Omega_t \Bigg\{ (1+\alpha B_{t+1})(\mu + g d_{t-1}) + \alpha A_{t+1} - p_t r + (1+\alpha B_{t+1}) V_{yt}\tau_{pt}\left[\frac{1}{C_t}(p_t - A_t - B_t d_{t-1})\right] \Bigg\} \\
&\quad + \frac{1}{2}\Bigg\{ \hat{\Omega}_t\left[(1+\alpha B_{t+1})(\mu + g d_{t-1}) + \alpha A_{t+1} - p_t r\right] + (1+\alpha B_{t+1})\left[\hat{\omega}_{pt}\frac{1}{C_t}\Big[p_t - A_t - B_t d_{t-1}\Big]\right] \Bigg\} \\
&\quad + (1+\alpha B_{t+1})\bar{\omega}_{yt}\tilde{y}_t - (1+\alpha B_{t+1})\frac{1}{\zeta_{1t}}\bar{\mathcal{Z}}_{1t}\tilde{x}_t \\
&\quad + \left[\alpha D_{t+1}\left(\bar{\Sigma}_t^{-1} - \bar{\chi}_{xt}\tau_x\right) - (1+\alpha B_{t+1})\frac{1}{\zeta_{2t}}\bar{\mathcal{Z}}_{2t}\right]\tilde{x}_{t+1},
\end{aligned}
\tag{B.38}
$$

where $\bar{\omega}_{yt} = \frac{1}{2r}\Omega_t V_{yt}\tau_{\epsilon yt} + \frac{1}{2}\hat{\omega}_{yt}$, $\bar{\chi}_{xt} = \frac{1}{2r}\Omega_t V_{xt}^{\xi} + \frac{1}{2}\hat{\chi}_{xt}$, $\bar{\mathcal{Z}}_{2t} = \frac{1}{2r}\Omega_t V_{yt}\tau_{pt}\left(1 - V_{xt}^{\xi}\tau_x\right) + \frac{1}{2}\hat{\mathcal{Z}}_{2t}$, and $\bar{\mathcal{Z}}_{1t} = \frac{1}{2r}\Omega_t V_{yt}\tau_{pt}V_{xt}\tau_{\epsilon xt} + \frac{1}{2}\hat{\mathcal{Z}}_{1t}$.

LHS only contains constants and $\tilde{x}_t$:

$$
\begin{aligned}
LHS = \rho(\bar{x}_t + \tilde{x}_t) &+ \frac{1}{2}\hat{\Omega}_{t+1}\frac{\mathcal{C}_t}{r\mathcal{V}_t} - \frac{1}{2}\hat{\Omega}_t\frac{\mathcal{C}_{t-1}}{r\mathcal{V}_{t-1}} + (1+\alpha B_t)\bar{\omega}_{yt-1}\tilde{y}_{t-1} \\
&+\frac{1}{2r}\Omega_{t-1}\Bigg\{(1+\alpha B_t)(\mu + gd_{t-2}) + \alpha A_t - p_{t-1}r \\
&+(1+\alpha B_t)V_{yt-1}\tau_{pt-1}\left[\frac{1}{C_{t-1}}(p_{t-1} - A_{t-1} - B_{t-1}d_{t-2})\right]\Bigg\} \\
&+\frac{1}{2}\Bigg\{\hat{\Omega}_{t-1}\left[(1+\alpha B_t)(\mu + gd_{t-2}) + \alpha A_t - p_{t-1}r\right] \\
&+(1+\alpha B_t)\left[\hat{\omega}_{pt-1}\frac{1}{C_{t-1}}\Big[p_{t-1} - A_{t-1} - B_{t-1}d_{t-2}\Big]\right]\Bigg\} \\
&-(1+\alpha B_t)\frac{1}{\zeta_{1t-1}}\bar{\mathcal{Z}}_{1t-1}\tilde{x}_{t-1} + \left[\alpha D_t\left(\bar{\Sigma}_{t-1}^{-1} - \bar{\chi}_{xt-1}\tau_x\right) - (1+\alpha B_t)\frac{1}{\zeta_{2t-1}}\bar{\mathcal{Z}}_{2t-1}\right]\tilde{x}_t.
\end{aligned}
$$

Notice that at the portfolio choice stage of period t, $\tilde{y}_{t-1}$ and $\tilde{x}_{t-1}$ are constants. For notational simplicity, let $\bar{A}_t = \frac{1}{2}\hat{\Omega}_{t+1}\frac{\mathcal{C}_t}{r\mathcal{V}_t} - \frac{1}{2}\hat{\Omega}_t\frac{\mathcal{C}_{t-1}}{r\mathcal{V}_{t-1}} + \frac{1}{2r}\Omega_{t-1}\Big\{(1+\alpha B_t)(\mu + gd_{t-2}) + \alpha A_t - p_{t-1}r + (1+\alpha B_t)V_{yt-1}\tau_{pt-1}\left[\frac{1}{C_{t-1}}(p_{t-1} - A_{t-1} - B_{t-1}d_{t-2})\right]\Big\} + \frac{1}{2}\Big\{\hat{\Omega}_{t-1}\left[(1+\alpha B_t)(\mu + gd_{t-2}) + \alpha A_t - p_{t-1}r\right] + (1+\alpha B_t)\left[\hat{\omega}_{pt-1}\frac{1}{C_{t-1}}\Big[p_{t-1} - A_{t-1} - B_{t-1}d_{t-2}\Big]\right]\Big\} + (1+\alpha B_t)\bar{\omega}_{yt-1}\tilde{y}_{t-1} - (1+\alpha B_t)\frac{1}{\zeta_{1t-1}}\bar{\mathcal{Z}}_{1t-1}\tilde{x}_{t-1}$.
Then

$$
LHS = \bar{A}_t + \rho\bar{x}_t + \rho\tilde{x}_t + \alpha D_t\left[\bar{\Sigma}_{t-1}^{-1} - \bar{\chi}_{xt-1}\tau_x\right]\tilde{x}_t - (1+\alpha B_t)\frac{1}{\zeta_{2t-1}}\bar{\mathcal{Z}}_{2t-1}\tilde{x}_t, \quad \text{(B.39)}
$$

Then we try to match coefficients of (B.38) and (B.39) with the following linear equilibrium price relation:

$$
p_t = A_t + B_t d_{t-1} + C_t\tilde{y}_t + D_t\tilde{x}_t + E_t\tilde{x}_{t+1}.
$$

Define the average market-wide posterior precision as $\bar{\Sigma}_t^{-1} \equiv \frac{1}{2}\left(\frac{1}{r}\Omega_t + \hat{\Omega}_t\right)$. Move (B.39) to RHS

and $\bar{\Sigma}_t^{-1} p_t r$ to LHS, and then multiply by $\frac{\bar{\Sigma}_t}{r}$ on both sides. This yields

$$
\begin{aligned}
& p_t r \\
= & \frac{1}{r}\bar{\Sigma}_t \Bigg\{ \frac{1}{2r}\Omega_t \left[(1+\alpha B_{t+1})(\mu + g d_{t-1}) + \alpha A_{t+1} + (1+\alpha B_{t+1}) V_{yt}\tau_{pt} \left[\frac{1}{C_t}(p_t - A_t - B_t d_{t-1}) \right] \right] \\
& + \frac{1}{2} \left[\hat{\Omega}_t \left[(1+\alpha B_{t+1})(\mu + g d_{t-1}) + \alpha A_{t+1} \right] + (1+\alpha B_{t+1}) \left[\hat{\omega}_{pt} \frac{1}{C_t} \left[p_t - A_t - B_t d_{t-1} \right] \right] \right] \\
& + (1+\alpha B_{t+1}) \bar{\omega}_{yt} \tilde{y}_t - (1+\alpha B_{t+1}) \frac{1}{\zeta_{1t}} \bar{\mathcal{Z}}_{1t} \tilde{x}_t + \left[\alpha D_{t+1} \left(\bar{\Sigma}_t^{-1} - \bar{\chi}_{xt}\tau_x \right) - (1+\alpha B_{t+1}) \frac{1}{\zeta_{2t}} \bar{\mathcal{Z}}_{2t} \right] \tilde{x}_{t+1} \\
& - \bar{A} - \rho \tilde{x}_t - \alpha D_t \left[\bar{\Sigma}_{t-1}^{-1} - \bar{\chi}_{xt-1}\tau_x \right] \tilde{x}_t + (1+\alpha B_t) \frac{1}{\zeta_{2t-1}} \bar{\mathcal{Z}}_{2t-1} \tilde{x}_t \Bigg\}.
\end{aligned}
$$

It is easy to see that at the coefficients of p_t and $B_t d_{t-1}$ on RHS only differ in sign. Define $\bar{\omega}_{pt} = \frac{1}{2r}\Omega_t V_{yt}\tau_{pt} + \frac{1}{2}\hat{\omega}_{pt}$. Again, move the p_t term to LHS and scale down so that $p_t r$ on LHS has coefficient 1:

$$
\begin{aligned}
p_t r = & \left(1 - \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1}) \bar{\omega}_{pt} \frac{1}{C_t} \right)^{-1} \times \frac{1}{r}\bar{\Sigma}_t \\
& \Bigg\{ \bar{\Sigma}_t^{-1} \left[(1+\alpha B_{t+1})(\mu + g d_{t-1}) + \alpha A_{t+1} \right] + (1+\alpha B_{t+1}) \bar{\omega}_{pt} \frac{1}{C_t} (-A_t - B_t d_{t-1}) \\
& + (1+\alpha B_{t+1}) \bar{\omega}_{yt} \tilde{y}_t - (1+\alpha B_{t+1}) \frac{1}{\zeta_{1t}} \bar{\mathcal{Z}}_{1t} \tilde{x}_t \\
& + \left[\alpha D_{t+1} \left(\bar{\Sigma}_t^{-1} - \bar{\chi}_{xt}^{\xi} \tau_x \right) - (1+\alpha B_{t+1}) \frac{1}{\zeta_{2t}} \bar{\mathcal{Z}}_{2t} \right] \tilde{x}_{t+1} \\
& - \bar{A} - \rho \tilde{x}_t - \alpha D_t \left[\bar{\Sigma}_{t-1}^{-1} - \bar{\chi}_{xt-1}^{\xi} \tau_x \right] \tilde{x}_t + (1+\alpha B_t) \frac{1}{\zeta_{2t-1}} \bar{\mathcal{Z}}_{2t-1} \tilde{x}_t \Bigg\}. \qquad \text{(B.40)}
\end{aligned}
$$

B_t:

From (B.40), it is easy to identify B_t, the coefficient before d_{t-1}:

$$
B_t = \left(1 - \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1}) \bar{\omega}_{pt} \frac{1}{C_t} \right)^{-1} \frac{1}{r}\bar{\Sigma}_t \left[\bar{\Sigma}_t^{-1} (1+\alpha B_{t+1}) g - (1+\alpha B_{t+1}) \bar{\omega}_{pt} \frac{1}{C_t} B_t \right].
$$

Moving all B_t terms to the left hand side, we have

$$B_t = \left[1 + \left(1 - \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)^{-1}\frac{1}{r}\bar{\Sigma}_t(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right]^{-1} \times \left(1 - \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)^{-1}\frac{1}{r}g(1+\alpha B_{t+1}).$$

Since $\left[1 + \left(1 - \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)^{-1}\frac{1}{r}\bar{\Sigma}_t(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right]^{-1}\left(1 - \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)^{-1} = 1$, it simplifies to

$$B_t = \frac{1}{r}g(1+\alpha B_{t+1}). \tag{B.41}$$

This is a linear difference equation and has a stationary solution[2]

$$B = \frac{g}{r - \alpha g}.$$

A_t:

From (B.40), the constant term in the equilibrium price relation is:

$$A_t = \left(1 - \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)^{-1} \times \frac{1}{r}\bar{\Sigma}_t\left[\bar{\Sigma}_t^{-1}(1+\alpha B_{t+1})\mu + \alpha A_{t+1} - (1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}A_t - \bar{A}_t\right].$$

Moving all A_t terms to the left hand side, we have

$$A_t = \frac{1}{r}\left[(1+\alpha B_{t+1})\mu - \rho\bar{x}_t - \bar{A}_t(A_{t+1}, A_{t+2}) + \alpha A_{t+1}\bar{\Sigma}_t\right]. \tag{B.42}$$

To emphasize the dependency of $\bar{A}_t$ on A_{t+1} and A_{t+2}, I write $\bar{A}_t(A_{t+1}, A_{t+2})$ instead of $\bar{A}_t$.

[2]The stationary solution is obtained by setting $B_t = B, \forall t$ in (B.41) and solving for B.

Plugging in the stationary solution of B_t, we have

$$A_t = \frac{\mu}{r-\alpha g} - \frac{\rho \bar{x}_t}{r} - \frac{1}{r}\bar{A}_t(A_{t+1}, A_{t+2}) + \frac{\bar{\Sigma}_t}{r}\alpha A_{t+1}\bar{\Sigma}_t.$$

$\underline{C_t}$:

The coefficient of $\tilde{y}_t$ in (B.40) is:

$$C_t = \left(1 - \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)^{-1}\frac{1}{r}\bar{\Sigma}_t(1+\alpha B_{t+1})\bar{\omega}_{yt}.$$

Multiplying $\left(1 - \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)$ on both sides and rearranging, we have

$$C_t = \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\left(\bar{\omega}_{yt} + \bar{\omega}_{pt}\right).$$

Since $V_{yt}(\tau_{\epsilon yt}+\tau_{pt}) = 1 - V_{yt}\tau_y$ and $\hat{\omega}_{yt}+\hat{\omega}_{pt} = 2\int_{i\in\mathcal{G}_{t-1}}\frac{\hat{V}_{yit}(\hat{\tau}_{\epsilon yit}+\hat{\tau}_{pit})}{\hat{\mathcal{V}}_{it}}\,\mathrm{d}i = \hat{\Omega}_t - 2\tau_y\int_{i\in\mathcal{G}_{t-1}}\frac{\hat{V}_{yit}}{\hat{\mathcal{V}}_{it}}\,\mathrm{d}i$, we have

$$C_t = \frac{1}{r-\alpha g}\left(1 - \bar{\Sigma}_t\bar{\chi}_{yt}\tau_y\right),$$

where

$$\bar{\chi}_{yt} \equiv \frac{1}{r}\Omega_t\frac{V_{yt}}{2} + \int_{i\in\mathcal{G}_{t-1}}\frac{\hat{V}_{yit}}{\hat{\mathcal{V}}_{it}}\,\mathrm{d}i\,.$$

$\underline{D_t}$:

The coefficient of the $\tilde{x}_t$ term in the price relation is:

$$D_t = \left(1 - \frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)^{-1}\frac{1}{r}\bar{\Sigma}_t\times$$
$$\left[-\rho - \alpha D_t\left(\bar{\Sigma}_{t-1}^{-1} - \bar{\chi}_{xt-1}^{\xi}\tau_x\right) - (1+\alpha B_{t+1})\frac{1}{\zeta_{1t}}\bar{\mathcal{Z}}_{1t} + (1+\alpha B_t)\frac{1}{\zeta_{2t-1}}\bar{\mathcal{Z}}_{2t-1}\right].$$

Multiplying $\left(1-\frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)$ on both sides, we have

$$D_t=\frac{\bar{\Sigma}_t}{r-\alpha g}\frac{1}{\zeta_{1t}}\left(\bar{\omega}_{pt}-\bar{\mathcal{Z}}_{1t}\right)-\frac{\rho\bar{\Sigma}_t}{r}+\frac{\bar{\Sigma}_t}{r-\alpha g}\frac{1}{\zeta_{2t-1}}\bar{\mathcal{Z}}_{2t-1}+\alpha D_t\frac{\bar{\Sigma}_t}{r}\left(\bar{\Sigma}_{t-1}^{-1}-\bar{\chi}_{xt-1}^{\xi}\tau_x\right).$$

We can then solve for D_t straightforwardly:

$$D_t=\left[1-\alpha\frac{\bar{\Sigma}_t}{r}\left(\bar{\Sigma}_{t-1}^{-1}-\bar{\chi}_{xt-1}^{\xi}\tau_x\right)\right]^{-1}\left[\frac{\bar{\Sigma}_t}{r-\alpha g}\frac{1}{\zeta_{1t}}\left(\bar{\omega}_{pt}-\bar{\mathcal{Z}}_{1t}\right)-\frac{\rho\bar{\Sigma}_t}{r}+\frac{\bar{\Sigma}_t}{r-\alpha g}\frac{1}{\zeta_{2t-1}}\bar{\mathcal{Z}}_{2t-1}\right].$$

For a better understanding of $\bar{\omega}_{pt}-\bar{\mathcal{Z}}_{1t}$, let $\chi_{yt}\equiv\Omega_t V_{yt}$, which characterizes the proportion of excess payoff uncertainty that can be explained by the uncertainty that comes from dividend innovation. When $\alpha=0$, the risky asset is short-lived and all the uncertainty comes from dividend innovation. Thus in this case $\chi_{yt}=1$. Notice that $\tau_{pt}=\frac{1}{\zeta_{1t}^{-2}V_{xt}+\zeta_{2t}^{-2}V_{xt}^{\xi}}=\frac{\zeta_{1t}^2}{V_{xt}}-\frac{\zeta_{1t}^4/V_{xt}^2}{\zeta_{1t}^2/V_{xt}+\zeta_{2t}^2/V_{xt}^{\xi}}$. Thus $\zeta_{1t}^{\dagger}\equiv\frac{\tau_{pt}V_{xt}}{\zeta_{1t}}=\zeta_{1t}-\frac{\zeta_{1t}^3/V_{xt}}{\zeta_{1t}^2/V_{xt}+\zeta_{2t}^2/V_{xt}^{\xi}}$ is the future-uncertainty-adjusted signal-to-current-noise ratio. Then define the market-wide average future-uncertainty-adjusted dividend-to-total-uncertainty-weighted signal-to-current-noise ratio to be

$$\bar{\zeta}_{1t}^{\dagger}\equiv\frac{1}{\zeta_{1t}}\left[\frac{1}{2r}\Omega_t V_{yt}\tau_{pt}V_{xt}+\int_{i\in\mathcal{G}_{t-1}}\frac{\hat{V}_{yit}\hat{\tau}_{pit}}{\hat{\mathcal{V}}_{it}}\hat{V}_{xit}\,\mathrm{d}i\right]. \tag{B.43}$$

When $\alpha=0$, $\bar{\zeta}_{1t}^{\dagger}=\frac{r+1}{2r}\zeta_t$. Then $(\bar{\omega}_{pt}-\bar{\mathcal{Z}}_{1t})/\zeta_{1t}=\bar{\zeta}_{1t}^{\dagger}\tau_x$.

E_t:

The coefficient of the term $\tilde{x}_{t+1}$ in the price relation is:

$$E_t=\left(1-\frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)^{-1}\frac{1}{r}\bar{\Sigma}_t\left[\alpha D_{t+1}\left(\bar{\Sigma}_t^{-1}-\bar{\chi}_{xt}^{\xi}\tau_x\right)-(1+\alpha B_{t+1})\frac{1}{\zeta_{2t}}\bar{\mathcal{Z}}_{2t}\right].$$

Multiplying both sides by $\left(1-\frac{\bar{\Sigma}_t}{r}(1+\alpha B_{t+1})\bar{\omega}_{pt}\frac{1}{C_t}\right)$ and rearranging, we have

$$E_t=\frac{\bar{\Sigma}_t}{r-\alpha g}\frac{1}{\zeta_{2t}}\left(\bar{\omega}_{pt}-\bar{\mathcal{Z}}_{2t}\right)+\frac{1}{r}\alpha D_{t+1}\left(1-\bar{\Sigma}_t\bar{\chi}_{xt}^{\xi}\tau_x\right).$$

Notice that $\tau_{pt} = \frac{1}{\zeta_{1t}^{-2}V_{xt}+\zeta_{2t}^{-2}V_{xt}^{\xi}} = \frac{\zeta_{2t}^2}{V_{xt}^{\xi}} - \frac{\zeta_{2t}^4/\left(V_{xt}^{\xi}\right)^2}{\zeta_{1t}^2/V_{xt}+\zeta_{2t}^2/V_{xt}^{\xi}}$. Thus $\zeta_{2t}^{\dagger} \equiv \frac{\tau_{pt}V_{xt}^{\xi}}{\zeta_{2t}} = \zeta_{2t} - \frac{\zeta_{2t}^3/V_{xt}^{\xi}}{\zeta_{1t}^2/V_{xt}+\zeta_{2t}^2V_{xt}^{\xi}}$ is the future-uncertainty-adjusted signal-to-current-noise ratio. Then define the market-wide average future-uncertainty-adjusted dividend-to-total-uncertainty-weighted signal-to-current-noise ratio as

$$\bar{\zeta}_{2t}^{\dagger} \equiv \frac{1}{\zeta_{2t}}\left[\frac{1}{2r}\Omega_t V_{yt}\tau_{pt}V_{xt}^{\xi} + \int_{i\in\mathcal{G}_{t-1}} \frac{\hat{V}_{yit}\hat{\tau}_{pit}}{\hat{\mathcal{V}}_{it}}\hat{V}_{xit}^{\xi}\,\mathrm{d}i\right]. \tag{B.44}$$

We can then write E_t as

$$E_t = \frac{1}{r}\alpha D_{t+1}\left(1 - \bar{\Sigma}_t\bar{\chi}_{xt}^{\xi}\tau_x\right) + \frac{\bar{\Sigma}_t}{r-\alpha g}\bar{\zeta}_{2t}^{\dagger}\tau_x. \tag{B.45}$$

In the static case, we have $E_t = \bar{\Sigma}_t\frac{E_t}{C_t}\tau_x$, which implies that $E_t = 0$.

B.1.3 Existence of Linear Equilibrium

The price coefficients are

$$A_t = \frac{\mu}{r-\alpha g} - \frac{\rho\bar{x}_t}{r} - \frac{1}{r}\bar{A}_t(A_{t+1}, A_{t+2}) + \frac{\bar{\Sigma}_t}{r}\alpha A_{t+1}\bar{\Sigma}_t, \tag{B.46}$$

$$B = \frac{g}{r-\alpha g}, \tag{B.47}$$

$$C_t = \frac{1}{r-\alpha g}\left(1 - \bar{\Sigma}_t\bar{\chi}_{yt}\tau_y\right), \tag{B.48}$$

$$D_t = \left[1 - \alpha\frac{\bar{\Sigma}_t}{r}\left(\bar{\Sigma}_{t-1}^{-1} - \bar{\chi}_{xt-1}^{\xi}\tau_x\right)\right]^{-1}\left[\frac{\bar{\Sigma}_t}{r-\alpha g}\bar{\zeta}_{1t}^{\dagger}\tau_x - \frac{\rho\bar{\Sigma}_t}{r} + \frac{\bar{\Sigma}_t}{r-\alpha g}\frac{1}{\zeta_{2t-1}}\bar{\mathcal{Z}}_{2t-1}\right], \tag{B.49}$$

$$E_t = \frac{1}{r}\alpha D_{t+1}\left(1 - \bar{\Sigma}_t\bar{\chi}_{xt}^{\xi}\tau_x\right) + \frac{\bar{\Sigma}_t}{r-\alpha g}\bar{\zeta}_{2t}^{\dagger}\tau_x. \tag{B.50}$$

B.2 Proof of Results

Proof of Proposition 15. In the benchmark case,

$$\frac{r+1}{r}\tau_{\epsilon xt}\zeta_t^2 + \rho\zeta_t + \frac{r+1}{r}\tau_{\epsilon yt} = 0.$$

Plugging in the solution for $\tau_{\epsilon xt}^B$ and $\tau_{\epsilon yt}^B$, we have

$$\rho\zeta_t = -\frac{r+1}{r}\sqrt{\frac{K}{\gamma_{\epsilon y}}\left[1 + \frac{\gamma_{\epsilon y}}{\gamma_{\epsilon x}}\zeta_t^4\right]} < 0, \tag{B.51}$$

or

$$\left(\frac{r+1}{r}\right)^2 \frac{K}{\gamma_{\epsilon x}}\zeta_t^4 - \rho^2\zeta_t^2 + \left(\frac{r+1}{r}\right)^2 \frac{K}{\gamma_{\epsilon y}} = 0.$$

This is a quadratic equation of ζ_t^2. We pick the solution which is consistent with the limit. Thus

$$\zeta_t^B = -\sqrt{\frac{\rho^2 - \sqrt{\rho^4 - 4\frac{\gamma_{\epsilon y}}{\gamma_{\epsilon x}}\left(\frac{K}{\gamma_{\epsilon y}}\right)^2\left(\frac{r+1}{r}\right)^4}}{2\frac{\gamma_{\epsilon y}}{\gamma_{\epsilon x}}\frac{K}{\gamma_{\epsilon y}}\left(\frac{r+1}{r}\right)^2}}.$$

□

Proof of Proposition 16. Recall that $\left(\bar{\Sigma}_t^B\right)^{-1} = \frac{r+1}{r}\left[\tau_y + \left(\zeta_t^B\right)^2\tau_x\right] - \rho\zeta_t^B$. Taking the partial derivative with respect to $|\zeta_t^B|$, we have

$$\frac{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}{\partial|\zeta_t^B|} = 2\frac{r+1}{r}|\zeta_t^B|\tau_x + \rho > 0.$$

By chain rule, we have

$$\begin{aligned}
\frac{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}{\partial K} &= \frac{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}{\partial|\zeta_t|}\frac{\partial|\zeta_t^B|}{\partial K} > 0,\\
\frac{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}{\partial \gamma_{\epsilon x}} &= \frac{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}{\partial|\zeta_t|}\frac{\partial|\zeta_t^B|}{\partial \gamma_{\epsilon x}} < 0,\\
\frac{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}{\partial \tau_x} &= \frac{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}{\partial|\zeta_t|}\frac{\partial|\zeta_t^B|}{\partial \tau_x} + \frac{r+1}{r}\left(\zeta_t^B\right)^2 > 0.
\end{aligned}$$

$$\begin{aligned}
&\frac{\mathbb{V}ar\left(p_t\right)}{\partial K}\\
&= 2\left[\frac{C_t^B}{\tau_y}\frac{\partial C_t^B}{\partial \bar{\Sigma}_t^B}\frac{\partial \bar{\Sigma}_t^B}{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}\frac{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}{\partial|\zeta_t|} + \frac{D_t^B}{\tau_x}\left(\frac{\partial D_t^B}{\partial \bar{\Sigma}_t^B}\frac{\partial \bar{\Sigma}_t^B}{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}\frac{\partial\left(\bar{\Sigma}_t^B\right)^{-1}}{\partial|\zeta_t^B|} + \frac{\partial D_t^B}{\partial|\zeta_t^B|}\right)\right]\frac{\partial|\zeta_t^B|}{\partial K}
\end{aligned}$$

Same as in the main model

$$\frac{\partial \tau_{\epsilon y t}^B}{\partial K} = \frac{\partial \tau_{\epsilon y t}^B}{\partial K} + \frac{\partial \tau_{\epsilon y t}^B}{\partial \zeta_t^B}\frac{\partial \zeta_t^B}{\partial K}.$$

The first term

$$\frac{\partial \tau_{\epsilon y t}^B}{\partial K} = \frac{1}{2\gamma_{\epsilon y}\left(\frac{\gamma_{\epsilon y}\zeta^4}{\gamma_{\epsilon x}}+1\right)\sqrt{\frac{K}{\gamma_{\epsilon y}\left(\frac{\gamma_{\epsilon y}\zeta^4}{\gamma_{\epsilon x}}+1\right)}}} > 0.$$

$$\begin{aligned}
\frac{\mathrm{d}\tau_{\epsilon x t}^B}{\mathrm{d}K} &= \underbrace{\frac{\partial \tau_{\epsilon x t}^B}{\partial K}}_{>0} + \underbrace{\frac{\partial \tau_{\epsilon x t}^B}{\partial \zeta_t^B}}_{<0}\underbrace{\frac{\partial \zeta_t^B}{\partial K}}_{<0} > 0,\\
\frac{\mathrm{d}\tau_{\epsilon y t}^B}{\mathrm{d}K} &= \underbrace{\frac{\partial \tau_{\epsilon y t}^B}{\partial K}}_{>0} + \underbrace{\frac{\partial \tau_{\epsilon y t}^B}{\partial \zeta_t^B}}_{>0}\underbrace{\frac{\partial \zeta_t^B}{\partial K}}_{<0};
\end{aligned}$$

$$\frac{\mathrm{d}\tau^B_{\epsilon xt}}{\mathrm{d}\gamma_{\epsilon x}} = \underbrace{\frac{\partial \tau^B_{\epsilon xt}}{\partial \gamma_{\epsilon x}}}_{>0} + \underbrace{\frac{\partial \tau^B_{\epsilon xt}}{\partial \zeta^B_t}}_{<0}\underbrace{\frac{\partial \zeta^B_t}{\partial \gamma_{\epsilon x}}}_{>0},$$
$$\frac{\mathrm{d}\tau^B_{\epsilon yt}}{\mathrm{d}\gamma_{\epsilon x}} = \underbrace{\frac{\partial \tau^B_{\epsilon yt}}{\partial \gamma_{\epsilon x}}}_{<0} + \underbrace{\frac{\partial \tau^B_{\epsilon yt}}{\partial \zeta^B_t}}_{>0}\underbrace{\frac{\partial \zeta^B_t}{\partial \gamma_{\epsilon x}}}_{>0};$$

$$\frac{\partial \zeta^B_t}{\partial \gamma_{\epsilon x}} = \frac{\frac{\rho^2 - \sqrt{\rho^4 - \frac{4\left(\frac{r+1}{r}\right)^4 K^2}{\gamma_{\epsilon x}\gamma_{\epsilon y}}}}{\left(\frac{r+1}{r}\right)^2 K} - \frac{2\left(\frac{r+1}{r}\right)^2 K}{\gamma_{\epsilon x}\gamma_{\epsilon y}\sqrt{\rho^4 - \frac{4\left(\frac{r+1}{r}\right)^4 K^2}{\gamma_{\epsilon x}\gamma_{\epsilon y}}}}}{2\sqrt{2}\sqrt{\frac{\gamma_{\epsilon x}\left(\rho^2 - \sqrt{\rho^4 - \frac{4\left(\frac{r+1}{r}\right)^4 K^2}{\gamma_{\epsilon x}\gamma_{\epsilon y}}}\right)}{\left(\frac{r+1}{r}\right)^2 K}}} < 0$$

always holds when $r, K, \rho, \gamma_{\epsilon x}, \gamma_{\epsilon y}$ and the equilibrium existence condition hold. □

Proof of Lemma 7 (Comovement of Σ_t^{-1} and $|\zeta_t|$). The signal-to-noise ratio is $\zeta_t = \frac{\bar{\Sigma}_t^{-1} - \frac{r+1}{r}\tau_y}{\frac{r+1}{r}\zeta_t\tau_x - \rho}$, which leads to

$$\bar{\Sigma}_t^{-1} = \frac{r+1}{r}\tau_x\zeta_t^2 - \rho\zeta_t + \frac{r+1}{r}\tau_y. \tag{B.52}$$

Taking partial derivative with respect to ζ_t, we have

$$\frac{\partial \bar{\Sigma}_t^{-1}}{\partial |\zeta_t|} = 2\frac{r+1}{r}|\zeta_t|\tau_x + \rho > 0.$$

The information processing capacity K, learning costs $\gamma_{\epsilon x}$, $\gamma_{\epsilon y}$ and $\gamma_{\nu x}$, and the prior precision of dividend innovation τ_y affect $\bar{\Sigma}_t^{-1}$ solely through $|\zeta_t|$. Thus any change in $|\zeta_t|$ caused by K, $\gamma_{\epsilon x}$, $\gamma_{\epsilon y}$, $\gamma_{\nu x}$ or τ_y will be accompanied by a change in $\bar{\Sigma}_t^{-1}$ in the same direction. □

Proof of Proposition 19 (Signal-to-noise ratio and average posterior precision of dividend innovation). Recall that $\bar{\Sigma}_t^{-1} = \frac{1}{r}\Sigma_t + \hat{\Sigma}_t$, which leads to $\bar{\Sigma}_t^{-1} = \frac{r+1}{r}\tau_y + \frac{1}{r}\tau_{\epsilon yt} + \hat{\tau}_{\epsilon yt} + \zeta_t^2\left(\frac{r+1}{r}\tau_x + \frac{1}{r}\tau_{\epsilon xt} + \hat{\tau}_{\epsilon xt} + \hat{\tau}_{\nu xt-1}\right)$.

Plug it into (B.52), we have

$$\left(\frac{1}{r}\tau_{\epsilon xt}+\hat{\tau}_{\epsilon xt}+\hat{\tau}_{\nu xt-1}\right)\zeta_t^2+\rho\zeta_t+\frac{1}{r}\tau_{\epsilon yt}+\hat{\tau}_{\epsilon yt}=0. \tag{B.53}$$

This is a quadratic equation of ζ_t, which has two solutions. We pick the larger solution, since when there is no demand information, the solution converges to the unique solution in the models where there is only dividend innovation acquisition.

$$\zeta_t=\frac{-\rho+\sqrt{\rho^2-4\left(\frac{1}{r}\tau_{\epsilon xt}+\hat{\tau}_{\epsilon xt}+\hat{\tau}_{\nu xt-1}\right)\left(\frac{1}{r}\tau_{\epsilon yt}+\hat{\tau}_{\epsilon yt}\right)}}{2\left(\frac{1}{r}\tau_{\epsilon xt}+\hat{\tau}_{\epsilon xt}+\hat{\tau}_{\nu xt-1}\right)}<0. \tag{B.54}$$

□

B.2.1 Information Choices

Proof of Proposition 13. (a) is obvious to see since $\hat{\tau}_{\nu xit}$ does not enter the *ex ante* utility.

(b) The Lagrangian is

$$\mathcal{L}=\hat{L}_{\epsilon y}\hat{\tau}_{\epsilon yit}+\hat{L}_{\epsilon y}\zeta_t^2\hat{\tau}_{\epsilon xit}-\lambda\left(\left[\gamma_{\epsilon x}\hat{\tau}_{\epsilon xit}^2+\gamma_{\epsilon y}\hat{\tau}_{\epsilon yit}^2\right]-K\right).$$

The first-order conditions are

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\hat{\tau}_{\epsilon yit}}&=\hat{L}_{\epsilon y}-2\lambda\gamma_{\epsilon y}\hat{\tau}_{\epsilon yit}=0,\\ \frac{\partial\mathcal{L}}{\partial\hat{\tau}_{\epsilon xit}}&=\hat{L}_{\epsilon y}\zeta_t^2-2\lambda\gamma_{\epsilon x}\hat{\tau}_{\epsilon xit}=0,\end{aligned}$$

which yields

$$\hat{\tau}_{\epsilon xit}=\frac{\gamma_{\epsilon y}}{\gamma_{\epsilon x}}\zeta_t^2\hat{\tau}_{\epsilon yit}. \tag{B.55}$$

Plugging (B.55) into the budget constraint yields the optimal information choices. □

Proof of Lemma 6. To simply calculations, one can do a strictly increasing transformation of the *ex ante* utility

$$V \overset{\text{def}}{=} -2\ln\left[\frac{-\mathbb{E}\left[-e^{-\rho W_{it+1^+}|\mathcal{I}^+_{t-1}}\right]}{e^{-\rho r^2 W_0}}\right] = \ln H_1 + \ln H_2 + H_3 + H_4, \tag{B.56}$$

where

$$H_1 = \frac{\mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right) + \Sigma_Z + \Sigma_a}{\mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right)}, \tag{B.57}$$

$$H_2 = \frac{\mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right) + \Sigma_S}{\mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right)}, \tag{B.58}$$

$$H_3 = \frac{\bar{\bar{\Sigma}}^2_{t+1}\rho^2\bar{x}^2}{\mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right) + \Sigma_Z + \Sigma_a}, \tag{B.59}$$

$$H_4 = \frac{\bar{\Sigma}^2_t\rho^2\bar{x}^2}{\mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right) + \Sigma_S}. \tag{B.60}$$

Notice that $\tau_{\epsilon yt}$ and $\tau_{\epsilon xt}$ enter only through $\mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right)$ and Σ_S via H_2 and H_4. Thus we only need to check whether $\frac{\partial \mathbb{V}ar\left(\tilde{d}_t|\mathcal{I}_{it}\right)}{\partial \tau_{\epsilon yt}} = \frac{\partial \mathbb{V}ar\left(\tilde{d}_t|\mathcal{I}_{it}\right)}{\partial \tau_{\epsilon xt}}\zeta_t^2$ and $\frac{\partial \Sigma_S}{\partial \tau_{\epsilon yt}} = \frac{\partial \Sigma_S}{\partial \tau_{\epsilon xt}}\zeta_t^2$ hold. Recall that

$$\mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right) = \frac{1}{\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)}.$$

Then

$$\frac{\partial \mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right)}{\partial \tau_{\epsilon yt}} = -\frac{1}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^2}, \tag{B.61}$$

$$\frac{\partial \mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right)}{\partial \tau_{\epsilon xt}} = -\frac{\zeta_t^2}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^2}. \tag{B.62}$$

Thus $\frac{\partial \mathbb{V}ar\left(\tilde{d}_t|\mathcal{I}_{it}\right)}{\partial \tau_{\epsilon yt}} = \frac{\partial \mathbb{V}ar\left(\tilde{d}_t|\mathcal{I}_{it}\right)}{\partial \tau_{\epsilon xt}}\zeta_t^2$ holds. One has

$$\Sigma_S = \frac{b_2^2}{\tau_y} + \frac{b_3^2}{\tau_x} + \frac{\tau_{\epsilon yt} + \zeta_t^2\tau_{\epsilon xt}}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^2} \tag{B.63}$$

Take partial derivative with respect to $\tau_{\epsilon yt}$ on both sides, one has

$$\begin{aligned}\frac{\partial \Sigma_S}{\partial \tau_{\epsilon yt}} = &\frac{2b_2}{\tau_y}\frac{\tau_y}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^2} - \frac{2b_3}{\tau_x}\frac{\zeta_t\tau_x}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^2} \\ &+\frac{1}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^2} - 2\frac{\tau_{\epsilon yt} + \zeta_t^2\tau_{\epsilon xt}}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^3},\end{aligned} \tag{B.64}$$

$$\begin{aligned}\frac{\partial \Sigma_S}{\partial \tau_{\epsilon xt}} = &\frac{2b_2}{\tau_y}\frac{\zeta_t^2\tau_y}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^2} - \frac{2b_3}{\tau_x}\frac{\zeta_t^3\tau_x}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^2} \\ &+\frac{\zeta_t^2}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^2} - 2\zeta_t^2\frac{\tau_{\epsilon yt} + \zeta_t^2\tau_{\epsilon xt}}{\left[\tau_y + \tau_{\epsilon yt} + \zeta_t^2\left(\tau_x + \tau_{\epsilon xt}\right)\right]^3}.\end{aligned} \tag{B.65}$$

Therefore $\frac{\partial \Sigma_S}{\partial \tau_{\epsilon yt}} = \frac{\partial \Sigma_S}{\partial \tau_{\epsilon xt}}\zeta_t^2$ holds. □

Proof of Proposition 14.

Step 1: Calculate learning indices

On the one hand, from the proof of Lemma 6, we know that the learning index associated with the signal for current fundamental shock is

$$L_{\epsilon yit} = \frac{\partial \ln H_2}{\partial \tau_{\epsilon yt}} + \frac{\partial H_4}{\partial \tau_{\epsilon yt}},$$

with

$$\begin{aligned}\frac{\partial \ln H_2}{\partial \tau_{\epsilon yt}} &= \frac{1}{\left[\mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right) + \Sigma_S\right]\mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right)}\left[\frac{\partial \Sigma_S}{\partial \tau_{\epsilon yt}}\mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right) - \frac{\partial \mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right)}{\partial \tau_{\epsilon yt}}\Sigma_S\right], \\ \frac{\partial H_4}{\partial \tau_{\epsilon yt}} &= -\frac{\bar{\Sigma}_t^2\rho^2\bar{x}^2}{\left[\mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right) + \Sigma_S\right]^2}\left[\frac{\partial \Sigma_S}{\partial \tau_{\epsilon yt}} + \frac{\partial \mathbb{V}ar\left(\tilde{d}_t \mid \mathcal{I}_{it}\right)}{\partial \tau_{\epsilon yt}}\right],\end{aligned}$$

where $\frac{\partial \Sigma_S}{\partial \tau_{\epsilon yt}}$ and $\frac{\partial \mathbb{V}ar\left(\tilde{d}_t|\mathcal{I}_{it}\right)}{\partial \tau_{\epsilon yt}}$ are calculated by (B.64) and (B.61).

On the other hand, notice that $\tau_{\nu xt}$ enter the ex-ante utility function only through $\mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right)$, Σ_Z and Σ_a via H_1 and H_3. Thus

$$L_{\nu xit} = \frac{\partial \ln H_1}{\partial \tau_{\nu xt}} + \frac{\partial H_3}{\partial \tau_{\nu xt}},$$

with

$$\frac{\partial \ln H_1}{\partial \tau_{\nu xt}} = \frac{\frac{\partial(\Sigma_Z+\Sigma_a)}{\partial \tau_{\nu xt}}\mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right) - \frac{\partial \mathbb{V}ar\left(\tilde{d}_{t+1}|\hat{\mathcal{I}}_{it+1}\right)}{\partial \tau_{\nu xt}}\left(\Sigma_Z + \Sigma_a\right)}{\left[\mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right) + \Sigma_Z + \Sigma_a\right]\mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right)},$$

$$\frac{\partial H_3}{\partial \tau_{\nu xt}} = -\frac{\bar{\bar{\Sigma}}^2_{t+1}\rho^2\bar{x}^2}{\left[\mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right) + \Sigma_Z + \Sigma_a\right]^2}\left[\frac{\partial\left(\Sigma_Z + \Sigma_a\right)}{\partial \tau_{\nu xt}} + \frac{\partial \mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right)}{\partial \tau_{\nu xt}}\right].$$

Recall that

$$\mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right) = \frac{1}{\tau_y + \hat{\tau}_{\epsilon y it+1} + \zeta^2_{t+1}\left(\tau_x + \hat{\tau}_{\epsilon x it+1} + \tau_{\nu xt}\right)}.$$

Thus

$$\frac{\partial \mathbb{V}ar\left(\tilde{d}_{t+1} \mid \hat{\mathcal{I}}_{it+1}\right)}{\partial \tau_{\nu xt}} = -\frac{\zeta^2_{t+1}}{\left[\tau_y + \hat{\tau}_{\epsilon y it+1} + \zeta^2_{t+1}\left(\tau_x + \hat{\tau}_{\epsilon x it+1} + \tau_{\nu xt}\right)\right]^2}.$$

One has

$$\Sigma_Z + \Sigma_a = \frac{a_2^2}{\tau_y} + \frac{a_3^2 + c^2}{\tau_x + \tau_{\nu xt}} + \frac{\hat{\tau}_{\epsilon y it+1} + \zeta^2_{t+1}\hat{\tau}_{\epsilon x it+1}}{\left[\tau_y + \hat{\tau}_{\epsilon y it+1} + \zeta^2_{t+1}\left(\tau_x + \hat{\tau}_{\epsilon x it+1} + \tau_{\nu xt}\right)\right]^2}.$$

Taking partial derivative with respect to $\tau_{\nu xt}$ on both sides, we have

$$
\begin{aligned}
&\frac{\partial\left(\Sigma_Z+\Sigma_a\right)}{\partial\tau_{\nu xt}}\\
=&\frac{2a_2}{\tau_y}\frac{\zeta_{t+1}^2\tau_y}{\left[\tau_y+\hat{\tau}_{\epsilon yit+1}+\zeta_{t+1}^2\left(\tau_x+\hat{\tau}_{\epsilon xit+1}+\tau_{\nu xt}\right)\right]^2}-\frac{a_3^2+c^2}{\left(\tau_x+\tau_{\nu xt}\right)^2}\\
&+\frac{2a_3}{\tau_x+\tau_{\nu xt}}\left[\frac{\zeta_{t+1}}{\tau_y+\hat{\tau}_{\epsilon yit+1}+\zeta_{t+1}^2\left(\tau_x+\hat{\tau}_{\epsilon xit+1}+\tau_{\nu xt}\right)}-\frac{\zeta_{t+1}^3\left(\tau_x+\tau_{\nu xt}\right)}{\left[\tau_y+\hat{\tau}_{\epsilon yit+1}+\zeta_{t+1}^2\left(\tau_x+\hat{\tau}_{\epsilon xit+1}+\tau_{\nu xt}\right)\right]^2}\right]\\
&+\frac{2c}{\tau_x+\tau_{\nu xt}}\left[\frac{\zeta_{t+1}}{\tau_y+\hat{\tau}_{\epsilon yit+1}+\zeta_{t+1}^2\left(\tau_x+\hat{\tau}_{\epsilon xit+1}+\tau_{\nu xt}\right)}-\frac{\zeta_{t+1}^3\tau_x}{\left[\tau_y+\hat{\tau}_{\epsilon yit+1}+\zeta_{t+1}^2\left(\tau_x+\hat{\tau}_{\epsilon xit+1}+\tau_{\nu xt}\right)\right]^2}\right]\\
&-2\zeta_{t+1}^2\frac{\hat{\tau}_{\epsilon yit+1}+\zeta_{t+1}^2\hat{\tau}_{\epsilon xit+1}}{\left[\tau_y+\hat{\tau}_{\epsilon yit+1}+\zeta_{t+1}^2\left(\tau_x+\hat{\tau}_{\epsilon xit+1}+\tau_{\nu xt}\right)\right]^3}.
\end{aligned}
$$

Step 2: Characterizing information choices with F.O.C.'s and information capacity constraint

The first-order conditions to the information choice problem are

$$
\begin{aligned}
\frac{\partial\mathcal{L}}{\partial\tau_{\epsilon yt}}&=L_{\epsilon yit}-2\lambda\gamma_{\epsilon y}\tau_{\epsilon yt}=0,\\
\frac{\partial\mathcal{L}}{\partial\tau_{\epsilon xt}}&=L_{\epsilon yit}\zeta_t^2-2\lambda\gamma_{\epsilon x}\tau_{\epsilon xt}=0,\\
\frac{\partial\mathcal{L}}{\partial\tau_{\nu xt}}&=L_{\nu xit}-2\lambda\gamma_{\nu x}\tau_{\nu xt}=0.
\end{aligned}
$$

Dividing them against each other, we have

$$
\begin{aligned}
\tau_{\epsilon yt}&=\frac{\gamma_{\nu x}}{\gamma_{\epsilon y}}\frac{L_{\epsilon yit}}{L_{\nu xit}}\tau_{\nu xt},\\
\tau_{\epsilon xt}&=\frac{\gamma_{\nu x}}{\gamma_{\epsilon x}}\frac{L_{\epsilon yit}}{L_{\nu xit}}\zeta_t^2\tau_{\nu xt}.
\end{aligned}
$$

Plug them back into the budget constraint (2.4), we have

$$
\frac{\tau_{\nu xt}^2}{L_{\nu xit}^2}\left[\frac{\gamma_{\nu x}}{\gamma_{\epsilon y}}L_{\epsilon yit}^2+\frac{\gamma_{\nu x}}{\gamma_{\epsilon x}}L_{\epsilon yit}^2\zeta_t^4+L_{\nu xit}^2\right]=\frac{K}{\gamma_{\nu x}}. \tag{B.66}
$$

Define

$$f_{it} \equiv \frac{L_{\epsilon yit}}{L_{\nu xit}} = \frac{\frac{\partial \ln H_2}{\partial \tau_{\epsilon yt}} + \frac{\partial H_4}{\partial \tau_{\epsilon yt}}}{\frac{\partial \ln H_1}{\partial \tau_{\nu xt}} + \frac{\partial H_3}{\partial \tau_{\nu xt}}}. \tag{B.67}$$

Then B.66 simplifies to

$$\tau_{\nu xt}^2 \left[\frac{\gamma_{\nu x}}{\gamma_{\epsilon y}} f_{it}^2 + \frac{\gamma_{\nu x}}{\gamma_{\epsilon x}} \zeta_t^4 f_{it}^2 + 1 \right] = \frac{K}{\gamma_{\nu x}}.$$

□

Proof of Theorem 3 (Price Volatility). The unconditional price volatility is $\mathbb{V}ar(p_t) = \frac{C_t^2}{\tau_y} + \frac{D_t^2}{\tau_x}$. Take partial derivative with respect to K, we have

$$\frac{\partial \mathbb{V}ar(p_t)}{\partial K} = 2 \left[\frac{C_t}{\tau_y} \frac{\partial C_t}{\partial \bar{\Sigma}_t} \frac{\partial \bar{\Sigma}_t}{\partial \zeta_t} + \frac{D_t}{\tau_x} \left(\frac{\partial D_t}{\partial \bar{\Sigma}_t} \frac{\partial \bar{\Sigma}_t}{\partial \zeta_t} + \frac{\partial D_t}{\partial \zeta_t} \right) \right] \frac{\partial \zeta_t}{\partial K}. \tag{B.68}$$

We then need to determine the sign of the term inside the square brackets. To do that, we first derive the partial derivatives of C_t and D_t with respect to $\bar{\Sigma}_t$ and ζ_t. We have $\frac{\partial C_t}{\partial \bar{\Sigma}_t} = -\frac{r+1}{r}\tau_y$, $\frac{\partial D_t}{\partial \bar{\Sigma}_t} = \frac{r+1}{r}\zeta_t \tau_x - \rho$, $\frac{\partial D_t}{\partial \zeta_t} = \frac{r+1}{r}\bar{\Sigma}_t \tau_x$ and $\frac{\partial \bar{\Sigma}_t}{\partial \zeta_t} = \bar{\Sigma}_t^2 \left(\rho - 2\frac{r+1}{r}\zeta_t \tau_x\right)$. Thus the term inside square brackets becomes

$$\begin{aligned}
&\frac{C_t}{\tau_y} \frac{\partial C_t}{\partial \bar{\Sigma}_t} \frac{\partial \bar{\Sigma}_t}{\partial \zeta_t} + \frac{D_t}{\tau_x} \left(\frac{\partial D_t}{\partial \bar{\Sigma}_t} \frac{\partial \bar{\Sigma}_t}{\partial \zeta} + \frac{\partial D_t}{\partial \zeta_t} \right) \\
&= \frac{1 - \frac{r+1}{r}\bar{\Sigma}_t \tau_y}{\tau_y} \left(-\frac{r+1}{r}\tau_y \right) \bar{\Sigma}_t^2 \left(\rho - 2\frac{r+1}{r}\zeta_t \tau_x \right) \\
&\quad + \frac{\frac{r+1}{r}\bar{\Sigma}_t \zeta_t \tau_x - \rho \bar{\Sigma}_t}{\tau_x} \left[\left(\frac{r+1}{r}\zeta_t \tau_x - \rho \right) \bar{\Sigma}_t^2 \left(\rho - 2\frac{r+1}{r}\zeta_t \tau_x \right) + \frac{r+1}{r}\bar{\Sigma}_t \tau_x \right] \\
&= \bar{\Sigma}_t^2 \Bigg\{ \frac{r+1}{r} \left(1 - \frac{r+1}{r}\bar{\Sigma}_t \tau_y \right) \left(2\frac{r+1}{r}\zeta_t \tau_x - \rho \right) \\
&\quad + \left(\frac{r+1}{r}\zeta_t \tau_x - \rho \right) \left[\left(\frac{r+1}{r}\zeta_t \tau_x - \rho \right) \frac{\bar{\Sigma}_t}{\tau_x} \left(\rho - 2\frac{r+1}{r}\zeta_t \tau_x \right) + \frac{r+1}{r} \right] \Bigg\}.
\end{aligned}$$

Let $\mathcal{A} \equiv \frac{r+1}{r} \left(1 - \frac{r+1}{r}\bar{\Sigma}_t \tau_y\right) \left(2\frac{r+1}{r}\zeta_t \tau_x - \rho\right) + \left(\frac{r+1}{r}\zeta_t \tau_x - \rho\right) \left[\left(\frac{r+1}{r}\zeta_t \tau_x - \rho\right) \frac{\bar{\Sigma}_t}{\tau_x} \left(\rho - 2\frac{r+1}{r}\zeta_t \tau_x\right) + \frac{r+1}{r} \right]$.

For a given parameter region, we determine the sign of $\mathcal{A}$.

For $\mathcal{A} <= 0$, $\zeta < 0$, $\rho, \tau_x, \tau_y > 0$ and $\bar{\Sigma}_t^{-1} = \frac{r+1}{r}\tau_y + \frac{r+1}{r}\tau_x\zeta_t^2 - \rho\zeta_t > 0$ to hold at the same time, we must have $0 < \bar{\Sigma}_t \leq \frac{-2\frac{r+1}{r}\tau_x\rho+3\left(\frac{r+1}{r}\right)^2\tau_x^2\zeta_t}{-\left(\frac{r+1}{r}\right)^2\tau_x\tau_y\rho-\rho^3+2\left(\frac{r+1}{r}\right)^3\tau_x^2\tau_y\zeta_t+4\frac{r+1}{r}\tau_x\rho^2\zeta_t-5\left(\frac{r+1}{r}\right)^2\tau_x^2\rho\zeta_t^2+2\left(\frac{r+1}{r}\right)^3\tau_x^3\zeta^3}$. Substitute in $\bar{\Sigma}_t$ and solve for ζ_t, we have $\{\zeta_t \mid \zeta_t < 0\}$ if $\tau_x \geq \bar{\Omega}_x$, $\{\zeta_t \mid \zeta_t \leq \underline{\zeta}_t \text{ or } \bar{\zeta}_t \leq \zeta_t < 0\}$ if $\underline{\Omega}_x < \tau_x < \bar{\Omega}_x$, and $\{\zeta_t \mid \zeta_t \leq \underline{\zeta}_t\}$ if $0 < \tau_x \leq \underline{\Omega}_x$. Since $\frac{\partial\zeta_t}{\partial K} < 0$, the sign of $\frac{\partial\mathbb{V}ar(p_t)}{\partial K}$ is:

(a) if $\tau_x \geq \bar{\Omega}_x$, then $\frac{\partial\mathbb{V}ar(p_t)}{\partial K} \geq 0$;

(b) if $\underline{\Omega}_x < \tau_x < \bar{\Omega}_x$, then

$$\frac{\partial\mathbb{V}ar(p_t)}{\partial K}\begin{cases} \geq 0, & \text{if } \zeta_t \leq \underline{\zeta}_t \text{ or } \bar{\zeta}_t \leq \zeta_t < 0, \\ < 0, & \text{if } \underline{\zeta}_t < \zeta_t < \bar{\zeta}_t; \end{cases}$$

(c) If $0 < \tau_x \leq \underline{\Omega}_x$, then

$$\frac{\partial\mathbb{V}ar(p_t)}{\partial K}\begin{cases} \geq 0, & \text{if } \zeta_t \leq \underline{\zeta}_t, \\ < 0, & \text{if } \underline{\zeta}_t < \zeta_t < 0; \end{cases}$$

where $\bar{\Omega}_x \equiv \frac{5\rho^2}{4\left(\frac{r+1}{r}\right)^2\tau_y}$, $\underline{\Omega}_x \equiv \frac{\rho^2}{\left(\frac{r+1}{r}\right)^2\tau_y}$, $\bar{\zeta}_t \equiv \frac{r}{2(r+1)\tau_x}\left[-\rho + \sqrt{-4\left(\frac{r+1}{r}\right)^2\tau_x\tau_y + 5\rho^2}\right]$, and $\underline{\zeta}_t \equiv \frac{r}{2(r+1)\tau_x}\left[-\rho - \sqrt{-4\left(\frac{r+1}{r}\right)^2\tau_x\tau_y + 5\rho^2}\right]$. □

Appendix C: Appendix of Chapter 3

C.1 General framework of optimal linear filtering

Generally, the unobservable state of the world V_t and the observable signal P_t are related by

$$\mathrm{d}P_t = A_t V_t \,\mathrm{d}t + B_t(P_t)\,\mathrm{d}Z_t = \left[A_t V_t - \eta_t^i\right]\mathrm{d}t + B_t(P_t)\,\mathrm{d}Z_t^i, \tag{C.1}$$

where A_t is deterministic at time t, and $B_t(P_t)$ is $\mathcal{F}_t$-measurable. We are interested in the optimal (minimum mean-square error) estimator for a $\mathcal{F}_t$-measurable function $h_t(V_t, P_t)$. Assume that h_t can be represented by

$$\mathrm{d}h_t = H_t\,\mathrm{d}t + b_t\,\mathrm{d}Z_t^f, \tag{C.2}$$

where H_t is a random process with $\mathbb{P}\left(\int_0^T |H_s|\,\mathrm{d}s < \infty\right) = 1$ and $\int_0^T \mathbb{E}^i\left[H_t^2\right]\mathrm{d}t < \infty$, b_t is a random process with $\mathbb{P}\left(\int_0^T |b_s|\,\mathrm{d}s < \infty\right) = 1$, and Z_t^f is a standard Wiener process over $(\Omega, \mathcal{F}, \mathbb{P})$, independent from Z_t. Let $\mathcal{F}_t^P = \{P_t; 0 \le s \le t\}$ denote the information filtration generated by the history of signal realizations, which is publicly observable. Agents also have a common prior $V_0 \mid \mathcal{F}_0^P \sim N(m_0, \gamma_0)$. Let $\pi_t^i(h) \equiv \mathbb{E}^i\left[h_t \mid \mathcal{F}_t^P\right]$. Theorem 4 shows how $\pi_t^i(h)$ is updated.

THEOREM 4. *Let the partially observable random process (V, P) permit the representation given by (C.1) and (C.2). If $\sup_{0\le t\le T}\mathbb{E}^i[h_t^2] < \infty$ and $\int_0^T \mathbb{E}^i\left[V_t^2\right]\mathrm{d}t < \infty$, then $\forall t \in [0, T]$, it holds almost surely with respect to probability measure $\mathbb{P}^i$ that*

$$\mathrm{d}\pi_t(h) = \pi_t(H)\,\mathrm{d}t + A_t\left[\pi_t(hV) - \pi_t(h)\pi_t(V)\right]\mathrm{d}Z_t^{Pi}, \tag{C.3}$$

where the innovation process $\mathrm{d}Z_t^{Pi} = \frac{1}{B_t(P_t)}\left[\mathrm{d}P_t - (A_t - \eta_t^i)\,\pi_t(V)\,\mathrm{d}t\right]$.

Theorem 4 can be proved by invoking the Liptser-Shiryaev optimal filtering equation[1] and change of measure. However, it does not explicitly specify the dynamics that V_t follows. In some special cases, though, we can write out its dynamics. If $h_t = V_t$ and $H_t = a_t V_t$, then

$$\mathrm{d}V_t = a_t V_t \,\mathrm{d}t + b_t \,\mathrm{d}Z_t^f \,. \tag{C.4}$$

Then equation (C.1) and (C.4) are the state-space representation that is widely studied in economics and finance literature. Applying Theorem 4, one can get the conditional mean of V_t from agent i's perspective, which is summarized by the following proposition.

Proposition 28. *Let* $m_t^i \equiv \mathbb{E}^i \left[V_t \mid \mathcal{F}_t^P \right]$. *Then*

$$\mathrm{d}m_t^i = a_t m_t^i \,\mathrm{d}t + A_t \gamma_t^i \,\mathrm{d}Z_t^{Pi}, \tag{C.5}$$

where $\mathrm{d}Z_t^{Pi} = \frac{1}{B_t(P_t)} \left[\mathrm{d}P_t - (A_t - \eta_t^i) m_t^i \,\mathrm{d}t \right]$.

By applying $h_t = V_t^2$ to Theorem 4 and Ito's Lemma to Proposition 28, one can get the tracking error (conditional variance) of V_t from agent i's perspective, which is summarized by the following proposition.

Proposition 29. *Let* $\gamma_t^i \equiv \mathbb{E}^i \left[(V_t - m_t^i)^2 \mid \mathcal{F}_t^P \right]$. *Then (C.5) and*

$$\pi_t^i(V_t) = m_0 + \int_0^t \left[2a_s \pi_s^i(V_t^2) + b_s^2 \right] \mathrm{d}s + \int_0^t A_s \left[\left(m_s^i \right)^3 + 2 m_s^i \gamma_s^i \right] \mathrm{d}Z_s^{Pi}, \tag{C.6}$$

$$\left(m_t^i \right)^2 = m_0^2 + \int_0^t \left[2a_s \left(m_s^i \right)^2 + A_s^2 \left(\gamma_s^i \right)^2 \right] \mathrm{d}s + 2 \int_0^t A_s m_s^i \gamma_s^i \,\mathrm{d}Z_s^{Pi}, \tag{C.7}$$

$$\gamma_t^i = \pi_t^i(V_t) - \left(m_t^i \right)^2 \tag{C.8}$$

jointly characterize γ_t^i.

[1] Liptser and Shiryaev (2001) proved a more general version of the optimal filtering equation.

C.2 Solutions to the main model

Proof of Proposition 29. We take $h_t = V_t^2$. Then by Ito's lemma,

$$\mathrm{d}V_t^2 = \left(2a_t V_t^2 + b_t^2\right)\mathrm{d}t + 2b_t V_t \,\mathrm{d}Z_t^f \,.$$

Apply (C.3) to V_t^2, we have

$$\mathrm{d}\pi_t^i(V^2) = \left[2a_t\pi_t^i(V^2) + b_t^2\right]\mathrm{d}t + A(t)\left[\pi_t^i(V^3) - \pi_t^i(V)\pi_t^i(V^2)\right]\mathrm{d}Z_t^{Pi} \,.$$

Due to the Gaussian behavior of the process $\{V_t, P_t\}$, the moments $\pi_t^i(V^n)$ for all $n \geq 3$ can be expressed in terms of $\pi_t^i(V)$ and $\pi_t^i(theta^2)$. In particular,

$$\pi_t^i(V^3) - \pi_t^i(V)\pi_t^i(V^2) = \left(m_t^i\right)^3 + 2m_t^i\gamma_t^i,$$

and therefore,

$$\mathrm{d}\pi_t^i(V^2) = \left[2a_t\pi_t^i(V^2) + b_t^2\right]\mathrm{d}t + A(t)\left[\left(m_t^i\right)^3 + 2m_t^i\gamma_t^i\right]\mathrm{d}Z_t^{Pi} \,.$$

Applying Ito's lemma to (C.5), we have

$$\mathrm{d}(m_t^i)^2 = \left[2a_t\left(m_t^i\right)^2 + A^2(t)\left(\gamma_t^i\right)^2\right]\mathrm{d}t + 2A(t)m_t^i\gamma_t^i\,\mathrm{d}Z_t^{Pi} \,.$$

Integrate the above two equations from 0 to t, we have

$$\begin{aligned}
\pi_t^i(V) &= m_0 + \int_0^t \left[2a_s\pi_s^i(V^2) + b_s^2\right]\mathrm{d}s + \int_0^t A_s\left[\left(m_s^i\right)^3 + 2m_s^i\gamma_s^i\right]\mathrm{d}Z_s^{Pi}, \\
\left(m_t^i\right)^2 &= m_0^2 + \int_0^t \left[2a_s\left(m_s^i\right)^2 + A_s^2\left(\gamma_s^i\right)^2\right]\mathrm{d}s + 2\int_0^t A_s m_s^i\gamma_s^i\,\mathrm{d}Z_s^{Pi} \,.
\end{aligned}$$

Then equations (C.5), (C.6), (C.7) and (C.8) jointly characterize γ_t^i. □

Proof of Proposition 24. The Bayesian agent has an objective perceptual set $\mathbb{P}$. The observable process P_t follows

$$\mathrm{d}P_t = V\,\mathrm{d}t + \sigma\,\mathrm{d}Z_t, \tag{C.9}$$

which has the solution $P_t = P_0 + Vt + \sigma\int_0^t \mathrm{d}Z_t$. At each t, the Bayesian agent compute the conditional mean and variance of V_t by $m_t = \tilde{\mathbb{E}}\left[V \mid \mathcal{F}_t^P\right]$ and $\gamma_t = \tilde{\mathbb{E}}\left[(V-m_t)^2 \mid \mathcal{F}_t^P\right]$, respectively. The innovation process becomes

$$\mathrm{d}Z_t^P = \frac{1}{\sigma}\left(\mathrm{d}P_t - m_t\,\mathrm{d}t\right). \tag{C.10}$$

It is a standard Wiener process with respect to $\mathcal{F}_t^P$ under $\mathbb{P}$. The Bayesian agent's belief updating process follows:

$$\mathrm{d}m_t = \frac{\gamma_t}{\sigma}\,\mathrm{d}Z_t^P, \tag{C.11}$$

$$\mathrm{d}\gamma_t = -\frac{(\gamma_t)^2}{\sigma^2}\,\mathrm{d}t\,. \tag{C.12}$$

By integrating (C.12), we obtain

$$\gamma_t = \frac{\gamma_0\sigma^2}{\gamma_0 t + \sigma^2}. \tag{C.13}$$

Since $\lim_{t\to\infty}\gamma_t = 0$, the unknown state V is eventually learned. To solve for the expected state of world, combine (C.10) and (C.11) to get

$$\begin{aligned}
\mathrm{d}m_t &= \frac{\gamma_t}{\sigma^2}\left(\mathrm{d}P_t - m_t\,\mathrm{d}t\right)\\
\implies \frac{1}{\gamma_t}\,\mathrm{d}m_t &= -\frac{m_t}{\sigma^2}\,\mathrm{d}t + \frac{1}{\sigma^2}\,\mathrm{d}P_t\\
\implies \frac{1}{\gamma_t}\,\mathrm{d}m_t + m_t\,\mathrm{d}\frac{1}{\gamma_t} &= \frac{1}{\sigma^2}\,\mathrm{d}P_t\\
\implies \mathrm{d}\left(\frac{m_t}{\gamma_t}\right) &= \frac{1}{\sigma^2}\,\mathrm{d}P_t\\
\implies m_t &= \frac{\gamma_t m_0}{\gamma_0} + \frac{\gamma_t}{\sigma^2}\left(P_0 + Vt + \sigma\int_0^t \mathrm{d}Z_t\right).
\end{aligned}$$

Plug in the solution of γ_t to finally arrive at

$$m_t = \frac{m_0\sigma^2 + P_0\gamma_0}{\gamma_0 t + \sigma^2} + V\frac{\gamma_0 t}{\gamma_0 t + \sigma^2} + \frac{\gamma_0\sigma}{\gamma_0 t + \sigma^2}\int_0^t \mathrm{d}Z_t\,. \tag{C.14}$$

Then it is obvious to see that $\lim_{t\to\infty} m_t = V$. So the Bayesian agent eventually learns the true state of the world. □

Behavioral agent i views the world through the lens of her subjective perceptual set $\mathbb{P}^i$. P_t has the following dynamics under measure $\mathbb{P}^i$:

$$\mathrm{d}P_t = (V - \eta^i)\,\mathrm{d}t + \sigma\,\mathrm{d}Z_t^i, \tag{C.15}$$

which has the solution $P_t = P_0 + (V-\eta^i)t + \sigma\sqrt{t}\int_0^t \mathrm{d}Z_t^i$. At time t, from agent i's perspective, V has conditional mean

$$m_t^i = \tilde{\mathbb{E}}^i\left[V \mid \mathcal{F}_t^P\right] \tag{C.16}$$

and variance

$$\gamma_t^i = \tilde{\mathbb{E}}^i\left[\left(V - m_t^i\right)^2 \mid \mathcal{F}_t^P\right]. \tag{C.17}$$

The innovation process

$$\mathrm{d}Z_t^i = \frac{1}{\sigma}\left(\mathrm{d}P_t - (m_t^i - \eta^i)\,\mathrm{d}t\right). \tag{C.18}$$

It is a standard Wiener process with respect to $\mathcal{F}_t^P$ under $\mathbb{P}^i$. Behavioral agent i updates belief as

follows:

$$\mathrm{d}m_t^i = \frac{\gamma_t^i}{\sigma}\,\mathrm{d}Z_t^i, \tag{C.19}$$

$$\mathrm{d}\gamma_t^i = -\frac{\left(\gamma_t^i\right)^2}{\sigma^2}\,\mathrm{d}t\,. \tag{C.20}$$

By integrating (C.20), we obtain

$$\gamma_t^i = \frac{\gamma_0\sigma^2}{\gamma_0 t + \sigma^2}, \tag{C.21}$$

which is independent from i and the same as that of the Bayesian agent. At a given time t, all behavioral agents and the Bayesian agent have the same posterior precision about V. To solve for the expected state of world, we combine (C.18), (C.19) and (C.20) to get

$$\mathrm{d}m_t^i = \frac{\gamma_t^i}{\sigma^2}\left(\mathrm{d}P_t - (m_t^i - \eta^i)\,\mathrm{d}t\right) \implies m_t^i = \frac{m_0\sigma^2 + P_0\gamma_0}{\gamma_0 t + \sigma^2} + (V+\eta^i)\frac{\gamma_0 t}{\gamma_0 t + \sigma^2} + \frac{\gamma_0\sigma}{\gamma_0 t + \sigma^2}\int_0^t \mathrm{d}Z_t\,. \tag{C.22}$$

Then it is obvious to see that $\lim_{t\to\infty} m_t^i = V + \eta^i$.