Learning Costs and Information Quality: Why Do We Disagree?

WEI WANG*

November 11, 2022
[Latest Version]

Abstract

I study asset prices and portfolio choice in an exchange economy with an information market. Investors are heterogeneous in learning costs when using signals of different qualities. Information providers decide the quality and price of signals. In equilibrium, ex-ante homogeneous information providers voluntarily differentiate in their signal quality. The market equilibrium is not socially optimal. The information quality gap is overly large. Under some conditions, low-quality information investors trade more actively and security price volatility is higher in the social optimum.

Keywords: learning costs; agree to disagree; information market; asset pricing.

JEL Codes: G11, G12, L86

*Wei Wang is with department of finance, Indiana University (Email: ww18@iu.edu). The views in this paper are the responsibility of the author, not the institutions he is affiliated with. I am indebted to my committee members Christian Heyerdahl-Larsen (Chair), Charles Trzcinka, Jun Yang, Alessandro Previtero for their support and advice. I also thank Jan Schneemeier, Eitan Goldman, Fotis Grigoris, Niklas Hüther, Alessio Piccolo, Jordan Martel, Kristoph Kleiner, Carter Davis, Preetesh Kantak, and other participants in Kelley Finance Seminar for helpful comments.

When investors make portfolio choices, they must form their beliefs about states, parameters, and models by learning from external information sources. As Veronesi (2000) argues, "... investors are flooded with a variety of information: corporations' earnings reports, revisions of macroeconomic indexes, policymakers' statements, and political news...several questions arise regarding the relationship between the quality of information that investors receive and asset returns...". There is a vast number of information sources with heterogeneous quality. Understanding the reasons for the heterogeneity in quality and how it impacts investors is important. For example, what types of tradeoffs are investors confronted with when choosing among information sources with different qualities? If producers determine the quality, which level would they choose in equilibrium? How would the decisions of quality affect investors' consumption plan and trading strategy? What are the implications for security markets and social welfare?

I extend the continuous-time Lucas tree model with heterogeneous beliefs and incorporate three features: first, information needs to be produced and their producers are profit-maximizing; second, the information industry is not monopolistic; third, the cost of using the same information is not heterogeneous across investors, especially high-quality information.

The model generates four interesting results: first, when there is a heterogeneity in investors' learning costs, ex-ante homogeneous producers will choose to be heterogeneous information producers (i.e., there are multiple information sources of different quality). The intuition for the first result comes from Hotelling (1929). In the original paper, consumers are uniformly distributed and sellers need to choose locations of their shops, which affects consumers' travelling costs to get to shops. Therefore, the game becomes similar to the Cournot model, where two sellers make choices strategically and thereby avoid competition. In our setting, the investors' learning costs are similar to the travelling costs in the original setting. The information quality is similar to the quality of goods. While investors want to avoid competition and get higher market power, they choose to produce information of different quality to capture different investors.

The second result speaks to the social welfare. In the market equilibrium, the information differentiation is higher than the social optimum (i.e., the information quality gap is too large). This could be explained in a similar way. Think about an investor who can learn to use good information with relatively low learning costs. He could have benefited from this high-quality information. But producers want to avoid competition and do not provide such options. Instead, they either produce information that is either low quality or takes high costs to learn. Social welfare could be higher if producers make their information quality gap smaller, which will reduce producers' profits.

The third result is about the disagreement across investor groups. The information quality gap does not always lead to more salient disagreement. Traditional wisdom naturally leads to the conclusion that the information quality gap generates volatility. However, this depends on the quality of signals. If the quality of one signal is very high, the disagreement volatility increases with the information gap. However, if the quality of one signal is not very high, the improvement of low-quality signals leads to higher volatility of disagreement because investors put more weight on signals, and by doing so bringing in extra noise from the signal.

The last result follows the previous one. I find that in some scenarios, in the social optimum where the low-quality information should be less noisier, the disagreement is less volatile while the security price is more volatile, which is conflicting with most results in the disagreement literature. The traditional wisdom argues that disagreement volatility leads to speculative trading. Speculation creates excess volatility in security prices (Dumas et al., 2009; Xiong and Yan, 2010). However, in my model, there are two effects: watermark effect and speculation effect. Watermark effect refers to how much weight investors put on signals when updating their beliefs. The weight will affect how much of the noise in a signal is incorporated into trading and prices. The speculation effect refers to the fact that investors take speculative positions against each other based on disagreement. The disagreement volatility increases the security prices volatility by inducing more frequent speculative trading.

This paper is related to the literature on disagreement and heterogeneous beliefs. There is a huge volume of studies about the source of heterogeneous beliefs, such asymmetric information (Wang, 1993; Grossman and Stiglitz, 1980; Verrecchia, 1982), psychological bias (Daniel et al., 2005; Scheinkman and Xiong, 2003; Dumas et al., 2009), model uncertainty (Buraschi and Jiltsov, 2006), and learning from experience Ehling et al. (2018). Other literatures study the impact of disagreement on stock markets, bond markets, foreign exchange markets, economic growth, etc (Chang et al., 2022; Xiong and Yan, 2010; Ehling et al., 2018; Illeditsch et al., 2022). My paper adds to the literature by endogenizing information production and introducing the competition between different producers.

The paper also contributes to the interplay between financial markets and information markets. Gentzkow and Shapiro (2006) discuss how information producers cater to the prior of consumers to build up reputation on quality, which lead to consistent bias among consumers and harm social welfare. Admati and Pfleiderer (1986) consider how sellers should design the quality of information to maximize the profits and prevent information leakage through informative prices. Admati and Pfleiderer (1988, 1990) investigate into two strategies of selling information, including "direct sale," which provides the raw signals to buyers, and "indirect sale", which provides a portfolio based on the raw signals to hide some details. Huang et al. (2022) extends this monopolistic setting by adding costly learning from price. These papers build their analysis on the assumption that investors learn information from prices, creating costs for information sellers to produce high-quality information. In my paper, the benchmark model assumes that investors agree to disagree, inducing speculative trading. Without information leakage from informative prices, heterogeneous signals emerge because information providers cater to the differential investors learning costs. Oligopolistic setting is more realistic and convenient for future empirical applications.

My paper has implications on research on social media and retail trading. Cookson and Niessner (2020) use data from Stocktwits, a social media trading platform, to highlight the role of information set difference in generating disagreement. Many retail traders rely on free online information sources to make portfolio decisions. The media could have impact on their portfolio choice. My paper conducts welfare analysis on how different media affect the financial markets. The policy implication is pretty counter-intuitive. Most policymakers see volatility of stock prices as a negative signal. My model shows that in some cases when low-quality information improves, security price volatility would increase. This increase is not due to speculative trading, but higher weight on the noisy signals.

The balance of the paper is arranged as follows. Section 1 sets up the model. Section 2 solves the model. Section 3 analyzes the security market behaviors. Section 4 compares the social optimum and market equilibrium. Section 5 contains concluding remarks.

1 Model Setup

The model is a sequential game. In the economy, there are two information providers and a continuum of atomistic investors. There are markets for information, consumption goods, and securities separately. The game is in two stages:

- Stage 0: information market clears
    1. Information providers decide information quality and prices
    2. Investors choose the unique information provider
- Stage 1: Consumption goods market and security markets clear
    1. Information providers produce signals
    2. Investors make portfolio choices to support their consumption plans

The remaining parts of the section solve the game backward starting with Stage 1 and then Stage 0.

1.1 Investors' Consumption and Investment Decisions

I adopt the continuous-time Lucas tree model with an unobservable long-run growth rate. To allow for disagreement, I also introduce two groups of investors with heterogeneous signals regarding the long-run growth rate. Investors choose and receive different signals. They "agree to disagree" and speculate in the capital markets. I study a competitive equilibrium in which each investor optimizes consumption and investment decisions based on her or his own expectation. Market clearing conditions determine the equilibrium short rate and asset prices given the information quality determined in stage 0.

1.1.1 Economy

Consider an economy with an aggregate dividend-flow diffusion process $\{\delta_t\}$. The dividend process follows:

$$\frac{d\delta_t}{\delta_t} = f_t dt + \sigma dZ_t \tag{1}$$

where f_t is the unobservable long-run growth rate and σ is constant and known. The fundamental uncertainty Z_t is a standard Brownian motion process. I assume the long-run growth rate f_t follows an *Ornstein-Uhlenbeck* process:

$$df_t = -\zeta(f_t - \overline{f})dt + \sigma_f dZ_t^f \tag{2}$$

where ζ is the mean-reverting parameter, $\overline{f}$ is the long-run mean of f_t, σ_f is a volatility parameter, and Z_t^f is a standard Brownian motion independent of Z_t. Intuitively, I interpret f_t as the potential economic growth and dZ_t^f as a random shock to actual economic growth.. The value of f_t needs to be estimated using economic data.

1.1.2 Heterogeneous Expectations

I analyze two groups of investors who self select into customers of different information sellers, who produce signals on the potential long-run growth rate f_t. In particular, I assume that once an investor become a member of one group, the investor only update beliefs based on the dividends and the signal to which they subscribe. More specifically, under the objective probability, the signal s_t^i sold by seller i ($i \in \{A, B\}$) follows:

$$ds_t^i = \phi_i \sigma_s dZ_t^f + \sqrt{1 - \phi_i^2} \sigma_s dZ_t^{s,i} \tag{3}$$

where the correlation $\phi_i \in [0, 1]$. ϕ_i is the correlation between the signal ds_t^i and the long-run growth rate shock dZ_t^f, which is the indicator of information quality in stage 0. The information market equilibrium in stage 0 determines the values of ϕ_i, and therefore, in stage 1, investors take them as given. $Z_t^{s,i}$ is the noise part in the signal, which is a standard Brownian motion independent of both Z_t^f and Z_t. Besides, $Z_t^{s,A}$ and $Z_t^{s,B}$ are independent of each other. The specification of the signal structure facilitates the measure of signal informativeness.

I assume that both agents can observe two signals. Group-i investors' information set at time t includes $\{s_\tau^A, s_\tau^B, \delta_\tau\}_{\tau=0}^t$. But it incurs learning costs to incorporate information sources into the investment processes, which are non-trivial and paid in stage 0. Under the subjective belief of group-i investors, the processes for signal i and signal $-i$ follow:

$$ds_t^i = \phi_i \sigma_s dZ_t^f + \sqrt{1 - \phi_i^2} \sigma_s dZ_t^{s,i} \tag{4}$$

$$ds_t^{-i} = \sigma_s dZ_t^{s,-i} \tag{5}$$

I denote group-i investors' posterior distribution about f_t at time t by

$$f_t | \{s_\tau^A, s_\tau^B, \delta_\tau\} \sim N(\hat{f}_t^i, \gamma_t^i),\ i \in \{A, B\} \tag{6}$$

where $\hat{f}_t^i$ is the mean of group-i investors' posterior distribution and γ_t^i is the estimation error, defined as $E_t((f_t - \hat{f}_t^i)^2)$. Hereafter I refer to $\hat{f}_t^i$ as their belief. From the filtering theory (see (theorem 12.7, Lipster and Shiryaev 2001)), the conditional expected values, $\hat{f}^A$ and the esitmation error γ_t^i obey the following stochastic differential equation:

$$d\hat{f}_t^i = -\zeta(\hat{f}_t^i - \overline{f})dt + \frac{\gamma_t^i}{\sigma} d\hat{Z}_t^i + \frac{\sigma_f \phi_i}{\sigma_s} ds_t^i \tag{7}$$

$$\frac{d\gamma_t^i}{dt} = -2\zeta\gamma_t + \sigma_f^2 - \frac{(\gamma_t^i)^2}{\sigma^2} - \sigma_f^2 \phi^2 \tag{8}$$

where $d\hat{Z}_t^i := \frac{1}{\sigma}\left(\frac{d\delta_t}{\delta_t} - \hat{f}_t^i\right)$ is the information shock in $\frac{d\delta_t}{\delta_t}$ to group-i investors. $\hat{Z}_t^i$ is a standard Brownian motion from group-i investors' point of view, i.e., under $\mathcal{F}_t^i := \sigma(\delta_\tau, s_\tau^i : 0 \leq \tau \leq t)$. γ_t^i converges to a constant level, the steady-state level γ^i, which is the positive root of equation $\frac{d\gamma_t^i}{dt} = 0$. The solution is $\gamma^i = -\zeta\sigma^2 + \sigma\sqrt{\zeta^2\sigma^2 + \sigma_f^2(1 - \phi_i^2)}$.

Note that since two groups update their beliefs based on two imperfectly correlated signals, they always disagree as long as one of ϕ_A and ϕ_B is not equal to one. Even if $\phi_A = \phi_B$, disagreement still exists. To clarify that, I define $g_t = \hat{f}_t^A - \hat{f}_t^B$ as disagreement. Proposition 1 displays the dynamic of disagreement.

Proposition 1 *The dynamic of disagreement follows an Ornstein-Uhlenbeck process:*

$$dg_t = -(\zeta + \frac{\gamma_A}{\sigma^2})\hat{g}_t dt + \frac{\gamma_A - \gamma_B}{\sigma} d\hat{Z}_t^B + \frac{\sigma_f}{\sigma_s}(\phi_A ds_t^A - \phi_B ds_t^B) \tag{9}$$

Proposition 1 has at least three implications. First, the disagreement is mean-reverting and the long-run mean is zero. There is no systematic difference in beliefs for two groups. Second, the dynamic of disagreement, while reverting to zero, could be exacerbated by both the dividend shock and the signal discrepancy. Third, the volatility of disagreement change, $E[(dg_t)^2] = \frac{(\gamma_A - \gamma_B)^2}{\sigma^2} + \sigma_f^2(\phi_A^2 + \phi_B^2 - 2\phi_A^2\phi_B^2)dt$, has a non-monotonic relationship with ϕ_A and ϕ_B, as shown in Figure 1. This means that whether increasing ϕ_B will mitigate or exacerbate the disagreement volatility depends on the value of ϕ_A and vice versa. The third implication, to the best of my knowledge, is new to the literature. The common setting is that ϕ_i is zero for one group versus nonzero for the other or as in Xiong and Yan (2010) it require that $\phi_A = -\phi_B = \phi$. However, setting both ϕ_i free on $[0, 1]$ will generate non-monotonicity not only in the disagreement volatility but also in many other calculation, which make the analysis non-trivial, not merely a simple extension.

[Place Figure 1 about here]

1.1.3 Equilibrium

Disagreement causes speculative trading. If $g_t > 0$, group A bets on the higher level but quicker decline of future f_t against group B. Note that from each group's perspective, there are three sources of risks. For group-i investors, the shocks are $d\hat{Z}_t^i$, ds_t^A, and ds_t^B. Thus, the markets are complete if investors can trade a risk-free asset and three risky assets that span these three sources of risks. In reality, financial markets offer many securities, such as derivations, for investors to construct their bets and to complete the markets. Buraschi and Jiltsov (2006) prove that the disagreement due to the model uncertainty creates market incompleteness and options can complete this incompleteness. Therefore, I analyze an equilibrium with dynamically complete financial markets.

Assuming a complete financial market, I can use the static martingale formulation (as in Cox and Huang (1989)). At time zero, group-i investors ($i \in \{A, B\}$) are endowed with $\overline{\theta}^i$ fraction of the total wealth of the economy, with $\overline{\theta}^i \in (0, 1)$ and $\overline{\theta}^A + \overline{\theta}^B = 1$. The objective of group i is to maximize their expected discounted lifetime utility from consumption c_t^i under the subjective belief:

$$\sup_c E_0^i \int_0^\infty e^{-\rho t} \frac{(c_t^i)^{1-\alpha}}{1-\alpha} dt \tag{10}$$

subject to the lifetime budget constraint

$$E^i \int_0^\infty \xi_t^i c_t^i dt = \overline{\theta}^i E^i \int_0^\infty \xi_t^i \delta_t^i dt \tag{11}$$

where ξ^i is the change of measure from group i's probability measure to the risk-neutralized measure and $\overline{\theta}^i$ is the the share of equity with which group i is initially endowed. [1]

In equilibrium, both groups of investors make their optimal consumption and portfolio choices based on their beliefs, and all financial and good markets clear. The detailed derivation of the equilibrium is in the Appendix B, and I summarize the properties in Theorem 1.

Theorem 1 *For the economy defined above,*

1 *Define* $\eta_t := \frac{dQ^A}{dQ^B}$ *where* Q^A *and* Q^B *are risk-neutral probability under the subjective measure of group A and B.* η *follows a*

$$\frac{d\eta_t}{\eta_t} = \frac{\hat{g}_t}{\sigma} d\hat{Z}_t^B \tag{12}$$

2 *The consumption of group A and B are* $c_t^A = (1-\omega(\eta_t))\delta_t$ *and* $c_t^B = \omega(\eta_t)\delta_t$ *where* $\omega(\eta_t) = \frac{(\lambda^A)^{\frac{1}{\alpha}}}{(\lambda^A)^{\frac{1}{\alpha}} + (\lambda^B)^{\frac{1}{\alpha}}(\eta_t)^{\frac{1}{\alpha}}}$. *The SDFs of group A and B are* $\zeta_t^A = \frac{1}{\lambda^A}\delta_t^{-\frac{1}{\alpha}}(1-\omega(\eta_t))^{-\frac{1}{\alpha}}$ *and* $\zeta_t^B = \frac{1}{\lambda^B}\delta_t^{-\frac{1}{\alpha}}\omega(\eta_t)^{-\frac{1}{\alpha}}$

3 *The equilibrium nominal short rate is given by:*

$$r_t = \underbrace{\rho + \alpha \hat{f}_t^B - \frac{(1+\alpha)\alpha}{2}\sigma^2}_{\text{short rate in the group-b economy}} + \underbrace{\alpha g[1-\omega(\eta_t)]}_{\text{markup from Group-A}} - \underbrace{\frac{\alpha-1}{2\alpha}\frac{g^2}{\sigma^2}\omega(\eta_t)[1-\omega(\eta_t)]}_{\text{disagreement-risk premium}} \tag{13}$$

Theorem 1 highlights the dynamics of consumption, SDF, and the nominal short rates as functions of Radon-Nikodym derivative. Radon-Nikodym derivative is also the ratio of

[1]Notice that $\eta_t = \frac{\zeta_t^B}{\zeta_A}$. This is because in equilibrium we have $E_t^A[\zeta_t^A x] = E_t^B[\zeta_t^B x]$. We have $E_t^A[x] = E_t^B[x * \eta_t]$. Therefore, η_t has to equal to $\frac{\zeta_B}{\zeta_A}$

two groups' SDFs, indicating the relative valuation of risks. A thought experiment helps understand the economic mechanism of the first two points in Theorem 1. Assume group B attributes the dividend increase to temporary shock instead of long-run growth rate (i.e., $d\hat{Z}_t^B > 0$). Group A has a higher expectation of f_t than group B (i.e.,$g_t > 0$). Due to the wealth effect, group A consumes more (i.e., $1 - \omega(\eta_t)$ increases). Point 3 in Theorem 1 decomposes the nominal short rate into three parts. The first part is the short rate in a hypothetical economy where only group b exists. The second part is due to the level of disagreement among two groups. Part 3 is the risk-premium compensated for the volatility of disagreement. In a homogeneous economy, the short rate volatility comes from the uncertainty in learning process. In a heterogeneous economy, disagreement is another source of volatility.

1.1.4 Investors' Value Functions

The setup of stage 0 has one critical element derived from stage 1: the lifetime utility of both groups under subjective belief: $V^i(\phi_A, \phi_B) = E_0^i\left[\int_0^\infty e^{-\rho t} u(c_t) dt\right]$. I assume that $\alpha - 1 \in \mathrm{N}$ to leverage the binomial expansion and to get an explicit solution of $V^i(\phi_A, \phi_B)$. For the non-integer case, fast Fourier transform can be used to calculate the numerical solution. I also assume that $\lambda_A = \lambda_B$ to simplify parametric complexity.

Lemma 1 presents the moment-generating function of δ_t and η_t under group B's subjective belief. The equation will appear recurrently in the lifetime utility, valuation of securities, etc. Dumas, Kurshev, and Uppal (2009) calculate the special case where $\phi_A = 0$ and ϕ_B is non-zero. Lemma 1 is a generalized version allowing $\phi_i \in [0, 1]$ $(i \in \{A, B\})$

Lemma 1 (Moment-generating Function)

$$E_{\hat{f}_t^B, g_t}[\delta_T^\epsilon \eta_T^\chi] = \delta_t^\epsilon \exp[A_0(t) + A_1(t)\hat{f}_t^B]\eta_t^\chi \exp[B_0(t) + B_1(t)g_t + B_2(t)g_t^2] \tag{14}$$

where ϵ is any real number and $\chi \in [0, 1]$

The proof of Lemma 1 is provided in Appendix D. The model setup belongs to the category of exponential affine-quadratic models. In this context, moment-generating functions have a quadratic exponential form. Proposition 2 shows that the lifetime utility can be decomposed into values that can be calculated using MGF.

Proposition 2 *Suppose $\lambda_A = \lambda_B$, the lifetime utility of both groups can be expressed as:*

$$V_A(\phi_A, \phi_B) = \frac{1}{1-\alpha} \sum_{n=0}^{\alpha-1} \binom{\alpha-1}{n} \int_0^{\infty} e^{-\rho t} E^B \left[(\delta_t)^{-(\alpha-1)} (\eta_t)^{1-\frac{n}{\alpha}} \right] dt \tag{15}$$

$$V_B(\phi_A, \phi_B) = \frac{1}{1-\alpha} \sum_{n=0}^{\alpha-1} \binom{\alpha-1}{n} \int_0^{\infty} e^{-\rho t} E^B (\delta_t^{-(\alpha-1)} \eta_t^{\frac{n}{\alpha}}) dt \tag{16}$$

Proposition 2 shows that under subjective belief of group B, lifetime utility of both groups can be decomposed into parts which can be explicitly calculated using Lemma 1. Figure 2 visualizes $V_A(\phi_A, \phi_B)$. Due to the symmetric nature of the model, this figure also conveys information of $V_B(\phi_A, \phi_B)$.

[Place Figure 2 about here]

Figure 2 shares the same patterns with Figure 1. Figure 2a displays the relationship between V_A and ϕ_A given different values of ϕ_B. 2b displays the relationship between V_A and ϕ_B given different values of ϕ_A. The intuition is that the volatility of disagreement measures the speculative trading intensity. Agents benefit from speculative trading. In other words, they want to trade under the subjective belief. And the relationship between the information quality difference and the speculative trading opportunities is quite complex, which generates many interesting results.

1.2 Information Market

Investors make choices of information providers considering the prices and qualities. Information providers also decide signal quality and prices strategically to attract investors. In period 0, information market clears with certain levels of prices and quality. Investors are the demand side and information providers are the supply side. The market is structured as a sequential game where the supply sides decide on quality and price by steps first and the demand side then choose the producers. Distinguished from the literature, my model accounts for investors' heterogeneity when they update beliefs using signals. To be more specific, investors need to incur a certain level of "learning cost" at period 0 after choosing the information provider. The magnitude of this learning cost is determined by investors' types as well as the quality of the signal. The learning cost heterogeneity generates incentives for producers to differentiate in quality levels.

1.2.1 Investors (Consumers) Problem

Investors' net surplus have three components: utility from investment, information price, and potential learning costs. The expected lifetime utility function, $V_i(\phi_A, \phi_B)$ ($i \in \{A, B\}$), is derived in Proposition 2. The information price is an endogenous variable determined in equilibrium.

I model learning costs in the spirits of Hotelling (1929). For a certain investor, k, she is endowed with a type, denoted as $x_k \in [0, 1]$. Given the information quality, ϕ_i, the learning costs incurred is equal to $\frac{t}{2}(x_k - \phi_i)^2$. The magnitude of t proxies for the learning cost heterogeneity. If $t = 0$, the learning costs are zero for any investor type x_k. Given two investor type, x_k and x_j, the gap between their learning costs will increases as t increases. [2]

The utility maximization problem for agent k with type x_k is to choose information seller i that maximize net surplus:

$$\max_{i \in \{A,B\}} V_i(\phi_A, \phi_B) - p_i - \frac{t}{2}(x_k - \phi_i)^2 \tag{17}$$

[2]The implicit assumption is the symmetry of learning costs. High-type investors incur the same learning costs using low-quality information as the low-type investors incur using high-quality information. While this assumption is counter-intuitive, it won't affect the qualitative results as shown in the toy model but provide great tractability when calculating market shares.

where $V_i(\phi_A, \phi_B)$ is the lifetime utility, p_i is the price of provider i and $\frac{t}{2}(x_k - \phi_i)^2$ is the learning costs.

The optimal choice is a threshold function of investor type. For notional convenience, I assume $\phi_A < \phi_B$ hereafter. The agent chooses A if $V_A - p_A - \frac{t}{2}(x_k - \phi_A)^2 > V_B - p_B - \frac{t}{2}(x_k - \phi_B)^2$. This generates a threshold:

$$\dot{x} = [\frac{1}{2}(\phi_A + \phi_B) + \frac{1}{t(\phi_B - \phi_A)}((P_B - P_A) - (V_B - V_A))]_0^1 \tag{18}$$

where $[]_0^1$ means being truncated between 0 and 1. If $x_k < \dot{x}$, the optimal choice is provider A. Otherwise, the investor should choose B. $\dot{x}$ becomes the boundary of market shares.

The distribution of type x_k also affects the learning costs heterogeneity. I assume that the distribution of type x_k over [0,1] is summarized by the probability distribution function $G(x)$ and the corresponding probability mass function $g(x)$. The market share for A and B are $G(\dot{x})$ and $1 - G(\dot{x})$. The statistical characteristics of the distribution is discussed further in later sections. Given my main focus is on characterization, I assume $G(x)$ and $g(x)$ are well-defined without explicitly stating the required regularity conditions.

1.2.2 Information Providers (Producers) Problem

I limit the analysis to the case with two information providers, with the same marginal costs (c) and the technology, i.e., the quality levels they can attain are the same.[3] Profits of information providers equal to the market share times the gap between price and marginal costs. To maximize profits, information providers make two strategic decisions sequentially: information qualities (ϕ_i) and price (p_i).

For the price decision, given the price of seller B and (ϕ_A, ϕ_B), provider A sets the optimal

[3] The implicit assumption is that there is an infinitely high entry of barrier. The entry barrier for information market could be reputations, regulations, etc. Even on the macro level, the competition is high in reality, the conclusions apply to niches. As long as the fixed costs are non-zero, local market power exists in the model.

price by solving:

$$P_A = \arg\max_{P_A} G(\dot{x})(P_A - c) \tag{19}$$

The optimal solution is $P_A = c + \frac{G(\dot{x})}{g(\dot{x})} t(\phi_B - \phi_A)$. P_A is the best response given the qualities and P_B. The best response of provider B is $P_B = c + \frac{1-G(\dot{x})}{g(\dot{x})} t(\phi_B - \phi_A)$ by solving a similar optimization problem. Nash equilibrium requires that the pricing strategies of both providers are the best response functions to each other. To solve the Nash equilibrium of pricing, substitute P_i into $\dot{x}$ in to get the boundary of market share x^* after the optimal pricing is imposed:

$$x^* = \frac{1}{2}(\phi_A + \phi_B) + \frac{1 - 2G(x^*)}{g(x^*)} - \frac{1}{t}\frac{V_B(\phi_A, \phi_B) - V_A(\phi_A, \phi_B)}{\phi_B - \phi_A} \tag{20}$$

x^* is a root of equation 20. In a special case where the distribution is uniformly distributed on $[a, b]$, the closed-form interior solution of x^* is $\frac{1}{3}(a + b) + \frac{1}{6}(\phi_A + \phi_B) - \frac{1}{3t}\frac{V_B - V_A}{\phi_B - \phi_A}$. Using Intermediate value theorem, it can be proved that when t is large enough, there must be at least one root existing on the interval $[0, 1]$. In the following analysis I focus only on cases where x^* is an interior solution.

The next step is to determine the quality. To do so, I follow a similar approach as to for the optimal pricing strategies. After inputting the optimal pricing strategies, the profits for group A and B are $\pi_A(\phi_A; \phi_B) = \frac{G(x^*)^2}{g(x^*)} t(\phi_B - \phi_A)$ and $\pi_B(\phi_B; \phi_A) = \frac{(1-G(x^*))^2}{g(x^*)} t(\phi_B - \phi_A)$. For provider i, the optimization problem is to choose ϕ_i to maximize profits. Assumption 1 describes the production sets defined by a technology constraint.

Assumption 1 *Denote* $\overline{x} := \sup\{x : g(x) > 0\}$*, which is the supremum of the support of the distribution* g*. When producing information, the quality cannot be higher than* $\overline{x}$*.*

An interpretation of assumption 1 is that the quality of information sold can't be higher than the maximum of the investor types. Intuitively, the rationale is that signals need to be produced by some agents. The producers should be able to use the information without any costs. If this assumption doesn't hold, it means the information has no producer.

Following the spirit of Nash equilibrium, the equilibrium quality levels (ϕ_A^*, ϕ_B^*) satisfy the following equation system:

$$\phi_A^* = \arg\max_{\phi_A \in [0,\overline{x}]} \pi_A(\phi_A; \phi_B^*) \tag{21}$$

$$\phi_B^* = \arg\max_{\phi_B \in [0,\overline{x}]} \pi_B(\phi_B; \phi_A^*) \tag{22}$$

Equation (21) and (22) correspond to best response function of provider A and B.Even without solving the equations, some qualitative results are straightforward. The next theorem shows why information sources are heterogeneous in quality even they have the same technology and access to markets.

Theorem 2 (Signal Differentiation) *If $t > 0$, $\phi_A^* \neq \phi_B^*$.*

Proof. If $\phi_A^* = \phi_B^*$, then the result is like in the Bertrand model, and each firm earns zero profits. By deviating from the level, each firm can earn positive profits through gaining some market power. Therefore, $\phi_A^* = \phi_B^*$ is not an equilibrium. ■

Theorem 2 demonstrates the effect of learning costs heterogeneity. When investors have different learning costs, ex-ante homogeneous firms enter non-price competition and produce different signal quality in equilibrium. Quality differentiation arises because it gives sellers local monopolistic power to maximize the profits.

2 Model Analysis

The complexity of solving the equilibrium is from three sources: (1) the functional form of g and G is general; (2) the critical market boundary x^* is an implicit function from nonlinear equation 20; (3) the equilibrium system (21) and (22) is highly nonlinear. To solve this equilibrium, I propose an iterative algorithm as summarized in Algorithm 1.

Algorithm 1 (Iterative Algorithm) *To solve the equilibrium system (21) and (22):*

1. *Propose an initial value of* $\phi^*_{B,0}$*;*

2. *Solve the optimization problem,*

   (a) $\phi^*_{A,n} = \arg\max_{\phi_A \in [0,\bar{x}]} \pi_A(\phi_A; \phi^*_{B,n-1})$

   (b) $\phi^*_{B,n} = \arg\max_{\phi_B \in [0,\bar{x}]} \pi_B(\phi_B; \phi^*_{A,n})$

3. *Repeat step (2) until* $|\phi^*_{A,n} - \phi^*_{A,n-1}|$ *is small enough.*

To make sure that this algorithm converges, some regularity conditions need to be imposed. However, once it converges, the limit is the equilibrium.

2.1 First-best Solution

To study the welfare implication, I need to define the social surplus. As shown in equation (23) The social surplus has two components: value from the investment and consumption, social learning costs.

$$\underbrace{V_A(\phi_A, \phi_B)G(\tilde{x}) + V_B(\phi_A, \phi_B)[1 - G(\tilde{x})]}_{\text{Value from Investment and Consumption}} - \underbrace{\int_0^{\tilde{x}} \frac{t}{2}(s - \phi_1)^2 g(s)ds + \int_{\tilde{x}}^1 \frac{t}{2}(s - \phi_2)^2 g(s)ds}_{\text{Social Learning Costs}} \tag{23}$$

where $\tilde{x} = \frac{1}{2}(\phi_1 + \phi_2) - \frac{1}{t}\frac{v_2 - v_1}{\phi_2 - \phi_1}$[4]. A benevolent social planner will balance the benefits from good information quality with the learning costs incurred. If the investor type is distributed over a continuous interval, the social optimum should be that information providers produce different qualities. Theorem 3 shows that the information quality gap in market equilibrium is overly high.

Theorem 3 (Quality Heterogeneity) *Under some regularity condition,* $|\phi^*_A - \phi^*_B| < |\tilde{\phi}_A - \tilde{\phi}_B|$*.*

[4]Given (ϕ_A, ϕ_B), agent k should be allocated signal A if $V_A(\phi_A, \phi_B) - \frac{t}{2}(x_k - \phi_A)^2 > V_B(\phi_A, \phi_B) - \frac{t}{2}(x_k - \phi_B)^2$, or equivalently, $x_k > \tilde{x}$. P_i is ignored because it is a surplus transfer from investors to sellers and therefore doesn't affect the total surplus.

Information quality differentiation affects relative competence of investor groups as well as the social learning costs. In the toy model, the social welfare losses come from the fact that high-quality information, even accounting for the learning costs, is more valuable to investors compared to low-quality information. The source of welfare losses could be decomposed along the two dimensions, i.e., investment value v.s. social learning costs. Different from the toy model, there is competition between two group of investors. Besides, learning costs are incurred when using low-quality information. This can be interpreted as the efforts to filter useful information from unrelated information. It can also be interpreted as the reputation losses if investors are known for using low-quality information.

In equilibrium, group A's decision to differentiate from B depends on two effects. First, lower ϕ_A generates market power (direct effects). Second, expecting group B has the same incentive, group A is not worried about competition (strategic effects). These two effects depend on the learning costs as well as the distribution of investor types. Next I will present three examples using uniform distribution and Beta distribution with different levels of concentration.

Example 1 (Standard Uniform Distribution) *$g(x)$ is an uniform distribution on $[0,1]$ and $t = 100$. $\Phi^* = (0,1)$ and $\tilde{\Phi} = (0.3042, 1)$.*

In the standard uniform distribution, market equilibrium generates corner solution. Low-quality information is purely noisy while the high-quality information is perfect. The strategic effect dominates and providers form a collusion to avoid competition as much as possible. In the next two examples, group A has incentive to stop differentiating before reaching $\phi_A = 0$.

Example 2 (Concentrated Uniform Distribution) *$g(x)$ is an uniform distribution on $[0.5\Delta, 1 - 0.5\Delta]$ where $\Delta = 0.6$ and $t = 100$. $\Phi^* = (0.1636, 0.7)$ and $\tilde{\Phi} = (0.435, 0.7)$.*

Example 2 present a non-corner solution where high-quality seller is binding but low-quality seller's product is not pure noise ($\phi_A = 0$). Group B hits the technology restriction. Group A stops differentiating to avoid losing too much market share.

Example 3 (Concentrated Beta Distribution) *$g(x)$ is a transformed Beta distribution with two parameters 4 and 2 on $[0.5\Delta, 1-0.5\Delta]$ where $\Delta = 0.6$ and $t = 100$. $\Phi^{=}(0.2882, 0.6513)$ and $\tilde{\Phi} = [0.5567, 0.7]$.*

Example 3 shows that non-corner solutions exist even the distribution is not symmetric and both groups are choosing interior ϕ_i. When investors' type is concentrated, heterogeneity is lower. Market power is lower and competition is higher because both providers want to capture market shares.

[Place Figure 3 about here]

2.2 Market Power and Welfare Losses

Learning costs heterogeneity gives rise to market power, which in turn causes social welfare losses. To measure the market power, I follow White (2012) and use Lerner index. Lerner index is defined as the ratio of markup (gaps between prices and marginal costs) to the price. Market power gives rise to the non-positive markup. This measure fits the duopoly setting employed in this paper. Using the results collected, the Lerner index of B is $LI_B = \left[\frac{cg(x^*)}{t(1-G(x^*))(\phi_B^* - \phi_A^*)} + 1\right]^{-1}$, where x^* is as in equation (20).

I focuses on two parameters: t and standard deviation of $G(x)$. Given two x_i and x_k, the difference in learning costs increases as t increases. On the aggregate level, the more disperse $G(x)$, the distribution of investors type is more stretched and a random set of the two investors has a higher likelihood to be far apart.

I limit the analysis to the uniform distribution on $[0.5\Delta, 1 - 0.5\Delta]$ for two reasons. First, the functional form of g and G are simple enough and the market boundary x^* is $\frac{1}{3} + \frac{1}{6}(\phi_A + \phi_B) - \frac{1}{3t}\frac{V_B - V_A}{\phi_B - \phi_A}$, which is not dependent on Δ. Second, Δ only changes the dispersion of distribution. The symmetry and mean are not affected. When $\Delta = 0$, it is the standard uniform distribution. The expectation and mode are both 0.5. The standard deviation is $\frac{1-\Delta}{2\sqrt{3}}$, skewness is zero, and kurtosis is $\frac{(1-\Delta)^4}{80}$.

What values of t are reasonable? Choosing different sources incurs different investment revenues. The gap is $V_B - V_A$, and the learning costs gap is $t(\phi_2 - \phi_1)[\frac{1}{2}(\phi_1 + \phi_2) - x]$. These two values should be compared in magnitudes. The term $\frac{1}{t}\frac{V_B - V_A}{\phi_B - \phi_A}$ should be around 1. Depending on (ϕ_A, ϕ_B), the value of $\frac{V_B - V_A}{\phi_B - \phi_A}$ can range from 9 to 200. I choose t to be 10, 20, 50, 100 and change Δ from 0.1 to 0.8 to show the Lerner index for two information providers.

[Place Figure 4 about here]

3 Securities Market Implications

Theorem 3 has two implications: (1) $\tilde{\phi}_B \leq \phi_B^*$; (2) $\tilde{\phi}_A \geq \phi_A^*$. In my framework, the two implications will generate different impacts on the security market outcomes. This section is dedicated to show how the security market behaves as ϕ_A or ϕ_B changes.

3.1 The Dynamics of Wealth

An agent's wealth is the net worth of her portfolio. It can be calculated as the sum of discounted future consumption. This is because the current wealth supports all future consumption plans. Wealth is the net worth of the portfolio. The dynamics of the wealth is a sufficient statistics of portfolio choice, and therefore I look at the dynamics of wealth to analyze investors' optimal choices of portfolio and risk exposure, i.e., how much risk they are willing to assume.

Using SDFs, the wealth of group i can be expressed as:

$$W_t^A(\delta_t, \eta_t, \hat{f}_t^B, g_t) = E_t^A\left[\int_t^{+\infty} \frac{\xi_s^A}{\xi_t^A} c_s^A ds\right] \tag{24}$$

$$W_t^B(\delta_t, \eta_t, \hat{f}_t^B, g_t) = E_t^B\left[\int_t^{+\infty} \frac{\xi_s^B}{\xi_t^B} c_s^B ds\right] \tag{25}$$

The wealth is a function of t, δ_t, η_t, $\hat{f}_t^B$, and g_t. The diffusion vectors decompose the change of wealth into the gradient vectors times the diffusion matrix of state variables, as shown in equation (26).

$$\frac{\text{diff } X_t}{X_t} = \frac{1}{X_t} \underbrace{\begin{bmatrix} \frac{\partial X_t^B}{\partial \delta_t} & \frac{\partial X_t^B}{\partial \eta_t} & \frac{\partial X_t^B}{\partial \hat{f}_t^B} & \frac{\partial X_t^B}{\partial g_t} \end{bmatrix}}_{\nabla X_t^\intercal} \underbrace{\begin{bmatrix} \sigma\delta_t & 0 & 0 \\ \frac{\eta_t g_t}{\sigma} & 0 & 0 \\ \frac{\gamma_B}{\sigma} & 0 & \frac{\sigma_f \phi_B}{\sigma_s} \\ \frac{\gamma_A - \gamma_B}{\sigma} & \frac{\sigma_f \phi_A}{\sigma_S} & -\frac{\sigma_f \phi_B}{\sigma_S} \end{bmatrix}}_{\Sigma_t^\intercal} \tag{26}$$

I focus on the exposure of wealth to disagreement, $\frac{\partial W_t^A}{\partial g_t}\frac{1}{W_t^A}$ and $\frac{\partial W_t^B}{\partial g_t}\frac{1}{W_t^B}$. This measure reflects investors' willingness to make portfolio decisions based on the disagreement. If the value is high, the portfolio choice tilts wealth so the disagreement volatility has high influence. If the value is low, it means that the portfolio choices make the net wealth insensitive to the disagreement and investors are not willing to take the disagreement risk.

[Place Figure 5 about here]

Figure 5 provides illustration on the mechanism how information quality determines the role of disagreement in portfolio choice and wealth. The top row shows how investors' wealth exposure to disagreement risk changes with information quality. Since both $\frac{\partial W_t^A}{\partial g_t}\frac{1}{W_t^A}$ and $\frac{\partial W_t^B}{\partial g_t}\frac{1}{W_t^B}$ are negative, I focus on the absolute value. There are two basic patterns: (1) both group A and B increases the exposure to disagreement; (2) group A's increase has larger magnitude than group B. The intuition is that as the low-quality information improves, group A puts larger weight on their signals to form portfolios and increases the exposure to disagreement. Higher exposure to disagreement reflects higher magnitude of speculative trading. From Figure 1, given $\phi_B = 0.9$, volatility of disagreement decreases with ϕ_A. Speculative trade can be negatively correlated with disagreement volatility. This happens when the low-quality information improves and its users are more active in trading on their signals.

The bottom row shows that when the high-quality information improves, both groups reduce the exposure to disagreement and the magnitude of reduction is more salient for group B. Group B knows their signal is of better quality and the volatility of disagreement is higher, they will reduce the exposure to disagreement because they don't want the volatility of wealth to be overly high, which is displayed in Figure 6.

[Place Figure 6 about here]

The top row in Figure 6 shows the instantaneous volatility of $\frac{dW_t^A}{W_t^A}$ and $\frac{dW_t^B}{W_t^B}$. In terms of magnitude, wealth B is higher than wealth A, which is the opposite in Figure 5. The improving low-quality information increases the volatility of wealth A more than wealth B. Part of the reason is that group A increases the exposure to disagreement and takes more speculative trades.

The bottom row in Figure 6 demonstrates that while group B reduces the exposure to disagreement volatility, the total volatility of wealth increases significantly. Part of the reasons is that disagreement has a higher level of volatility when the gap in information quality of two sources enlarges.

3.2 Security Valuation Volatility

In the previous section, I show how the information quality affects investors' exposure to disagreement and wealth volatility. But it remains a question how the information quality affects the prices of risks. Theorem 1 gives the expression of the stochastic discount factors for both groups. The value of risks under subjective beliefs is obtained. The equity, whose value is denoted as F_t, pays the aggregate dividend δ_t perpetually. The bond, whose price is denoted as $Q(t,T)$, pays a coupon that is equal to one unit of consumption goods at time T.

$$F_t(\delta_t, \eta_t, \hat{f}_t^B, g_t) = E_t^B\left[\int_t^{+\infty} \frac{\xi_s^B}{\xi_t^B}\delta_s ds\right] \tag{27}$$
$$Q_t(T, \delta_t, \eta_t, \hat{f}_t^B, g_t) = E_t^B\left[\frac{\xi_T^B}{\xi_t^B}\right] \tag{28}$$

Proposition 3 *Security valuation*

$$F_t(\delta_t, \eta_t, \hat{f}_t^B, g_t) = \omega(\eta_t)^\alpha \delta_t^\alpha \sum_{n=0}^{\alpha} \binom{\alpha}{n} \int_t^{+\infty} E_t^B[\delta_s^{1-\alpha}\eta_s^{\frac{n}{\alpha}}]\exp(-\rho(s-t))ds \tag{29}$$
$$Q_t(T, \delta_t, \eta_t, \hat{f}_t^B, g_t) = \omega(\eta_t)^\alpha \delta_t^\alpha \sum_{n=0}^{\alpha} \binom{\alpha}{n} E_t^B[\delta_s^{-\alpha}\eta_s^{\frac{n}{\alpha}}]\exp(-\rho(T-t)) \tag{30}$$

The exposures of security prices to different variables have different meanings. I focus on $\frac{\partial F_t}{\partial g_t}\frac{1}{F_t}$ and $\frac{\partial P_t}{\partial g_t}\frac{1}{P_t}$, which are informative of speculative parts in prices. The top two panels in Figure 7 demonstrate the exposure of stock prices to disagreement. The left one shows that given high-quality information, the stock price exposure to disagreement increases in the absolute value as the low-quality information improves. Consistent with the behaviors of low-quality information users, they trade more on their signals. The right one shows that given low-quality information, the stock price exposure to disagreement decreases in the absolute value as the high-quality information improves. In this case, the disagreement volatility increases and high-quality users reduce their exposure to disagreement, which contributes to the lower exposure of stock prices to disagreement. The bottom row of Figure 7 present the exposure of a bond maturing in one year to disagreement. The basic pattern is the same as stocks but the magnitude is much smaller.

[**Place Figure 7 about here**]

The next is to look at the volatility of security prices and their correlation, as shown in Figure 8. The top row shows that as the low-quality information improves, the security price volatility increases. The second row shows that as the high-quality information, the security price volatility increases as well. One thing worth noting is that the volatility of disagreement decreases in the first scenario but increases in the second scenario. The relationship between the disagreement volatility and the security price volatility is ambiguous. The bottom row is the correlation between stock and bond price change, i.e., $corr(\frac{dF_t}{F_t}, \frac{dP_t}{P_t})$. The correlation greatly increases as the quality improves for low-quality information.

[**Place Figure 8 about here**]

3.3 Portfolio Choices

How do the investors form portfolios? In a complete market, the implementation is doable but the composition depends on the security pool to choose from. I want to construct the portfolio in a specific scenario to generate implications of trading volume.

There are three Brownian motions that agents care about, Z_t^B, s_t^A, and s_t^B, four securities that are linearly independent are required to complete financial markets and implement the equilibrium. The choice of securities is arbitrary. I assume that there is a riskless, instantaneous bank deposit with a rate of interest r. The second security, equity or total wealth F_t, pays the aggregate dividend δ_t perpetually. The third security is a bond, $Q(t, T)$, paying a coupon that is equal to one unit of consumption goods at time T. For simplicity, I assume that there is another security whose price dynamic $E(t)$ has a unit loading of s_t^A. An example is that low-quality signal is an index and the unit loading security is an ETF. In reality, ETFs can also be shorted.

I construct the portfolio using diffusion vectors of wealth and security prices. Suppose $(\theta_t^B)^\intercal = [\theta_{t,1}^B, \theta_{t,2}^B, \theta_{t,3}^B]$, which is the weight invested on stock, bond, and unit-loading security respectively. $(\theta_t^B)^\intercal$ replicates the uncertainty exposure of wealth of group B. The linear equation is:

$$\begin{bmatrix} \frac{\partial W_t^B}{\partial \delta_t}\frac{1}{\partial W_t^B} & \frac{\partial W_t^B}{\partial \eta_t}\frac{1}{\partial W_t^B} & \frac{\partial W_t^B}{\partial \hat{f}_t^B}\frac{1}{\partial W_t^B} & \frac{\partial W_t^B}{\partial g_t}\frac{1}{\partial W_t^B} \end{bmatrix} \begin{bmatrix} \sigma\delta_t & 0 & 0 \\ \frac{\eta_t g_t}{\sigma} & 0 & 0 \\ \frac{\gamma_B}{\sigma} & 0 & \frac{\sigma_f \phi_B}{\sigma_s} \\ \frac{\gamma_A-\gamma_B}{\sigma} & \frac{\sigma_f\phi_A}{\sigma_S} & -\frac{\sigma_f\phi_B}{\sigma_S} \end{bmatrix}$$

$$= (\theta_t^B)^\intercal \begin{bmatrix} \frac{\partial F_t}{\partial \delta_t}\frac{1}{\partial F_t} & \frac{\partial F_t}{\partial \eta_t}\frac{1}{\partial F_t} & \frac{\partial F_t}{\partial \hat{f}_t^B}\frac{1}{\partial F_t} & \frac{\partial F_t}{\partial g_t}\frac{1}{\partial F_t} \\ \frac{\partial P_t}{\partial \delta_t}\frac{1}{\partial P_t} & \frac{\partial P_t}{\partial \eta_t}\frac{1}{\partial P_t} & \frac{\partial P_t}{\partial \hat{f}_t^B}\frac{1}{\partial P_t} & \frac{\partial P_t}{\partial g_t}\frac{1}{\partial P_t} \\ \frac{\partial E_t}{\partial \delta_t}\frac{1}{\partial E_t} & \frac{\partial E_t}{\partial \eta_t}\frac{1}{\partial E_t} & \frac{\partial E_t}{\partial \hat{f}_t^B}\frac{1}{\partial E_t} & \frac{\partial E_t}{\partial g_t}\frac{1}{\partial E_t} \end{bmatrix} \begin{bmatrix} \sigma\delta_t & 0 & 0 \\ \frac{\eta_t g_t}{\sigma} & 0 & 0 \\ \frac{\gamma_B}{\sigma} & 0 & \frac{\sigma_f \phi_B}{\sigma_s} \\ \frac{\gamma_A-\gamma_B}{\sigma} & \frac{\sigma_f\phi_A}{\sigma_S} & -\frac{\sigma_f\phi_B}{\sigma_S} \end{bmatrix} \tag{31}$$

where the abbreviation is $\nabla(W_t^B)^\intercal\Sigma_t^\intercal = (\theta_t^B)^\intercal\nabla[F_t, P_t, E_t]^\intercal\Sigma_t^\intercal$.[5] The arguments apply to portfolio choices of both groups. Theorem 4 introduces the calculation of portfolios.

Proposition 4 *The dynamic of portfolio θ_t^A and θ_t^B is*

$$\theta_t^A = (\Sigma_t\nabla[F_t, P_t, E_t])^{-1}(\Sigma_t\nabla W_t^A) \tag{32}$$

$$\theta_t^B = (\Sigma_t\nabla[F_t, P_t, E_t])^{-1}(\Sigma_t\nabla W_t^B) \tag{33}$$

The top row is the demand for stocks from two groups when the low-quality information improves. The y-axis is the weight of wealth invested in stocks. Without short constraint, the demand of stock dramatically increases for group A compared to B, who has a smaller increase. This shows that when group A has better information, they hold more stocks. The bottom row are the demand for stock given the high-quality information is improved. The change of demand depends on whether the disagreement is positive or negative.

[Place Figure 9 about here]

[5]The common usage of notation ∇ is gradients multivariate functions. To keep the compactness of notation, I use it for change rates.

Bonds provide a safer asset. Figure 10 shows the demand for bonds from two groups when the quality of information changes. The top row, which presents the scenarios when the low-quality information improves, demonstrates that at least for group B, the demand for bond deceases. The bottom row, which presents the scenarios when the high-quality information improves, demonstrates the the demand for bonds increases from both groups.

[Place Figure 10 about here]

4 Social Optimum v.s. Market Equilibrium

The previous section presents the potential security market implications when the information quality is different. In this section, I study the difference in the social optimum and market equilibrium. Table 1 summarizes the parameters' values that need to be calibrated. What remains to be calibrated include (1) the distribution $G(x)$; (2) the learning cost parameter t. For illustrating purpose, I use Example 2 where $\Phi^* = (0.1636, 0.7)$ and $\tilde{\Phi} = (0.435, 0.7)$.

4.1 Security Market Patterns

I present the volatility of stock and bond prices under the market equilibrium and social optimum in Figure 11. In the market equilibrium, the volatility of stock and bond prices are lower than in the social optimum. This result might feel counter-intuitive. In the social optimum the information quality gap is smaller. The volatility of disagreement is lower (0.0214 versus 0.0217 in the market equilibrium). From the traditional wisdom, disagreement is the source of speculative trades. Less disagreement means fewer trades and therefore lower volatility. The factor in my setting that makes differences is that both informative trading and speculative trading increase. If the low-quality information improves, under the users beliefs, their estimation error on long-run growth rate is lower and they will put more weights on the external signals. This increases both the informative trades and noisy trades because signals bring in both the information part and noise part. Investors cannot fully separate them. Therefore, on average, two groups of investors have closer expectations on the unobserved long-run growth rate and the disagreement is less volatile. But more noise trades push up the security price volatility. This gives important policy implications that excess volatility should be carefully measured. Disagreement, speculation, and excess volatility do not always appear together. It is not always the social welfare improving to dampen market volatility.

[Place Figure 11 about here]

4.2 Survival of Investors

The subjective beliefs affect the security market outcomes and portfolio choices. However, the relative wealth is under objective beliefs. In this section, I study the evolution of the share of the total dividend that will be consumed by group A or B. The probability distribution of this share is computed under the objective, or true, probability measure rather than under the measure of either Group A or B. I use Monte-Carlo simulation to show how the consumption shares evolve over time.

[Place Figure 12 about here]

Figure 12 produces three observations. First, in the long run, high-quality information users dominate in terms of their consumption share. Second, using the current calibrated value, it takes more than a century for high-quality information users to increase 10 percent in the total consumption, which is consistent withDumas et al. (2009). Finally, in social optimum, the speed at which high-quality information users' consumption increases is lower than in the market equilibrium.

5 Conclusion

In this paper, I present a dynamic equilibrium model with two groups of investors and an information market. Investors are heterogeneous in their learning costs and the heterogeneity causes information providers to differentiate. The market equilibrium is not socially optimal and information quality gap is too large. I show that once the information quality gap is endogeneized investors agree to disagree and the disagreement volatility could be bigger or smaller. I also show that in some scenarios, stock prices volatility does not pair with disagreement volatility, which is against the traditional wisdom.

References

Admati, A. R. and P. Pfleiderer (1986). A monopolistic market for information. *Journal of Economic Theory 39*(2), 400–438.

Admati, A. R. and P. Pfleiderer (1988). Selling and trading on information in financial markets. *The American Economic Review 78*(2), 96–103.

Admati, A. R. and P. Pfleiderer (1990). Direct and indirect sale of information. *Econometrica: Journal of the Econometric Society*, 901–928.

Buraschi, A. and A. Jiltsov (2006). Model uncertainty and option markets with heterogeneous beliefs. *The Journal of Finance 61*(6), 2841–2897.

Chang, Y.-C., P.-J. Hsiao, A. Ljungqvist, and K. Tseng (2022). Testing disagreement models. *The Journal of Finance 77*(4), 2239–2285.

Cookson, J. A. and M. Niessner (2020). Why don't we agree? evidence from a social network of investors. *The Journal of Finance 75*(1), 173–228.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (2005). Investor psychology and security market under-and overreaction. *Advances in Behavioral Finance, Volume II*, 460–501.

Dumas, B., A. Kurshev, and R. Uppal (2009). Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. *The Journal of Finance 64*(2), 579–629.

Ehling, P., M. Gallmeyer, C. Heyerdahl-Larsen, and P. Illeditsch (2018). Disagreement about inflation and the yield curve. *Journal of Financial Economics 127*(3), 459–484.

Ehling, P., A. Graniero, and C. Heyerdahl-Larsen (2018). Asset prices and portfolio choice with learning from experience. *The Review of Economic Studies 85*(3), 1752–1780.

Gentzkow, M. and J. M. Shapiro (2006). Media bias and reputation. *Journal of political Economy 114*(2), 280–316.

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

Hotelling, H. (1929). Stability in competition. *The Economic Journal 39*(153), 41–57.

Huang, S., Y. Xiong, and L. Yang (2022). Skill acquisition and data sales. *Management Science 68*(8), 6116–6144.

Illeditsch, P., H. Kung, and C. Heyerdahl-Larsen (2022). Growth through diversity. *Working Paper 231*.

Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal of Political Economy 111*(6), 1183–1220.

Veronesi, P. (2000). How does information quality affect stock returns? *The Journal of Finance 55*(2), 807–837.

Verrecchia, R. E. (1982). Information acquisition in a noisy rational expectations economy. *Econometrica: Journal of the Econometric Society*, 1415–1430.

Wang, J. (1993). A model of intertemporal asset prices under asymmetric information. *The Review of Economic Studies 60*(2), 249–282.

White, L. J. (2012). Market power: How does it arise? how is it measured?

Xiong, W. and H. Yan (2010). Heterogeneous expectations and bond markets. *The Review of Financial Studies 23*(4), 1433–1466.

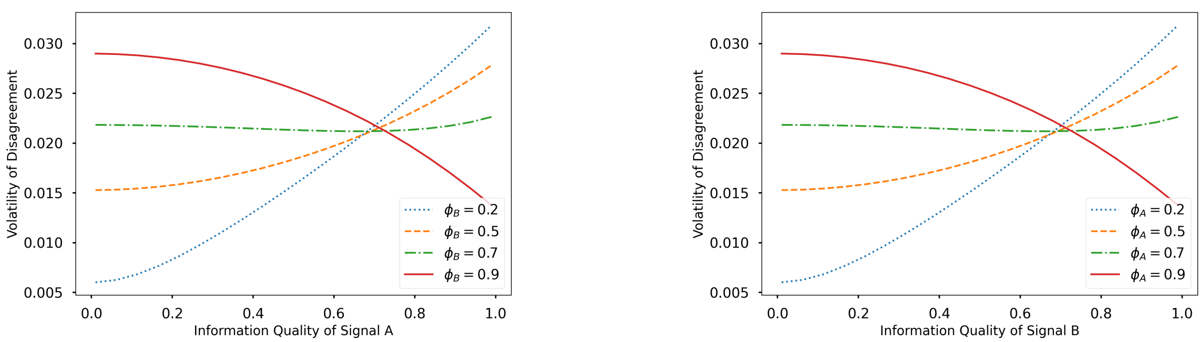
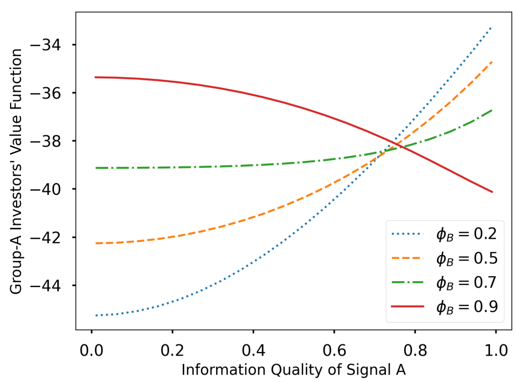
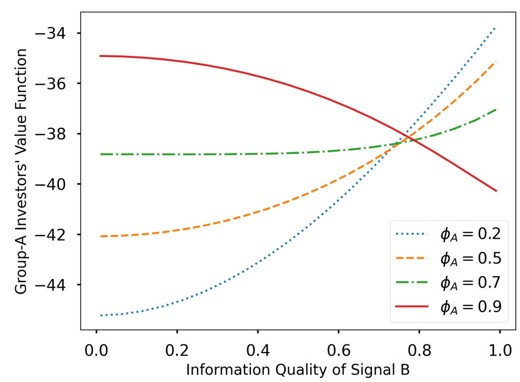


Figure 1: **Volatility of Disagreement**

The figure shows how instantaneous disagreement volatility changes as the information quality changes. The instantaneous volatility is $\frac{(\gamma_A - \gamma_B)^2}{\sigma^2} + \sigma_f^2(\phi_A^2 + \phi_B^2 - 2\phi_A^2\phi_B^2)$. The left panel reports the relationship between volatility and ϕ_A, which is the quality level of low-quality information. The right panel reports the relationship between volatility and ϕ_B, which is the quality level of high-quality information.



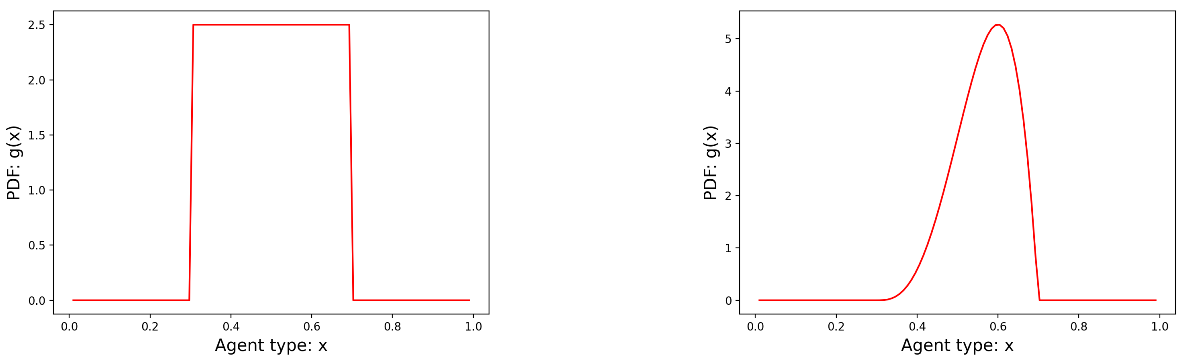
(a)



(b)

Figure 2: **Group-A Lifetime Utility Function**

The figure shows how lifetime discounted utility of Group-A changes as the information quality changes. The left panel reports the relationship between lifetime discounted utility and ϕ_A, which is the quality level of low-quality information. The right panel reports the relationship between lifetime discounted utility and ϕ_B, which is the quality level of high-quality information.

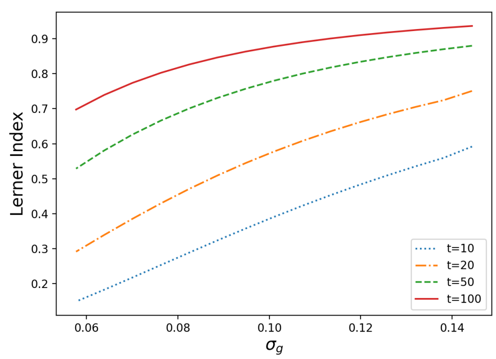


(a) Example 2

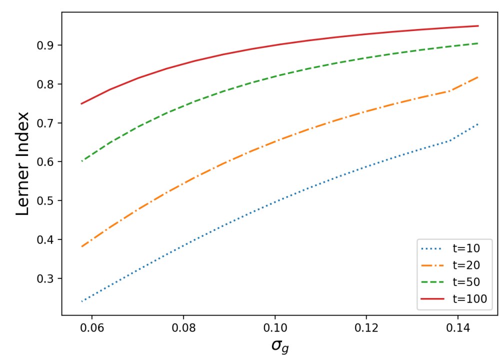
(b) Example 3

Figure 3: **Probability Density Function in Example 2 and 3**

The figure displays the probability density functions in example 2 and 3. The left panel shows the density function of a probability measure uniformly distributed on $[0.3, 0.7]$. The right panel shows the density function of a beta distribution concentrated on $[0.3, 0.7]$.



(a) Information Provider A



(b) Information Provider B

Figure 4: **Lerner Index**

The figure displays the relationship between Lerner index of two information providers and the standard deviation of investors type distribution. The left panel is the Lerner index of information provider A. The right panel is the Lerner index of information provider B. The investor type is uniformly distributed over $[0.5\Delta, 1 - 0.5\Delta]$ and $\sigma_g = \frac{1-\Delta}{2\sqrt{3}}$.

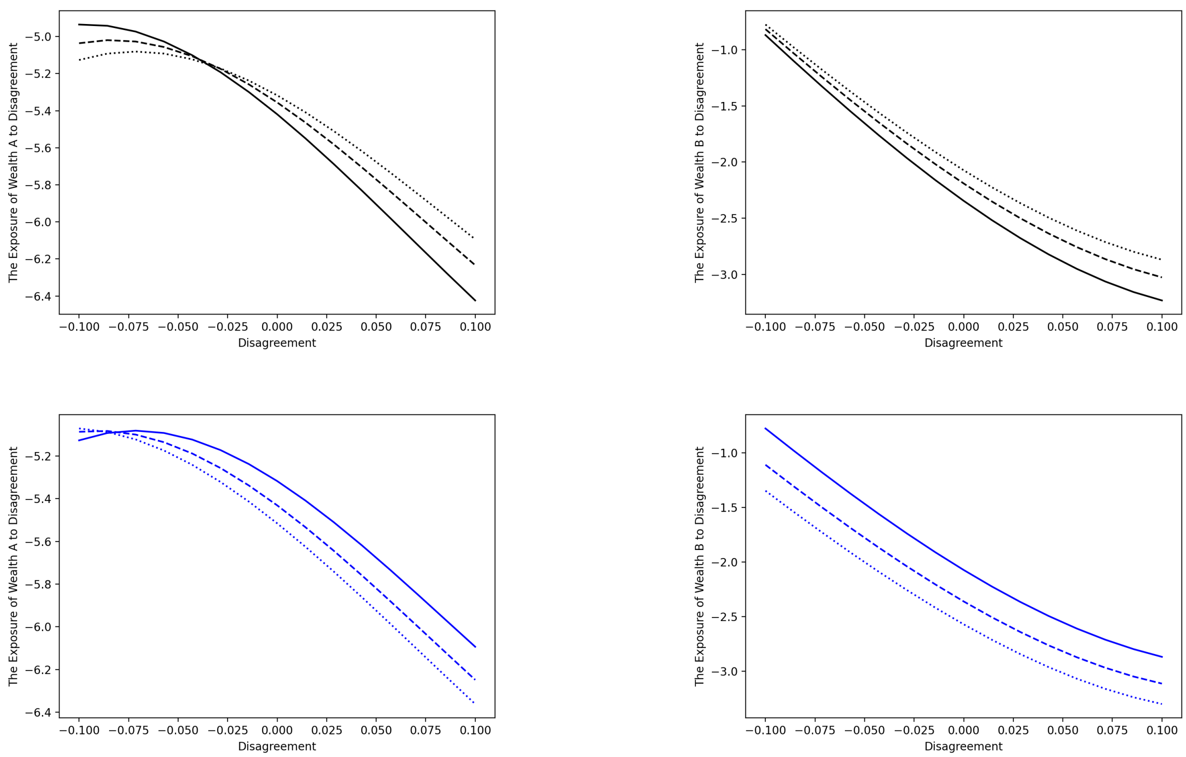


Figure 5: **Wealth Exposure to Disagreement**

This figure displays the exposure of investors' wealth to disagreement. The top row reports the relationship between the exposure of wealth to disagreement and disagreement as the low-quality information improves. The bottom row reports the same relationship as the high-quality information improves. Different values of (ϕ_A, ϕ_B) correspond to different line styles with black and blue. $(0.2, 0.9)$ is the black dotted line and blue solid line. $(0.5, 0.9)$ is the black dashed line. $(0.7, 0.9)$ is the black solid line. $(0.2, 0.5)$ is the blue dotted line. $(0.2, 0.7)$ is the blue dashed line.

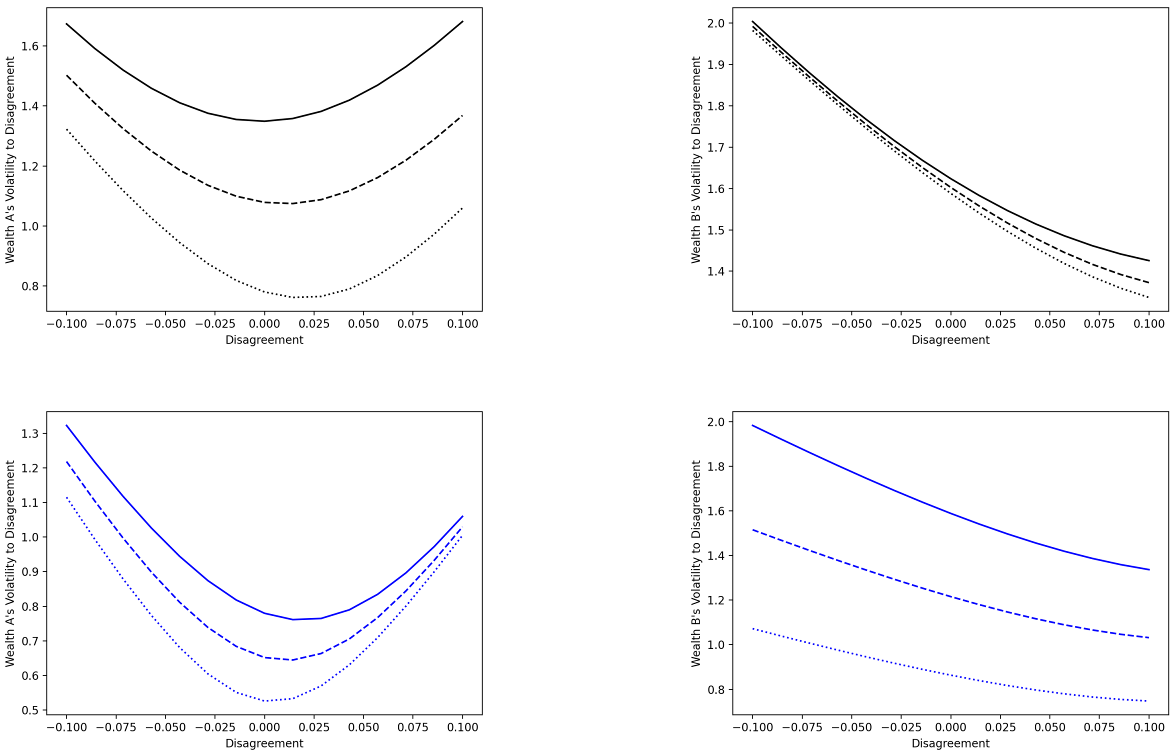


Figure 6: **Wealth Volatility**

This figure displays the volatility of investors' wealth. The top row reports the relationship between wealth volatility and disagreement as the low-quality information improves. The bottom row reports the same relationship as the high-quality information improves. Different values of (ϕ_A, ϕ_B) correspond to different line styles with black and blue. $(0.2, 0.9)$ is the black dotted line and blue solid line. $(0.5, 0.9)$ is the black dashed line. $(0.7, 0.9)$ is the black solid line. $(0.2, 0.5)$ is the blue dotted line. $(0.2, 0.7)$ is the blue dashed line.

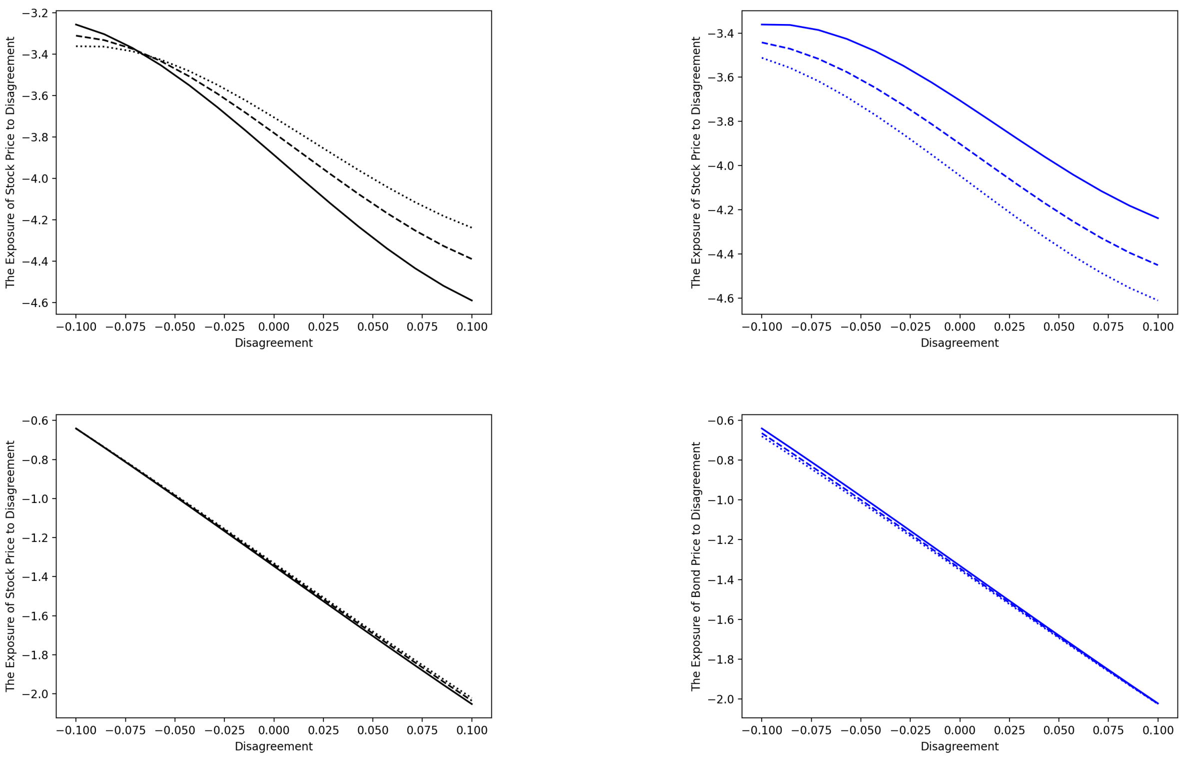


Figure 7: **Stock/Bond Price Exposure to Disagreement**

This figure displays the exposure of stock or bond prices to disagreement. The top row reports the relationship between stock price exposure and disagreement. The bottom row reports the same relationship between bond price exposure and disagreement. Different values of (ϕ_A, ϕ_B) correspond to different line styles with black and blue. $(0.2, 0.9)$ is the black dotted line and blue solid line. $(0.5, 0.9)$ is the black dashed line. $(0.7, 0.9)$ is the black solid line. $(0.2, 0.5)$ is the blue dotted line. $(0.2, 0.7)$ is the blue dashed line.

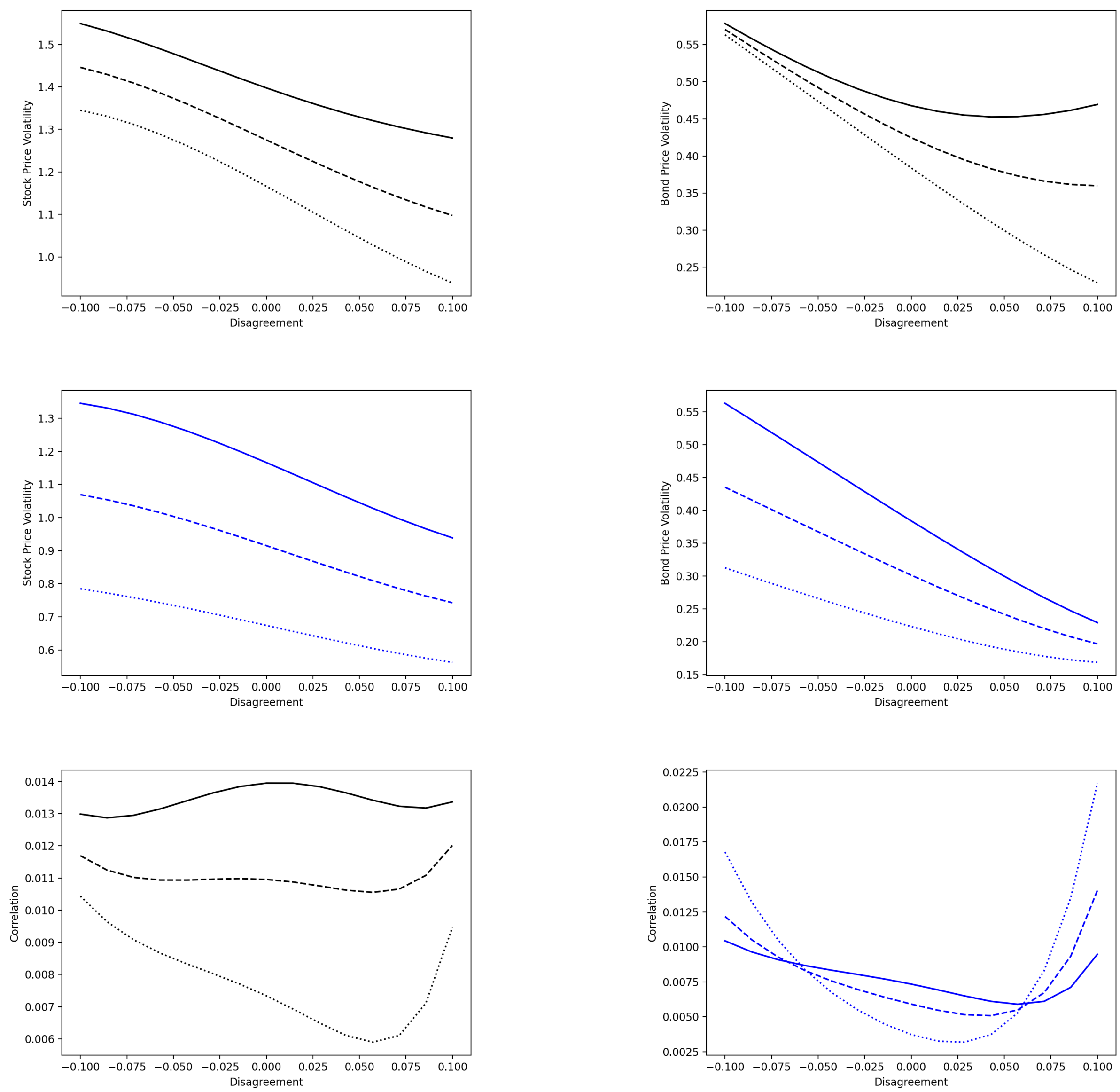


Figure 8: **Stock/Bond Price Volatility and Correlation**

This figure displays the volatility of stock prices and bond prices and their correlation. The top row reports the relationship between stock price volatility and disagreement. The middle row reports the relationship between bond price volatility and disagreement. The bottom row reports the correlation between stock prices and bond prices. Different values of (ϕ_A, ϕ_B) correspond to different line styles with black and blue. $(0.2, 0.9)$ is the black dotted line and blue solid line. $(0.5, 0.9)$ is the black dashed line. $(0.7, 0.9)$ is the black solid line. $(0.2, 0.5)$ is the blue dotted line. $(0.2, 0.7)$ is the blue dashed line.

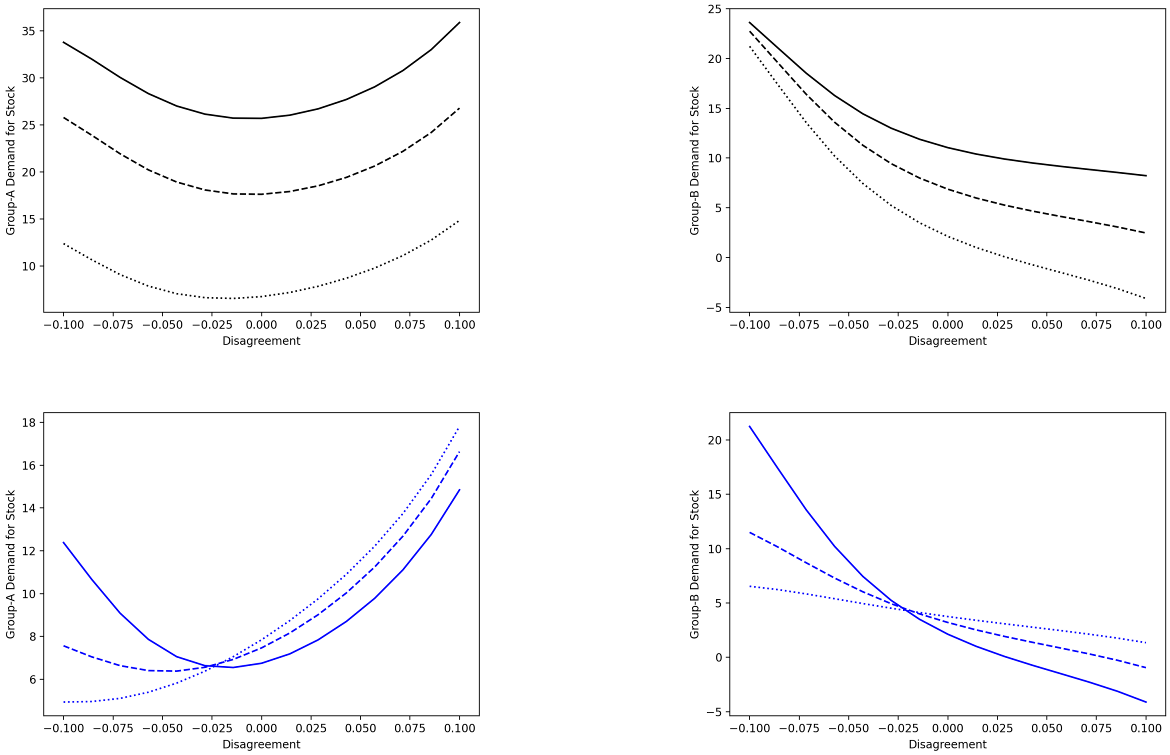


Figure 9: **Demand For Stocks**

This figure displays the weight of stocks in investors' portfolio. The top row reports the relationship between portfolio weight on stock and disagreement as the low-quality information improves. The bottom row reports the same relationship as the high-quality information improves. Different values of (ϕ_A, ϕ_B) correspond to different line styles with black and blue. $(0.2, 0.9)$ is the black dotted line and blue solid line. $(0.5, 0.9)$ is the black dashed line. $(0.7, 0.9)$ is the black solid line. $(0.2, 0.5)$ is the blue dotted line. $(0.2, 0.7)$ is the blue dashed line.

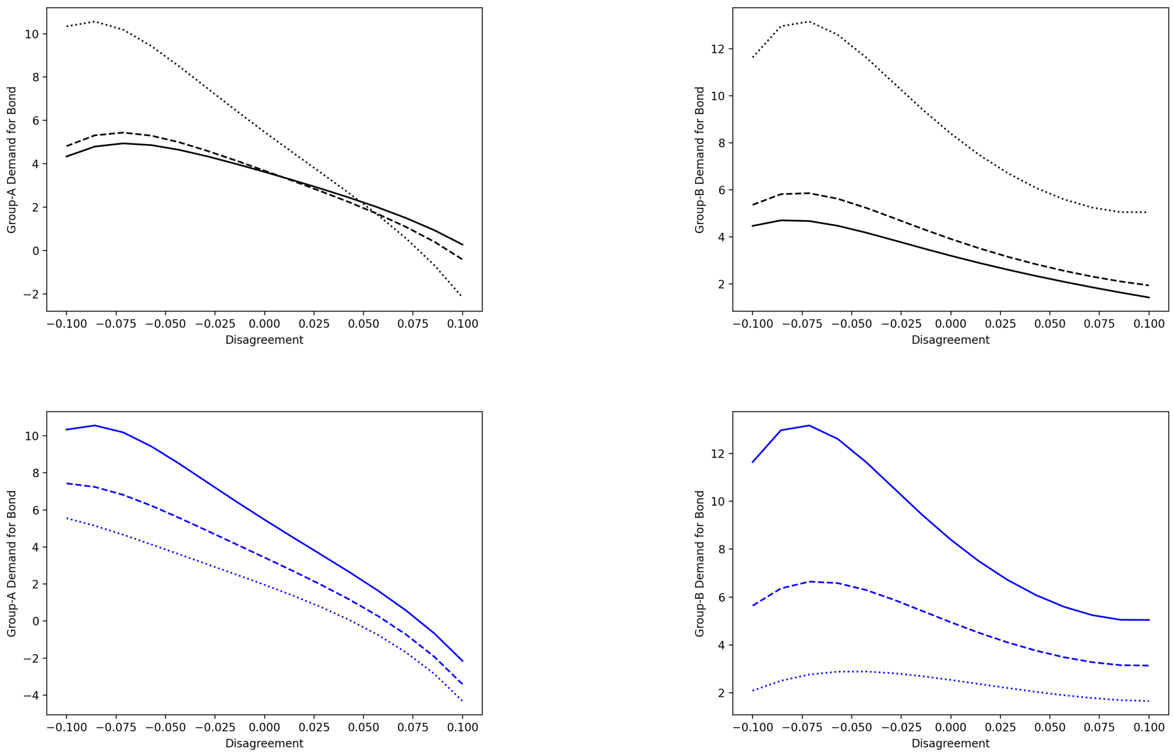


Figure 10: **Demand For Bonds**

This figure displays the weight of bonds in investors' portfolio. The top row reports the relationship between portfolio weight on bond and disagreement as the low-quality information improves. The bottom row reports the same relationship as the high-quality information improves. Different values of (ϕ_A, ϕ_B) correspond to different line styles with black and blue. $(0.2, 0.9)$ is the black dotted line and blue solid line. $(0.5, 0.9)$ is the black dashed line. $(0.7, 0.9)$ is the black solid line. $(0.2, 0.5)$ is the blue dotted line. $(0.2, 0.7)$ is the blue dashed line.

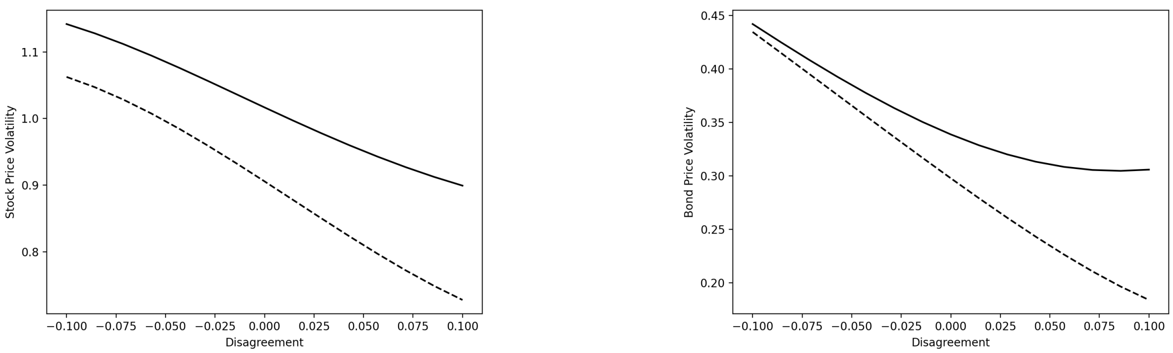


Figure 11: **Stock/Bond Price Volatility**

This figure displays stock and bond price volatility in social optimum $((0.435, 0.7))$ and market equilibrium $((0.1636, 0.7))$. The solid line is the case of social optimum. The dashed line is the case of market equilibrium. The left panel reports the relationship between stock price volatility and disagreement. The right panel reports the relationship between bond price volatility and disagreement.

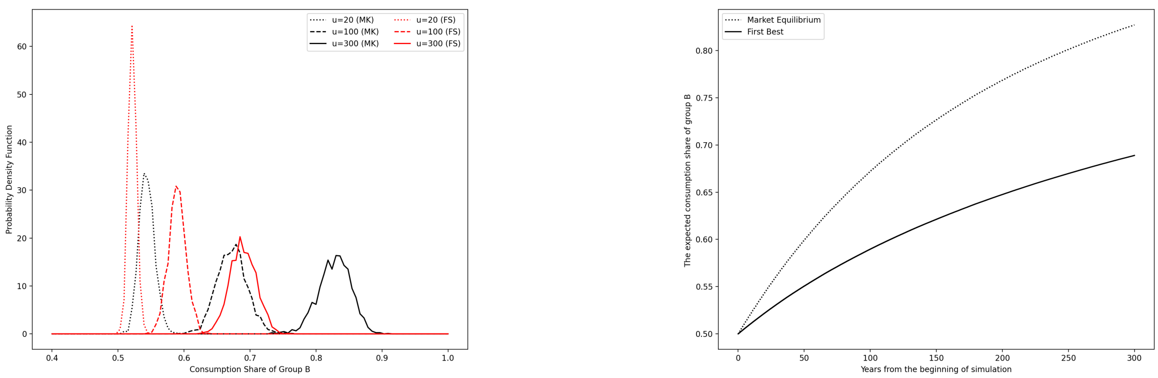


Figure 12: **Survival of Group B**

This figure displays the dynamics of group B's consumption share. The left panel reports the probability density function of group B's consumption share $\omega(\eta)$ at different time points. The dotted line is 20 years since the beginning of simulation. The dashed line is 100 years. The solid line is 300 years. The red line is in the social optimum. The black line is in the market equilibrium. The right panel reports the expected consumption share of group B at different time points. The Monte-Carlo simulation starts from the consumption share equals to 0.5, $\eta_0 = 1$, $\delta_0 = 1$, and $f_0 = 0.015$. Other parameters values are listed in Table 1.

A Calibration Parameters

Table 1: **Choice of Parameter Values**

Name	Symbol	Value
Parameters for aggregate endowment		
Long-term average growth rate of aggregate endowment	$\overline{f}$	0.015
Volatility of expected growth rate of endowment	σ_f	0.03
Volatility of aggregate endowment	σ	0.13
Mean reversion parameter	η	0.2
Parameters for the agents		
Relative Lagrange multiplier $\frac{\lambda_A}{\lambda_B}$	λ_A/λ_B	1
Time preference parameter for both agents	ρ	0.10
Relative risk aversion for both agents	$1-\alpha$	3

B Proof of Theorem 1

The proof is the use of Girsanov Theorem. Define $\eta_t := \frac{dQ^A}{dQ^B}$ where Q^A and Q^B are risk-neutral probability under the subjective measure of group A and B. Then we have:

$$\eta_t = \exp\left\{\int_0^t \frac{\hat{g}_t}{\sigma} d\hat{Z}_t^B - \frac{1}{2}\int_0^t \left(\frac{\hat{g}_s}{\sigma}\right)^2 ds\right\} \tag{34}$$

or equivalently by Ito's lemma we have $\frac{d\eta_t}{\eta_t} = \frac{\hat{g}_t}{\sigma} d\hat{Z}_t^B$

C Proof of Proposition 2

Given that $\lambda_A = \lambda_B$ and $\alpha - 1 \in \mathrm{N}$, binomial expansion can decompose the instantaneous utility function as:

$$u(c_t^A) = \frac{1}{1-\alpha} \sum_{n=0}^{\alpha-1} \binom{\alpha-1}{n} (\delta_t)^{-(\alpha-1)} (\eta_t)^{-\frac{n}{\alpha}} \tag{35}$$

And based on Fubini Theorem, the lifetime utility function can be expressed as:

$$\begin{aligned} V_A(\phi_A, \phi_B) &= E^i \int_0^\infty e^{-\rho t} \frac{(c_t^i)^{1-\alpha}}{1-\alpha} dt \\ &= \frac{1}{1-\alpha} \sum_{n=0}^{\alpha-1} \binom{\alpha-1}{n} \int_0^\infty e^{-\rho t} E^A \left[(\delta_t)^{-(\alpha-1)} (\eta_t)^{-\frac{n}{\alpha}}\right] dt \end{aligned}$$

Similarly,

$$V_B(\phi_A, \phi_B) = \frac{1}{1-\alpha} \sum_{n=0}^{\alpha-1} \binom{\alpha-1}{n} \int_0^\infty e^{-\rho t} E^B (\delta_t^{-(\alpha-1)} \eta_t^{\frac{n}{\alpha}}) dt$$

Notice that using Radon-Nikodym derivative, $E^A\left[(\delta_t)^{-(\alpha-1)}(\eta_t)^{-\frac{n}{\alpha}}\right] = E^B\left[(\delta_t)^{-(\alpha-1)}(\eta_t)^{1-\frac{n}{\alpha}}\right]$.

D Proof of Lemma 1

The moment-generating functions of the four variables are critical for asset pricing. The goal is to calculate $E_t[\delta_T^{\varepsilon} \eta_T^{\chi}]$, which is a joint inverse fourier tranformation. It would be better to start with simple cases where only one variable is involved. Let's do δ_t first. $E_t[\delta_T^{\varepsilon}] = E_t[\exp(\varepsilon ln(\delta_T)]$. If δ_t follows $d\delta_t/\delta_t = \mu dt + \sigma dZ_t$, we can take out the term δ_t^{ε} and rearrange it: $\delta_t^{\varepsilon} E_t[\varepsilon(ln(\delta_T) - ln(\delta_t)]$, but $dln(\delta_t) = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dZ_t$, which can be recovered explicitly. So the calculation becomes very straightforward. This trick doesn't work in our case because the drift term is also random. In that case, we can't write $ln(\delta_T) - ln(\delta_t)$ explicitly.

Another way to do this is to generalize it. Notice that $E_t[\delta_T^\varepsilon]$ is a martingale. And $dln(\delta_t) = (\hat{f}_t^B - 0.5 * \sigma^2)dt + \sigma d\hat{Z}_t^B$. So $ln(\delta_T) - ln(\delta_t)$ has a drift term that depends on f_t's future path and a Brownian motion diffusion term. This drift term should be dependent on f_t. Therefore, we put a conjecture that $E_t[\delta_T^\varepsilon] = \delta_t^\varepsilon \overline{h}(t, f_t)$. $\overline{h}(t, f_t)$ is the expectation of drift terms. Brownian motion is independent of now. Diffusion process has strong markov property. That is why $E_t[ln(\delta_T) - ln(\delta_t)]$ is only a function of f_t and T.

Since $E_t[\delta_T^\varepsilon]$ is a martingale, the coefficient of the drift term in $dE_t[\delta_T^\varepsilon]$ will be zero. If we input the conjecture of the functional form into the SDE and set the coefficient of drift term to zero, we will have after the term δ^ε is cancelled on both sides:

$$\frac{\partial \overline{h}}{\partial t} + \varepsilon \hat{f}_t^B \overline{h} - \zeta(\hat{f}_t^B - \overline{f})\frac{\partial \overline{h}}{\partial f} + \frac{1}{2}\varepsilon(\varepsilon-1)\sigma^2 \overline{h} + \frac{1}{2}(\gamma_B^2/\sigma^2 + \sigma_f^2\phi_B^2)\frac{\partial^2 \overline{h}}{\partial f^2} + \varepsilon\gamma_B\frac{\partial \overline{h}}{\partial f} = 0$$

with the terminal condition $h(T) = \delta_T^\varepsilon$. We impose another conjecture that $\overline{h}(t, f) = exp(A_0(t) + A_1(t)f + A_2(t)f^2)$ and replace $\overline{h}$ in the equation above with this conjected form. Then we can get three oridinary differential equations:

$$A_2'(t) - 2\zeta A_2(t) + 2(\gamma_B^2/\sigma^2 + \sigma_f^2\phi_B^2)A_2^2(t) = 0 \quad (36)$$

$$A_1'(t) + \varepsilon - \zeta A_1(t) + 2\zeta\overline{f}A_2(t) + 2(\gamma_B^2/\sigma^2 + \sigma_f^2\phi_B^2)A_1(t)A_2(t)f + 2\varepsilon\sigma\gamma_B A_2(t) = 0 \quad (37)$$

$$A_0'(t) + \zeta\overline{f}A_1(t) + \frac{1}{2}\varepsilon(\varepsilon-1)\sigma^2 + \frac{1}{2}(\gamma_B^2/\sigma^2 + \sigma_f^2\phi_B^2)[2A_2(t) + A_1(t)^2] + \varepsilon\gamma_B A_1(t) = 0 \quad (38)$$

with the terminal conditions $A_0(T) = A_1(T) = A_2(T) = 0$. The solutions are:

$$A_2(t) = 0 \quad (39)$$

$$A_1(t) = \frac{\varepsilon}{\zeta}[1 - exp(-\zeta(T-t))] \quad (40)$$

$$A_0(t) = \left[(\zeta\overline{f} + \varepsilon\gamma_B)\frac{\varepsilon}{\zeta} + \frac{1}{2}(\gamma_B^2/\sigma^2 + \sigma_f^2\phi_B^2)\left(\frac{\varepsilon}{\zeta}\right)^2 + \frac{1}{2}\varepsilon(\varepsilon-1)\sigma^2\right](T-t) \quad (41)$$

$$- \left[(\zeta\overline{f} + \varepsilon\gamma_B)\frac{\varepsilon}{\zeta} + (\gamma_B^2/\sigma^2 + \sigma_f^2\phi_B^2)\left(\frac{\varepsilon}{\zeta}\right)^2\right]\frac{1}{\zeta}(1 - \exp(-\zeta(T-t))) \quad (42)$$

$$+ \frac{1}{2}(\gamma_B^2/\sigma^2 + \sigma_f^2\phi_B^2)\left(\frac{\varepsilon}{\zeta}\right)^2 * \frac{1}{2\zeta}(1 - \exp(-2\zeta(T-t))) \quad (43)$$

Following the same step, we can find that $E_t[\eta_T^\chi] = \eta_T^\chi$, which means that η_t^χ is a martingale. If δ_t and η_t are independent, then the joint MGF is the product of their own MGF. However, the common factor $d\hat{Z}_t^B$ creates correlation. This correlation will become a drift term when we investigate the dynamic of $\delta^\varepsilon\eta^\chi$. Suppose $E_t[\delta_T^\varepsilon\eta_T^\chi] = H(t, \delta_t, \hat{f}_t^B, \eta_t, g_t)$, since H is a martingale, then coefficient of the drift term should be zero. Notice that once we put them together, the interaction between δ_t and η_t or δ_t and g_t will appear in the PDE. This could have resulted huge challenges. But the tricky part is that their diffusion terms have constant coefficients. This help us to separate δ_t and $\hat{f}_t^B$ from η_t and g_t. Suppose $H(t,\delta,f,\eta,g) = \delta^\varepsilon \exp(A_0(t) + A_1(t)f)\eta^\varepsilon \overline{\overline{h}}(t,g)$, then the terms related to δ and f can be cancelled because $\delta^\varepsilon \exp(A_0(t) + A_1(t)f)$ follows the PDE only with respect to δ_t. Take out $\delta^\varepsilon \exp(A_0(t) + A_1(t)f)$:

$$\frac{\partial\overline{\overline{h}}}{\partial t} - (\zeta + \frac{\gamma_A}{\sigma^2})g\frac{\partial\overline{\overline{h}}}{\partial g} + \frac{1}{2}\chi(\chi-1)\frac{g^2}{\sigma^2}\overline{\overline{h}} \tag{44}$$

$$+ \frac{1}{2}\left[\frac{(\gamma_A-\gamma_B)^2}{\sigma^2} + \sigma_f^2(\phi_A^2 + \phi_B^2 - 2\phi_A^2\phi_B^2)\right]\frac{\partial^2\overline{\overline{h}}}{\partial g^2} + \varepsilon(\gamma_A - \gamma_B)\frac{\partial\overline{\overline{h}}}{\partial g} \tag{45}$$

$$+ \varepsilon\chi g\overline{\overline{h}} + \chi A_1(t)\frac{\gamma_B}{\sigma^2}g\overline{\overline{h}} - A_1(t)\left[\frac{\gamma_B(\gamma_B-\gamma_A)}{\sigma^2} + \sigma_f^2\phi_B^2(1-\phi_A^2)\right]\frac{\partial\overline{\overline{h}}}{\partial g} + \chi\frac{(\gamma_A-\gamma_B)}{\sigma^2}g\frac{\partial\overline{\overline{h}}}{\partial g} = 0 \tag{46}$$

Suppose $\overline{\overline{h}}(t,g) = \exp(B_0(t) + B_1(t)g + B_2(t)g^2)$, replace $\overline{\overline{h}}$ in the equation above and we can generate three ODEs:

$$B_2'(t) - 2\left[(\zeta + \frac{\gamma_A}{\sigma^2}) + \frac{\gamma_B-\gamma_A}{\sigma^2}\chi\right]B_2(t) + 2\left[\frac{(\gamma_A-\gamma_B)^2}{\sigma^2} + \sigma_f^2(\phi_A^2+\phi_B^2-2\phi_A^2\phi_B^2)\right]B_2^2(t) \tag{47}$$

$$+ \frac{1}{2}\chi(\chi-1)\frac{1}{\sigma^2} = 0 \tag{48}$$

$$B_1'(t) - (\zeta + \frac{\gamma_A}{\sigma^2})B_1(t) + 2\left[\frac{(\gamma_A-\gamma_B)^2}{\sigma^2} + \sigma_f^2(\phi_A^2+\phi_B^2-2\phi_A^2\phi_B^2)\right]B_2(t)B_1(t) \tag{49}$$

$$- 2\varepsilon(\gamma_B-\gamma_A)B_2(t) + \varepsilon\chi + \chi\frac{\gamma_B}{\sigma^2}A_1(t) - 2\left[\frac{\gamma_B(\gamma_B-\gamma_A)}{\sigma^2} + \sigma_f^2\phi_B^2(1-\phi_A^2)\right]A_1(t)B_2(t) \tag{50}$$

$$- \frac{\chi(\gamma_B-\gamma_A)}{\sigma^2}B_1(t) = 0 \tag{51}$$

$$B_0'(t) + \frac{1}{2}\left[\frac{(\gamma_A-\gamma_B)^2}{\sigma^2} + \sigma_f^2(\phi_A^2+\phi_B^2-2\phi_A^2\phi_B^2)\right][2B_2(t) + B_1^2(t)] - \varepsilon(\gamma_B-\gamma_A)B_1(t) \tag{52}$$

$$- A_1(t)\left[\frac{\gamma_B(\gamma_B-\gamma_A)}{\sigma^2} + \sigma_f^2\phi_B^2(1-\phi_A^2)\right]B_1(t) = 0 \tag{53}$$

with the terminal conditions $B_2(T) = B_1(T) = B_0(T) = 0$.

The first ode is Raccati equation. Let's define $\kappa_1 = \frac{(\gamma_A-\gamma_B)^2}{\sigma^2} + \sigma_f^2(\phi_A^2+\phi_B^2-2\phi_A^2\phi_B^2)$, $\kappa_2 = \chi\frac{\gamma_B-\gamma_A}{\sigma^2} + (\zeta + \frac{\gamma_A}{\sigma^2})$, $\kappa_3 = \frac{\gamma_B(\gamma_B-\gamma_A)}{\sigma^2} + \sigma_f^2\phi_B^2(1-\phi_A^2)$, $q = \sqrt{\kappa_2^2 - \frac{\chi(\chi-1)}{\sigma^2}\kappa_1}$.

$$B_2(t) = \frac{\chi(\chi-1)}{2\sigma^2}\frac{1-\exp(-2q(T-t))}{\kappa_2 + q - (\kappa_2 - q)\exp(-2q(T-t))} \tag{54}$$

The second ode has a structure of $y' = p(t)y + q(t)$ where the analytic solution exists: $y(t) = (C_0 + \int \exp(q(t)\int -p(t)dt)dt)\exp(\int p(t)dt)$. If we organize and collect the terms in the second ode:

$$p(t) : (\zeta + \frac{\gamma_A}{\sigma^2}) + \frac{\chi(\gamma_B - \gamma_A)}{\sigma^2} - 2\kappa_1 B_2(t) \tag{55}$$

$$q(t) : 2\varepsilon(\gamma_B - \gamma_A)B_2(t) - \varepsilon\chi - \chi\frac{\gamma_B}{\sigma^2}A_1(t) + 2\kappa_3 A_1(t)B_2(t) \tag{56}$$

$$\tag{57}$$

A beautiful intermediate result is $exp(\int -p(t)dt) = \exp(q(T-t))[(\kappa_2+q) - (\kappa_2 - q)\exp(-2q(T-t))]$. This will greatly simplify the expression of the result.

$$B_1(t) = \frac{\sum_{i=1}^{5} \nu_i \exp(-v_i(T-t))}{\kappa_2 + q - (\kappa_2 - q)\exp(-2q(T-t))} \tag{58}$$

where $\nu_1 = -\frac{1}{q}[\varepsilon(\gamma_B - \gamma_A)\frac{\chi(\chi-1)}{\sigma^2} - \varepsilon\chi(\kappa_2+q) - \frac{\chi\gamma_B\varepsilon}{\zeta\sigma^2}(\kappa_2+q) + \kappa_3\frac{\varepsilon\chi(\chi-1)}{\zeta\sigma^2}]$, $\nu_2 = \frac{1}{q}[-\varepsilon(\gamma_B - \gamma_A)\frac{\chi(\chi-1)}{\sigma^2} + \varepsilon\chi(\kappa_2 - q) + \frac{\chi\gamma_B\varepsilon}{\zeta\sigma^2}(\kappa_2 - q) - \kappa_3\frac{\varepsilon\chi(\chi-1)}{\zeta\sigma^2}]$, $\nu_3 = \frac{-1}{q-\zeta}[\frac{\chi\gamma_B\varepsilon}{\zeta\sigma^2}(\kappa_2+q) - \kappa_3\frac{\varepsilon\chi(\chi-1)}{\zeta\sigma^2}]$, and $\nu_4 = \frac{1}{q+\zeta}[-\frac{\chi\gamma_B\epsilon}{\zeta\sigma^2}(\kappa_2 - q) + \kappa_3\frac{\varepsilon\chi(\chi-1)}{\zeta\sigma^2}]$, $\nu_5 = -\sum_{i=1}^4 \nu_i$, $v_1 = 0$, $v_2 = 2q$, $v_3 = \zeta$, $v_4 = 2q+\zeta$, $v_5 = q$.

The third ODE is straightforward to solve but the challenge is that it has so many terms. The algebric operation is very vulnerable to errors. We will decompose these terms separately.

$$B_0(t) = -\kappa_1 \int B_2(t)dt - \frac{1}{2}\kappa_1 \int B_1^2(t)dt + \varepsilon(\gamma_B - \gamma_A)\int B_1(t)dt + \kappa_3 \int A_1(t)B_1(t)dt \tag{59}$$

These terms contain integration of one common function:

$$D_1(p,t) = \int \frac{\exp(-p(T-t))}{\kappa_2 + q - (\kappa_2 - q)\exp(-2q(T-t))}dt \tag{60}$$

$$D_2(p,t) = \int \frac{\exp(-p(T-t))}{[\kappa_2 + q - (\kappa_2 - q)\exp(-2q(T-t))]^2}dt \tag{61}$$

The calculation of $D_1(p,t)$, we want to calculate a useful formula to simplify our calculation: $\int \frac{x^a}{(x+\theta)^b}dx = \frac{x^{a+1}}{\theta^b(a+1)}\mathcal{H}(b, a+1, a+2, -\frac{x}{\theta})$. The proof is as follows:

$$\begin{aligned}
\int \frac{x^a}{(x+\theta)^b}dx &= \int x^a \sum_{n=0}^{\infty} \theta^{-b-n}(-1)^n (b)_n \frac{x^n}{n!}dx \\
&= \sum_{n=0}^{\infty} \theta^{-b-n}(b)_n \frac{(-1)^n}{n+a+1}\frac{x^{n+a+1}}{n!} \\
&= \frac{x^{a+1}}{\theta^b(a+1)}\sum(-1)^n(b)_n \frac{a+1}{n+a+1}\frac{(x/\theta)^n}{n!} \\
&= \frac{x^{a+1}}{\theta^b(a+1)}\sum(-1)^n(b)_n \frac{(a+1)_n}{(a+2)_n}\frac{(x/\theta)^n}{n!} \\
&= \frac{x^{a+1}}{\theta^b(a+1)}\mathcal{H}(b, a+1, a+2, -\frac{x}{\theta}) \qquad (62)
\end{aligned}$$

If $p > 0$, then we have

$$\int \frac{\exp(-p(T-t))}{\kappa_2 + q - (\kappa_2 - q)\exp(-2q(T-t))}dt \qquad (63)$$

$$= -\frac{1}{2q(\kappa_2 - q)}\int_t^T \frac{x^{\frac{p}{2q}-1}}{x-\theta}dx \qquad (64)$$

$$= -\frac{1}{2q(\kappa_2 - q)}\frac{2qx^{\frac{p}{2q}}}{(-\theta)p}\mathcal{H}(1, \frac{p}{2q}, \frac{p}{2q}+1, \frac{x}{\theta}) + C \qquad (65)$$

$$= \frac{\exp(-p(T-t))}{p(\kappa_2+q)}\mathcal{H}(1, \frac{p}{2q}, \frac{p}{2q}+1, \frac{\kappa_2 - q}{\kappa_2 + q}\exp(-2q(T-t))) + C \qquad (66)$$

The second step is to use $x = \exp(-2q(T-t))$ to change the variable. $D_1(p,t) = -\frac{1}{p(\kappa_2+q)}\left[\mathcal{H}(1, \frac{p}{2q}, \frac{p}{2q}+1, \frac{\kappa_2-q}{\kappa_2+q}) - \exp(-p(T-t))\mathcal{H}(1, \frac{p}{2q}, \frac{p}{2q}+1, \frac{\kappa_2-q}{\kappa_2+q}\exp(-2q(T-t)))\right]$

If $p = 0$, then $D_1(p,t) = -\frac{1}{2q(\kappa_2+q)}[2q(T-t) - ln(2q) + ln((\kappa_2+q) - (\kappa_2-q)\exp(-2q(T-t)))]$.

To calculate $D_2(p,t)$, transform it into an expression of $D_1(p,t)$ could save the distinction between the case $p = 0$ and $p > 0$.

$$
\begin{aligned}
D_2(p,t) &= \int \frac{\exp(-p(T-t))}{[\kappa_2+q-(\kappa_2-q)\exp(-2q(T-t))]^2}dt && (67)\\
&= \frac{1}{2q(\kappa_2-q)^2}\int \frac{x^{\frac{p}{2q}-1}}{(x-\theta)^2}dx && (68)\\
&= \frac{1}{2q(\kappa_2-q)^2}\int \frac{1}{\theta}\frac{\theta-x+x}{(x-\theta)^2}x^{\frac{p}{2q}-1}dx && (69)\\
&= \frac{1}{2q(\kappa_2-q)^2}\left[-\frac{1}{\theta}\int \frac{x^{\frac{p}{2q}-1}}{x-\theta}dx + \int_t^T \frac{1}{\theta}\frac{x^{\frac{p}{2q}}}{(x-\theta)^2}dx\right] && (70)\\
&= \frac{1}{2q(\kappa_2-q)(\kappa_2+q)}\left[-\int \frac{x^{\frac{p}{2q}-1}}{x-\theta}dx + \int \frac{x^{\frac{p}{2q}}}{(x-\theta)^2}dx\right] && (71)\\
&= \frac{1}{2q(\kappa_2-q)(\kappa_2+q)}\left[-\int \frac{x^{\frac{p}{2q}-1}}{x-\theta}dx - \int x^{\frac{p}{2q}}d\frac{1}{x-\theta}\right] && (72)\\
&= \frac{1}{2q(\kappa_2-q)(\kappa_2+q)}\left[-\int \frac{x^{\frac{p}{2q}-1}}{x-\theta}dx - \frac{x^{\frac{p}{2q}}}{x-\theta} + \frac{p}{2q}\int \frac{x^{\frac{p}{2q}-1}}{x-\theta}dx\right] && (73)\\
&= \frac{1}{(\kappa_2+q)}(1-\frac{p}{2q})D_1(p,t) + \frac{1}{2q(\kappa_2+q)}\frac{\exp(-p(T-t))}{\kappa_2+q-(\kappa_2-q)\exp(-2q(T-t))} + C && (74)\\
&= \frac{1}{2q(\kappa_2+q)}\left[(2q-p)D_1(p,t) + \frac{\exp(-p(T-t))}{\kappa_2+q-(\kappa_2-q)\exp(-2q(T-t))} - \frac{1}{2q}\right] && (75)
\end{aligned}
$$

Another way to make this representation neat is to re-parameterize the equation

The idea is to chop the representation into a combination of homogeneous functions, like in this case, there are expressed as $\exp(-p(T-t))$ with different ps. This idea has been adopted for computational simplicity like polynomial ststorage, etc. With this representation, the calculation of B_0 is also highly simplified.

$$
\begin{aligned}
B_0(t) = &-\kappa_1\frac{\chi(\chi-1)}{2\sigma^2}D_1(0,t) + \kappa_1\frac{\chi(\chi-1)}{2\sigma^2}D_1(2q,t) && (76)\\
&-\frac{\kappa_1}{2}\sum_{i=1}^{5}\sum_{j=1}^{5}\nu_i\nu_j D_2(v_i+v_j,t) && (77)\\
&+[\varepsilon(\gamma_B-\gamma_A)+\kappa_3\frac{\varepsilon}{\zeta}]\sum_{i=1}^{5}\nu_i D_1(v_i,t) && (78)\\
&-\kappa_3\frac{\varepsilon}{\zeta}\sum_{i=1}^{5}\nu_i D_1(v_i+\zeta,t) && (79)
\end{aligned}
$$

E Proof of Theorem 3

I need Assumption 2 is to provide an analytical proof.

Assumption 2 *$G(x)$ is symmetric: $g(x) = g(1-x)$.*

In the market equilibrium, the quality choice can be constructed as a Cournot model:

$$\phi_1 = \arg\max_{\phi_1} \frac{G(x^*)^2}{g(x^*)} t(\phi_2 - \phi_1) \tag{80}$$

$$\phi_2 = \arg\max_{\phi_2} \frac{(1-G(x^*))^2}{g(x^*)} t(\phi_2 - \phi_1) \tag{81}$$

where $P_1 = c + \frac{G(x^*)}{g(x^*)} t(\phi_2 - \phi_1)$ and $P_2 = c + \frac{1-G(x^*)}{g(x^*)} t(\phi_2 - \phi_1)$ and $x^* = \frac{1}{2}(\phi_1 + \phi_2) + \frac{1}{t}\frac{P_2 - P_1)}{\phi_2 - \phi_1} - \frac{1}{t}\frac{v_2 - v_1}{\phi_2 - \phi_1}$.

The social optimum is to maximize

$$vG(\tilde{x}) + v_2[1 - G(\tilde{x})] - \int_0^{\tilde{x}} \frac{t}{2}(s - \phi_1)^2 g(s) ds - \int_{\tilde{x}}^1 \frac{t}{2}(s - \phi_2)^2 ds \tag{82}$$

where $\tilde{x} = \frac{1}{2}(\phi_1 + \phi_2) - \frac{1}{t}\frac{v_2 - v_1}{\phi_2 - \phi_1}$. To get the analytic solution, we start from the simplest case. Suppose $\frac{v_2 - v_1}{\phi_2 - \phi_1} = 0$, the question degenerates to:

$$\int_0^{\tilde{x}} (s - \phi_1)^2 g(s) ds + \int_{\tilde{x}}^1 (s - \phi_2)^2 g(s) ds \tag{83}$$

where $\tilde{x} = \frac{1}{2}(\phi_1 + \phi_2)$.

Frist, let's make assumption that the distribution $G(x)$ is symmetric: $g(x) = g(1-x)$. The implication is that $E(x) = 0.5$ ($\int_0^1 sg(s)ds = \int_0^{0.5} sg(s)ds + \int_{0.5}^1 sg(s)ds = \int_0^{0.5} sg(s)ds + \int_0^{0.5}(1-s)g(s)ds = G(0.5) = 0.5$. The FOC conditions above are:

$$\int_0^{\tilde{x}} (s - \phi_1) g(s) ds = 0 \tag{84}$$

$$\int_{\tilde{x}}^1 (s - \phi_2) g(s) ds = 0 \tag{85}$$

The optimal solution is $\tilde{\phi}_1 = 2\int_0^{0.5} sg(s)ds$ and $\tilde{\phi}_2 = 2\int_{0.5}^1 sg(s)ds$.

I impose this strong assumption on probability distribution since it is important to get the analytical solution of social optimum. But it is flexible enough to study the effect of different distribution. Our analysis starts from this benchmark.

Another assumption is about the upper limit of ϕ_B. $\phi_B <= \sup\{x \in [0,1] : g(x) > 0\}$. This assumption will limit the differentiation of producer B. But it is reasonable. It means that if no one in the economy can understand the information without learning, it cannot be produced. This limits the information production technology.

Now we use the second-order approximation to solve out the analytic solution of social optimum. Let's denote $K_1(\phi_A, \phi_B) = \frac{1}{t}[v_A(\phi_A, \phi_B)G(\tilde{(x)}) + v_B(\phi_A, \phi_B)(1 - G(\tilde{x}))]$ and $K_2(\phi_A, \phi_B) = \int_0^{\tilde{x}} \frac{1}{2}(s-\phi_A)^2 g(s)ds + \int_{\tilde{x}}^1 \frac{1}{2}(s-\phi_B)^2 g(s)ds$. The objective function of the social optimal solution can be rearranged as $V(\phi_A, \phi_B) = K_1(\phi_A, \phi_B) - K_2(\phi_A, \phi_B)$. Since we know the minimum of K_2, denoted as (ϕ_A^0, ϕ_B^0). Then we have $(\Phi := (\phi_A, \phi_B))$:

$$V(\Phi) = V(\Phi_0) + \frac{1}{t}\nabla H(\Phi_0)^T(\Phi - \Phi_0) + \frac{1}{2}(\Phi - \Phi_0)^T(\frac{1}{t}\nabla^2 K_1(\Phi_0) - \nabla^2 K_2(\Phi_0))(\Phi - \Phi_0) + o(||\Phi - \Phi_0||^2) \tag{86}$$

Lemma 2 provides an approximate analytical solution, which provides some insights about the trading off.

Lemma 2 *The social optimum, $\tilde{\Phi} := (\tilde{\phi}_A, \tilde{\phi}_B)$, is*

$$\tilde{\Phi} = \Phi_0 - \left[\frac{1}{t}\nabla^2 K_1(\Phi_0) - \nabla^2 K_2(\Phi_0)\right]^{-1} \frac{1}{t}\nabla K_1(\Phi_0) \tag{87}$$

where $K_1(\phi_A, \phi_B) = v_A(\phi_A, \phi_B)G(\tilde{x}) + v_B(\phi_A, \phi_B)(1 - G(\tilde{x}))$, $K_2(\phi_A, \phi_B) = \int_0^{\tilde{x}} \frac{1}{2}(s-\phi_A)^2 g(s)ds + \int_{\tilde{x}}^1 \frac{1}{2}(s - \phi_B)^2 g(s)ds$, and $\Phi_0 = (2\int_0^{0.5} sg(s)ds, 2\int_{0.5}^1 sg(s)ds)$

Φ_0 is the minimizer of social learning costs. $\tilde{\Phi}$ deviates from Φ_0 because of the benefits from different information quality levels.

The approximation of the optimal solution is $\Phi^* = \Phi_0 + \left[\frac{1}{t}\nabla^2 K_1(\Phi_0) - \nabla^2 K_2(\Phi_0)\right]^{-1} \frac{1}{t}\nabla K_1(\Phi_0)$. Notice that $\nabla F(\Phi_0) = 0$ and $det(\nabla^2 F(\Phi_0)) < 0$ from our special case.

Now to prove that $|\phi_A^{**} - \phi_B^{**}| > |\phi_A^* - \phi_B^*|$, we needto prove that at the social optimum, at least one producer has incentive to deviate. We start from the simple case: $v_1 = v_2$, which are constant. The derivative of producer 1's profits with respect to ϕ_A is:

$$\frac{\phi_A}{\partial \phi_A} = -\frac{G^2(x^*)}{g(x^*)}t + t(\phi_2 - \phi_1)\left[2G(x^*) - \frac{G^2(x^*)}{g^2(x^*)}g'(x^*)\right]\frac{\partial x^*}{\partial \phi_A} \tag{88}$$

$$= -\frac{1}{4g(x^*)}t + t(\phi_B^* - \phi_A^*)\frac{\partial x^*}{\partial \phi_A} \tag{89}$$

$$\frac{\partial x^*}{\partial \phi_A} = \frac{1}{2} - 2t(\phi_B^* - \phi_A^*) \tag{90}$$

$\frac{\partial x^*}{\partial \phi_A} < 0$ when t is large enough. Therefore, producer A will have incentive to deviate.