ESSAYS ON INFORMATION ACQUISITION AND ASSET PRICING

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

By
Paul Marmora
July 2015

Examining Committee Members:

Dr. Moritz Ritter, Advisory Chair, Economics
Dr. Oleg Rytchkov, Finance
Dr. Charles Swanson, Economics
Dr. Yan Li, External Member, Finance

ABSTRACT

In this dissertation, I explore different mechanisms by which information is generated in financial markets, and whether these mechanisms can account for empirical anomalies that models without information choice have difficulty explaining.

In the first chapter, I survey the theoretical literature on perfectly competitive asset markets, with a particular focus on rational expectations models with endogenous information acquisition.

In the second chapter, "The Distribution of Information, the Market for Financial News, and the Cost of Capital", I present a rational expectations model with a competitive market for financial news that provides an explanation for why stocks with a higher degree of information asymmetry tend to earn higher expected returns. I demonstrate that when a small fraction of investors hold a large fraction of a firm's private information, few investors demand a copy of firm-specific news in equilibrium. As a result, each investor must incur a larger share of the fixed cost of news production to obtain a copy, which deters investors from learning more about the firm and therefore raises their required risk premium. This result hinges crucially on the ability of investors to share in the fixed cost of news production, which suggests that the financial news media plays an important role in determining how the cost of capital varies with the inequality of information across investors.

In the third chapter, "Learning About Noise" (with Oleg Rytchkov), we study theoretical implications of endogenous acquisition of non-fundamental information in financial markets. We develop a rational expectations model with heterogeneous information and multidimensional costly learning and demonstrate that i) investors specialize in information acquisition, that is, those who are endowed with high (low) quality information about fundamentals learn only about fundamentals (noise), ii) learning about fundamentals increases the asymmetry of information, whereas learning about noise decreases it, and iii) the opportunity to learn about noise unambiguously increases price informativeness.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation and gratitude to my advisor, Dr. Moritz Ritter, for his guidance and supervision throughout the dissertation process. His patience and willingness to help greatly contributed to the completion of this work.

I would also like to thank my committee members, Dr. Oleg Rytchkov and Dr. Charles Swanson, for their help and support. Dr. Rytchkov played a pivotal role in improving the theoretical foundations and helping me derive the proofs, while Dr. Swanson gave me useful direction when I experienced conceptual and expositional difficulties during the early stages of this project. I would also like to thank Dr. Diamantaras for his input and Dr. Yan Li for being the external member of my defense.

Finally, I would like to thank my friends and family for their encouragement and patience while I pursued my graduate degree.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

A central question in the study of financial markets is whether asset prices accurately reflect information dispersed among a large number of market participants. In recent years, a vast theoretical literature has used the rational expectations framework to study how well prices aggregate and communicate private information to other traders in competitive markets. In this chapter, I review this literature and briefly discuss the contributions of this work.

One of the first papers to analyze price informativeness within a rational expectations framework is Grossman (1976). Specifically, Grossman (1976) describes a perfectly competitive market for a single risky asset, where each agent is endowed with a mutually independent signal about the risky asset's payoff. After observing his private signal, as well as the asset's price (which can be used to infer other agents' signals), each agent then decides how many shares to purchase. In this framework, Grossman (1976) does find that the asset's price perfectly aggregates and reveals all private information. However, since Grossman (1976) assumes that agents take into account that their private signals and the asset's price are statistically dependent, in equilibrium agents only condition their demand on price, which gives rise to a paradox: if no single agent incorporates private information into their investment decisions, how can prices reveal any private information in the first place?

To address the paradox presented by Grossman (1976), a second strand of papers, beginning with Hellwig (1980) and Diamond and Verrecchia (1981), introduces an additional ingredient: agents are not only uncertain about asset payoffs, but also about the aggregate endowment of asset shares. Randomness in the supply of shares, usually attributed to the presence of noise traders or uncertain liquidity preferences, serves to prevent the asset's price from perfectly revealing all relevant information.[1] In other words, due to this second source of uncertainty, agents cannot discern whether fluctuations in price arise from agents' private information or from randomness in supply. As a result, prices are no longer perfectly revealing, and agents no longer neglect their private signals when forming portfolios, as they do in Grossman (1976). Admati (1985) provides a generalization of these results to the case with multiple risky assets, and finds that an asset's price can decrease in its own payoff or increase in its supply when asset payoffs are correlated.

While much of the seminal discussion is focused on the case where participants' information sets are exogenous, an important step in understanding the workings of financial markets is to recognize that investors have some choice over the quantity and quality of information at their disposal. This distinction is particularly important when one considers that learning choices and investment choices are mutually dependent: investors' portfolio decisions are driven by their beliefs, but beliefs are determined by learning choices, which are themselves a function of portfolio expectations.

Beginning with Grossman and Stiglitz (1980) and Verrecchia (1982), a strand of literature considers endogenous information acquisition within a perfectly competitive market for a risky asset (or risky assets), where asset payoffs and number of outstanding asset shares are initially unknown to agents, as in Hellwig (1980). In these models, before agents form their portfolios, they can reduce their uncertainty about the risky asset by acquiring a signal for some specified cost. In a subsequent period, agents observe their acquired signals and the asset's price and use their updated beliefs to form portfolios. In the final period, payoffs are revealed and agents consume.

Although these models share the same basic framework, their details often differ. For example, in the canonical model of Grossman and Stiglitz (1980), agents choose whether to pay an exogenous cost to acquire the same payoff signal. In contrast, Verrecchia (1982) allows agents to acquire payoff signals that are independent across agents, where the cost of a signal is increasing in its precision. A general result in this literature is that the informativeness of price with respect

[1] While Hellwig (1980) assumes random aggregate supply, Diamond and Verrecchia (1981) assumes that each agent's endowment is random. The two approaches are functionally equivalent as long as the number of agents is finite.

to fundamentals is increasing in the average amount of signal precision possessed by agents. Since agents observe asset prices for free, this implies that learning is a strategic substitute: the value of collecting private information is decreasing in the amount of private information possessed by others. This also gives rise to the well-known Grossman-Stiglitz paradox: if learning is costly, prices cannot fully reveal private information in equilibrium because agents do not have an incentive to pay for information that they can simply infer from prices for free.

Since the pioneering efforts of Grossman and Stiglitz (1980) and Verrecchia (1982), this framework has been extended in several dimensions. One of the more popular extensions is to endogenize the cost of learning by introducing a market for financial news. Admati and Pfleiderer (1986) consider a monopolistic information supplier who can observe private signals before deciding how to sell them. Alternatively, Admati and Pfleiderer (1990) consider the case where, rather than selling information directly, a monopolist creates a portfolio based on his private signals and sells shares of this portfolio to traders. A central message of this literature is that profit-maximizing information suppliers have an incentive to dilute the quality of the signals they observe before selling in order to diminish the leakage of these signals through price.

Rather than assume a monopolistic supplier, Veldkamp (2006) considers a competitive market for financial news by replacing the constant price of information in Grossman and Stiglitz (1980) with an endogenous price characterized by increasing returns to scale in production. Specifically, Veldkamp (2006) considers a setting where each agent can discover the same signal about payoffs and sell the signal to other agents at no marginal cost. Given free entry into the news market and that the market is perfectly contestable, Veldkamp (2006) finds that the news signal is priced at average cost. This serves to counteract the strategic substitutability result of Grossman and Stiglitz (1980): the more agents who purchase news, the lower the price of news and the greater the incentive for other agents to purchase news as well. In the model described in chapter 2, I consider the same competitive information market, except I assume that agents interpret the acquired news signal with idiosyncratic noise, so that agents receive independent payoff signals (as in Hellwig (1980) and Verrecchia (1982)) instead of the same signal (as in Grossman and Stiglitz (1980)).

In addition to Veldkamp (2006), several other mechanisms have been analyzed that overturn the strategic substitutability result of Grossman and Stiglitz (1980).

Barlevy and Veronesi (2000) show that learning can be a strategic complement when the number of outstanding shares is not normally distributed. Ganguli and Yang (2009) find that allowing agents to learn about both payoffs and asset supply can also generate strategic complementarity in information acquisition, as well as multiple equilibria.[2] Finally, Goldstein and Yang (2014) consider multiple fundamentals that affect an asset's value, and find that learning about one fundamental can decrease uncertainty related to trading on information about the other fundamental, thus increasing the incentive to learn.

Instead of introducing a market for information, another strand of research models agents as being " rationally inattentive" in the sense that they are limited in their ability to process available information. The type of learning constraint used in this literature, first introduced in Sims (2006) and motivated by the information-theoretic concept of entropy, restricts how much agents can reduce their uncertainty about the true state of the world. Each agent must then rationally choose the joint distribution between observed signals and the true state. Variations of this entropy-based learning constraint in rational expectations models have been used to explain asset return comovement (Peng and Xiong, 2006), home bias (Van Nieuwerburgh and Veldkamp, 2009; Mondria and Wu, 2010), and income inequality (Kacperczyk et al., 2014). An implicit feature of this learning constraint is that the maximum allowable posterior precision of beliefs is increasing in prior precision. In other words, it is easier to process information about risks one is already well-informed about, a type of increasing returns to learning that is consistent with evidence that concentrated portfolios tend to earn higher returns (Van Nieuwerburgh and Veldkamp, 2010).

The term increasing returns to learning can also refer to a different phenomena: the value of acquiring information is rising in the expected scale of investment. While this effect is present in all asset-pricing models, it is less salient when agents have constant absolute risk aversion (CARA) utility (Van Nieuwerburgh and Veldkamp, 2010). Consequently, several recent papers have departed from the more common CARA preferences used in the literature in order to exploit returns to scale in the value of information. For example, Peress (2004) considers agents with constant relative risk aversion and uses this feature to explain why wealthier households tend to hold riskier assets. Peress (2010) employs the mean-variance preferences

[2] In chapter 3, we also consider a model where agents are allowed to learn about both payoffs and asset supply, but the mechanics and motivation differ.

first used in Epstein and Zin (1989) and finds that increasing a firm's shareholder base reduces each shareholder's incentive to acquire information (an effect that may be mitigated by the mechanism I describe in chapter 2). Van Nieuwerburgh and Veldkamp (2009) and Mondria and Wu (2010) both find that when combining Epstein-Zin preferences with the entropy-based learning constraint described in the previous paragraph, agents choose to learn about, and therefore hold, assets they are already familiar with, a result they then use to explain the persistence of home bias in household portfolios.

Recent studies have also extended the framework of Grossman and Stiglitz (1980) and Hellwig (1980) to examine what role, if any, asymmetric information plays in the determination of prices. Easley and O'Hara (2004) consider a rational expectations model where agents exogenously receive a mix of private signals (i.e, signals mutually independent across agents) and public signals (i.e, signals correlated across a subset of agents). They find that increasing the ratio of private signals to public signals decreases price informativeness and raises the risk premium, thus concluding that asymmetric information raises the cost of capital. However, as noted by Lambert et al. (2012), changing the composition of private signals to public signals in these models affects price only because it alters the *average* quality of signals that traders possess, not because it changes how a given quality is distributed *across* traders.[3] In chapter 2, I build a rational expectations model where differences in information quality across traders can lower price informativeness, but only because it makes an asset more difficult to learn about, thus lowering the average quality of information at the time of investment.

[3]It is important to note that the term asymmetric information here strictly refers to differences in the quality of signals across traders. This stands in contrast to the type of asymmetric information frequently attributed to Merton (1987), where a subset of agents refrain from trading an asset entirely (potentially because they are unaware that the asset exists). Reducing the number of agents who trade in a risky asset only affects the risk premium in Merton (1987) because it raises the number of shares, and therefore the risk, that the average trader expects to hold. Therefore, in Merton (1987) it is still true that only the average amount of risk faced by traders, and not how this risk is distributed, affects the cost of capital.

Chapter 2

THE DISTRIBUTION OF INFORMATION, THE MARKET FOR FINANCIAL NEWS, AND THE COST OF CAPITAL

2.1. Introduction

In this chapter, I develop a three-period noisy rational expectations model with a continuum of risk-averse agents that competitively trade one risky asset. The total quality, or precision, of prior signals about the risky asset's payoff is heterogeneously distributed across the population. Before agents decide how much to invest, they each have an opportunity to purchase an additional news signal about the asset's payoff. I make two assumptions on the form of this news signal. First, agents purchase this piece of news on a competitive information market characterized by increasing returns to scale in production. Second, an agent's ability to interpret the news signal is increasing in the precision of their initial information.

In this setting, I find that a higher inequality in the distribution of initial (prior) precision reduces the average (posterior) precision of agents' beliefs at the time of investment, which increases the asset's expected risk premium. Notably, this result is contingent on both information markets and increasing returns to learning. In fact, when the cost of acquiring the news signal is exogenously fixed, I reach exactly the opposite conclusion: a sufficiently high inequality of initial information actually decreases excess returns. The reason is that when initial information about a firm is highly concentrated, a small fraction of agents expect to hold so many shares that they are willing to acquire news even when the expected return per share is extremely low, which results in a smaller equilibrium risk premium.

These results have several implications. First, they suggest that whether the cost of capital rises or falls with a higher inequality of information in competitive equity markets rests crucially on the extent to which individuals are able to share in the cost of news production. For example, in financial markets with few media outlets, a more equal distribution of information across many investors should be associated with higher costs of capital, as no single investor expects to hold enough shares to pay the entire fixed cost of producing additional information themselves. On the other hand, a more equal distribution of information should be associated with a lower cost of capital in markets with a sophisticated news industry because stocks that many investors expect to hold will be cheaper to acquire news on.

Second, the model has implications for the reporting preferences of the financial press. Indeed, one way to interpret the mechanism driving the main result is that widely-circulated, low-cost news providers are more likely to cover stocks with a more equal distribution of information, i.e., stocks in which many investors each expect to hold a smaller number of shares. This interpretation is consistent with evidence that the mass media is more likely to cover stocks primarily owned by individual investors (e.g., Fang and Peress, 2009; Solomon, 2012). Because a more equal distribution of information is associated with more extensive media coverage, the negative relation between the equality of information and the cost of capital predicted by the model is also consistent with the growing empirical literature linking greater media coverage to higher asset prices (e.g., Huberman and Regev, 2001; Fang and Peress, 2009; Tetlock, 2010, 2011; Engelberg et al., 2012).

This chapter is related to several strands of research. First, it contributes to the recent debate about whether information asymmetry among investors affects the cost of capital in perfectly competitive markets.[1] O'Hara (2003), Easley and O'Hara (2004), and Hughes et al. (2007) conclude that firms with a greater degree

[1] Lambert et al. (2012) shows that asymmetric information can have a separate effect on the cost of capital, but only in models with imperfect competition, like Kyle (1985).

of asymmetric information have higher costs of capital. However, Lambert et al. (2012) point out that these theoretical results are entirely driven by changes to the average precision of information, not by changes to how this precision is distributed across the population. Importantly, this debate is concerned with how an exogenous distribution *directly* affects returns, whereas the distributional effects I document arise *indirectly* via the learning decision. That is, my results imply that asymmetric information increases the risk premium, but only because it lowers the total amount of additional information produced, thereby lowering average posterior precision. The end result is that firms with a less equal distribution of precision experience a higher cost of capital, even though, consistent with Lambert et al. (2012), the risk premium is only affected by the average posterior precision of beliefs.

The learning decision that agents face blends two different approaches used in the literature. In the first, agents learn by acquiring additional noisy signals about future payoffs subject to an exogenous cost (Grossman and Stiglitz, 1980; Verrecchia, 1982). In my model, the precision of an agent's acquired signal is increasing in the precision of his prior beliefs, a type of increasing returns to learning that is also present in the learning constraints motivated by the information-theoretic concept of entropy, such as those used in Peng and Xiong (2006), Van Nieuwerburgh and Veldkamp (2009), Mondria and Wu (2010), and Kacperczyk et al. (2014).[2] In the second approach, an agent can sell noisy signals about future payoffs to other agents (Admati and Pfleiderer, 1986, 1990; Allen, 1990). Within this strand of research, my model is most closely related to Veldkamp (2006), who uses the same cost structure and competitive information market as the one used here to explain periodic "media frenzies" during times of heightened market volatility. However, in her paper, all agents are ex-ante homogenous and acquire news signals with equal precision, so that her framework cannot address how the cost of capital is affected by the distribution of information.

Hong and Stein (2007) provide an alternative mechanism through which news suppliers can affect asset returns. They argue that media coverage captures the attention of investors with heterogeneous beliefs, which, according to Miller (1977),

[2]As Sims (2006) asserts, the entropy-based learning technology is better suited to address the cost of "processing" readily available information, not the cost associated with how information is made available in the first place. Most finance-oriented papers in this literature assume that investors can reduce all asset uncertainty to arbitrarily low levels if endowed with sufficient information-processing capacity, which, in light of Sim's criticism, implies that every investor is exogenously supplied access to payoff information on all stocks.

drives asset prices up in the presence of short-sales constraints. In contrast, the mechanism presented here does not rely on short-sales constraints: risk-averse investors are aware that all assets exist and are free to take any position on them, but are simply less willing to hold an asset whose payoff they are more uncertain of.[3]

Finally, the analysis builds on Peress (2010), who shows that expanding a firm's investor base without raising capital induces incumbent shareholders to conduct less research because they each expect to hold a smaller stake in the firm. His results are driven by the same scaling effects to the benefit of learning responsible for the results here, but there are two crucial differences. First, while Peress (2010) exogenously limits the size of an asset's investor base (as in ?), in this chapter an asset's ownership structure is determined endogenously from the distribution of prior information. More importantly, I include a competitive market for news, which partially offsets the risk-sharing effect he documents. That is, those who invest in a firm with a small shareholder base each expects to hold more risk, but they must also bear a higher cost of producing news.

The rest of this section is organized as follows. Section 2 develops a noisy rational expectations model with both heterogeneous agents and endogenous information acquisition. Section 3 characterizes the model's equilibrium. Section 4 presents the model's main results. Section 5 reviews existing empirical literature that supports the model's predictions. Section 6 concludes and offers direction for future research.

2.2. Model Setup

2.2.1. Timeline and Assets

In this chapter, I analyze how the initial distribution of asset information across investors affects how much additional asset-specific news is produced. To do so, I develop a general equilibrium noisy rational expectations model with a continuum

[3]In heterogeneous noisy rational expectations models, each agent perceives a different risk-return tradeoff based on their individual information sets. Consequently, unlike the homogenous information setting of the CAPM, agents differ in their assessment of what constitutes an optimally diversified portfolio.

of heterogeneously informed agents of measure one and two assets: one risky asset (stock) and one riskless asset (bond).

The riskless asset has a price and payoff normalized to one and is in perfectly elastic supply. The per-capita supply of the risky asset is $\bar{x} + x$, where $x \sim N(0, \sigma_x^2)$. Random supply, usually attributed to the existence of noise traders or liquidity needs, prevents the risky asset price p, which is determined in equilibrium, from perfectly revealing all aggregate information.

The static model is divided into three periods. In period one, each agent chooses whether to purchase an additional piece of news about the risky asset that can be used to reduce payoff uncertainty. In period two, each agent chooses their optimal portfolio after observing the realization of both acquired news and the risky asset price. In period three, agents receive their payoffs.

2.2.2. Preferences

Agent i, with risk aversion ρ, makes information and portfolio decisions to maximize mean-variance utility over terminal wealth, $W_{i,3}$:

$$U_{i,1} = E_{i,1}[\rho E_{i,2}(W_{i,3}) - \frac{\rho^2}{2}\text{Var}_{i,2}(W_{i,3})]. \tag{2.1}$$

$E_{i,1}[\cdot]$ is an individual's expectation in period one, before he observes prices and acquired signals. $E_{i,2}[\cdot]$ and $\text{Var}_{i,2}[\cdot]$ are the expectation and variance of terminal wealth in period two, after individuals have observed signal realizations but before the investment decision has been made.

Given that terminal wealth is normally distributed, equation (2.1) is functionally equivalent to:

$$U_{i,1} = E_{i,1}[\log\left(E_{i,2}[e^{-\rho W_{i,3}}]\right)]. \tag{2.2}$$

This formulation of utility, whose axiomatic foundations are first provided in Kreps and Porteus (1978), represents a preference for early resolution of uncertainty, as in Epstein and Zin (1989). That is, investors with these preferences are not adverse to the risk resolved in period two: after private signals are realized but before payoffs are known. This property allows for the benefit of information to rise in the scale of investment, a feature that is absent in the case of standard CARA preferences (Peress, 2010).

2.2.3. The Distribution of Initial Information

The risky asset payoff f is not known with certainty, but each agent is endowed with some private prior knowledge about it. Formally, agent i is endowed with an independent private signal: $\widetilde{f}_i \sim N\left(f, \widetilde{\tau}_i^{-1}\right)$.

In order to isolate how the initial distribution of information affects the learning decision, I assume that the prior precision of agent i is:

$$\widetilde{\tau}_i = \alpha \bar{\tau} i^{\alpha-1}, \tag{2.3}$$

where $\bar{\tau} = \int_0^1 \widetilde{\tau}_i di$ is the average quality of initial information and $\alpha \geq 1$. Given this assumption, α can readily be interpreted as the "inequality" in the distribution of initial information. Specifically, a higher α corresponds to a greater concentration of prior precision among a smaller percentage of the population. Consistent with this interpretation, it can be shown that α is directly proportional to standard measures of inequality of a distribution used in other literatures, such as a Gini coefficient, often used to assess the concentration of income in a given population.[4]

Figure 2.1 plots the distribution of precision for different values of α. The dotted line corresponds to the distribution when $\alpha = 1$, where all agents are equally familiar about the risky asset's payoff. The dashed and solid lines correspond to the distribution with higher values of α, where a smaller segment of agents hold a greater proportion of the risky asset's prior information.

This inequality in the precision of initial information could originate from many sources. For example, a large body of theoretical research argues that the quality of corporate disclosure, which is known to differ widely across firms (Sengupta, 1998), can affect the degree of information asymmetry among shareholders (Diamond and Verrecchia, 1991; Kim and Verrecchia, 1997). Alternatively, many empirical studies have found that investors exhibit a strong preference for local firms (Coval and Moskowitz, 1999; Huberman, 2001; Grinnblatt and Keloharju, 2001), which is usually attributed to information advantages emanating from their

[4] In this context, the Gini coefficient is equal to $1 - \frac{2}{\alpha+1}$. For an elementary exposition on the Gini coefficient and how it is used to measure income inequality, see Sen (1973).

geographic proximity (Brennen and Cao, 1997; Van Nieuwerburgh and Veldkamp, 2009; Mondria and Wu, 2010). Information differences across investors could also arise due to firm-specific characteristics that affect how quickly its news diffuses across the population, such as how much media coverage the firm has previously received (Peress, 2014).

2.2.4. News Production

In the first period, each agent decides whether to acquire an additional news signal, which can be used to reduce payoff uncertainty. Let χ be the per-capita fixed cost of discovering news about the risky asset's payoff, which can be interpreted as the cost of hiring a reporter that covers an upcoming announcement. Following Veldkamp (2006), I introduce competitive information markets by assuming that once news is discovered, a copy can be sold to others at no marginal cost.[5] Importantly, an agent can freely enter the market for news even after other agents have announced the prices they will charge. In other words, the market for news is perfectly contestable. One way to ensure this contestability would be to assume that there are two sub-periods: in the first sub-period, each agent announces the price they will charge on the news item, and in the second sub-period, agents decide whether to enter the market by paying the fixed cost. This is a natural assumption for information markets, where prices are often set well before media outlets decide whether to report on a particular story.

Let d_j be an indicator variable equal to 1 if agent j discovers news and let $l_i(c_j, c_{-j}) = 1$ if agent i chooses to purchase news for price c_j, given other announced prices c_{-j}. Then, agent j chooses c_j and d_j to maximize profit:

$$\pi_j = \max_{d_j, c_j} d_j \left(c_j \int_0^1 l_i(c_j, c_{-j}) di - \chi \right), \tag{2.4}$$

where π_j enters agent j's utility function through his terminal wealth, $W_{j,3}$. Pricing and entry decisions are a sub-game perfect Nash equilibrium. Since agents are strictly better off buying news at the lowest possible price, I economize on notation by denoting l_i as an indicator for whether agent i purchases news at the lowest offered price, c. Once purchased, a copy of news cannot be resold to other agents.

[5] High fixed costs and low marginal costs appear to be a fundamental aspect of the news industry (Hamilton, 2004).

If agent i buys a copy of news (i.e., $l_i = 1$), the signal agent i observes is:

$$s_i = f + e_i, \tag{2.5}$$

where $e_i \sim N\left(0, \eta_i^{-1}\right)$ is noise in interpretation. Importantly, I assume that η_i is equal to:

$$\eta_i = k\widetilde{\tau}_i, \tag{2.6}$$

where k can be interpreted as an agent's "capacity" to process financial information. A similar form of signal precision can be explicitly derived by the entropy-motivated learning constraints used in recent asset-pricing literature, such as Van Nieuwerburgh and Veldkamp (2009), Mondria and Wu (2010), and Kacperczyk et al. (2014).[6] However, unlike those models, I assume that before an agent can allocate any of their processing-capacity k towards learning about an asset, they must first pay a separate cost to gain access to its news.[7]

While agents must pay for access to the news signal, all agents receive a free public signal about f from the price itself. Agents do not know the realization of this price signal before period two, but they can infer its precision in period one by knowing the learning decisions of other investors.[8]

2.2.5. Portfolio Selection

In period two, individuals use their prior beliefs $\widetilde{f}_i$, their private signal s_i (if news was purchased), and price p to update posterior beliefs, which are then used to choose portfolios subject to a budget constraint. Let $q_{i,0}$ denote agent i's demand for the riskless asset and let q_i denote agent i's demand for the risky asset. If $W_{i,0}$

[6]The assumption that investors' ability to interpret asset news is increasing in their prior precision is supported by evidence that portfolio managers who specialize in select stocks tend to earn higher returns (Ivkovic et al., 2008).

[7]Requiring agents to pay a separate cost to acquire news addresses the point raised by Sims (2006) that entropy-based learning constraints are not well-suited to address "costly investigation", which includes the production of news.

[8]By not incorporating the price signal into (2.6), I am implicitly assuming that no processing-capacity is required to process the information in price. Mondria (2010) shows that subjecting the price signal to entropy constraints generates multiple equilibria. To avoid this issue, one can assume that investors are endowed with enough capacity to process the price signal, as in Van Nieuwerburgh and Veldkamp (2009).

is agent i's initial wealth, then the budget constraint is:

$$W_{i,0} = q_{i,0} + q_i p + l_i c. \tag{2.7}$$

Terminal wealth $W_{i,3}$ is equal to:

$$W_{i,3} = q_{i,0} + q_i f + \pi_i. \tag{2.8}$$

Combining (2.7) and (2.8), agent i's terminal wealth can be written as:

$$W_{i,3} = W_{i,0} + q_i(f - p) + \pi_i - l_i c. \tag{2.9}$$

2.2.6. Equilibrium Definition

A rational expectations equilibrium consists of agent decisions (l_i, q_i, c_i, d_i), asset price p, news price c, and exogenous asset supply $\bar{x} + x$ such that:

- Given prices, each agent i chooses asset demand q_i and whether or not to purchase news l_i to maximize (2.1) subject to (2.6) and (2.9).
- Information supply entry decisions d_i and pricing strategies c_i are a sub-game perfect Nash equilibrium that maximize (2.4).
- Asset markets clear $\int_0^1 q_i di = \bar{x} + x$.
- Rational expectations: Beliefs about payoffs, prices, and the optimal asset demands are consistent with their true distribution.

2.3. Solution

The model is solved through backward induction. First, each agent decides on their optimal asset demand given arbitrary signals and prices. Second, knowing how optimal asset demand relates to every possible signal/price combination, each agent makes their information choices given the information choices of other agents.

2.3.1. Optimal Asset Demand

In period two, agents use the public price signal and acquired private signals to update beliefs. Let $z(p)$ denote the public price signal and let ζ denote its precision. Then with normally distributed random variables, posterior beliefs are a weighted average of the signals observed:

$$\mathrm{Var}_{i,2}(f|\widetilde{f}_i, l_i s_i, p) = \widehat{\tau}_i^{-1} = \frac{1}{\widetilde{\tau}_i + l_i\eta_i + \zeta}, \tag{2.10}$$

$$E_{i,2}(f|\widetilde{f}_i, l_i s_i, p) = \widehat{f}_i = \frac{\widetilde{\tau}_i\widetilde{f}_i + l_i\eta_i s_i + \zeta z(p)}{\widetilde{\tau}_i + l_i\eta_i + \zeta}, \tag{2.11}$$

They make their investment decisions to maximize period two utility:

$$U_{i,2} = \rho E_{i,2}(W_{i,3}) - \frac{\rho^2}{2}\mathrm{Var}_{i,2}(W_{i,3}). \tag{2.12}$$

After substituting (2.9) into (2.12), the period two problem becomes:

$$U_{i,2} = \max_{q_i} W_{i,0} + \rho q_i(\widehat{f}_i - p) - \frac{\rho^2 q_i^2}{2\widehat{\tau}_i} - l_i c + \pi_i. \tag{2.13}$$

Taking first order conditions with respect to q_i leads to optimal share-holdings:

$$q_i = \frac{\widehat{\tau}_i}{\rho}(\widehat{f}_i - p). \tag{2.14}$$

Asset demand is increasing in both posterior precision $\widehat{\tau}_i$ and the expected excess return per share $\widehat{f}_i - p$. Intuitively, the more certain I am about payoffs or the more my personal evaluation of the payoff is above the opportunity cost of purchasing a share, the more risk I will be willing to undertake.

2.3.2. Market Clearing Price

Aggregating optimal asset demand across agents and imposing the market-clearing condition determines the asset's equilibrium price. Following Hellwig (1980) and Admati (1985), price is a linear function of the risky asset's payoff.

Proposition 2.1. *Given information choices, there exists a unique linear rational expectations equilibrium. The equilibrium asset price is given by:*

$$p = f - \left(\frac{\rho\bar{x}}{\theta + \zeta} + \frac{\rho x}{\theta}\right), \tag{2.15}$$

where:

$$\theta = \bar{\tau} + \int_0^1 l_i \eta_i di, \tag{2.16}$$

$$\zeta = \left(\frac{\theta}{\sigma_x \rho}\right)^2. \tag{2.17}$$

Proof. See Appendix.

The term θ represents the average precision of priors and purchased signals, so that $\theta+\zeta$ is equal to the posterior precision of the "average agent": $\int_0^1 \widehat{\tau}_i di$. Substituting this term into equation (2.15) and taking expectations reveals that the expected excess return per share in period one is:

$$E_1(f - p) = \frac{\rho\bar{x}}{\int_0^1 \widehat{\tau}_i di}. \tag{2.18}$$

The realization of price is only observed in period two, so this unconditional expected return is the same across all agents, and is equivalent to the expected return that an outside observer would estimate. Note that $E_1(f-p)$ is increasing in both risk aversion and average number of outstanding shares, while it is decreasing in average posterior precision. The intuition is that, in order for the market to clear, asset prices must be lower to compensate an average investor for holding more risk. Since $\bar{x}$, ρ, and $\bar{\tau}$ are exogenous, equation (2.18) implies that a change in the distribution of initial information can only affect expected returns though the average precision of purchased signals, $\int_0^1 l_i \eta_i di$.

2.3.3. Optimal Information Acquisition

The appendix shows that substituting optimal asset demand (2.14) into the period two objective function (2.13) and taking expectations over period two indirect utility yields the period one problem:

$$U_{i,1} = \max_{l_i,c_i,d_i} W_{i,0} + R\left(\theta\right)\left(\widetilde{\tau}_i + l_i\eta_i + \zeta\right) - l_i c + \pi_i, \tag{2.19}$$

where:

$$R\left(\theta\right) = \frac{1}{2}\left(\left(\frac{\rho\sigma_x}{\theta}\right)^2 + \left(\frac{\rho\bar{x}}{\theta + \left(\frac{\theta}{\rho\sigma_x}\right)^2}\right)^2\right). \tag{2.20}$$

Agents are price-takers, so they take the term $R(\theta)$ as fixed. The first step in solving an agent's period one problem is to identify how much suppliers will charge for a copy of news. The following proposition states that, given increasing returns to scale in information production, the equilibrium cost of news is declining in the number of purchasers.

Proposition 2.2. *In equilibrium, one agent supplies news to the entire market:* $d_v^* = 1$ *and* $d_j^* = 0$ *for all* $j \neq v$*. Furthermore, this agent will charge at average cost:*

$$c^* = \frac{\chi}{\lambda^*}, \tag{2.21}$$

where $\lambda^* = \int_0^1 l_i^* di$*.*

Proof. See Appendix.

The more agents who buy a copy of news, the smaller the share of the fixed cost of discovery each agent incurs. For this reason, high demand news is less expensive to purchase. A crucial aspect of Proposition 2.2 for the ensuing analysis is that equation (2.21) is only a function of how many agents purchase a copy and not how much each agent is willing to pay conditional on having purchased. The reason is tied to the assumption of free entry: if any supplier charges above average cost, another agent can discover news, charge slightly below the incumbent, and take the entire market.[9]

[9] Although this result (i.e., the price of news declining in the number of purchasers) does rely on both the free entry assumption and the fixed-cost production technology, it is robust to alternative forms of competition, such as Cournot or monopolistically competitive frameworks (Veldkamp, 2006).

Because agents are not ex-ante identical, equilibrium news demand not only depends on how many agents purchase, but also depends on which agents are doing the purchasing. Before describing how λ^* is determined, the following Lemma greatly reduces the number of permissible allocations.

Lemma 2.3. *In equilibrium, agent i purchases news only if all agents with a higher prior precision do the same: $l_i^* = 1$ only if $l_v^* = 1$ for all $v > i$. Therefore:*

$$\theta(\lambda^*, \alpha) = \bar{\tau} + \int_{1-\lambda^*}^{1} k\bar{\tau}\alpha i^{\alpha-1}. \tag{2.22}$$

Proof. See Appendix.

Two separate factors are responsible for Lemma 2.3. First, recall that agents who are initially more familiar with the risky asset can more accurately interpret additional asset news. A second effect, independent of any assumption placed on η_i, is that agents who are initially more familiar with the risky asset expect to hold more shares ex-ante. Since the benefit of information is rising in the expected scale of investment—one piece of news can be used to evaluate many shares—agents with a higher prior precision find adding one unit of precision more valuable. The combination of both these effects mean that, given a particular $R(\cdot)$ and c, an agents' willingness to purchase is strictly increasing in their prior precision, which in turn means that if a certain fraction of agents are purchasing, it must be the fraction who are most initially informed.

Given Proposition 2.2 and Lemma 2.3, equilibrium news production is fully characterized by λ^*.

Proposition 2.4. *Let $B(\lambda, \alpha)$ be the net benefit of news to the marginal purchaser:*

$$B(\lambda, \alpha) = R\left(\theta(\lambda, \alpha)\right) k\bar{\tau}\alpha(1-\lambda)^{\alpha-1} - \frac{\chi}{\lambda}. \tag{2.23}$$

Then λ^* is defined by the following conditions:

- *If $B(\lambda, \alpha) < 0$ for all $\lambda \in (0, 1]$, then $\lambda^* = 0$.*
- *If $B(1, \alpha) \geq 0$, then $\lambda^* = 1$.*

- *If* $B(\lambda, \alpha) > 0$ *for some* $\lambda \in (0, 1)$ *and* $B(1, \alpha) < 0$*, then:*

$$B(\lambda^*, \alpha) = 0, \tag{2.24}$$

$$\frac{\partial B(\lambda^*, \alpha)}{\partial \lambda^*} < 0, \tag{2.25}$$

Proof. See Appendix.

There are three factors driving the net benefit of news to the marginal purchaser. First, the benefit of news is increasing in $R(\cdot)$, which is positively correlated with the risky asset's expected excess return per share. Also of note is that, for a given α, $R(\cdot)$ is strictly decreasing in λ. This is the canonical strategic substitutability in information acquisition result of Grossman and Stiglitz (1980): the more agents who acquire news, the lower the expected return (or equivalently, the more informative price is). An additional feature here stemming from my assumption on η_i is that this negative effect is amplified if the investors who acquire news have a high prior precision because these investors expect to trade more aggressively on the news item, which makes price more informative and lowers the required risk premium further.

Second, the benefit of news to the marginal purchaser is decreasing in the marginal purchaser's signal precision $k\bar{\tau}\alpha(1-\lambda)^{\alpha-1}$, which is itself a decreasing function of λ. The latter relation is a direct consequence of Lemma 1: the more agents who purchase news, the lower the signal precision of the agent who derives the least benefit from doing so.

Finally, the benefit curve is decreasing in the price of news, which is decreasing in λ by Proposition 2.2. Overall, λ has three effects on the benefit curve: two negative and one positive. When λ is arbitrarily close is close to 0, the positive effect dominates, so that the benefit curve generally takes the shape illustrated in Figure 2.2. Note that while $B(\lambda, \alpha)$ crosses 0 at two different points in Figure 2.2, only the higher point can be an equilibrium for supply of information. Namely, at any point where $B(\lambda, \alpha) = 0$ and $\frac{\partial B(\lambda,\alpha)}{\partial \lambda} > 0$, any agent can enter the market for news, charge slightly below average cost and make a small profit. The same cannot be said for the point where $B(\lambda, \alpha) = 0$ and $\frac{\partial B(\lambda,\alpha)}{\partial \lambda} < 0$, making it the unique equilibrium.

2.4. Main Results

In this section, I present the main results, most of which are a direct implication of the following proposition.

Proposition 2.5. *Assume a positive fraction of agents purchase news in equilibrium. Then:*

$$\frac{\partial c^*}{\partial \alpha} > 0. \tag{2.26}$$

Proof. See Appendix.

For intuition, consider how an incremental increase in the level of inequality α affects the benefit of news to the marginal agent, holding the fraction of purchasers λ^* fixed. First, an increase in α redistributes prior precision from less informed agents to more informed agents. Due to increasing returns to learning, this redistribution increases average posterior precision for any value of λ admissible by Lemma 1, which lowers the expected return per share by equation (2.18). Second, an increase in α changes the marginal agent's signal precision η_i. Given an endogenous cost of news, this agent becomes less familiar with the risky asset, which reduces how aggressively he plans to trade on the story. Since the marginal agent was indifferent to purchasing before the change (by definition), and both the expected return per share and expected shareholdings are lower after the change, he is strictly better off not purchasing. This makes news more expensive for the other purchasers, who are now left with having to cover a larger share of the fixed cost of discovery. Figure 2.3 illustrates how the benefit curve is affected by changes to α.[10] Note that higher values of α are associated with a smaller fraction of news purchasers and consequently higher news prices.

For the purposes here, Proposition 2.5 suggests that news markets reduce the incentive of rational investors to learn about stocks whose information is highly concentrated. This is most clearly evident by considering how the market for news alters the learning decision of a hypothetical investor whose prior precision is unaffected by α. Without information markets, an increase in α can only affect this

[10] In Figure 2.3, when $\alpha = 1$, $\lambda^* = 1$ and no agent is indifferent between purchasing and not purchasing. However, with an infinitesimally small increase in α, $\eta_0 = 0$, so that Proposition 2.5 still applies in this case.

hypothetical investor's willingness to acquire news via a change in the expected return per share. However, with information markets, the price of news is also increasing in α, meaning that the investor may be less inclined to purchase even if the expected return per share goes up, which gives rise to this chapter's main result.

Proposition 2.6. *Assume a positive fraction of agents purchase news in equilibrium. Then:*

$$\frac{\partial E_1(f-p)^*}{\partial \alpha} > 0. \tag{2.27}$$

Proof. See Appendix.

From Proposition 2.5, an increase in the inequality of prior precision reduces the number of agents who expect to hold enough shares of the risky asset to warrant paying for additional asset-related news. As a result, the remaining prospective purchasers must bear a higher share of the fixed cost of discovery, which discourages them from buying and therefore increases the risky asset's expected return.

Importantly, the cross-sectional return pattern implied by Proposition 2.5, i.e., a higher concentration of prior information leads to higher excess returns, hinges crucially on *both* an endogenous news price arising from the information market *and* increasing returns to learning implied by my assumption on η_i. To emphasize how both features factor into the result, the next two propositions illustrate how excess returns vary with α when either feature is absent.

Proposition 2.7. *Assume there are no information markets and the price of news is exogenously fixed. Then there exists an $\bar{\alpha}$ such that if $\alpha > \bar{\alpha}$:*

$$\frac{\partial E_1(f-p)^*}{\partial \alpha} < 0. \tag{2.28}$$

Proof. See Appendix.

Proposition 2.8. *Assume signal precision is independent of prior precision: $\eta_i = k$. Then:*

$$\frac{\partial E_1(f-p)^*}{\partial \alpha} = 0. \tag{2.29}$$

Proof. See Appendix.

Proposition 2.7 states if the price of news is exogenously fixed and α is sufficiently high, one reaches the exact opposite conclusion: in equilibrium the excess return actually *decreases* in α. To understand why, recall that agent i's terminal wealth from the risky asset is the return per share times the number of shares purchased, $q_i(f-p)$. With a low level of α, each agent holds a low amount of prior precision, and hence expects to hold a low number of shares. Consequently, agents are only willing to purchase news if each share expects to pay a relatively large amount. On the other hand, with a high level of α, a small fraction of agents expect to hold so many shares that they are willing to purchase news even if the payoff per share is low. Because agents are willing to keep purchasing news even at low returns, the equilibrium risk premium is smaller.

While this effect is still present when there are information markets, it is outweighed by an increasing cost of news. In other words, for high values of α and correspondingly low values of λ^*, the marginal agent still expects to hold a large stake in the risky asset, and is therefore willing to pay for information even at extremely low excess returns per share. Nevertheless, the high price of news brought on by low number of purchasers makes it prohibitively expensive to do so. Thus, the ability of agents to share the fixed cost of information production reverses how the risk premium moves with the concentration of prior precision.

The impact that an endogenous news price has on the relationship between equilibrium returns and the inequality of initial information is captured by Figure 2.4. In Figure 2.4a, the price of news is endogenous, i.e., $c = \frac{\chi}{\lambda^*}$, so that the risk premium is monotonically increasing in α by Proposition 2.6. In contrast, Figure 2.4b presents the case where the price of news is exogenously fixed at $\frac{\chi}{.5}$, so that the risk premium monotonically decreases for a high enough level of α by Proposition 2.7.[11]

Similarly, Proposition 2.8 reveals that when there are information markets but an agent's ability to process news is independent of his priors, the equilibrium

[11] It is worth noting that Figure 2.4b contains an increasing portion for small values of α. Recall that an increase in α amounts to a redistribution of precision from less informed to more informed agents. At a high λ^* (brought on by a low α), the marginal agent is losing prior precision after an incremental increase in α, and hence more willing to forgo purchasing. Eventually, prior precision becomes so concentrated that the marginal purchaser is actually on the receiving end of this redistribution, which gives rise to Proposition 2.7. While Proposition 2.7 holds for any parameters, the increasing portion of Figure 2.4b only occurs for a low a exogenous news cost.

expected return per share is independent of α. In fact, it can be shown that if there were multiple risky assets, in this case an agent's learning choice is completely independent of the quality of his initial information. The reason is that, on one hand, an investor can better diversify if he chooses to learn about a less familiar asset. On the other hand, as discussed in the previous section, an investor expects to hold a higher share of his portfolio in more familiar assets. Given mean-variance utility over wealth and a news signal independent of prior precision, these two effects completely offset.[12]

2.5. Empirical Implications

The model described in this chapter makes testable predictions as to what kind of firms will be inexpensive for investors to learn about and how this affects a firm's information environment and cost of capital. In this section, I review some existing empirical literature that supports the model's conclusions.

2.5.1. Media Coverage and Individual Investors

A key insight of this model is that assets in which many investors each expect to hold a small fraction of outstanding shares will be cheaper to learn about than assets in which few investors each expect to hold a large fraction of outstanding shares. Insomuch as mass media outlets charge cheaper prices and individual investors hold smaller portfolios than institutional investors do, this suggests that the mass media is more likely to cover stocks with high individual ownership.

This prediction is corroborated by several recent studies. In their analysis of the no-coverage premium, Fang and Peress (2009) also measure the determinants of media coverage for 4 nationally circulated daily newspapers: *New York Times*, *USA Today*, *Wall Street Journal*, and *Washington Post*, which together account for nearly 11% of daily circulation in the United States. They find that after controlling for other firm characteristics, most notably firm size and idiosyncratic volatility, a 1% increase in the fraction of shares owned by individual investors increases the annual number of articles published about a firm by .18. Similarly,

[12] The same would not be true in the case of agents with constant absolute risk aversion (CARA). For an in-depth discussion of alternative learning technologies and information preferences within a partial equilibrium setting, see Van Nieuwerburgh and Veldkamp (2010).

Solomon (2012) reports that a 1% increase in the fraction of shares owned by institutions decreases the number of articles about a firm announcement by between 27% to 30% in the Factiva news archive. Both papers also find that firm size has a overwhelmingly positive effect on the probability of news coverage, lending further support to the notion that the media covers firms with a large shareholder base, even if each shareholder holds a small stake in the firm.

This insight is also consistent with the growing body of evidence suggesting the mass media can influence the buying behavior of individual investors. Barber and Odean (2008) find that individual investors are much more likely to be net buyers of stocks in the news than those that are not. Engelberg et al. (2012) find that stocks receiving recommendations on the television show *Mad Money* experience large overnight price increases and subsequent reversals. Engelberg and Parsons (2011) find that local newspaper coverage increases the daily trade volume of local retail investors by anywhere between 8% to almost 50%. In the latter study, the authors exploit the exact timing of newspaper delivery relative to an observed spike in local trading, making it unlikely that media coverage and market reactions are both driven by some unobserved characteristic, such as how much a story peaks the public interest (Manela, 2014).

2.5.2. Asymmetric Attention and the Cross-Section of Returns

Apart from having consequences for the decisions of news suppliers, the model also makes a macro-level prediction concerning how the distribution of prior precision varies with the cost of capital. Although investors' information is not directly observable, a central tenet of the learning literature with inattentive agents is that prior precision should be highly correlated with the amount of attention paid to a stock (Van Nieuwerburgh and Veldkamp, 2009; Mondria and Wu, 2010). Therefore, one way to test the model's main result is to estimate whether stocks garnering a larger fraction of national attention from a smaller segment of the population earn higher excess returns.

Using a direct measure of investor attention first suggested by Da et al. (2011), Mondria and Wu (2012) find just that. Namely, they find that stocks with a greater degree of asymmetric attention, defined as the fraction of abnormal Google search volume for a stock generated by local investors, earn higher returns.

This finding is difficult to reconcile with models that do not include both information markets and increasing returns to learning. First, the distribution of information in and of itself should not affect the cost of capital in markets well-approximated by perfect competition (Lambert et al., 2012). Second, as per Proposition 2.7, a higher concentration of prior precision can actually decrease excess returns when there are no information markets and the cost of news is exogenous. Finally, models that incorporate information suppliers, such as Veldkamp (2006) and Admati and Pfleiderer (1986), typically assume that investors are ex-ante homogenous, and therefore do not address how the distribution of prior precision affects the cost of capital. Moreover, Proposition 2.8 shows that a heterogeneous distribution of prior precision and an information market alone cannot generate this relationship if all investors interpret a piece of news with equal precision.

2.6. Conclusion

The model presented in this chapter predicts that firms with a higher fraction of its private information concentrated within a smaller fraction of the population will be more expensive to learn about, which deters investors from following these firms and increases their cost of capital. Although the model is static, in a dynamic setting with multiple risky assets I conjecture that this effect can persist even while new market participants who are equally familiar with all risky assets continuously enter the market. This is because new investors are more likely to be initially exposed to stocks whose news is cheaper to purchase. This initial exposure leads them to prefer learning about the same stocks in subsequent periods. The end result is that cross-sectional variation in expected returns can persist even in the absence of any cross-sectional persistence in volatility. A dynamic framework could also explain the puzzling observation that higher media coverage is associated with lower returns, despite the fact that media coverage appears to be a stable firm characteristic (Fang and Peress, 2009).

Another extension involves allowing investors the ability to purchase news of various qualities. Given the results documented here, a reasonable hypothesis is that the demand for news mimics the distribution of prior information: when prior information is concentrated, the market for news is represented by a small contingent of investors demanding high-quality news, whereas if prior information were uniformly distributed across the population, the market for news is represented by many investors demanding low-quality news.

Finally, this chapter has an important implication for empirical efforts that use the amount of media coverage as a proxy for learning preferences: while media coverage may proxy for the information demands of many consumers, the direction of causality may indeed be reversed for a large segment of the population given increasing returns to scale in news production, especially if subscribers' demands are heterogeneous and the fixed cost of production is high. That is, it is not that media coverage reflects the learning preferences of all its subscribers, but rather subscribers' learning choices are restricted to the subset of stories covered substantially by the news.

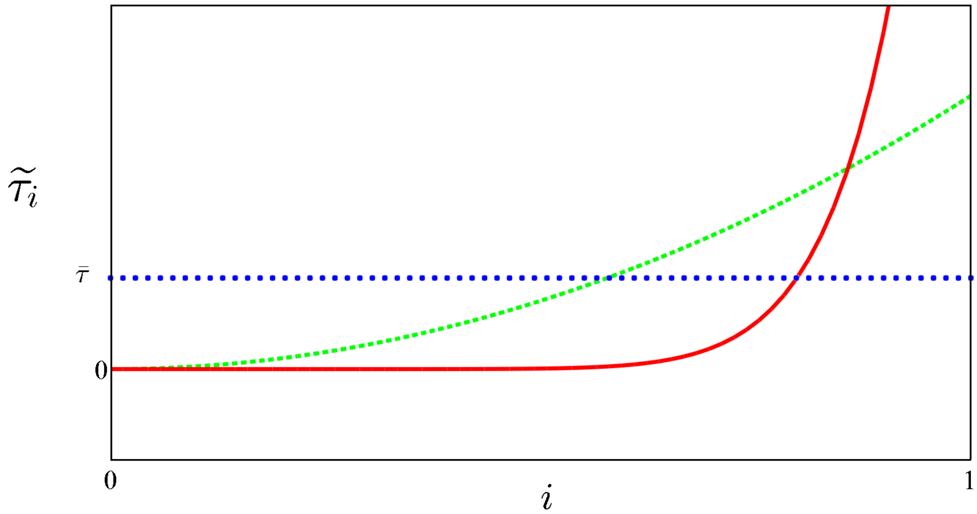


FIGURE 2.1: **The distribution of prior precision with varying levels of inequality.** This figure presents how the distribution of prior precision varies with α. The dotted line corresponds to $\alpha = 1$, where the distribution of precision is uniformly distributed across agents. The dashed curve and solid curve correspond to $\alpha = 3$ and $\alpha = 20$, respectively, whereby prior precision is more concentrated within a smaller fraction of the population. $\bar{\tau} = 1$.

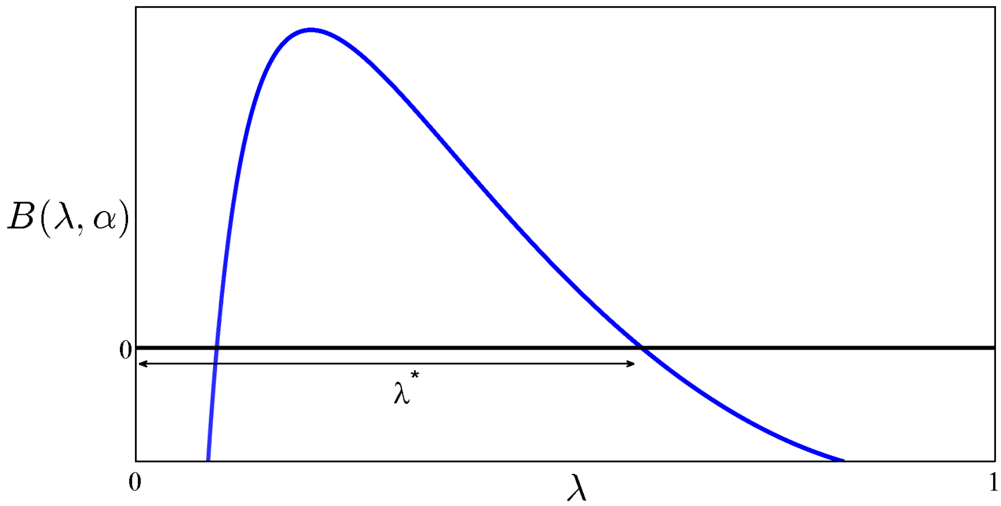


FIGURE 2.2: **Equilibrium news production.** This figure presents the benefit of news to the marginal purchaser as a function of λ. $k = 2$, $\rho = 2.5$, $\bar{\tau} = 1$, $\bar{x} = 2$, $\sigma_x^2 = 2$, $\chi = 1.1$, $\alpha = 3$.

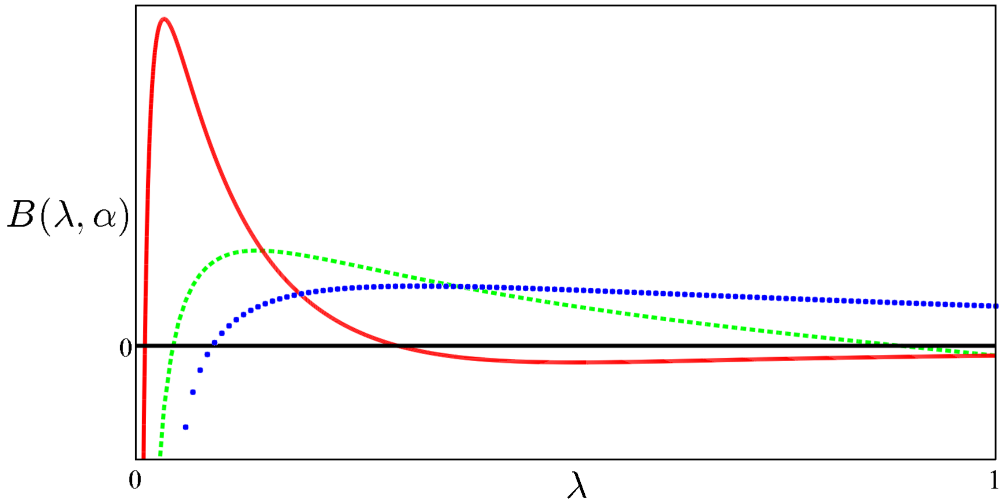
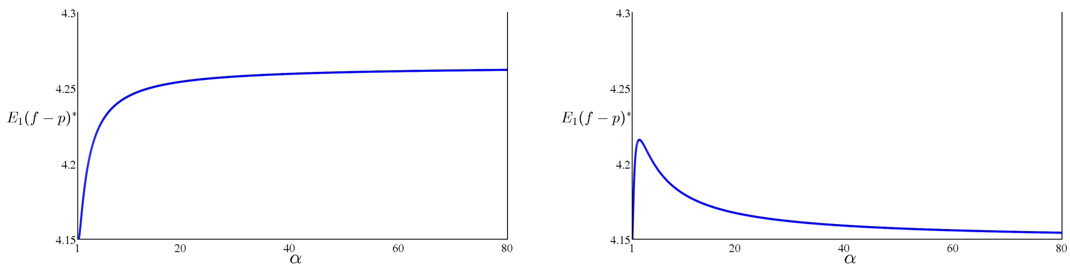


FIGURE 2.3: **The benefit of purchasing news with varying levels of inequality.** This figure presents how the benefit of news to the marginal purchaser varies with α. $k = .6$, $\rho = 2.5$, $\bar{\tau} = 1$, $\bar{x} = 2$, $\sigma_x^2 = 2$, $\chi = 1.1$, α(dotted)=1, α(dashed)=2, α(solid)=8.



(A) Endogenous Price of News

(B) Exogenous Fixed Price of News

FIGURE 2.4: **Equilibrium returns with and without an endogenous price of news.** This figure presents how the equilibrium return per share $E_1(f-p)^*$ is affected by α with and without an endogenous price of news. $k = .2$, $\rho = 2.5$, $\bar{\tau} = 1$, $\bar{x} = 2$, $\sigma_x^2 = 2$, $\chi = 1.1$, c(figure a)=$\frac{\chi}{\lambda^*}$, c(figure b)=$\frac{\chi}{.5}$.

Chapter 3

LEARNING ABOUT NOISE

3.1. Introduction

The objective of this chapter is to study theoretically how the opportunity to acquire both fundamental and non-fundamental information affects the distribution of information across investors, their incentives to acquire information, and price informativeness. We develop a three-period rational expectations model with one risky asset and two types of investors: noise traders, who trade for exogenous reasons, and an infinite number of competitive rational investors, who maximize constant absolute risk aversion (CARA) utility over the third-period consumption and who are initially endowed with information of heterogeneous quality about the future asset payoff. Before trading, any rational investor can obtain additional information in the form of two signals with optimally chosen precisions: one is about the asset payoff (about fundamentals) and the other is about the asset supply by noise traders (about noise). The information acquisition is costly: to receive the signals, the investor should buy learning capacity, which determines the maximum possible reduction in his uncertainty about fundamentals and noise (quantified by the entropy) and effectively bounds the precisions of the signals from above. After choosing the amount of purchased learning capacity, each investor decides how to allocate it between fundamental and non-fundamental information. The obtained information along with the initial information and the information contained in the price is used in the trading period for developing the optimal portfolio strategy. The equilibrium price is determined by the market clearing condition.

Our analysis delivers several insights. First, our model captures the tradeoff between acquiring fundamental and non-fundamental information and shows that in equilibrium investors specialize in information acquisition, that is, each investor decides to learn either only about the future asset payoff or only about noise. In particular, those investors who are endowed with relatively precise information about fundamentals choose to learn more about them, whereas the investors who have poor initial information about fundamentals use all their resources to discover non-fundamental information. If the marginal cost of learning is large enough for any amount of information, the investors with fundamental information of intermediate quality do not participate in information acquisition at all. The optimality of acquiring only one type of information stems from the interaction between the assumed entropy-based learning technology, which implies increasing returns to initial information, and the informativeness of the equilibrium price about fundamentals. The effect resembles the specialization in acquisition of fundamental information about individual stocks demonstrated, for example, by Van Nieuwerburgh and Veldkamp (2009), Kacperczyk et al. (2012), and Kacperczyk et al. (2014), although in our model investors choose between fundamentals and noise and the mechanics of the effect are different.

Second, we demonstrate that learning about fundamentals increases the asymmetry of information among investors, whereas learning about noise decreases it. These effects result from interplay of three model ingredients: the cost of learning capacity, the entropy-based information acquisition technology, and the assumption that the initial fundamental information is dispersed across investors more unevenly than the information about noise. The latter assumption captures the idea that empirically investors' sentiment and liquidity needs are less persistent than firm fundamentals, so it is harder to preserve information advantage about noise. The increasing returns to initial information implied by the entropy constraint amplify the heterogeneity in information precisions among those investors who decide to learn about fundamentals, and this makes the asymmetry of ex post information more pronounced. Meanwhile, those who learn about noise end up with similar information precisions because initially better informed investors spend less on learning about noise than those who are initially less informed. These effects highlight two possible uses of initial information: to facilitate the acquisition of additional fundamental information and to reduce the information acquisition costs when the investor learns about noise.

Third, our analysis shows that the opportunity to obtain non-fundamental information increases the price informativeness about fundamentals. This result may look counterintuitive since learning about noise could be a waste of information processing resources: some investors who otherwise would pay for learning about fundamentals switch to learning about noise. However, the investors who are not well informed about the payoff learn about noise because this helps them to extract more information about fundamentals from the price and when they trade on that information they make the price more informative for all other investors. The effect is magnified by complementarity between the non-fundamental information and price, that is, a more informative price makes the signal about noise more valuable for investors. We also find that the magnitude of the effect depends on the initial distribution of information across investors and appears to be particularly large when only few investors possess high quality fundamental information and others are almost uninformed. In that case, the possibility to learn about noise does not change the information acquisition strategy of the informed minority (they still learn about fundamentals) but opens up a valuable source of information for the uninformed majority. When information is dispersed more evenly across investors, more of them acquire information about fundamentals and the additional opportunity to learn about noise has a weaker effect on the equilibrium strategies and price informativeness.

Our results have several interesting implications. In particular, the specialization in information acquisition predicted by the model can explain the anecdotal evidence that some market participants (like mutual funds) gather mostly fundamental information about firms, whereas others (like some hedge funds) deal mostly with temporary market disbalances. Our finding that learning about fundamentals increases the asymmetry of information across investors implies that even a small difference in the initial quality of information about firms may ultimately result in a highly unequally informed population of investors. This effect is partially mitigated by acquisition of non-fundamental information, which decreases the informational inequality among those who learn about noise, but it still can produce a high concentration of information and, as a result, increase the cost of capital (Marmora, 2014). Our conclusion that learning about noise improves price informativeness implies that speculative traders, who are often blamed for counterproductive trading, make markets more informationally efficient. Without additional market frictions, investors realize their competitive advantage in information acquisition of a particular type and by choosing what to learn about they improve information aggregation by prices.

This chapter belongs to a large literature on endogenous information acquisition in rational expectations models pioneered by Grossman and Stiglitz (1980) and Verrecchia (1982) and surveyed in Vives (2010) and Veldkamp (2011). In contrast to the majority of existing models in which investors can learn only about fundamentals, our analysis assumes that investors can acquire both fundamental and non-fundamental information. The closest to our study is the paper by Ganguli and Yang (2009), who also assume that market participants learn about random asset supply. However, Ganguli and Yang (2009) focus on the multiplicity of equilibria and complementarities in information acquisition when the signal is two-dimensional and, in contrast to this chapter, do not allow investors to decide on the amount of information in the signal and its composition.

There are also several studies on the role of non-fundamental information in the Kyle (1985) framework. Madrigal (1996) demonstrates that the entry of agents with non-fundamental information (non-fundamental speculators) reduces price informativeness and market liquidity. Yu (1999) examines the value of non-fundamental information and finds that it can be negative when the information is imprecise. Park (2010) studies the acquisition of non-fundamental information and argues that it hurts informed traders. In contrast to these papers, we find advantages of learning about noise and highlight the role of competitive markets and learning from prices in efficient aggregation of different kinds of information.

There are also several related studies in which investors, along with the uncertainty about fundamentals, deal with the uncertainty about various aspects of their trading environment. Using the Kyle (1985) framework, Hong and Rady (2002) model learning about the variance of the noise traders' supply, which evolves over time according to a two-state Markov chain. Gao et al. (2013) study the impact of uncertainty about the number of informed investors and conclude that price informativeness can be a non-monotonic function of the proportion of informed traders. In Banerjee and Green (2014) some traders are uncertain about whether other investors trade on informative signals or noise. Banerjee et al. (2014) construct a model in which investors learn about fundamentals and intensity of feedback trading and argue that the price informativeness can decrease with the precision of public information and information processing capacity of investors. In Easley et al. (2014) some ambiguity-averse investors (mutual funds) are uncertain about the risk tolerance of other investors (hedge funds).

The rest of the chapter is organized as follows. In Section 3.2 we present the main model. Section 3.3 describes the construction of the equilibrium. Section 3.4 discusses the most interesting properties of the equilibrium. Section 3.5 concludes the main part of this chapter by summarizing the main results. The Appendix contains the proofs of all propositions.

3.2. Model

3.2.1. Main Setup

Our model extends the classic framework of Grossman and Stiglitz (1980), Hellwig (1980), and Verrecchia (1982). There are two assets in the economy: a risk-free asset with the exogenously fixed zero rate of return and $\bar{\theta}$ shares of a risky asset with a random payoff f per share. The economy is populated by two types of investors: competitive rational investors and noise traders. The rational investors are indexed by $i \in \mathcal{I}$, where $\mathcal{I}$ is an infinite countable set with a measure μ such that it weights all investors equally and $\mu(\mathcal{I}) = 1$.[1] The investors live in three periods. In the first period ($t = 1$) the rational investors acquire information, in the second period ($t = 2$) all investors trade, and in the third period ($t = 3$) they observe the payoff f and consume it. Each rational investor i has constant absolute risk aversion (CARA) preferences with the same coefficient of risk aversion γ and derives utility from terminal wealth $W_{3,i}$. Also, investor i is endowed with initial information about the payoff f: in his information set $f \sim \mathcal{N}(f_i, \tau_i^{-1})$, where the estimates f_i are unbiased in the sense that $\int_{\mathcal{I}} f_i d\mu(i) = f$ and τ_i, $i \in \mathcal{I}$, are their precisions.[2] Without losing generality, we assume that τ_i is a measurable function on $\mathcal{I}$. Note that in contrast to many models with asymmetric information in which investors can be either informed or uninformed, we allow for a much reacher heterogeneity of information precisions across investors. As a result, we can determine not only how many investors decide to become informed but also

[1] This approach to modeling an infinite number of investors is promoted in Feldman and Gilles (1985) and used by He and Wang (1995) and Constantinides and Duffie (1996), among many others. The assumption that there is a countable number of investors facilitates aggregation using the law of large numbers. See technical details in Feldman and Gilles (1985) and Judd (1985).

[2] This specification admits an alternative interpretation: each investor has an improper uniform prior about f and before the first period receives a normally distributed signal $f_i \sim \mathcal{N}(f, \tau_i^{-1})$.

who acquires information and how the optimal learning strategy is affected by prior information.

The noise traders participate in the market due to exogenous reasons such as sentiment, liquidity needs, willingness to hedge a variation in private endowments or private investment opportunities, etc. and they supply θ shares of the risky asset at $t = 2$. It is natural to assume that the average supply of shares by noise traders is zero and liquidity needs as well as other reasons for uninformed trading are short-lived, at least compared to the information about fundamentals. Thus, no one initially has superior information about noise and in the information set of the rational investors at the beginning of the first period $\theta \sim \mathcal{N}(0, \tau_\theta^{-1})$, where τ_θ^{-1} can be interpreted as the unconditional historical variability of noise trading.

3.2.2. Information Acquisition

A unique element of our model is the specification of information acquisition in the period $t = 1$. In contrast to the vast majority of existing studies, we assume that each investor can learn not only about fundamentals (the future payoff of the risky asset f) but also about the supply of shares by noise traders θ. Because the information choice is two-dimensional, each investor effectively makes two decisions: how much information to acquire in total and what this information is about. Thus, it is natural to model the learning process as a two-step procedure.

In the first step, each investor buys information processing capacity K for which he pays $c(K)$ from his initial wealth. The cost function $c(\cdot)$ is twice continuously differentiable, positive, increasing, and non-concave: $c(K) > 0$, $c'(K) > 0$, and $c''(K) \geq 0$ for $K > 0$. The information processing capacity can be interpreted as availability of powerful computers that investors use to process information, human capital involved in the information analysis, etc.

In the second step, each investor allocates the information processing capacity between fundamentals and noise. The information takes a form of two individual signals received by investor i: the signal $s_{f,i} \sim \mathcal{N}(f, \tau_{sf,i}^{-1})$ is about fundamentals and the signal $s_{\theta,i} \sim \mathcal{N}(\theta, \tau_{s\theta,i}^{-1})$ is about noise. The signals are uncorrelated with each other and their precisions $\tau_{sf,i}$ and $\tau_{s\theta,i}$ are the investor's choice variables. This specification implies that learning about fundamentals and noise are separate and distinct tasks. For example, the investor can scour accounting statements, subscribe for analyst reports, conduct private surveys of customers and suppliers, etc. to obtain information on fundamentals and use technical analysis, measure market sentiment, identify fire sales, etc. to obtain information on noise. The signals that the investor receives represent the outcome of his information gathering activity.

To model the allocation of learning capacity between fundamentals and noise, we assume that the capacity K allows the investor to decrease the entropy of the joint distribution of f and θ by $\log(1+K)$. The entropy is the standard measure of uncertainty in information theory and for an n-variate normal distribution it is equal to $\log|\Omega|/2 + n\log(2\pi e)/2$, where Ω is the variance-covariance matrix (e.g., Cover and Thomas, 2006).[3] Denoting the variance-covariance matrices of the vector $(f \;\; \theta)$ in the information set of agent i before and after information acquisition as Σ_i and $\hat{\Sigma}_i$, respectively, the entropy reduction constraint can be written as

$$\frac{|\Sigma_i|}{|\hat{\Sigma}_i|} \leq 1+K. \tag{3.1}$$

Because fundamentals and noise are uncorrelated and the signals $s_{f,i}$ and $s_{\theta,i}$ are drawn independently, the matrices Σ_i and $\hat{\Sigma}_i$ are diagonal and have the following form:

$$\Sigma_i = \begin{pmatrix} \tau_{f,i}^{-1} & 0 \\ 0 & \tau_\theta^{-1} \end{pmatrix}, \qquad \hat{\Sigma}_i = \begin{pmatrix} (\tau_{f,i}+\tau_{sf,i})^{-1} & 0 \\ 0 & (\tau_\theta+\tau_{s\theta,i})^{-1} \end{pmatrix},$$

so the entropy reduction constraint in equation (3.1) reduces to

$$\left(1+\frac{\tau_{sf,i}}{\tau_{f,i}}\right)\left(1+\frac{\tau_{s\theta,i}}{\tau_\theta}\right) \leq 1+K. \tag{3.2}$$

Note that in contrast to many papers that use the entropy to describe bounds on investors' information processing ability that result from limited attention and other purely psychological factors and consider K as an exogenous parameter, we use the entropy to model limits of technology that investors employ to learn about fundamentals and noise. As a results, the level of the capacity constraint in our model is endogenous as investors decide how much costly technology to devote to information discovery. Our interpretation of the learning process also implies

[3] Sims (2003) proposes to use entropy bounds to measure rational inattention of economic agents. The entropy-based constraints on learning are also used by Peng (2005), Peng and Xiong (2006), Van Nieuwerburgh and Veldkamp (2009), Mondria (2010), Van Nieuwerburgh and Veldkamp (2010), Banerjee et al. (2014), Kacperczyk et al. (2014), among others.

that market prices convey information without error and extraction of information from them does not consume information processing capacity as it would when constraints on learning result from limited attention (Sims, 2006).

3.3. Equilibrium

3.3.1. Equilibrium Definition

We define the equilibrium in the model as a standard rational expectations equilibrium, which is characterized by the following conditions.

1. At time $t = 2$ each rational investor chooses an optimal trading strategy that maximizes his expected utility using all available information, which includes the price of the risky asset.

2. The market for the risky asset clears, that is, the aggregate demand of rational investors is equal to the sum of the total supply of the asset and the supply by noise traders.

3. The informativeness of the price implied by the market clearing condition coincides with its informativeness assumed by rational investors.

4. At time $t = 1$ each rational investor maximizes his expected utility by deciding how much information processing capacity to buy and choosing the precisions of the signals about fundamentals and noise that he receives.

As usual, we construct the equilibrium starting from the second period. Taking the precisions of the signals as given, we find first the equilibrium price function and then, assuming that investors can commit to a particular portfolio strategy, compute their optimal information acquisition strategies in the first period.

3.3.2. Trading Period

Denote the information set of investor $i \in \mathcal{I}$ at time $t = 2$ as $\mathcal{F}_{2,i}$. It contains the prior information of the investor, the signals that the investor received at $t = 1$, and the price of the asset p. Conditional on this information, the expectation of the asset payoff is $\hat{f}_i = E[f|\mathcal{F}_{2,i}]$ and its precision is $\hat{\tau}_i = Var[f|\mathcal{F}_{2,i}]^{-1}$. Each investor i decides how many shares of the risky asset x_i to buy by solving the standard utility maximization problem

$$\max_{x_i} E\left[-e^{-\gamma W_{3,i}}|\mathcal{F}_{2,i}\right] \tag{3.3}$$

subject to the budget constraint

$$W_{3,i} = W_{2,i} + x_i(f - p),$$

where $W_{2,i}$ is the wealth of investor i at $t = 2$. The well-known solution to this optimization problem is

$$x_i = \gamma^{-1}\hat{\tau}_i(\hat{f}_i - p). \tag{3.4}$$

The aggregation of demands of all investors together with the market clearing condition

$$\int_{\mathcal{I}} x_i d\mu(i) = \bar{\theta} + \theta \tag{3.5}$$

yields the equilibrium price. The rational expectations equilibrium in the trading period is described by Proposition 3.1.

Proposition 3.1. *Assume that $\gamma^2 \geq 4(\bar{\tau}_f + \bar{\tau}_{sf})\bar{\tau}_{s\theta}$, where*

$$\bar{\tau}_f = \int_{\mathcal{I}} \tau_{f,i} d\mu(i), \qquad \bar{\tau}_{sf} = \int_{\mathcal{I}} \tau_{sf,i} d\mu(i), \qquad \bar{\tau}_{s\theta} = \int_{\mathcal{I}} \tau_{s\theta,i} d\mu(i). \tag{3.6}$$

Then, there exists a linear rational expectations equilibrium with the price function

$$p = -\frac{\gamma\bar{\theta}}{\bar{\tau}} + f + p_\theta\theta, \tag{3.7}$$

where $\bar{\tau}$ is the aggregate ex post precision of investors' information about fundamentals

$$\bar{\tau} = \int_{\mathcal{I}} \hat{\tau}_i d\mu(i), \qquad \hat{\tau}_i = \tau_{f,i} + \tau_{sf,i} + \frac{1}{p_\theta^2}(\tau_\theta + \tau_{s\theta,i}). \tag{3.8}$$

The sensitivity of the price to noise p_θ solves the following equation:

$$(\bar{\tau}_f + \bar{\tau}_{sf})p_\theta + \frac{\bar{\tau}_{s\theta}}{p_\theta} + \gamma = 0. \tag{3.9}$$

Proof. See Appendix.

As in the vast majority of the rational expectations equilibria constructed in the literature, the equilibrium price in equation (3.7) is linear in fundamentals f and noise θ. Following the literature (e.g., Vives, 1995; Goldstein and Yang, 2014), we identify price informativeness about fundamentals with the precision of the payoff f conditional only on the price, that is, with $Var[f|p]^{-1}$. Equation (3.7) implies that $Var[f|p]^{-1} = \tau_\theta/p_\theta^2$, so the price informativeness is solely determined by p_θ and inversely proportional to its square. Because of that, our analysis focuses on the parameter p_θ and, noting that $p_\theta < 0$ (this immediately follows from equation (3.9)), we interpret prices with higher p_θ as more informative.

Proposition 3.1 implies that additional information about noise can produce multiple equilibria in the trading period. When investors do not receive signals about θ, we have $\bar{\tau}_{s\theta} = 0$ and equation (3.9) has a unique solution $p_\theta = -\gamma/(\bar{\tau}_f + \bar{\tau}_{sf})$, which corresponds to the equilibrium constructed in Hellwig (1980) and Admati (1985).[4] However, the equation for p_θ becomes quadratic when $\bar{\tau}_{s\theta} > 0$. The solutions to this equation are

$$p_\theta = \frac{-\gamma \pm \sqrt{\gamma^2 - 4(\bar{\tau}_f + \bar{\tau}_{sf})\bar{\tau}_{s\theta}}}{2(\bar{\tau}_f + \bar{\tau}_{sf})} \tag{3.10}$$

and they exist only when $\gamma^2 \geq 4(\bar{\tau}_f + \bar{\tau}_{sf})\bar{\tau}_{s\theta}$, that is, when the information about fundamentals and noise is relatively imprecise. The multiplicity of equilibria produced by multidimensionality of private signals is consistent with the results of Ganguli and Yang (2009) and Manzano and Vives (2011).

Note that for both solutions from equation (3.10) we have $p_\theta > -\gamma/(\bar{\tau}_f + \bar{\tau}_{sf})$, so in both equilibria the prices are more informative about fundamentals than in the case with $\bar{\tau}_{s\theta} = 0$ and the same aggregate precision of the information about fundamentals $\bar{\tau}_f + \bar{\tau}_{sf}$. This is not surprising because any additional information about noise reduces the error in the price as a signal about fundamentals. However, when the precisions of both signals $s_{f,i}$ and $s_{\theta,i}$ are choice variables, investors may decide to learn more about noise (this increases the price informativeness) but less about fundamentals (this decreases the price informativeness), so the overall effect of the possibility to learn about noise could be ambiguous. The comparison of the price informativeness in models with and without learning about noise that takes

[4] The coefficients of the price function in Hellwig (1980) and Admati (1985) are different because they assume that all investors have an identical prior about fundamentals.

into account the endogeneity of information acquisition is performed in Section 3.4.2.

3.3.3. Information Acquisition Period

In the first period, each investor decides how much information to acquire and what this information is about. Denote by $\mathcal{F}_{1,i}$ the information set of investor i that includes only the prior about the payoff f. The optimal information acquisition strategy maximizes the investor's utility function in the period $t = 1$, which is

$$U_{1,i} = E\left[-e^{-\gamma W_{3,i}}|\mathcal{F}_{1,i}\right]. \tag{3.11}$$

Proposition 3.2 presents $U_{1,i}$ as a function of signal precisions when the investor follows the optimal trading strategy from equation (3.4).

Proposition 3.2. *The expected utility of investor i for the chosen precisions of the signals $\tau_{sf,i}$ and $\tau_{s\theta,i}$ before the information is acquired is*

$$U_{1,i} = -\sqrt{\frac{Var[f-p|\mathcal{F}_{2,i}]}{Var[f-p|\mathcal{F}_{1,i}]}}\exp\left(-\gamma W_{1,i} + \gamma c(K) - \frac{E[f-p|\mathcal{F}_{1,i}]^2}{2Var[f-p|\mathcal{F}_{1,i}]}\right). \tag{3.12}$$

Proof. See Appendix.

The expression for $U_{1,i}$ in equation (3.12) is a generalization of the utility computed in Grossman and Stiglitz (1980) for the case of arbitrary information sets $\mathcal{F}_{1,i}$ and $\mathcal{F}_{2,i}$, although it is still assumed that conditional on these information sets all variables are normally distributed.

Using the equilibrium price function from equation (3.7), $Var[f-p|\mathcal{F}_{1,i}] = p_\theta^2\tau_\theta^{-1}$ and $E[f-p|\mathcal{F}_{1,i}] = \gamma\bar{\theta}/\bar{\tau}$. Also noting that $Var[f-p|\mathcal{F}_{2,i}]^{-1} = \tau_{f,i} + \tau_{sf,i} + (\tau_\theta + \tau_{s\theta,i})/p_\theta^2$, the investor's problem at $t = 1$ can be written as

$$\max_{\{K\geq 0,\ \tau_{sf,i}\geq 0,\ \tau_{s\theta,i}\geq 0\}} \frac{U_{0,i}}{\sqrt{\tau_{f,i} + \tau_{sf,i} + (\tau_\theta + \tau_{s\theta,i})/p_\theta^2}} e^{\gamma c(K)} \tag{3.13}$$

subject to

$$\left(1 + \frac{\tau_{sf,i}}{\tau_{f,i}}\right)\left(1 + \frac{\tau_{s\theta,i}}{\tau_\theta}\right) \leq 1 + K, \tag{3.14}$$

where

$$U_{0,i} = -\sqrt{\frac{\tau_\theta}{p_\theta^2}} \exp\left(-\gamma W_{1,i} - \frac{\gamma^2\bar{\theta}^2\tau_\theta}{2\bar{\tau}^2 p_\theta^2}\right).$$

Because the information processing capacity is costly and $c'(K) > 0$, the constraint (3.14) must hold in equilibrium as equality. Noting that $U_{0,i}$ is a negative constant, the maximization problem from equation (3.13) is equivalent to the following minimization problem:

$$\min_{\{\tau_{sf,i}\geq 0,\ \tau_{s\theta,i}\geq 0\}} \frac{1}{\sqrt{\tau_{f,i} + \tau_{sf,i} + (\tau_\theta + \tau_{s\theta,i})/p_\theta^2}} \exp\left(\gamma c\left(\left(1 + \frac{\tau_{sf,i}}{\tau_{f,i}}\right)\left(1 + \frac{\tau_{s\theta,i}}{\tau_\theta}\right) - 1\right)\right). \tag{3.15}$$

The properties of its solution are described by Proposition 3.3.

Proposition 3.3.

1. *The optimization problem (3.15) has only boundary solutions, that is, either $\tau_{s\theta,i} = 0$ or $\tau_{sf,i} = 0$.*

2. *Assume that $c'(0) = 0$. Then there exists τ_f^* such that i) for $\tau_{f,i} < \tau_f^*$ the optimization problem (3.15) has a unique solution with $\tau_{sf,i} = 0$ and $\tau_{s\theta,i} = \tau_{s\theta,i}^*(\tau_{f,i}) > 0$, where $\tau_{s\theta,i}^*(\tau_{f,i})$ is defined as a solution to the following equation*

$$\frac{2\gamma}{\tau_\theta} c'\left(\frac{\tau_{s\theta,i}^*}{\tau_\theta}\right) = \frac{1}{p_\theta^2\tau_{f,i} + \tau_\theta + \tau_{s\theta,i}^*}, \tag{3.16}$$

and ii) for $\tau_{f,i} > \tau_f^*$ the optimization problem (3.15) has a unique solution with $\tau_{sf,i} = \tau_{sf,i}^*(\tau_{f,i}) > 0$ and $\tau_{s\theta,i} = 0$, where $\tau_{sf,i}^*(\tau_{f,i})$ is defined as a solution to the following equation

$$\frac{2\gamma}{\tau_{f,i}} c'\left(\frac{\tau_{sf,i}^*}{\tau_{f,i}}\right) = \frac{p_\theta^2}{p_\theta^2(\tau_{f,i} + \tau_{sf,i}^*) + \tau_\theta}. \tag{3.17}$$

The threshold τ_f^* is uniquely determined by the following equation:

$$c\left(\frac{\tau_{sf,i}^*(\tau_f^*)}{\tau_f^*}\right) - c\left(\frac{\tau_{s\theta,i}^*(\tau_f^*)}{\tau_\theta}\right) = \frac{1}{2\gamma}\log\frac{p_\theta^2(\tau_f^* + \tau_{sf,i}^*(\tau_f^*)) + \tau_\theta}{p_\theta^2\tau_f^* + \tau_\theta + \tau_{s\theta,i}^*(\tau_f^*)}, \tag{3.18}$$

where $\tau_{sf,i}^*(\tau_f^*)$ and $\tau_{s\theta,i}^*(\tau_f^*)$ are defined above. The investors with $\tau_{f,i} = \tau_f^*$ are indifferent between learning about fundamentals and learning about noise.

3. *Assume that* $c(K) = c_0 K$ *and define*

$$\tilde{\tau}_{sf,i} = \left(\frac{1}{2\gamma c_0} - 1\right)\tau_{f,i} - \frac{\tau_\theta}{p_\theta^2}, \qquad \tilde{\tau}_{s\theta,i} = \left(\frac{1}{2\gamma c_0} - 1\right)\tau_\theta - p_\theta^2 \tau_{f,i}. \quad (3.19)$$

Then the optimal information acquisition strategy is as follows:

(a) *if* $c_0 \leq 1/(4\gamma)$, *then*

$$(\tau^*_{sf,i}, \tau^*_{s\theta,i}) = \begin{cases} (0, \tilde{\tau}_{s\theta,i}), & \text{if } \tau_{f,i} < \tau_\theta/p_\theta^2; \\ (0, \tilde{\tau}_{s\theta,i}) \text{ or } (\tilde{\tau}_{sf,i}, 0), & \text{if } \tau_{f,i} = \tau_\theta/p_\theta^2; \\ (\tilde{\tau}_{sf,i}, 0), & \text{if } \tau_{f,i} > \tau_\theta/p_\theta^2. \end{cases} \quad (3.20)$$

(b) *if* $1/(4\gamma) < c_0 < 1/(2\gamma)$, *then*

$$(\tau^*_{sf,i}, \tau^*_{s\theta,i}) = \begin{cases} (0, \tilde{\tau}_{s\theta,i}), & \text{if } \tau_{f,i} < \frac{\tau_\theta}{p_\theta^2}\left(\frac{1}{2\gamma c_0} - 1\right); \\ (0, 0), & \text{if } \frac{\tau_\theta}{p_\theta^2}\left(\frac{1}{2\gamma c_0} - 1\right) \leq \tau_{f,i} \leq \frac{\tau_\theta}{p_\theta^2}\left(\frac{1}{2\gamma c_0} - 1\right)^{-1}; \\ (\tilde{\tau}_{sf,i}, 0), & \text{if } \tau_{f,i} > \frac{\tau_\theta}{p_\theta^2}\left(\frac{1}{2\gamma c_0} - 1\right)^{-1}. \end{cases} \quad (3.21)$$

(c) *if* $c_0 \geq 1/(2\gamma)$, *then investors do not learn at all and* $(\tau^*_{sf,i}, \tau^*_{s\theta,i}) = (0, 0)$.

Proof. See Appendix.

The first statement of Proposition 3.3 implies that it is optimal for investors to specialize in acquisition of a particular type of information and no one simultaneously acquires signals about fundamentals and noise. This result follows from an interaction of two aspects of the model: i) the entropy-based learning technology implies increasing returns to the prior information and ii) the signals about the asset payoff and noise are substitutes when investors learn from prices. The first statement is evident from the entropy reduction constraint (3.2): investors with a higher precision of prior information can obtain a more precise signal using the same information processing capacity. This property of entropy-based learning is also exploited by Van Nieuwerburgh and Veldkamp (2009), who explain the home bias puzzle by the choice of investors to specialize in learning about domestic firms. However, the entropy-based learning alone does not necessarily produce the specialization in information acquisition. For example, Van Nieuwerburgh and Veldkamp (2010) study the optimal acquisition of information about fundamentals of multiple assets and find that an investor with CARA preferences and an entropy-based information processing constraint is indifferent between any allocations of his learning capacity. The insufficiency of increasing returns to the initial information for producing the specialization in learning is also pointed out by Van Nieuwerburgh and Veldkamp (2009), who illustrate this by considering a setting in which investors ignore the influence of acquired information on their portfolios.

The second element of our model responsible for the specialization in information acquisition is the substitutability of signals about fundamentals and noise. It is specific to our analysis and arises because investors extract information from prices, which are determined by both fundamentals and noise. For example, if an investor receives a signal about noise he uses it to extract more precise information about fundamentals from the price, which is equivalent to receiving a direct signal about fundamentals. The dependence of prices on fundamentals and noise is a very general property of rational expectations models, so our specialization result does not rely on the linearity of the price function and is likely to be highly robust to alternative model specifications.

Next, Proposition 3.3 characterizes the information acquisition strategy of an investor with prior precision $\tau_{f,i}$ when the marginal cost of information processing capacity is small for low capacities. The latter condition implies that it is suboptimal for investors not to participate in information production and each of them learns about either noise or fundamentals. Proposition 3.3 states that those investors who have a relatively precise prior about fundamentals ($\tau_f > \tau_f^*$) choose to receive a signal about fundamentals, whereas those with imprecise priors ($\tau_f < \tau_f^*$) learn about noise. For each value of the parameter p_θ, the threshold τ_f^* is uniquely defined. Thus, Proposition 3.3 not only establishes the specialization of investors but also describes how they specialize.

In general, equations (3.16) and (3.17) cannot be resolved for $\tau^*_{s\theta,i}$ and $\tau^*_{sf,i}$ and only indirectly characterize the optimal precisions of the signals. Nevertheless, they still allow us to identify the relation between price informativeness and incentives of investors to acquire additional information. Taking derivatives with respect to p_θ, we get that

$$\frac{\partial \tau^*_{s\theta,i}}{\partial p_\theta} = -\frac{2p_\theta \tau_{f,i}}{(p_\theta^2\tau_{f,i} + \tau_\theta + \tau^*_{s\theta,i})^2}\left[\frac{2\gamma}{\tau_\theta^2}c''\left(\frac{\tau^*_{s\theta,i}}{\tau_\theta}\right) + \frac{1}{(p_\theta^2\tau_{f,i} + \tau_\theta + \tau^*_{s\theta,i})^2}\right]^{-1} > 0, \tag{3.22}$$

$$\frac{\partial \tau^*_{sf,i}}{\partial p_\theta} = \frac{2p_\theta \tau_{f,i}}{(p_\theta^2(\tau_{f,i} + \tau^*_{sf,i}) + \tau_\theta)^2}\left[\frac{2\gamma}{\tau_{f,i}^2}c''\left(\frac{\tau^*_{sf,i}}{\tau_{f,i}}\right) + \frac{p_\theta^4}{(p_\theta^2(\tau_{f,i} + \tau^*_{sf,i}) + \tau_\theta)^2}\right]^{-1} < 0 \tag{3.23}$$

because $p_\theta < 0$ and $c'' \geq 0$. Equation (3.23) indicates that when the price is more informative, investors choose a lower precision of the signal about fundamentals, so the price and the exogenous signal are informational substitutes (Grossman and Stiglitz, 1980). However, equation (3.22) shows that a more informative price makes the information about noise more valuable to investors and they increase the precision of the signal $\tau^*_{s\theta,i}$. Thus, the price and signal about noise are informational complements.[5] The complementarity opens up a possibility for a positive feedback loop: if learning about noise makes the price more informative, investors rationally obtain more precise non-fundamental signals, trade on them, and further increase price informativeness. This effect is further discussed in Section 3.4.3.

The third part of Proposition 3.3 considers a specification with a linear capacity cost function. In this special case, the optimization problem from equation (3.15) admits a closed-form solution. In particular, when the marginal cost of the information processing capacity is relatively small ($c_0 \leq 1/(4\gamma)$), investors with $\tau_{f,i} > \tau_\theta/p_\theta^2$ learn about fundamentals and investors with $\tau_{f,i} < \tau_\theta/p_\theta^2$ learn about noise. In both cases, the optimal precisions $\tau^*_{sf,i}$ and $\tau^*_{s\theta,i}$ are linear functions of the prior precision $\tau_{f,i}$, and this is the main source of analytical tractability. When the marginal cost is higher ($1/(4\gamma) < c_0 < 1/(2\gamma)$), investors with intermediate information quality do not participate in information acquisition at all and only those who have relatively precise (relatively imprecise) prior information choose to receive signals about fundamentals (noise). When the marginal cost of learning capacity is high ($c_0 \geq 1/(2\gamma)$), investors abstain from acquiring information of any type. Also note that $\tilde{\tau}_{sf,i}$ ($\tilde{\tau}_{s\theta,i}$) in equation (3.19) is a decreasing (increasing) function of p_θ when $p_\theta < 0$, so again the price and the signal about fundamentals (noise) are informational substitutes (complements).

[5] This conclusion echoes the results of Goldstein and Yang (2014), who identify strategic complementarity in the acquisition of information about different fundamentals, and Avdis (2014), who demonstrate complementarities in information acquisition in the presence of two trading periods.

3.4. Analysis

In this section, we discuss several remarkable properties of the model equilibrium presented in Section 3.3. In particular, we investigate how endogenous learning changes the asymmetry of information among investors, how the possibility to learn about noise changes price informativeness, and how the degree of initial information asymmetry among investors affects the impact of learning about noise on the equilibrium.

3.4.1. Ex ante and Ex post Asymmetry of Information Across Investors

Our model assumes that the precision of initial information about fundamentals varies across investors, so there is an information asymmetry among them. The information acquisition strategy described in Proposition 3.3 changes the uncertainty of each investor about the payoff f and, hence, the distribution of information across investors and the degree of information asymmetry. To quantify the impact of learning, it is convenient to use the derivative $\partial\hat{\tau}_i/\partial\tau_{f,i}$. Indeed, consider two investors with prior precisions $\tau_{f,i}$ and $\tau_{f,j}$, where $\Delta\tau_f = \tau_{f,j} - \tau_{f,i}$ is small. In the first order approximation, the difference in the ex post precisions of these investors is $\hat{\tau}_{f,j} - \hat{\tau}_{f,i} \approx (\partial\hat{\tau}_i/\partial\tau_{f,i}) \times \Delta\tau_f$. Thus, when the derivative is below 1 (above 1) the difference in the ex post precisions is smaller (larger) than before learning and the asymmetry of information between the investors decreases (increases). Note that the derivative measures a local change in the asymmetry of information, that is, a relative change in information of investors who initially had comparable precisions of their priors. Thus, the asymmetry of information can increase for some investors (for a range of the prior precisions $\tau_{f,i}$) and decrease for others.

Proposition 3.4. *Assume that $c'(0) = 0$. Then,*

$$\frac{\partial\hat{\tau}_i}{\partial\tau_{f,i}} < 1 \;\; for \;\; \tau_{f,i} < \tau_f^*, \qquad \frac{\partial\hat{\tau}_i}{\partial\tau_{f,i}} > 1 \;\; for \;\; \tau_{f,i} > \tau_f^*, \tag{3.24}$$

where τ_f^* is defined in Proposition 3.3.

Proof. See Appendix.

Proposition 3.4 characterizes the derivative $\partial\hat{\tau}_i/\partial\tau_{f,i}$ in the equilibrium described in the second part of Proposition 3.3.[6] Because investors with $\tau_{f,i} < \tau_f^*$ learn about noise and investors with $\tau_{f,i} > \tau_f^*$ learn about fundamentals, Proposition 3.4 implies that learning about fundamentals increases the asymmetry of information, whereas learning about noise decreases it.

To understand the intuition behind this effect, consider two investors with initial precisions $\tau_{f,i}$ and $\tau_{f,j}$, where $\tau_{f,i} < \tau_{f,j}$, who both learn either about fundamentals or noise. In general, an investor with more precise prior information can use his informational advantage in two different ways: i) increase the quality of the signal about the asset payoff by exploiting the increasing returns to initial information implied by the entropy-based learning technology or ii) decrease the quality of the obtained signal and save on the cost of information processing capacity. For those investors who find it optimal to learn about noise, the first option is unavailable because their learning efficiency is determined by the precision of their information about noise, not fundamentals, which is the same for all investors.[7] As a result, the investor with a higher precision of initial information $\tau_{f,j}$ saves more on learning compared to the investor with the precision $\tau_{f,i}$ and the asymmetry of information between the investors decreases.

The outcome is different when both investors decide to receive additional signals about fundamentals. In this case, both effects can be at work and the investor with the more precise prior faces a non-trivial tradeoff between the consequences of increasing and decreasing the precision of his signal. Proposition 3.4 shows that under relatively general conditions the effect of the increasing returns to initial information dominates and the investor who is more informed from the beginning acquires a more precise signal. This increases the heterogeneity in information precisions among those who learn about fundamentals and makes the asymmetry of information more pronounced.

Proposition 3.4 establishes the change in the asymmetry of information due to learning for a wide class of cost functions $c(K)$ with $c'(0) = 0$. However, the latter condition is not crucial and the result is even more transparent when the cost function is linear and the information acquisition strategies can be found in a closed form as described in the third part of Proposition 3.3. In this case, the

[6]Note that equation (3.8) for $\hat{\tau}_i$, together with equations (3.16) and (3.17), implies that the derivative $\partial\hat{\tau}_i/\partial\tau_{f,i}$ is well defined.

[7]Recall that information about noise is assumed to be short-lived and the uncertainty of all investors about noise is determined by the same unconditional prior.

ex post precision of investors' information about fundamentals is

$$\hat{\tau}_i = \begin{cases} \frac{1}{2\gamma c_0}\tau_{f,i}, & \text{if } \tau^*_{sf,i} > 0,\ \tau^*_{s\theta,i} = 0; \\ \frac{1}{2\gamma c_0}\frac{\tau_\theta}{p_\theta^2}, & \text{if } \tau^*_{sf,i} = 0,\ \tau^*_{s\theta,i} > 0. \end{cases} \tag{3.25}$$

Again, equation (3.25) shows that learning about fundamentals and learning about noise have different effects on the asymmetry of information among investors: the precisions of those who choose to learn about fundamentals become more diverse (their initial precisions are rescaled by the factor $1/(2\gamma c_0) > 1$) but those who decide to learn about noise end up with identical ex post precisions even though they have diverse prior precisions $\tau_{f,i}$. The latter implies that the initial information advantage is used to decrease the precision of the obtained signal about noise and reduce the cost of learning.

Finally, note that our results pertain to a local change in information asymmetry, that is, the asymmetry among those investors who initially are comparably informed and who follow similar information acquisition strategies. The effect of learning on the asymmetry of information among all investors is ambiguous: it can increase when many investors specialize in learning about fundamentals or decrease when many investors learn about noise. The latter possibility challenges the conclusion of Van Nieuwerburgh and Veldkamp (2009) that learning amplifies information asymmetry and emphasizes the importance of considering the possibility to learn about noise.

3.4.2. Price Informativeness

The possibility to learn about noise affects the equilibrium price function and, hence, the informativeness of prices about fundamentals. In general, there are two opposite effects. On the one hand, some investors who would learn about fundamentals in the absence of an opportunity to receive a signal about noise switch to learning about noise when the opportunity appears. As a result, the precision of their information about fundamentals decreases, they trade less aggressively on it, and the informativeness of the price decreases. On the other hand, an additional signal about noise allows an investor to extract more information about fundamentals from the price and decrease uncertainty about them. This makes the demand for the asset more sensitive to their information about fundamentals, so investors trade more aggressively on it and the informativeness of the price increases.

To make the analysis concrete and as transparent as possible, we only consider the case with the linear capacity cost function $c(K) = c_0 K$, where $c_0 \leq 1/(4\gamma)$. As follows from Proposition 3.3, in this case each investor learns either about fundamentals or noise and the optimal precisions of the signals $\tau^*_{sf,i}$ and $\tau^*_{s\theta,i}$ are piecewise-linear functions of $\tau_{f,i}$ given by equations (3.19) and (3.20). The substitution of these functions into equation (3.9) yields an equation for the parameter p_θ of the price function:

$$\frac{p_\theta}{2\gamma c_0}\left(\int\limits_{\left\{i:\, \tau_{f,i}\geq\frac{\tau_\theta}{p_\theta^2}\right\}} \tau_{f,i} d\mu(i) + \frac{\tau_\theta}{p_\theta^2} I\left(i:\ \tau_{f,i} \leq \frac{\tau_\theta}{p_\theta^2}\right)\right) - \frac{\tau_\theta}{p_\theta} + \gamma = 0, \tag{3.26}$$

where $I(\,\cdot\,)$ denotes the measure of the investors whose initial precisions $\tau_{f,i}$ satisfy the stated condition. Note that equation (3.26) is nonlinear and in general may have multiple solutions. This observation is in line with Ganguli and Yang (2009), who also recognize the possibility of equilibrium multiplicity at the information acquisition stage when investors receive multidimensional signals.

To find the impact of learning about noise on price informativeness, we compare price functions in the economies with and without the possibility to learn about noise. The latter specification serves as a benchmark model and the rational expectations equilibrium in it is described by Proposition 3.5.

Proposition 3.5. *Without the opportunity to learn about noise, there can exist not more than one equilibrium in which the price is given by*

$$p^0 = -\frac{\gamma\bar{\theta}}{\bar{\tau}^0} + f + p^0_\theta \theta, \tag{3.27}$$

where

$$\bar{\tau}^0 = \bar{\tau}_f + \bar{\tau}^*_{sf} + \frac{\tau_\theta}{(p^0_\theta)^2}, \qquad \bar{\tau}^*_{sf} = \int_{\mathcal{I}} \tau^*_{sf,i} d\mu(i). \tag{3.28}$$

The optimal information acquisition strategy of an investor with the prior precision $\tau_{f,i}$ is

$$\tau^*_{sf,i} = \max\left(h^{-1}\tau_{f,i} - \frac{\tau_\theta}{(p^0_\theta)^2}, 0\right) \tag{3.29}$$

and the equilibrium parameter p_θ^0 *solves the following equation:*

$$p_\theta^0 \left(\bar{\tau}_f + h^{-1} \int\limits_{\left\{i:\, \tau_{f,i} \geq \frac{\tau_\theta h}{(p_\theta^0)^2}\right\}} \tau_{f,i} d\mu(i) + \frac{\tau_\theta}{(p_\theta^0)^2} I \left(i :\, \tau_{f,i} \leq \frac{\tau_\theta h}{(p_\theta^0)^2} \right) \right) - \frac{\tau_\theta}{p_\theta^0} + \gamma = 0, \tag{3.30}$$

where

$$h = \left(\frac{1}{2\gamma c_0} - 1 \right)^{-1}.$$

Proof. See Appendix.

Proposition 3.5 shows that although the price function in the benchmark economy has exactly the same functional form as in the main setting, the parameter p_θ^0 is determined by a different equation, which in contrast to equation (3.26) can have only one solution. Not surprisingly, the optimal information acquisition strategy in the benchmark case is also different: investors with the precision of initial information above a certain threshold ($\tau_\theta h/(p_\theta^0)^2$) receive a signal about fundamentals, whereas the investors below the threshold do not learn at all.

The next proposition compares the equilibrium parameters p_θ and p_θ^0 in the models with and without learning about noise and states one of the most important results of this chapter.

Proposition 3.6. *If* p_θ *and* p_θ^0 *are solutions to equations (3.26) and (3.30), respectively, then* $p_\theta > p_\theta^0$.

Proof. See Appendix.

Proposition 3.6 implies that in any equilibrium in the model with learning about noise the price is more informative about fundamentals than in the benchmark model. Indeed, given that the parameters p_θ and p_θ^0 are negative, Proposition 3.6 implies that $Var[f|p]^{-1} = \tau_\theta/p_\theta^2 > \tau_\theta/(p_\theta^0)^2 = Var[f|p^0]^{-1}$. Note that the statement of the proposition pertains to all possible equilibria that can exist in the model with learning about noise and the uniqueness of the equilibrium in the benchmark model simplifies the proof.

Proposition 3.6 has another interesting implication. Recall that in the main economy the agents learn about fundamentals (noise) when $\tau_{f,i} > \tau_\theta/p_\theta^2$ ($\tau_{f,i} < \tau_\theta/p_\theta^2$) and in the benchmark economy they learn about fundamentals when $\tau_{f,i} > \tau_\theta h/(p_\theta^0)^2$. Given that the condition $c_0 \leq 1/(4\gamma)$ implies $0 < h \leq 1$ and Proposition 3.6 together with the negative signs of p_θ and p_θ^0 yields $p_\theta^2 < (p_\theta^0)^2$, the thresholds appear to be unambiguously ordered as $\tau_\theta h/(p_\theta^0)^2 < \tau_\theta/p_\theta^2$. Thus, some investors who would learn about fundamentals in the benchmark model decide to learn about noise in the main model. This observation illustrates that in the main model investors rationally exploit their comparative informational advantages: those who are relatively less informed about fundamentals optimally acquire information about noise leveraging the knowledge of its unconditional distribution and use the price to infer fundamental information. Trading on the fundamental information obtained in this way increases price informativeness.

3.4.3. Information Distribution Across Investors and Learning about Noise

Finally, we investigate how the effect of learning about noise varies with the degree of information asymmetry among investors. As in the previous section, we focus on the specification with the linear cost function $c(K) = c_0 K$ with $c_0 \leq 1/(4\gamma)$ in which the parameters p_θ and p_θ^0 solve equations (3.26) and (3.30) in the cases with and without learning about noise, respectively.

To quantify the heterogeneity in the investors' initial information about fundamentals, we use the parametrization of the precisions $\tau_{f,i}$ proposed in Marmora (2014). In particular, we assume that the investors are labeled by rational numbers in the unit interval, that is, $i \in Q([0,1])$ and $\tau_{f,i} = \bar{\tau}_f \alpha i^{\alpha-1}$, where $\bar{\tau}_f > 0$ and the parameter $\alpha \geq 1$ measures inequality in the distribution of information across investors. When $\alpha = 1$, all investors have the same precision of information about fundamentals; when α is high, a very small fraction of investors possess highly precise information about fundamentals and others know almost nothing about them. Note that for any value of α the total amount of information that the investors have is the same and $\int_0^1 \tau_{f,i} d\mu(i) = \bar{\tau}_f$, so the parameter $\bar{\tau}_f$ coincides with the aggregate precision of prior information, which is disentangled from the distribution of information.

To illustrate the discussed properties of the equilibrium and get new insights, we fix the model parameters and find the equilibrium in the specifications with and without learning about noise numerically. In particular, we set $\gamma = 5$ and $c_0 = 0.03$, where the value of c_0 is largely determined by the necessity to satisfy the condition $c_0 \leq 1/(4\gamma)$. The aggregate precisions of information about fundamentals and noise are assumed to be identical and set as $\bar{\tau}_f = \tau_\theta = 1$. This choice ensures that fundamentals and noise have comparable effects on the price and neither component dominates simply due to the choice of the parameters. The value of the parameter α is considered in the range from 1 to 20.

The parameters p_θ and p_θ^0 that solve equations (3.26) and (3.30) are plotted in Panel A of Figure 3.1 as functions of α. These graphs provide three observations. First, for all values of α we have $p_\theta^0 < p_\theta$ when some investors learn about noise, so the price in the model with learning about noise is unambiguously more informative. This is an illustration of Proposition 3.6. Second, in both specifications the price informativeness increases with the inequality in the distribution of information across investors. Indeed, keeping the aggregate precision of the fundamental information $\bar{\tau}_f$ fixed, inequality in the distribution of information implies that a small fraction of investors is well informed and, because the entropy-based learning technology implies increasing returns to initial information, such investors can cheaply obtain fundamental signals of a high quality. Trading on this information, they make the price more informative. Third, and most importantly, the relation between price informativeness and α is much stronger in the presence of learning about noise. For example, when $\alpha = 20$ we find that $p_\theta = -0.97$, whereas $p_\theta^0 = -1.55$. In contrast, when α is close to 1 both p_θ and p_θ^0 are close to -1.68. Intuitively, the opportunity to learn about noise substantially increases the price informativeness when α is high due to two effects. First, because high α implies that the vast majority of investors have poor prior information about fundamentals, they largely benefit from learning about noise, extracting information about the asset payoff from the price, and trading on it. This makes the price more informative. Second, a more informative price further increases the value of information about noise due to the complementarity of information from the price and signal about noise discussed in Section 3.3.3. As a result, the investors choose to receive a more precise signal about noise and trading on it further increases the price informativeness.

According to Proposition 3.3, in the model with learning about noise the investors with $\tau_{f,i} > \tau_\theta/p_\theta^2$ ($\tau_{f,i} < \tau_\theta/p_\theta^2$) acquire fundamental (non-fundamental) information. Panel B of Figure 3.1 plots the threshold $i_* = \{i : \tau_{f,i} = \tau_\theta/p_\theta^2\}$, which for the assumed distribution of precisions $\tau_{f,i}$ is equal to $i_* = (\tau_\theta/(\alpha\bar{\tau}_f p_\theta^2))^{1/(\alpha-1)}$. As a benchmark, the panel also presents the threshold $i_*^0 = \{i : \tau_{f,i} = \tau_\theta h/(p_\theta^0)^2\}$ from Proposition 3.5, which separates those who learn about fundamentals and those who do not acquire information at all in the benchmark model. For the assumed $\tau_{f,i}$ it is equal to $i_*^0 = (\tau_\theta h/(\alpha\bar{\tau}_f(p_\theta^0)^2))^{1/(\alpha-1)}$.

The graphs illustrate two facts. First, they demonstrate that the inequality in the distribution of information reduces the number of investors who receive signals about fundamentals and the effect is quite strong: the fraction of those who learn about fundamentals decreases from 100% for $\alpha = 1$ to 14% (22%) in the main (benchmark) specification for $\alpha = 20$. This is a direct consequence of increasing returns to prior information implied by the entropy reduction learning technology: when α is high, only a small number of investors have fundamental information that is sufficiently precise to justify acquisition of additional signals about the payoff. Second, the comparison of i_* and i_*^0 shows that $i_* \geq i_*^0$ for any α, so some investors who would learn about fundamentals in the benchmark model switch to learning about noise when such a possibility appears. This effect was discussed in Section 3.4.2.

3.5. Conclusion

The main objective of this chapter is to study the effect of learning about noise when investors are endowed with fundamental information of diverse quality. We find that even though investors can obtain two separate signals about fundamentals and noise, under quite general conditions no one chooses to receive both of them: those who have high quality prior information learn more about fundamentals, whereas those who are endowed with information of low quality learn about noise. Moreover, those investors who acquire signals about fundamentals amplify their initial information advantage, so the asymmetry of information among such investors increases. In contrast, the asymmetry of information among those who learn about noise decreases because they use their prior knowledge about fundamentals to decrease the cost of learning (those who know more about fundamentals learn less about noise). Also, we find that the possibility to learn about noise increases the informativeness of prices and the effect is particularly strong when the dispersion of the prior information precisions across investors is large.

Our analysis can be potentially extended in several directions. Our model has one trading period and investors learn about noise only because this helps them to extract more precise information about fundamentals from the price. In the presence of multiple trading periods, the information acquisition by investors is likely to be more complicated because the possibility to exploit price deviations produced by noise traders in multiple periods opens up new speculative opportunities. Thus, investors may have additional incentives to learn about noise and can do this either directly (e.g., Froot et al., 1992) or indirectly by getting more precise information about fundamentals and extracting information about noise from prices (e.g., Avdis, 2014). Unfortunately, multiple periods in the model with heterogeneous precisions of the prior fundamental information technically complicate the aggregation of investors' asset demands opening up the possibility for existence of the "forecasting the forecasts of others" problem (e.g., Townsend, 1983; Makarov and Rytchkov, 2012) and, thus, make the analysis much less tractable.

Also, it may be interesting to explore the optimal multidimensional information acquisition and price informativeness in the presence of markets for financial information, that is, when investors can sell fundamental and non-fundamental information. In the Kyle (1985) framework, Foucault and Lescourret (2003) examine the effects of sharing fundamental and non-fundamental information between two speculators and Cheynel and Levine (2012) study the sale of non-fundamental information and its consequences. A similar analysis in the rational expectations framework with diversely informed investors and endogenous learning could be an interesting direction for future research.

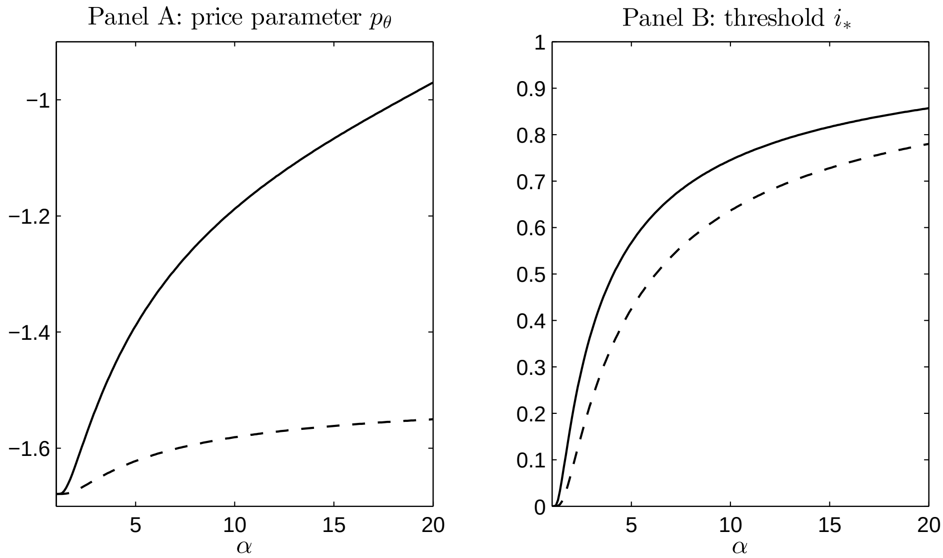


FIGURE 3.1: **Price parameter p_θ and information acquisition threshold i_*.** Panel B presents the information acquisition thresholds $i_* = (\tau_\theta/(\alpha\bar{\tau}_f p_\theta^2))^{1/(\alpha-1)}$ (solid line) and $i_*^0 = (\tau_\theta h/(\alpha\bar{\tau}_f (p_\theta^0)^2))^{1/(\alpha-1)}$ (dashed line). All variables are plotted as functions of the information distribution inequality parameter α in a model with a linear cost function and the prior precisions of the fundamental information $\tau_{f,i} = \bar{\tau}_f \alpha i^{\alpha-1}$. $\gamma = 5$, $c_0 = 0.03$, $\bar{\tau}_f = 1$, $\tau_\theta = 1$.

BIBLIOGRAPHY

Admati, Anat (1985), "A Noisy Rational Expectations Equilibrium for Multi-Asset Securities." *Econometrica*, 53(3), 629–658.

Admati, Anat and Pauli Pfleiderer (1986), "A Monopolistic Market For Information." *Journal of Economic Theory*, 39, 400–438.

Admati, Anat and Pauli Pfleiderer (1990), "Direct and Indirect Sales of Information." *Econometrica*, 58(4), 901–928.

Allen, Franklin (1990), "The Market For Information and the Origin of Financial Intermediation." *Jorunal of Financial Intermediation*, 1(1), 3–30.

Avdis, Efstathios (2014), "Information Trade-offs in Dynamic Financial Markets." Working paper, University of Alberta.

Banerjee, Snehal, Jesse Davis, and Naveen Gondhi (2014), "When Transparency Improves, Must Prices Reflect Fundamentals Better?" Working paper, Northwestern University.

Banerjee, Snehal and Brett S. Green (2014), "Signal or Noise? Uncertainty and Learning About Whether Other Traders are Informed." Working paper, Northwestern University.

Barber, Brad and Terrance Odean (2008), "All that Glitters: The Effect of Attention on the Buying Behavior of Individual and Institutional Investors." *Review of Financial Studies*, 21, 785–818.

Barlevy, Gadi and Pietro Veronesi (2000), "Information Acquisition in Financial Markets." *Review of Economic Studies*, 67, 79–90.

Brennen, Michael and Henry Cao (1997), "International Portfolio Investment Flows." *Journal of Finance*, 52(5), 1851–1879.

Cheynel, Edwige and Carolyn B. Levine (2012), "Analysts' Sale and Distribution of Non Fundamental Information." *Review of Accounting Studies*, 17, 352–388.

Constantinides, George M. and Darrell Duffie (1996), "Asset Pricing with Heterogeneous Consumers." *Journal of Political Economy*, 104, 219–240.

Coval, Joshua and Tobias Moskowitz (1999), "Home Bias at Home: Local Equity Preference in Domestic Portfolios." *Journal of Finance*, 54(6), 42045–2073.

Cover, Thomas M. and Joy A. Thomas (2006), *Elements of Information Theory*, 2nd edition. Wiley-Interscience.

Da, Zhi, Joseph Engelberg, and Pengjie Gao (2011), "In Search of Attention." *Journal of Finance*, 66(5), 1461–1499.

Diamond, Douglas and Robert Verrecchia (1991), "Disclosure, Liquidity, and the Cost of Capital." *Journal of Finance*, 46(4), 1325–1359.

Diamond, Douglas W. and Robert E. Verrecchia (1981), "Information Aggregation in a Noisy Rational Expectations Economy." *Journal of Financial Economics*, 221–235.

Easley, David and Maureen O'Hara (2004), "Information and the Cost of Capital." *Journal of Finance*, 58(4), 1553–1583.

Easley, David, Maureen O'Hara, and Liyan Yang (2014), "Opaque Trading, Disclosure, and Asset Prices: Implications For Hedge Fund Regulation." *Review of Financial Studies*, 27, 1190–1237.

Engelberg, Joseph, Caroline Sasseville, and Jared Williams (2012), "Market Madness? the Case of Mad Money." *Management Science*, 58(2), 351–364.

Engelberg, Joseph and Christopher Parsons (2011), "The Causal Impact of Media in Financial Markets." *Journal of Finance*, 66(1)), 67–97.

Epstein, Larry and Stanley Zin (1989), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns; a Theoretical Framework." *Econometrica*, 57(4), 937–969.

Fang, Lily and Joel Peress (2009), "Media Coverage and the Cross-section of Stock Returns." *Journal of Finance*, 64(5), 2021–2052.

Feldman, Mark and Christian Gilles (1985), "An Expository Note on Individual Risk Without Aggregate Uncertainty." *Journal of Economic Theory*, 35, 26–32.

Foucault, Thierry and Laurence Lescourret (2003), "Information sharing, liquidity and transaction costs." *Finance*, 24, 45–78.

Froot, Kenneth A., David S. Scharfstein, and Jeremy C. Stein (1992), "Herd on the Street: Informational Inefficiencies in a Market With Short-Term Speculation." *Journal of Finance*, 47, 1461–1484.

Ganguli, Jayant V. and Liyan Yang (2009), "Complementarities, Multiplicity, and Supply Information." *Journal of the European Economic Association*, 7, 90–115.

Gao, Feng, Fengming Song, and Jun Wang (2013), "Rational Expectations Equilibrium With Uncertain Proportion of Informed Traders." *Journal of Financial Markets*, 16, 387–413.

Goldstein, Itay and Liyan Yang (2014), "Information Diversity and Complementarities in Trading and Information Acquisition." Forthcoming in *Journal of Finance*.

Grinnblatt, Mark and Matti Keloharju (2001), "How Distance, Language, and Culture Influences Stockholdings and Trades." *Journal of Finance*, 56(3), 1053–1074.

Grossman, Sanford (1976), "On the Efficiency of Competitive Stock Markets Where Agents Have Diverse Information." *Journal of Finance*, 31, 573–585.

Grossman, Sanford and Joseph Stiglitz (1980), "On the Impossibility of Informationally Efficient Markets." *The American Economic Review*, 70(3), 393–408.

Hamilton, James T. (2004), *All the News That's Fit To Sell: How the Market Transforms Information Into News.* Princeton University Press.

He, Hua and Jiang Wang (1995), "Differential Informational and Dynamic Behavior of Stock Trading Volume." *Review of Financial Studies*, 8, 919–972.

Hellwig, Martin (1980), "On the Aggregation of Information in Competitive Markets." *Journal of Economic Theory*, 22(3), 477–498.

Hong, Harrison and Sven Rady (2002), "Strategic Trading and Learning About Liquidity." *Journal of Financial Markets*, 5, 419–450.

Hong, Harrison and Jeremy Stein (2007), "Disagreement and the Stock Market." *Journal of Economic Perspectives*, 21(2), 109–128.

Huberman, Gur (2001), "Familiarity Breeds Investment." *Review of Financial Studies*, 14(3), 659–680.

Huberman, Gur and Tomer Regev (2001), "Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar." *Journal of Finance*, 56(1), 387–396.

Hughes, John, Jing Liu, and Jun Liu (2007), "Information Asymmetry, Diversification, and the Cost of Capital." *Accounting Review*, 82(3), 702–729.

Ivkovic, Zoran, Clemens Sialm, and Scott Weisbenner (2008), "Portfolio Concentration and the Performance of Individual Investors." *Journal of Financial and Quantitative Analysis*, 43(3), 613–656.

Judd, Kenneth L. (1985), "The Law of Large Numbers With a Continuum of iid Random Variables." *Journal of Economic Theory*, 35, 19–25.

Kacperczyk, Marcin, Jaromir Nosal, and Luminita Stevens (2014), "Investor Sophistication and Capital Income Inequality." Working Paper.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp (2012), "Rational Attention Allocation Over the Business Cycle." Working paper.

Kim, Oliver and Robert Verrecchia (1997), "Pre-Announcement and Event-Period Private Information." *Journal of Accounting and Economics*, 24(3), 395–419.

Kreps, David and Evan Porteus (1978), "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica*, 46(1), 185–200.

Kyle, Albert (1985), "Continuous Auctions and Insider Trading." *Econometrica*, 53(6), 1315–1336.

Lambert, Richard, Christian Leuz, and Robert Verrecchia (2012), "Information Precision, Information Asymmetry, and the Cost of Capital." *Review of Finance*, 16, 1–29.

Madrigal, Vicente (1996), "Non-Fundamental Speculation." *Journal of Finance*, 51, 553–578.

Makarov, Igor and Oleg Rytchkov (2012), "Forecasting the Forecasts of Others: Implications for Asset Pricing." *Journal of Economic Theory*, 147, 941–966.

Manela, Asaf (2014), "The Value of Diffusing Information." *Journal of Financial Economics*, 111(5), 181–199.

Manzano, Carolina and Xavier Vives (2011), "Public and Private Learning From Prices, Strategic Substitutability and Complementarity, and Equilibrium Multiplicity." *Journal of Mathematical Economics*, 47, 346–369.

Marmora, Paul (2014), "The Distribution of Information, the Market for Financial News, and the Cost of Capital." Working paper, Temple University.

Merton, Robert (1987), "A Simple Model of Capital Market Equilibrium with Incomplete Information." *Journal of Finance*, 42(3), 483–510.

Miller, Edward (1977), "Risk, Uncertainty, and Divergence of Opinion." *Journal of Finance*, 32, 1151–1168.

Mondria, Jordi (2010), "Portfolio Choice, Attention Allocation, and Price Comovement." *Journal of Economic Theory*, 145(5), 1837–64.

Mondria, Jordi and Thomas Wu (2010), "The Puzzling Evolution of the Home Bias, Information Processing, and Financial Openess." *Journal of Economic Dynamics and Control*, 34(5), 876–896.

Mondria, Jordi and Thomas Wu (2012), "Asymmetric Attention and Stock Returns." Working Paper.

O'Hara, Maureen (2003), "Presidential Address: Liquidity and Price Discovery." *Journal of Finance*, 584, 1335–1354.

Park, Cheol (2010), "Acquisition of Non-Fundamental Information in the Kyle Model of Speculation." *Asian Review of Financial Research*, 23, 159–212.

Peng, Lin (2005), "Learning With Information Capacity Constraints." *Journal of Financial and Quantitative Analysis*, 40, 307–329.

Peng, Lin and Wei Xiong (2006), "Investor Attention, Overconfidence, and Category Learning." *Journal of Financial Economics*, 80, 563–602.

Peress, Joel (2004), "Wealth, information acquisition, and portfolio choice." *Review of Financial Studies*, 17, 879–914.

Peress, Joel (2010), "The Tradeoff between Risk Sharing and Information Production in Financial Markets." *The Journal of Economic Theory*, 145(1), 124–155.

Peress, Joel (2014), "The Media and the Diffusion of Information in Financial Markets: Evidence from Newspaper Strikes." *Journal of Finance*.

Sen, Amartya (1973), *On Economic Inequality*. Clarendon Press, Oxford.

Sengupta, Partha (1998), "Corporate Disclosure Quality and the Cost of Debt." *The Accounting Review*, 73(4), 459–474.

Sims, Christopher (2003), "Implications of Rational Inattention." *Journal of Monetary Economics*, 50, 665–690.

Sims, Christopher (2006), "Rational Inattention: Beyond the Linear-Quadratic Case." *American Economic Review Papers and Proceedings*, 96(2), 158–163.

Solomon, David (2012), "Selective Publicitity and Stock Prices." *Journal of Finance*, 67(2), 599–638.

Tetlock, Paul C (2010), "Does Public Financial News Resolve Asymmetric Information?" *Review of Financial Studies*, 23, 3520–557.

Tetlock, Paul C (2011), "All the News That's Fit to Reprint: Do Investors React to Stale Information?" *Review of Financial Studies*, 24(5), 1481–1512.

Townsend, Robert M. (1983), "Forecasting the Forecasts of Others." *Journal of Political Economy*, 91, 546–588.

Van Nieuwerburgh, Stign and Laura Veldkamp (2009), "Information Immobility and the Home Bias Puzzle." *Journal of Finance*, 64(3).

Van Nieuwerburgh, Stign and Laura Veldkamp (2010), "Information Acquisition and Under-Diversification." *Review of Economic Studies*, 77(2).

Veldkamp, Laura (2006), "Media Frenzies in Markets for Financial Information." *American Economic Review*, 96(3), 577–601.

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

Verrecchia, Robert (1982), "Information Acquisition in a Noisy Rational Expectations Economy." *Econometrica*, 50(6), 1415–1430.

Vives, Xavier (1995), "The Speed of Information Revelation in a Financial Market Mechanism." *Journal of Economic Theory*, 67, 178–204.

Vives, Xavier (2010), *Information and Learning in Markets: The Impact of Market Microstructure*. Princeton University Press.

Yu, Fan (1999), "What is the Value of Knowing Uninformed Trades?" *Economics Letters*, 64, 87–98.

APPENDIX

Proof of Proposition 2.1. The proof is almost identical to the one found in Hellwig (1980) and Admati (1985), but with independent prior signals. First conjecture that p is of the form:

$$p = g_1 + g_2 f + g_3 x. \tag{31}$$

Next, use the market-clearing conditions to solve for g_1, g_2, and g_3:

$$\bar{x} + x = \int_0^1 q_i di. \tag{32}$$

Substituting optimal asset demand (2.14) gives:

$$\bar{x} + x = \int_0^1 \frac{\widehat{f}_i - p}{\rho \widehat{\tau}_i^{-1}} di. \tag{33}$$

Substituting in equation (2.11) for investor i's posterior mean $\widehat{f}_i$ gives:

$$\bar{x} + x = \int_0^1 \frac{\widetilde{\tau}_i \widetilde{f}_i + l_i \eta_i s_i + \zeta z(p)}{\rho \widehat{\tau}_i^{-1} (\widetilde{\tau}_i + l_i \eta_i + \zeta)} - \frac{p}{\rho \widehat{\tau}_i^{-1}} di. \tag{34}$$

Since $\widehat{\tau}_i = \widetilde{\tau}_i + l_i \eta_i + \zeta$, this reduces to:

$$\begin{aligned} \rho(\bar{x} + x) &= \int_0^1 \widetilde{\tau}_i \widetilde{f}_i + l_i \eta_i s_i + \zeta z(p) - \frac{p}{\widehat{\tau}_i^{-1}} di \\ &= \int_0^1 \widetilde{\tau}_i \widetilde{f}_i + l_i \eta_i s_i di + \zeta z(p) - \int_0^1 p \left(\widetilde{\tau}_i + l_i \eta_i + \zeta \right) di. \end{aligned} \tag{35}$$

From (31), the public signal of f that each investor gleans from price is $\frac{p-g_1}{g_2}$. Substituting this in for $z(p)$ gives:

$$\rho(\bar{x}+x) = \int_0^1 \widetilde{\tau}_i \widetilde{f}_i + l_i \eta_i s_i di + \frac{\zeta p - \zeta g_1}{g_2} - \int_0^1 p\left(\widetilde{\tau}_i + l_i \eta_i + \zeta\right) di. \tag{36}$$

Since $\widetilde{f}_i$ and s_i equal f in expectation:

$$\begin{aligned}\rho(\bar{x}+x) &= \int_0^1 \widetilde{\tau}_i + l_i \eta_i di f + \frac{\zeta p - \zeta g_1}{g_2} - \int_0^1 p\left(\widetilde{\tau}_i + l_i \eta_i + \zeta\right) di \\ &= \left(\bar{\tau} + \int_0^1 l_i \eta_i di\right) f - \frac{\zeta g_1}{g_2} + p\left(\frac{\zeta}{g_2} - \bar{\tau} - \int_0^1 l_i \eta_i di - \zeta\right).\end{aligned} \tag{37}$$

Let $\theta = \bar{\tau} + \int_0^1 l_i \eta_i di$, which is the average precision of the priors and private signals. Substituting θ into equation (37) and rearranging gives:

$$p\left(\frac{\zeta}{g_2} - \theta - \zeta\right) = \rho\bar{x} + \frac{\zeta g_1}{g_2} - \theta f + \rho x. \tag{38}$$

Next, solve for g_1, g_2, and g_3 by matching the coefficients of (31) with the coefficients of (38). First, solving for g_2:

$$g_2 = -\left(\frac{\zeta}{g_2} - \theta - \zeta\right)^{-1} \theta = 1. \tag{39}$$

Solving for g_3:

$$g_3 = \left(\frac{\zeta}{g_2} - \theta - \zeta\right)^{-1} \rho = -\frac{\rho}{\theta}. \tag{40}$$

Solving for g_1:

$$g_1 = \left(\frac{\zeta}{g_2} - \theta - \zeta\right)^{-1} \left(\rho\bar{x} + \frac{\zeta g_1}{g_2}\right) = -\frac{\rho\bar{x}}{\theta+\zeta}. \tag{41}$$

Substituting these coefficients into equation (31) yields:

$$p = f - \left(\frac{\rho\bar{x}}{\theta+\zeta} + \frac{\rho x}{\theta}\right), \tag{42}$$

which is exactly equation (2.15). Finally, each agent can invert this price function into a normally distributed signal about f, where the signal's precision is:

$$\zeta = \left(\frac{\theta}{\sigma_x \rho}\right)^2, \tag{43}$$

which completes the proof.

Deriving Period One Utility. The period two objective function is:

$$U_{i,2} = \rho E_{i,2}(W_{i,3}) - \frac{\rho^2}{2}\mathrm{Var}_{i,2}(W_{i,3}), \tag{44}$$

subject to:

$$W_{i,3} = W_{i,0} + q_i(f - p) + \pi_i - cl_i. \tag{45}$$

Substituting the terminal wealth equation (45) into the period two utility function (44) gives:

$$U_{i,2} = \rho E_{i,2}(W_{i,0} + q_i(f - p)) + \pi_i - cl_i) - \frac{\rho^2}{2}Var_{i,2}(W_{i,0} + q_i(f - p) + \pi_i - cl_i). \tag{46}$$

Prices are observed in period 2, $E_{i,2}(f) = \widehat{f}_i$, and $\mathrm{Var}_{i,2}(f) = \widehat{\tau}_i^{-1}$. Distributing period two expectation and variance operators gives:

$$U_{i,2} = W_{i,0} + \rho q_i(\widehat{f}_i - p) - \frac{\rho^2 q_i^2}{2\widehat{\tau}_i} + \pi_i - cl_i. \tag{47}$$

Next, substitute in optimal asset demand (2.14) to find period two indirect utility and simplify:

$$\begin{aligned} U_{i,2} &= W_{i,0} + \rho\left(\frac{\widehat{\tau}_i(\widehat{f}_i - p)}{\rho}\right)(\widehat{f}_i - p) - \frac{\rho^2}{2\widehat{\tau}_i}\left(\frac{\widehat{\tau}_i(\widehat{f}_i - p)}{\rho}\right)^2 + \pi_i - cl_i \\ &= W_{i,0} + \frac{\widehat{\tau}_i}{2}\left(\widehat{f}_i - p\right)^2 + \pi_i - cl_i. \end{aligned} \tag{48}$$

Signal and price realizations are unknown to investors in period one, so expectations must be taken over $(\widehat{f}_i - p)^2$ in order to derive period one utility:

$$E_{i,1}(U_{i,2}) = U_{i,1} = W_{i,0} + \frac{\widehat{\tau}_i}{2}E_{i,1}\left[\left(\widehat{f}_i - p\right)^2\right] + \pi_i - cl_i. \tag{49}$$

Since $E(X^2) = Var(X) + E(X)^2$:

$$E_{i,1}\left[(\widehat{f_i} - p)^2\right] = Var_{i,1}(\widehat{f_i} - p) + E_{i,1}(\widehat{f_i} - p)^2. \tag{50}$$

From equation (2.15):

$$E_{i,1}(\widehat{f_i} - p) = E_{i,1}\left(\widehat{f_i} - f - \frac{\rho\bar{x}}{\theta + \zeta} - \frac{\rho x}{\theta}\right). \tag{51}$$

Since $E_{i,1}(x) = 0$:

$$E_{i,1}(\widehat{f_i} - p) = -\frac{\rho\bar{x}}{\theta + \zeta}. \tag{52}$$

The period one expected return is the same across all agents. Next, using the Law of Total Variance:

$$\begin{aligned}\mathrm{Var}_{i,1}(\widehat{f_i} - p) &= \mathrm{Var}_{i,1}(f - p) - E_{i,1}(\mathrm{Var}_{i,2}(\widehat{f_i} - p)) \\ &= \mathrm{Var}_{i,1}\left(f - f - \frac{\rho\bar{x}}{\theta + \zeta} - \frac{\rho x}{\theta}\right) - E_{i,1}(\mathrm{Var}_{i,2}(\widehat{f_i} - p)).\end{aligned} \tag{53}$$

The terms θ, $\bar{x}$, and ζ are known in period one, while p is known in period two, implying:

$$\mathrm{Var}_{i,1}(\widehat{f_i} - p) = \frac{\rho^2\sigma_x^2}{\theta^2} - \widehat{\tau}_i^{-1}. \tag{54}$$

Substituting equations (54) and (52) into equation (50), and substituting this equation into the period one utility function (49) yields:

$$U_{i,1} = W_{i,0} + \frac{\widehat{\tau}_i}{2}\left(\frac{\rho^2\sigma_x^2}{\theta^2} - \widehat{\tau}_i^{-1} + \left(\frac{\rho\bar{x}}{\theta + \zeta}\right)^2\right) + \pi_i - cl_i. \tag{55}$$

Finally, dropping the constant term, substituting $\frac{\theta^2}{\rho^2\sigma_x^2}$ in for ζ by equation (2.15), and substituting in the equation for posterior precision delivers the period one problem captured by equation (2.19):

$$U_{i,1} = W_{i,0} + \frac{1}{2}\left(\left(\frac{\rho\sigma_x}{\theta}\right)^2 + \left(\frac{\rho\bar{x}}{\theta + \left(\frac{\theta}{\rho\sigma_x}\right)^2}\right)^2\right)(\widetilde{\tau}_i + l_i\eta_i + \zeta) + \pi_i - cl_i. \tag{56}$$

Proof of Proposition 2.2. The proof follows directly from the free entry assumption. If c^* is above average cost, another agent can enter the market for news, charge slightly below c^*, take the entire market and make a positive profit. If c^* is below average cost, any agent supplying news is making negative profit and thus strictly better off exiting the market. Finally, if c^* is priced at average cost and there is more than one supplier, then both suppliers must also be making negative profit and are strictly better off exiting.

Proof of Lemma 2.3. Lemma 2.3 is proven by contradiction. It will first be convenient to derive the (net) benefit of purchasing news to each agent, which is equal to expected utility conditional on purchasing minus expected utility conditional on not purchasing, given a particular cost c and average signal precision θ.

After substituting equation (2.10) into equation (2.19), expected utility conditional on $l_i = 1$ is:

$$U_{i,1} = W_{i,0} + R(\theta)\left(\widetilde{\tau}_i + \eta_i + \zeta\right) + \pi_i - c. \tag{57}$$

The expected utility conditional on $l_i = 0$ is:

$$U_{i,1} = W_{i,0} + R(\theta)\left(\widetilde{\tau}_i + \zeta\right) + \pi_i. \tag{58}$$

Taking the difference between equations (57) and (58) gives:

$$R(\theta)\eta_i - c. \tag{59}$$

Substituting in for η_i yields the net benefit of purchasing:

$$R(\theta)k\bar{\tau}\alpha i^{\alpha-1} - c. \tag{60}$$

Without loss of generality, assume that agent a purchases news in equilibrium, but agent b does not, where $b > a$. Then agent a must weakly prefer purchasing, whereas agent b must weakly prefer not purchasing:

$$R(\theta^*)k\bar{\tau}\alpha a^{\alpha-1} - c^* \geq 0, \tag{61}$$

$$R(\theta^*)k\bar{\tau}\alpha b^{\alpha-1} - c^* \leq 0. \tag{62}$$

Equations (61) and (62) imply that $a \geq b$, which is a contradiction. Therefore, an agent cannot purchase news in equilibrium unless all agents with a higher prior precision do the same. This in turn means that a given equilibrium fraction of purchasers λ^* must be the λ^* agents with the highest prior precision, i.e., agents from $1 - \lambda^*$ to 1, so that:

$$\theta(\lambda^*, \alpha) = \bar{\tau} + \int_{1-\lambda^*}^{1} \eta_i di. \tag{63}$$

Substituting in for η_i gives:

$$\theta(\lambda^*, \alpha) = \bar{\tau} + \int_{1-\lambda^*}^{1} k\bar{\tau}\alpha i^{\alpha-1} di, \tag{64}$$

which completes the proof.

Proof of Proposition 2.4. First, given a particular $\lambda > 0$, define the (net) benefit of purchasing news to the agent who has the least to gain from doing so:

$$B(\lambda, \alpha) = R\left(\theta(\lambda, \alpha)\right) \eta_{1-\lambda} - \frac{\chi}{\lambda}, \tag{65}$$

where news is priced at average cost from Proposition 2.2 and the purchaser who benefits the least is agent $1 - \lambda$ from Lemma 2.3. Substituting in signal precision of agent $1 - \lambda$ gives:

$$B(\lambda, \alpha) = R\left(\theta(\lambda, \alpha)\right) k\bar{\tau}\alpha(1 - \lambda)^{\alpha-1} - \frac{\chi}{\lambda}. \tag{66}$$

There are three cases:

If $B(\lambda, \alpha) < 0$ for all $\lambda \in (0, 1]$: Assume $\lambda^* > 0$. Then all agents from $1 - \lambda^*$ to 1 must weakly prefer purchasing. However, the benefit of purchasing to agent $1 - \lambda$ is strictly negative for any $\lambda > 0$ by assumption. Therefore, $\lambda^* > 0$ cannot be an equilibrium. By the same logic, if $\lambda^* = 0$, no agent can enter the market for news and charge a price that makes them positive profit. Therefore, $\lambda^* = 0$.

If $B(1, \alpha) \geq 0$: Assume $\lambda^* = 1$. Since the agent with the smallest prior precision (i.e., agent 0) weakly prefers to purchase, all other agents with a higher prior precision must as well. Furthermore, if any agent tries to enter the market for news and charge a small enough price to undermine the incumbent supplier, he must be charging less than χ and thus be making negative profit. Therefore $\lambda^* = 1$ is an equilibrium.

If $B(\lambda, \alpha) = 0$ for some $\lambda \in (0, 1)$: Assume that $\lambda^* \in (0, 1)$ and $B(\lambda^*, \alpha) = 0$. Then the marginal agent (agent $(1 - \lambda^*)$)is indifferent to purchasing by assumption, and all agents with high prior precisions strictly prefer doing so. Therefore, λ^* is an equilibrium for news demand. However, consider the case where $\frac{\partial B(\lambda^*, \alpha)}{\partial \lambda^*} > 0$. In this case, an agent can enter the market for news, charge $\frac{\chi}{\lambda^* + e}$ for an arbitrarily small $e > 0$ and get at least $\lambda^* + e$ to purchase, making a positive profit. Therefore, only a point where $\frac{\partial B(\lambda, \alpha)}{\partial \lambda} \leq 0$ can be an interior equilibrium for both the supply and demand for news, which completes the proof.

Proof of Proposition 2.5. Assume that the equilibrium fraction of purchasers is positive ($\lambda^* > 0$). First, note that $\eta_0 = 0$ when $\alpha > 1$. Therefore, an equilibrium where all agents purchase ($\lambda^* = 1$) can only occur when $\alpha = 1$, which in turn implies that $\lambda^* < 1$ for any incremental increase in α. Next, note that the cost of news c^* is strictly decreasing in the equilibrium fraction of purchasers λ^* by Proposition 2.2. Given these two statements, proving Proposition 2.5 amounts to showing that:

$$\frac{\partial \lambda^*}{\partial \alpha} < 0, \tag{67}$$

when $\lambda^* \in (0, 1)$. Observe that:

$$\frac{d\theta(\lambda^*, \alpha)}{d\alpha} = \frac{\partial \theta(\lambda^*, \alpha)}{\partial \lambda^*} \frac{\partial \lambda^*}{\partial \alpha} + \frac{\partial \theta(\lambda^*, \alpha)}{\partial \alpha}, \tag{68}$$

which implies:

$$\frac{\partial \lambda^*}{\partial \alpha} = \frac{\frac{d\theta(\lambda^*, \alpha)}{d\alpha} - \frac{\partial \theta(\lambda^*, \alpha)}{\partial \alpha}}{\frac{\partial \theta(\lambda^*, \alpha)}{\partial \lambda^*}}. \tag{69}$$

Recall that $\theta(\lambda^*, \alpha) = \bar{\tau} + \int\limits_{1-\lambda^*}^{1} k\bar{\tau}\alpha i^{\alpha-1} di$, so by the Leibnitz rule:

$$\frac{\partial \theta(\lambda^*, \alpha)}{\partial \lambda^*} = \eta_{1-\lambda^*} > 0, \quad \frac{\partial \theta(\lambda^*, \alpha)}{\partial \alpha} = -k\bar{\tau}(1 - \lambda^*)^{\alpha} \log(1 - \lambda^*) > 0. \tag{70}$$

Therefore, equation (69) must be strictly negative if $\frac{d\theta(\lambda^*,\alpha)}{d\alpha} < 0$. From Proposition 3, if $\lambda^* \in (0,1)$:

$$B(\lambda^*,\alpha) = R\left(\theta(\lambda,\alpha)\right)\eta_{1-\lambda^*} - \frac{\chi}{\lambda^*} = 0. \tag{71}$$

Taking the total derivative of both sides yields:

$$\frac{\partial B(\lambda^*,\alpha)}{\partial \lambda^*}\frac{\partial \lambda^*}{\partial \alpha} + \frac{\partial B(\lambda^*,\alpha)}{\partial \alpha} = 0. \tag{72}$$

which implies:

$$\frac{\partial \lambda^*}{\partial \alpha} = -\frac{\frac{\partial B(\lambda^*,\alpha)}{\partial \alpha}}{\frac{\partial B(\lambda^*,\alpha)}{\partial \lambda^*}}. \tag{73}$$

Next, note that:

$$\frac{\partial B(\lambda^*,\alpha)}{\partial \alpha} = \frac{\partial R(\theta(\lambda^*,\alpha))}{\partial \theta(\lambda^*,\alpha)}\frac{\partial \theta(\lambda^*,\alpha)}{\partial \lambda^*}\eta_{1-\lambda^*} + R(\theta(\lambda^*,\alpha))\frac{\partial \eta_{1-\lambda^*}}{\partial \alpha}, \tag{74}$$

$$\frac{\partial B(\lambda^*,\alpha)}{\partial \lambda^*} = \frac{\partial R(\theta(\lambda^*,\lambda^*))}{\partial \theta(\lambda^*,\alpha)}\frac{\partial \theta(\lambda^*,\alpha)}{\partial \lambda^*}\eta_{1-\lambda^*} + R(\theta(\lambda^*,\alpha))\frac{\partial \eta_{1-\lambda^*}}{\partial \lambda^*} + \frac{\chi}{\lambda^{*2}}. \tag{75}$$

Substituting equations (73), (74), and (75) into equation (68) and canceling terms gives:

$$\frac{d\theta(\lambda^*,\alpha)}{d\alpha} = \frac{1}{\frac{\partial B(\lambda^*,\alpha)}{\partial \lambda^*}}\left[-\frac{\partial \theta(\lambda^*,\alpha)}{\partial \lambda^*}\left(R(\theta(\lambda^*,\alpha))\frac{\partial \eta_{1-\lambda^*}}{\partial \alpha}\right) + \frac{\partial \theta(\lambda^*,\alpha)}{\partial \alpha}\left(R(\theta(\lambda^*,\alpha))\frac{\partial \eta_{1-\lambda^*}}{\partial \lambda^*} + \frac{\chi}{\lambda^{*2}}\right)\right. \tag{76}$$

Since $R\left(\theta(\lambda,\alpha)\right)\eta_{1-\lambda^*} = \frac{\chi}{\lambda^*}$ in equilibrium:

$$\begin{aligned}\frac{d\theta(\lambda^*,\alpha)}{d\alpha} &= \frac{1}{\frac{\partial B(\lambda^*,\alpha)}{\partial \lambda^*}}\left[-\frac{\partial \theta(\lambda^*,\alpha)}{\partial \lambda^*}\left(\frac{\frac{\chi}{\lambda^*}}{\eta_{1-\lambda^*}}\frac{\partial \eta_{1-\lambda^*}}{\partial \alpha}\right) + \frac{\partial \theta(\lambda^*,\alpha)}{\partial \alpha}\left(\frac{\frac{\chi}{\lambda^*}}{\eta_{1-\lambda^*}}\frac{\partial \eta_{1-\lambda^*}}{\partial \lambda^*} + \frac{\chi}{\lambda^{*2}}\right)\right] \\ &= \frac{\frac{\chi}{\lambda^*}}{\frac{\partial B(\lambda^*,\alpha)}{\partial \lambda^*}}\left[-\frac{\partial \theta(\lambda^*,\alpha)}{\partial \lambda^*}\left(\frac{\frac{\partial \eta_{1-\lambda^*}}{\partial \alpha}}{\eta_{1-\lambda^*}}\right) + \frac{\partial \theta(\lambda^*,\alpha)}{\partial \alpha}\left(\frac{\frac{\partial \eta_{1-\lambda^*}}{\partial \lambda^*}}{\eta_{1-\lambda^*}} + \frac{\frac{\chi}{\lambda^{*2}}}{\frac{\chi}{\lambda^*}}\right)\right].\end{aligned} \tag{77}$$

Next:

$$\left(\frac{\frac{\partial \eta_{1-\lambda^*}}{\partial \lambda^*}}{\eta_{1-\lambda^*}} + \frac{\frac{\chi}{\lambda^{*2}}}{\frac{\chi}{\lambda^*}}\right) = -\frac{(\alpha-1)\bar{\tau}\alpha(1-\lambda^*)^{\alpha-2}}{\bar{\tau}\alpha(1-\lambda^*)^{\alpha-1}} + \frac{1}{\lambda^*}$$
$$= \frac{1-\alpha\lambda^*}{\lambda^*(1-\lambda^*)}. \tag{78}$$

Substituting the equation for $\frac{\partial\theta(\lambda^*,\alpha)}{\partial\lambda^*}$ and $\frac{\partial\theta(\lambda^*,\alpha)}{\partial\alpha}$ along with equation (78) into (77) gives:

$$\frac{d\theta(\lambda^*,\alpha)}{d\alpha} = \frac{\frac{\chi}{\lambda^*}}{\frac{\partial B(\lambda^*,\alpha)}{\partial\lambda^*}}\left[-\eta_{1-\lambda^*}\left(\frac{\frac{\partial\eta_{1-\lambda^*}}{\partial\alpha}}{\eta_{1-\lambda^*}}\right) - k\bar{\tau}(1-\lambda^*)^{\alpha}\log(1-\lambda^*)\left(\frac{1-\alpha\lambda^*}{\lambda^*(1-\lambda^*)}\right)\right]. \tag{79}$$

Since $\frac{\partial\eta_{1-\lambda^*}}{\partial\alpha} = k\bar{\tau}(1-\lambda^*)^{\alpha-1}(1+\alpha\log(1-\lambda^*))$:

$$\frac{d\theta(\lambda^*,\alpha)}{d\alpha} = \frac{\frac{\chi}{\lambda^*}}{\frac{\partial B(\lambda^*,\alpha)}{\partial\lambda^*}}\left[-k\bar{\tau}(1-\lambda^*)^{\alpha-1}(1+\alpha\log(1-\lambda^*)) - k\bar{\tau}(1-\lambda^*)^{\alpha}\log(1-\lambda^*)\left(\frac{1-\alpha\lambda^*}{\lambda^*(1-\lambda^*)}\right)\right]$$
$$= \frac{-\lambda^* k\bar{\tau}(1-\lambda^*)^{\alpha-1}\frac{\chi}{\lambda^*}}{\frac{\partial B(\lambda^*,\alpha)}{\partial\lambda^*}}\left[\lambda^* + \log(1-\lambda^*)\right]. \tag{80}$$

The term $\lambda^* + \log(1-\lambda^*)$ is strictly negative for all $\lambda^* \in (0,1)$, and $\frac{\partial B(\lambda^*,\alpha)}{\partial\lambda^*}$ cannot be positive for any interior solution by Proposition 2.4. Therefore, $\frac{d\theta(\lambda^*,\alpha)}{d\alpha}$ is strictly negative, which implies $\frac{\partial\lambda^*}{\partial\alpha} < 0$ when $\lambda^* > 0$ by equation (69), thus completing the proof.

Proof of Proposition 2.6. Assume that the equilibrium fraction of purchasers is positive ($\lambda^* > 0$). From equation (2.18), the period one excess return per share $E_1(f-p)$ is:

$$E_1(f-p) = \frac{\rho\bar{x}}{\int_0^1 \widehat{\tau}_i di}, \tag{81}$$

where $\int_0^1 \widehat{\tau}_i di$ is strictly increasing in $\theta(\lambda^*,\alpha)$. Therefore, it is sufficient to show that:

$$\frac{d\theta(\lambda^*,\alpha)}{d\alpha} < 0, \tag{82}$$

when $\lambda^* > 0$. But this follows directly from equation (80). Therefore the equilibrium risk premium $E_1(f-p)^*$ is strictly increasing in α, which completes the proof.

Proof of Proposition 2.7. If the cost of news c is exogenously fixed, the net benefit of purchasing to the marginal agent is:

$$B(\lambda, \alpha) = R\left(\theta(\lambda^*, \alpha)\right) k\bar{\tau}\alpha(1-\lambda^*)^{\alpha-1} - c, \tag{83}$$

which is now strictly decreasing in λ because the endogenous cost is absent. Since the period one excess return per share $E_1(f-p)^*$ is strictly decreasing in θ, it is sufficient to show that there exists a cutoff value $\bar{\alpha}$ such that:

$$\frac{d\theta(\lambda^*, \alpha)}{d\alpha} > 0, \tag{84}$$

for all $\alpha > \bar{\alpha}$. First, note that for a sufficiently high α, equation (83) is strictly positive when $\lambda = 0$, which in turn implies that $\lambda^* > 0$. Therefore, for a sufficiently high α, we can apply the characterization of an interior solution found in Proposition 2.4 and repeat the same steps as the proof of Proposition 2.5, which yields:

$$\begin{aligned}\frac{d\theta(\lambda^*, \alpha)}{d\alpha} &= \frac{c}{\frac{\partial B(\lambda^*,\alpha)}{\partial \lambda^*}}\left[-k\eta_{1-\lambda^*}\left(\frac{\frac{\partial \eta_{1-\lambda^*}}{\partial \alpha}}{\eta_{1-\lambda^*}}\right) + k\bar{\tau}(1-\lambda^*)^{\alpha}\log(1-\lambda^*)\left(\frac{\alpha-1}{1-\lambda^*}\right)\right] \\ &= \frac{-k\bar{\tau}(1-\lambda^*)^{\alpha-1}c}{\frac{\partial B(\lambda^*,\alpha)}{\partial \lambda^*}}\left[1+\log(1-\lambda^*)\right].\end{aligned} \tag{85}$$

The term $1+\log(1-\lambda^*)$ is strictly positive for all λ^* less than 0.632. Since $\lambda^* \in (0,1)$ for a sufficiently high α by the logic above, and λ^* is strictly decreasing for a sufficiently large increase in α by Proposition 2.5, $1+(1-\lambda^*)$ must be strictly positive above some cutoff α, which makes equation (85) strictly positive as well, thus completing the proof.

Proof of Proposition 2.8. Assume that the precision of news is independent of prior precision: $\eta_i = k$. In this case, the average precision of priors and purchased signals is not influenced by which agents are purchasing, so that:

$$\theta = \bar{\tau} + \int_0^{\lambda^*} k\,di, \tag{86}$$

which is independent of α. Since both θ and η_i are independent of α, a change in α cannot affect the equilibrium fraction of buyers, which completes the proof.

Proof of Proposition 3.1. Following the literature on the rational expectations equilibrium, we look for the equilibrium price that has a linear form

$$p = p_0 + p_f f + p_\theta \theta. \tag{87}$$

Equation (87) implies that the price contains information about f and observing the price is equivalent to receiving a signal $s_p = f + (p_\theta/p_f)\theta$, which is equal to $s_p = (p - p_0)/p_f$. The investor i's expectation of the payoff $\hat{f}_i = E[f|\mathcal{F}_{2,i}]$ based on all exogenous signals as well as on the price is given by

$$\hat{f}_i = \frac{1}{\hat{\tau}_i}\left[\tau_{f,i} f_i + \frac{p_f^2}{p_\theta^2}(\tau_\theta + \tau_{s\theta,i})s_p + \tau_{sf,i}s_{f,i} - \frac{p_f}{p_\theta}\tau_{s\theta,i}s_{\theta,i}\right] \tag{88}$$

and the conditional variance of the payoff is $Var[f|\mathcal{F}_{2,i}] = \hat{\tau}_i^{-1}$, where

$$\hat{\tau}_i = \tau_{f,i} + \tau_{sf,i} + \frac{p_f^2}{p_\theta^2}(\tau_\theta + \tau_{s\theta,i}). \tag{89}$$

Indeed, since all variables are normal, we can use the standard formulas for conditional expectation of a multivariate normal distribution $E[Y|X] = E[Y] + \beta(X - E[X])$, $\beta = Cov[Y, X]Var[X]^{-1}$ and the conditional variance $Var[Y|X] = Var[Y] - Cov[Y, X]Var[X]^{-1}Cov[X, Y]$. In our case, $Y = f$, $X = (s_p \;\; s_{f,i} \;\; s_{\theta,i})$, so

$$E[f|\mathcal{F}_{2,i}] = f_i + \beta_p(s_p - f_i) + \beta_{sf}(s_{f,i} - f_i) + \beta_{s\theta}s_{\theta,i},$$

where

$$\begin{pmatrix}\beta_p & \beta_{sf} & \beta_{s\theta}\end{pmatrix} = \begin{pmatrix}\tau_{f,i}^{-1} & \tau_{f,i}^{-1} & 0\end{pmatrix}\begin{pmatrix}\tau_{f,i}^{-1} + \frac{p_\theta^2}{p_f^2}\tau_\theta^{-1} & \tau_{f,i}^{-1} & \frac{p_\theta}{p_f}\tau_\theta^{-1} \\ \tau_{f,i}^{-1} & \tau_{f,i}^{-1} + \tau_{sf,i}^{-1} & 0 \\ \frac{p_\theta}{p_f}\tau_\theta^{-1} & 0 & \tau_\theta^{-1} + \tau_{s\theta,i}^{-1}\end{pmatrix}^{-1}.$$

A straightforward computation yields

$$\begin{pmatrix}\beta_p & \beta_{sf} & \beta_{s\theta}\end{pmatrix} = \frac{1}{\hat{\tau}_i}\begin{pmatrix}\frac{p_f^2}{p_\theta^2}(\tau_\theta + \tau_{s\theta,i}) & \tau_{sf,i} & -\frac{p_f}{p_\theta}\tau_{s\theta,i}\end{pmatrix},$$

where $\hat{\tau}_i$ is given by equation (89). Hence,

$$\begin{aligned} E[f|\mathcal{F}_{2,i}] &= (1-\beta_p-\beta_{sf})f_i + \beta_p s_p + \beta_{sf} s_{sf,i} + \beta_{s\theta} s_{s\theta,i} \\ &= \frac{1}{\hat{\tau}_i}\left(\tau_{f,i} f_i + \frac{p_f^2}{p_\theta^2}(\tau_\theta + \tau_{s\theta,i}) s_p + \tau_{sf,i} s_{f,i} - \frac{p_f}{p_\theta}\tau_{s\theta,i} s_{\theta,i}\right) \end{aligned}$$

and this is equation (88). Similarly,

$$Var[f|\mathcal{F}_{2,i}] = \tau_{f,i}^{-1} - \frac{1}{\hat{\tau}_i}\begin{pmatrix} \frac{p_f^2}{p_\theta^2}(\tau_\theta + \tau_{s\theta,i}) & \tau_{sf,i} & -\frac{p_f}{p_\theta}\tau_{s\theta,i} \end{pmatrix} \begin{pmatrix} \tau_{f,i}^{-1} \\ \tau_{f,i}^{-1} \\ 0 \end{pmatrix} = \hat{\tau}_i^{-1}.$$

The market clearing condition (3.5) together with equation (3.4) yields

$$\int_{\mathcal{I}} \gamma^{-1}\hat{\tau}_i(\hat{f}_i - p)d\mu(i) = \bar{\theta} + \theta.$$

Using the expectation of the payoff from equation (88),

$$\begin{aligned} &\int_{\mathcal{I}} \tau_{f,i} f_i d\mu(i) + \frac{p_f^2}{p_\theta^2}\left(\tau_\theta + \int_{\mathcal{I}} \tau_{s\theta,i} d\mu(i)\right) s_p + \int_{\mathcal{I}} \tau_{sf,i} s_{f,i} d\mu(i) \\ &\qquad - \frac{p_f}{p_\theta}\int_{\mathcal{I}} \tau_{s\theta,i} s_{\theta,i} d\mu(i) - p\int_{\mathcal{I}} \hat{\tau}_i d\mu(i) = \gamma(\bar{\theta} + \theta). \end{aligned}$$

Introducing the notation for average precisions

$$\bar{\tau}_f = \int_{\mathcal{I}} \tau_{f,i} d\mu(i), \quad \bar{\tau}_{s\theta} = \int_{\mathcal{I}} \tau_{s\theta,i} d\mu(i), \quad \bar{\tau}_{sf} = \int_{\mathcal{I}} \tau_{sf,i} d\mu(i), \quad \bar{\tau} = \int_{\mathcal{I}} \hat{\tau}_i d\mu(i)$$

and using the law of large numbers we get

$$\int_{\mathcal{I}} \tau_{f,i} f_i d\mu(i) = \bar{\tau}_f f, \quad \int_{\mathcal{I}} \tau_{sf,i} s_{f,i} d\mu(i) = \bar{\tau}_{sf} f, \quad \int_{\mathcal{I}} \tau_{s\theta,i} s_{\theta,i} d\mu(i) = \bar{\tau}_{s\theta}\theta.$$

Hence, the market clearing condition yields

$$p = \bar{\tau}^{-1}\left(-\gamma\bar{\theta} + \left(\bar{\tau}_f + \bar{\tau}_{sf} + \frac{p_f^2}{p_\theta^2}(\tau_\theta + \bar{\tau}_{s\theta})\right) f + \left(\frac{p_f}{p_\theta}\tau_\theta - \gamma\right)\theta\right)$$

or

$$p = -\frac{\gamma\bar{\theta}}{\bar{\tau}} + f + \frac{1}{\bar{\tau}}\left(\frac{p_f}{p_\theta}\tau_\theta - \gamma\right)\theta. \tag{90}$$

Comparing equations (87) and (90) we get that $p_0 = -\gamma\bar{\theta}/\bar{\tau}$, $p_f = 1$, and p_θ solves the following equation:

$$\left(\bar{\tau}_f + \bar{\tau}_{sf} + \frac{1}{p_\theta^2}(\tau_\theta + \bar{\tau}_{s\theta})\right) p_\theta = \frac{\tau_\theta}{p_\theta} - \gamma,$$

which can be rewritten as

$$(\bar{\tau}_f + \bar{\tau}_{sf})p_\theta + \frac{\bar{\tau}_{s\theta}}{p_\theta} + \gamma = 0.$$

This is equation (3.9). Q.E.D.

Proof of Proposition 3.2. The optimal trading strategy of investor i at $t = 2$ is $x_i = \gamma^{-1}\hat{\tau}_i(\hat{f}_i - p)$, so the expected utility at $t = 2$ is

$$U_{2,i} = E\left[-e^{-\gamma W_{3,i}}|\mathcal{F}_{2,i}\right] = -e^{-\gamma W_{2,i}} E\left[e^{-\hat{\tau}_i(\hat{f}_i - p)(f-p)}|\mathcal{F}_{2,i}\right] = -e^{-\gamma W_{2,i} - \frac{1}{2}\hat{\tau}_i(\hat{f}_i - p)^2}.$$

For computation of $U_{1,i} = E[U_{2,i}|\mathcal{F}_{1,i}]$ it is convenient to introduce the variable $z_i = \hat{f}_i - p - \mu_i$, where $\mu_i = E[\hat{f}_i - p|\mathcal{F}_{1,i}] = E[f - p|\mathcal{F}_{1,i}]$ is the expected return computed at $t = 1$. Because $\hat{f}_i$ and p are normally distributed, $z_i|\mathcal{F}_{1,i} \sim \mathcal{N}(0, \sigma_{z,i}^2)$. To find $\sigma_{z,i}^2 = Var[z_i|\mathcal{F}_{1,i}]$, note that

$$\sigma_{z,i}^2 = Var[\hat{f}_i - f + f - p|\mathcal{F}_{1,i}] = Var[\hat{f}_i - f|\mathcal{F}_{1,i}] + Var[f - p|\mathcal{F}_{1,i}] + 2Cov[\hat{f}_i - f, f - p|\mathcal{F}_{1,i}].$$

Using the law of iterated expectations and the definition of $\hat{\tau}_i$, we get that $Var[\hat{f}_i - f|\mathcal{F}_{1,i}] = \hat{\tau}_i^{-1}$ and $Cov[\hat{f}_i - f, f - p|\mathcal{F}_{1,i}] = -\hat{\tau}_i^{-1}$. The term $Var[f - p|\mathcal{F}_{1,i}]$ measures the precision of the price as a signal about the asset payoff, so it is natural to denote it as $\tau_{p,i}^{-1}$. Thus, $\sigma_{z,i}^2 = \tau_{p,i}^{-1} - \hat{\tau}_i^{-1}$.

Using the variable z and noting that $W_{2,i} = W_{1,i} - c(K)$, the expected utility at $t = 1$ is

$$U_{1,i} = -e^{-\gamma(W_{1,i} - c(K))} E\left[e^{-\frac{1}{2}\hat{\tau}_i(\hat{f}_i - p)^2}|\mathcal{F}_{1,i}\right] = -e^{-\gamma(W_{1,i} - c(K))} E\left[e^{-\frac{1}{2}\hat{\tau}_i(z_i^2 + 2\mu_i z_i + \mu_i^2)}|\mathcal{F}_{1,i}\right].$$

To compute the expectation, we use that for a normally distributed random variable $z \sim \mathcal{N}(0, \sigma^2)$

$$E\left[\exp(az^2 + bz + c)\right] = \frac{1}{\sqrt{1 - 2a\sigma^2}} \exp\left(\frac{b^2\sigma^2}{2(1 - 2a\sigma^2)} + c\right).$$

Hence,

$$U_{1,i} = -\frac{1}{\sqrt{1+\hat{\tau}_i(\tau_{p,i}^{-1} - \hat{\tau}_i^{-1})}} \exp\left(-\gamma(W_{1,i} - c(K)) + \frac{\mu_i^2 \hat{\tau}_i^2 (\tau_{p,i}^{-1} - \hat{\tau}_i^{-1})}{2(1+\hat{\tau}_i(\tau_{p,i}^{-1} - \hat{\tau}_i^{-1}))} - \frac{1}{2}\hat{\tau}_i \mu_i^2\right)$$
$$= -\sqrt{\frac{\tau_{p,i}}{\hat{\tau}_i}} \exp\left(-\gamma(W_{1,i} - c(K)) - \frac{1}{2}\tau_{p,i}\mu_i^2\right).$$

Noting that $\tau_{p,i}^{-1} = Var[f-p|\mathcal{F}_{1,i}]$, $\hat{\tau}_i^{-1} = Var[f-p|\mathcal{F}_{2,i}]$, and $\mu_i = E[f-p|\mathcal{F}_{1,i}]$, we arrive at equation (3.12). Q.E.D.

Proof of Proposition 3.3. To simplify notation, we omit the investor index i. The minimization problem in equation (3.15) is equivalent to minimization of

$$\tilde{U}(\tau_{sf}, \tau_{s\theta}) = -\frac{1}{2}\log\left(\tau_f + \tau_{sf} + \frac{1}{p_\theta^2}(\tau_\theta + \tau_{s\theta})\right) + \gamma c\left(\left(1+\frac{\tau_{sf}}{\tau_f}\right)\left(1+\frac{\tau_{s\theta}}{\tau_\theta}\right) - 1\right).$$

Assume that $\tilde{U}(\tau_{sf}, \tau_{s\theta})$ has a local internal minimum $(\tau_{sf}^*, \tau_{s\theta}^*) = \arg\min_{\{\tau_{sf}>0,\ \tau_{s\theta}>0\}} \tilde{U}(\tau_{sf}, \tau_{s\theta})$. The function $\tilde{U}(\tau_{sf}, \tau_{s\theta})$ is twice continuously differentiable for $\tau_{sf} \geq 0$ and $\tau_{s\theta} \geq 0$, so the point $(\tau_{sf}^*, \tau_{s\theta}^*)$ satisfies the first order conditions

$$\frac{\gamma}{\tau_f}\left(1+\frac{\tau_{s\theta}^*}{\tau_\theta}\right)c' = \frac{1}{2\hat{\tau}^*}, \qquad \frac{\gamma}{\tau_\theta}\left(1+\frac{\tau_{sf}^*}{\tau_f}\right)c' = \frac{1}{2p_\theta^2\hat{\tau}^*}, \tag{91}$$

where $\hat{\tau}^* = \tau_f + \tau_{sf}^* + (\tau_\theta + \tau_{s\theta}^*)/p_\theta^2$ and the argument of the function c is suppressed. The Hessian matrix of $\tilde{U}(\tau_{sf}, \tau_{s\theta})$ is

$$H = \begin{pmatrix} \frac{1}{2\hat{\tau}^2} + \frac{\gamma}{\tau_f^2}\left(1+\frac{\tau_{s\theta}}{\tau_\theta}\right)^2 c'' & \frac{1}{2p_\theta^2\hat{\tau}^2} + \frac{\gamma}{\tau_f\tau_\theta}\left(1+\frac{\tau_{sf}}{\tau_f}\right)\left(1+\frac{\tau_{s\theta}}{\tau_\theta}\right)c'' + \frac{\gamma}{\tau_f\tau_\theta}c' \\ \frac{1}{2p_\theta^2\hat{\tau}^2} + \frac{\gamma}{\tau_f\tau_\theta}\left(1+\frac{\tau_{sf}}{\tau_f}\right)\left(1+\frac{\tau_{s\theta}}{\tau_\theta}\right)c'' + \frac{\gamma}{\tau_f\tau_\theta}c' & \frac{1}{2p_\theta^4\hat{\tau}^2} + \frac{\gamma}{\tau_\theta^2}\left(1+\frac{\tau_{sf}}{\tau_f}\right)^2 c'' \end{pmatrix}.$$

Because $c'' \geq 0$, the element $H_{11} > 0$. A straightforward computation gives that at the point where the first order conditions (91) are satisfied

$$\det H = -\frac{\gamma}{p_\theta^2\tau_f\tau_\theta(\hat{\tau}^*)^2}c' - \frac{\gamma^2}{\tau_f^2\tau_\theta^2}(c')^2 - \frac{2\gamma}{p_\theta^2\tau_f\tau_\theta(\hat{\tau}^*)^2}\left(1+\frac{\tau_{sf}^*}{\tau_f}\right)\left(1+\frac{\tau_{s\theta}^*}{\tau_\theta}\right)c'' < 0.$$

Thus, $(\tau_{sf}^*, \tau_{s\theta}^*)$ is a saddle point and it cannot be a local minimum. This contradiction proves that $\tilde{U}(\tau_{sf}, \tau_{s\theta})$ is minimized on the boundary $\tau_{sf} = 0$ or $\tau_{s\theta} = 0$. This is the first statement of Proposition 3.3.

To prove the second statement of Proposition 3.3, note first that the point $\tau_{sf} = \tau_{s\theta} = 0$ cannot be a solution to the minimization problem (3.15) because

$$\frac{\partial \tilde{U}}{\partial \tau_{sf}}(0,0) = -\frac{p_\theta^2}{2(p_\theta^2\tau_f + \tau_\theta)} + \frac{\gamma}{\tau_f}c'(0) < 0, \qquad \frac{\partial \tilde{U}}{\partial \tau_{s\theta}}(0,0) = -\frac{1}{2(p_\theta^2\tau_f + \tau_\theta)} + \frac{\gamma}{\tau_\theta}c'(0) < 0.$$

Hence, the solution is either on the boundary $\tau_{s\theta} = 0$ and $\tau_{sf} > 0$ or $\tau_{sf} = 0$ and $\tau_{s\theta} > 0$. The first order conditions (91) on the boundaries $\tau_{s\theta} = 0$ and $\tau_{sf} = 0$ yield equations (3.17) and (3.16), respectively. Each of these equations has a unique solution because the left hand sides are non-decreasing functions of the unknown variable and the right hand sides are strictly decreasing functions of the unknown variable. Thus, the functions $\tau_{sf}^*(\tau_f)$ and $\tau_{s\theta}^*(\tau_f)$ introduced in Proposition 3.3 are well defined. Noting that

$$\frac{\partial^2 \tilde{U}}{\partial \tau_{sf}^2}(\tau_{sf}^*, 0) = \frac{p_\theta^4}{2(p_\theta^2(\tau_f + \tau_{sf}^*) + \tau_\theta)^2} + \frac{\gamma}{\tau_f^2}c''\left(\frac{\tau_{sf}^*}{\tau_f}\right) > 0,$$

$$\frac{\partial^2 \tilde{U}}{\partial \tau_{s\theta}^2}(0, \tau_{s\theta}^*) = \frac{1}{2(p_\theta^2\tau_f + \tau_\theta + \tau_{s\theta}^*)^2} + \frac{\gamma}{\tau_\theta^2}c''\left(\frac{\tau_{s\theta}^*}{\tau_\theta}\right) > 0$$

we conclude that the points $(\tau_{sf}^*, 0)$ and $(0, \tau_{s\theta}^*)$ are local minimums of the function $\tilde{U}$. To find which point is the global minimum, we need to compare $\tilde{U}(\tau_{sf}^*, 0)$ and $\tilde{U}(0, \tau_{s\theta}^*)$ and the result depends on τ_f. To identify the form of this dependence, consider the difference $\Delta\tilde{U}(\tau_f) = \tilde{U}(\tau_{sf}^*, 0) - \tilde{U}(0, \tau_{s\theta}^*)$ as a function of τ_f. Using the envelope theorem and the first order conditions for τ_{sf}^*,

$$\begin{aligned}
\frac{\partial \Delta\tilde{U}}{\partial \tau_f} &= -\frac{p_\theta^2}{2(p_\theta^2(\tau_f + \tau_{sf}^*) + \tau_\theta)} - \frac{\gamma\tau_{sf}^*}{\tau_f^2}c'\left(\frac{\tau_{sf}^*}{\tau_f}\right) + \frac{p_\theta^2}{2(p_\theta^2\tau_f + \tau_\theta + \tau_{s\theta}^*)} \\
&= -\frac{p_\theta^2(1 + \tau_{sf}^*/\tau_f)}{2(p_\theta^2(\tau_f + \tau_{sf}^*) + \tau_\theta)} + \frac{p_\theta^2}{2(p_\theta^2\tau_f + \tau_\theta + \tau_{s\theta}^*)} \\
&= -\frac{p_\theta^2}{2(p_\theta^2(\tau_f + \tau_{sf}^*) + \tau_\theta)(p_\theta^2\tau_f + \tau_\theta + \tau_{s\theta}^*)}\left(\tau_{s\theta}^* + \frac{\tau_{sf}^*}{\tau_f}(\tau_\theta + \tau_{s\theta}^*)\right) < 0.
\end{aligned}$$

Therefore, there could exist at most one point τ_f^* such that $\Delta\tilde{U}(\tau_f^*) = 0$. For $\tau_f > \tau_f^*$ we have $\tilde{U}(\tau_{sf}^*, 0) < \tilde{U}(0, \tau_{s\theta}^*)$ and τ_{sf}^* is the global minimum of $\tilde{U}$. For $\tau_f < \tau_f^*$ the inequality is reversed: $\tilde{U}(\tau_{sf}^*, 0) > \tilde{U}(0, \tau_{s\theta}^*)$ and $\tau_{s\theta}^*$ is the global minimum of $\tilde{U}$. Because

$$\tilde{U}(\tau_{sf}^*, 0) = -\frac{1}{2}\log\left(\tau_f + \tau_{sf}^* + \frac{\tau_\theta}{p_\theta^2}\right) + \gamma c\left(\frac{\tau_{sf}^*}{\tau_f}\right),$$

$$\tilde{U}(0,\tau_{s\theta}^*) = -\frac{1}{2}\log\left(\tau_f + \frac{1}{p_\theta^2}(\tau_\theta + \tau_{s\theta}^*)\right) + \gamma c\left(\frac{\tau_{s\theta}^*}{\tau_\theta}\right),$$

the condition $\tilde{U}(\tau_{sf}^*, 0) = \tilde{U}(0, \tau_{s\theta}^*)$ yields equation (3.18) that determines τ_f^* and this equation has a solution when considered on a sufficiently large interval for τ_f.

To prove the third statement of Proposition 3.3, consider again the first order conditions on the boundaries $\tau_{s\theta} = 0$ and $\tau_{sf} = 0$, which for the linear cost function take the following form:

$$\frac{2\gamma c_0}{\tau_f} = \frac{p_\theta^2}{p_\theta^2(\tau_f + \tau_{sf}^*) + \tau_\theta}, \qquad \frac{2\gamma c_0}{\tau_\theta} = \frac{1}{p_\theta^2\tau_f + \tau_\theta + \tau_{s\theta}^*}. \tag{92}$$

Thus, the optimal precisions of the signals are

$$\tau_{sf}^* = \left(\frac{1}{2\gamma c_0} - 1\right)\tau_f - \frac{\tau_\theta}{p_\theta^2}, \qquad \tau_{s\theta}^* = \left(\frac{1}{2\gamma c_0} - 1\right)\tau_\theta - p_\theta^2\tau_f. \tag{93}$$

Note that for $2\gamma c_0 \geq 1$ both τ_{sf}^* and $\tau_{s\theta}^*$ are negative, the derivatives

$$\frac{\partial\tilde{U}}{\partial\tau_{sf}}(0,0) = -\frac{p_\theta^2}{2(p_\theta^2\tau_f + \tau_\theta)} + \frac{\gamma c_0}{\tau_f}, \qquad \frac{\partial\tilde{U}}{\partial\tau_{s\theta}}(0,0) = -\frac{1}{2(p_\theta^2\tau_f + \tau_\theta)} + \frac{\gamma c_0}{\tau_\theta}$$

are positive, so the point $(0,0)$ is the global minimum. This is statement 3(c).

In the case $2\gamma c_0 < 1$, an investor would consider acquiring the fundamental (non-fundamental) information when $\tau_{sf}^* > 0$ ($\tau_{s\theta}^* > 0$). These conditions are satisfied when

$$\tau_f > \frac{\tau_\theta}{p_\theta^2}\left(\frac{1}{2\gamma c_0} - 1\right)^{-1}, \qquad \tau_f < \frac{\tau_\theta}{p_\theta^2}\left(\frac{1}{2\gamma c_0} - 1\right).$$

When $1/(2\gamma c_0) - 1 < 1$, these inequalities cannot be satisfied simultaneously and characterize the optimal information acquisition described by statement 3(b). When $1/(2\gamma c_0) - 1 \geq 1$, the function $\tilde{U}$ can have two local minima at the points $(\tau_{sf}^*, 0)$ and $(0, \tau_{s\theta}^*)$ with

$$\tilde{U}(\tau_{sf}^*, 0) = -\frac{1}{2}\log\left(\frac{\tau_f}{2\gamma c_0}\right) + \gamma c_0\left(\frac{1}{2\gamma c_0} - 1\right) - \gamma c_0\frac{\tau_\theta}{p_\theta^2\tau_f},$$

$$\tilde{U}(0, \tau_{s\theta}^*) = -\frac{1}{2}\log\left(\frac{\tau_\theta}{2\gamma c_0 p_\theta^2}\right) + \gamma c_0\left(\frac{1}{2\gamma c_0} - 1\right) - \gamma c_0\frac{p_\theta^2\tau_f}{\tau_\theta}.$$

Obviously, $\tilde{U}(\tau_{sf}^*, 0) = \tilde{U}(0, \tau_{s\theta}^*)$ when $\tau_f = \tau_\theta/p_\theta^2$. It is not difficult to show that the function $\tilde{U}(\tau_{sf}^*, 0) - \tilde{U}(0, \tau_{s\theta}^*)$ strictly decreases when $\tau_{sf}^* > 0$ and $\tau_{s\theta}^* > 0$, that is, when

$$\frac{\tau_\theta}{p_\theta^2}\left(\frac{1}{2\gamma c_0}-1\right)^{-1} \le \tau_f \le \frac{\tau_\theta}{p_\theta^2}\left(\frac{1}{2\gamma c_0}-1\right).$$

As a result, the global minimum of $\tilde{U}$ is $(0, \tau_{s\theta}^*)$ when $\tau_f < \tau_\theta/p_\theta^2$ and is $(\tau_{sf}^*, 0)$ when $\tau_f > \tau_\theta/p_\theta^2$. This gives statement 3(a). Q.E.D.

Proof of Proposition 3.4. According to Proposition 3.3, an investor with $\tau_{f,i} < \tau_f^*$ learns only about noise and his ex post precision is $\hat{\tau}_i = \tau_{f,i} + (\tau_\theta + \tau_{s\theta,i}^*)/p_\theta^2$, where $\tau_{s\theta,i}^*$ solves equation (3.16). Taking the derivative of both sides of this equation with respect to $\tau_{f,i}$, we get

$$\frac{1}{p_\theta^2}\frac{\partial \tau_{s\theta,i}^*}{\partial \tau_{f,i}} = -\left(\frac{2\gamma}{\tau_\theta^2}c''\left(\frac{\tau_{s\theta,i}^*}{\tau_\theta}\right) + \frac{1}{(p_\theta^2\tau_{f,i} + \tau_\theta + \tau_{s\theta,i}^*)^2}\right)^{-1} \frac{1}{(p_\theta^2\tau_{f,i} + \tau_\theta + \tau_{s\theta,i}^*)^2}. \tag{94}$$

Because $c''(K) \ge 0$, equation (94) implies that

$$-1 \le \frac{1}{p_\theta^2}\frac{\partial \tau_{s\theta,i}^*}{\partial \tau_{f,i}} < 0.$$

Noting that

$$\frac{\partial \hat{\tau}_i}{\partial \tau_{f,i}} = 1 + \frac{1}{p_\theta^2}\frac{\partial \tau_{s\theta,i}^*}{\partial \tau_{f,i}},$$

we obtain the statement of the proposition for $\tau_{f,i} < \tau_f^*$. When $\tau_{f,i} > \tau_f^*$, the investor learns about fundamentals and the precision of the obtained signal is given by equation (3.17). Again, taking the derivative of both sides of this equation with respect to $\tau_{f,i}$, we get

$$\begin{aligned}\frac{\partial \tau_{sf,i}^*}{\partial \tau_{f,i}} &= \left(\frac{2\gamma}{\tau_{f,i}^2}c''\left(\frac{\tau_{sf,i}^*}{\tau_{f,i}}\right) + \frac{1}{(\tau_{f,i} + \tau_{sf,i}^* + \tau_\theta/p_\theta^2)^2}\right)^{-1} \\ &\quad \times \left(\frac{2\gamma}{\tau_{f,i}^2}\left(c'\left(\frac{\tau_{sf,i}^*}{\tau_{f,i}}\right) + \frac{\tau_{sf,i}^*}{\tau_{f,i}}c''\left(\frac{\tau_{sf,i}^*}{\tau_{f,i}}\right)\right) - \frac{1}{(\tau_{f,i} + \tau_{sf,i}^* + \tau_\theta/p_\theta^2)^2}\right). \end{aligned} \tag{95}$$

Given that $c''(K) \ge 0$, the first factor in equation (95) is always positive. Omitting the argument of the cost function and using equation (3.17), the second factor can be written as

$$\frac{2\gamma}{\tau_{f,i}^2}\left(c' + \frac{\tau_{sf,i}^*}{\tau_{f,i}}c'' - 2\gamma(c')^2\right),$$

which is also positive because $c'' \geq 0$ and $0 \leq c' < 1/(2\gamma)$. The last inequality follows from equation (3.17). Thus, the derivative in equation (95) is positive and

$$\frac{\partial \hat{\tau}_i}{\partial \tau_{f,i}} = 1 + \frac{\partial \tau^*_{sf,i}}{\partial \tau_{f,i}} > 1.$$

This is the statement of the proposition for $\tau_{f,i} > \tau^*_f$. Q.E.D.

Proof of Proposition 3.5. Equations (3.27) and (3.28) immediately follow from equations (3.7) and (3.8) after setting $\tau_{s\theta,i} = 0$. The optimal information acquisition strategy in equation (3.29) is a solution to the following optimization problem, which is the optimization problem (3.15) in the absence of learning about noise:

$$\min_{\{\tau_{sf,i}\geq 0\}} \frac{1}{\sqrt{\tau_{f,i} + \tau_{sf,i} + \tau_\theta/(p^0_\theta)^2}} \exp\left(\gamma c_0 \frac{\tau_{sf,i}}{\tau_{f,i}}\right). \tag{96}$$

Thus,

$$\bar{\tau}^*_{sf} = h^{-1} \int_{\left\{i:\ \tau_{f,i}\geq \frac{\tau_\theta h}{(p^0_\theta)^2}\right\}} \tau_{f,i} d\mu(i) - \frac{\tau_\theta}{(p^0_\theta)^2} I\left(i:\ \tau_{f,i} \geq \frac{\tau_\theta h}{(p^0_\theta)^2}\right)$$

and using equation (3.9) we get equation (3.30).

To prove that the equilibrium is unique, note first that equation (3.9) implies that $p^0_\theta < 0$. Next, rewrite equation (3.30) as

$$\bar{\tau}_f + h^{-1} \int_{\left\{i:\ \tau_{f,i}\geq \frac{\tau_\theta h}{(p^0_\theta)^2}\right\}} \tau_{f,i} d\mu(i) - \frac{\tau_\theta}{(p^0_\theta)^2} I\left(i:\ \tau_{f,i} \geq \frac{\tau_\theta h}{(p^0_\theta)^2}\right) = -\frac{\gamma}{p^0_\theta}. \tag{97}$$

The right hand side of this equation is a continuous and increasing function of p^0_θ for $p^0_\theta < 0$. The left hand side can be represented as $g^0(\tau_\theta/(p^0_\theta)^2)$, where the function $g^0(z)$ is

$$g^0(z) = \bar{\tau}_f + h^{-1} \int_{\{i:\ \tau_{f,i}\geq hz\}} \tau_{f,i} d\mu(i) - zI\left(i:\ \tau_{f,i} \geq hz\right).$$

The function $g^0(z)$ is non-increasing for $z > 0$. Indeed, for $\Delta z > 0$

$$\begin{aligned} & g^0(z+\Delta z) - g^0(z) \\ & = -h^{-1} \int\limits_{\{i:\, hz \le \tau_{f,i} \le h(z+\Delta z)\}} \tau_{f,i} d\mu(i) - (z+\Delta z) I\left(i:\ \tau_{f,i} \ge h(z+\Delta z)\right) + zI\left(i:\ \tau_{f,i} \ge hz\right) \\ & \quad \le -h^{-1} \cdot hz \cdot I\left(i:\ hz \le \tau_{f,i} \le h(z+\Delta z)\right) - \Delta z I\left(i:\ \tau_{f,i} \ge h(z+\Delta z)\right) \\ & \qquad + zI\left(i:\ hz \le \tau_{f,i} \le h(z+\Delta z)\right) = -\Delta z I\left(i:\ \tau_{f,i} \ge h(z+\Delta z)\right) \le 0. \end{aligned}$$

Because $\tau_\theta/(p_\theta^0)^2$ increases with p_θ^0 for $p_\theta^0 < 0$, $g^0(\tau_\theta/(p_\theta^0)^2)$ is a non-increasing function of p_θ^0. Its graph can intersect a graph of an increasing function (the right hand side of equation (97)) only once, so equation (97) has maximum one solution. Q.E.D.

Proof of Proposition 3.6. Equations (3.26) and (3.30) that determine p_θ and p_θ^0 can be written as

$$g\left(\frac{\tau_\theta}{p_\theta^2}\right) = -\frac{\gamma}{p_\theta}, \qquad g^0\left(\frac{\tau_\theta}{p_\theta^2}\right) = -\frac{\gamma}{p_\theta}, \tag{98}$$

where

$$g(z) = (1+h^{-1})\left(\int\limits_{\{i:\, \tau_{f,i} \ge z\}} \tau_{f,i} d\mu(i) + zI\left(i:\ \tau_{f,i} \le z\right)\right) - z$$

and

$$g^0(z) = \bar{\tau}_f + h^{-1} \int\limits_{\{i:\, \tau_{f,i} \ge hz\}} \tau_{f,i} d\mu(i) - zI\left(i:\ \tau_{f,i} \ge hz\right).$$

Note first that for any $z > 0$ we have $g(z) > g^0(z)$. Indeed, because $0 < h \leq 1$

$$g(z) - g^0(z) = (1 + h^{-1}) \int\limits_{\{i:\ \tau_{f,i} \geq z\}} \tau_{f,i} d\mu(i) + (1 + h^{-1}) z I\left(i:\ \tau_{f,i} \leq z\right) - z$$

$$- \bar{\tau}_f - h^{-1} \int\limits_{\{i:\ \tau_{f,i} \geq hz\}} \tau_{f,i} d\mu(i) + z I\left(i:\ \tau_{f,i} \geq hz\right)$$

$$= - \int\limits_{\{i:\ \tau_{f,i} \leq z\}} \tau_{f,i} d\mu(i) - h^{-1} \int\limits_{\{i:\ hz \leq \tau_{f,i} \leq z\}} \tau_{f,i} d\mu(i) + z I\left(i:\ hz \leq \tau_{f,i} \leq z\right) + h^{-1} z I\left(i:\ \tau_{f,i} \leq z\right)$$

$$= \left[- \int\limits_{\{i:\ \tau_{f,i} \leq z\}} \tau_{f,i} d\mu(i) + h^{-1} z I\left(i:\ \tau_{f,i} \leq z\right) \right]$$

$$+ \left[-h^{-1} \int\limits_{\{i:\ hz \leq \tau_{f,i} \leq z\}} \tau_{f,i} d\mu(i) + z I\left(i:\ hz \leq \tau_{f,i} \leq z\right) \right]$$

The expressions in the brackets yield

$$- \int\limits_{\{i:\ \tau_{f,i} \leq z\}} \tau_{f,i} d\mu(i) + h^{-1} z I\left(i:\ \tau_{f,i} \leq z\right) \geq -z I\left(i:\ \tau_{f,i} \leq z\right) + h^{-1} z I\left(i:\ \tau_{f,i} \leq z\right)$$

$$= (h^{-1} - 1) z I\left(i:\ \tau_{f,i} \leq z\right),$$

$$- h^{-1} \int\limits_{\{i:\ hz \leq \tau_{f,i} \leq z\}} \tau_{f,i} d\mu(i) + z I\left(i:\ hz \leq \tau_{f,i} \leq z\right) \geq -h^{-1} z I\left(i:\ hz \leq \tau_{f,i} \leq z\right)$$

$$+ z I\left(i:\ hz \leq \tau_{f,i} \leq z\right) = (1 - h^{-1}) z I\left(i:\ hz \leq \tau_{f,i} \leq z\right),$$

so we get that

$$g(z) - g^0(z) \geq (h^{-1} - 1) z I\left(i:\ \tau_{f,i} \leq z\right) + (1 - h^{-1}) z I\left(i:\ hz \leq \tau_{f,i} \leq z\right)$$

$$= (h^{-1} - 1) z \left(I\left(i:\ \tau_{f,i} \leq z\right) - I\left(i:\ hz \leq \tau_{f,i} \leq z\right)\right) = (h^{-1} - 1) z I\left(i:\ \tau_{f,i} \leq hz\right) > 0$$

when $I\left(i:\ \tau_{f,i} \leq hz\right) > 0$.

From the proof of Proposition 3.5 we know that the function $g^0(\tau_\theta / p_\theta^2)$ is non-increasing for $p_\theta < 0$ and equation (3.30) has a unique solution p_θ^0, so for $p_\theta \leq p_\theta^0$

we have $g^0(\tau_\theta/p_\theta^2) \geq -\gamma/p_\theta$. Denote the smallest solution to equation (3.26) as $p_{\theta,min}$ and assume that $p_{\theta,min} \leq p_\theta^0$. Then,

$$g\left(\frac{\tau_\theta}{p_{\theta,min}^2}\right) = -\frac{\gamma}{p_{\theta,min}} \leq g^0\left(\frac{\tau_\theta}{p_{\theta,min}^2}\right) < g\left(\frac{\tau_\theta}{p_{\theta,min}^2}\right). \tag{99}$$

The obtained contradiction proves the statement of the proposition. Q.E.D.