Skill Acquisition and Data Sales

Shiyang Huang, Yan Xiong, and Liyan Yang*

July 2019

Abstract

We analyze a data-sales model in which investors acquire uncertain skills to interpret purchased data, thereby changing the data seller's behavior. When the seller owns accurate data, she optimally adds noise to the sold data to dampen information leakage via asset prices. If skill acquisition of investors is uncertain, the seller cannot fully control this information leakage. As a result, price informativeness can increase with skill-acquisition costs and decrease with the average level of investor skills. Our analysis helps explain some empirical regularities and highlights fundamental interactions between the asset management industry and the data industry.

Key words: Data sales, skill acquisition, gold rush, clarity, price informativeness, cost of capital, return volatility, alternative data

*Huang: Faculty of Business and Economics, University of Hong Kong; Email: huangsy@hku.hk. Xiong: Hong Kong University of Science and Technology; E-mail: yanxiong@ust.hk. Yang: Rotman School of Management, University of Toronto; E-mail: liyan.yang@rotman.utoronto.ca. This paper was previously titled "Skill Acquisition and Information Sales." We thank Pat Akey, Kerry Back, Robert Battalio, Bruno Biais, Philip Bond, Ling Cen, Hui Chen, Alexander Chinco, Peter Christoffersen, David Easley, Thierry Foucault, Diego García, David Hirshleifer, Jennifer Huang, Ron Kaniel, Hong Liu, Tom McCurdy, Joshua Mollner, Christine A. Parlour, Alessio Piccolo, Francesco Sangiorgi, Joel Shapiro, Chester Spatt, Laura Veldkamp, Xavier Vives, Ansgar Walther, Yajun Wang, Yizhou Xiao, Hongjun Yan, Mindy Zhang, and the audiences at the 2016 CMU Credit Ratings Conference, the 2017 China International Conference in Finance (CICF), the 2017 Financial Intermediation Research Society (FIRS) Conference, the 2017 Northern Finance Association Annual Conference (NFA), the 2017 Oxford Financial Intermediation Theory (OxFIT) Conference, the 2017 FTG Summer School on Liquidity in Financial Markets and Institutions, the 2017 SFS Cavalcade Asia-Pacific, and the seminar in CUHK Shenzhen for comments and suggestions. Liyan Yang thanks Social Sciences and Humanities Research Council of Canada (SSHRC) for financial support.

1 Introduction

Much data in financial markets is provided by professional data vendors (e.g., Bloomberg L.P., Thomson Reuters, and alternative data providers such as DataMinr and Interactive Data). The data covers all aspects of the economy, ranging from conventional variables (e.g., asset prices and macro data) to customer/investor sentiment, or even to as specific as crop yields calculated by satellite images or survey data on construction permits. One recent Wall Street Journal article makes a vivid description of the role of data in financial markets: "A new species is prowling America's most obscure industry conferences: the data hunter...Hedge funds and other sophisticated investors are increasingly relying on intermediaries like Mr. Haines, 35 years old, as they seek insights into a company's sales and health that aren't readily available from conventional sources." ("Wall Street's insatiable lust: data, data, data," *Wall Street Journal*, September 12, 2016).

Data providers sell their data to asset management companies such as hedge funds who in turn trade on the purchased data. However, interpreting and trading on data require skills.[1] Data is often called the new oil ("Data is giving rise to a new economy," *The Economist*, May 6, 2017), and like oil, data must go through a similar refinement process. Asset management companies hire analysts and data-science staff to decode the data to make better trading strategies. Only those institutions with high-powered skills can generate profits from the purchased data. Thus, both data itself and the skill to analyze data are scarce resources and so both should earn economic rents.[2] In addition, modern data analysis is often likened as gold rush in the sense that the process usually involves great uncertainty. This is particularly

[1] As Admati and Pfleiderer (1986, p.400–401) state, "(i)t is most convenient to envision information as a signal, a random variable that is jointly distributed with the state of the world." However, "(a) signal doesn't become a signal by chance. It takes an economist, scientist, or entrepreneur to unlock the value of an emerging data source before it truly becomes as a source of signals. The commercial value of discovering new signals, whether in business, science, or finance, has led to several instances of discovery for new and innovative data source." (Brown, 2010, p.167) Throughout the paper, we use the words "data", "information", and "signals" interchangeably, all of which refer to the products sold by data providers. We use the term "skills" to refer to the ability to convert the data to a signal that predicts the asset payoff.

[2] Echoing on this view, the recent empirical work by Berk and van Binsbergen (2015) shows that active mutual fund managers add \$3.2 million per year in Y2000 dollars. Meanwhile, according to a recent survey ("Global market data demand," Burton-Taylor International Consulting, 2016), with the fast-growing demand for data among financial institutions, global spending on information/analysis has grown to USD27.5 billion in 2016, with an annual growth rate of 2.6%.

true for the increasingly popular alternative data.[3] For instance, as a chief investment officer observes, "there's a gold rush in (alternative) data mining; most people that went off to prospect for gold came back penniless, but that doesn't mean there wasn't any gold." ("Hedge funds see a gold rush in data mining," *Financial Times*, August 29, 2017).

Despite the expanding availability of commercial databases, as well as the development of technology in processing the data, financial market prices have not necessarily become more informative. Figure 1 plots the time trend of price informativeness using the measure constructed by Bai, Philippon, and Savov (2016). Panel A presents price informativeness for long-lived firms. Consistent with Figure 7 in Farboodi, Matray, and Veldkamp (2017), price informativeness is almost stagnant in a relatively composition-bias-free sample. Panel B plots price informativeness for all listed firms. As in Bai, Philippon, and Savov (2016) and Farboodi, Matray, and Veldkamp (2017), price informativeness for the average listed firm is declining over time.

[FIGURE 1 ABOUT HERE]

In this paper, we provide a theoretical model to conceptualize the aforementioned facts regarding data sales and data processing, and to understand their implications for financial markets. Specifically, we aim to address the following questions: How do the skill properties of the buy side (e.g., the cost to acquire skills, skill levels) affect the sell side's data-sales decisions (e.g., the clarity and price of the sold data)? How do the interactions between the buy side (e.g., hedge funds, mutual funds) and the sell side (e.g., data vendors) affect market variables such as price informativeness, the cost of capital, and return volatility? What are the implications of the "gold rush" feature of skill acquisition for financial markets? How does skill acquisition affect the performance of the active asset management industry?

Our model extends the standard information-sales setting of Admati and Pfleiderer (1986) to allow for skill acquisition. One data seller sells data that is subsequently used by investors

[3] Alternative data is information gathered from non-traditional information sources. It can be compiled from various sources such as financial transactions, sensors, mobile devices, satellites, public records, and the Internet. Alternative data is being used by institutional investors such as hedge funds and mutual funds to generate alpha. For instance, in 2015, some hedge funds purchased satellite-imagery based traffic data from RS Metrics to successfully gauge JC Penney's quarterly earnings. Because alternative data is raw or unstructured, investors need to make large investments to acquire the necessary skills and infrastructure to trade on alternative data. See Appendix B for a detailed description about the markets for alternative data.

to speculate in a financial market. Investors spend resources to acquire skills that are useful for interpreting the purchased data. Those investors who do not acquire skills cannot understand the data and thus cannot trade on it. Skill-acquisition costs are heterogeneous across investors, which reflects the fact that some investors are better talented than others. To capture the gold rush feature of data analysis, we assume that after paying the costs to acquire skills, investors may become high-skilled or low-skilled with some probabilities.[4] The price of data is generated from bilateral Nash-bargaining games played by the seller and skilled investors, and in equilibrium, the total trading gains are split between the two sides (i.e., both data and skills earn economic rents).

Following the literature (e.g., Admati and Pfleiderer, 1986; García and Sangiorgi, 2011), we consider a signal structure that allows the seller to add conditionally independent noise (personalized noise) to the data being sold.[5] The precision of the added noise is jointly determined by data clarity (which is chosen by the seller) and investors' skills of decoding the purchased data. As in Admati and Pfleiderer (1986), the seller uses data clarity to control the information leakage effect via asset prices (that is, investors can free-ride on the information possessed by others through observing the equilibrium asset price). When the seller's original data is accurate, the seller adds personalized noise into the sold data in equilibrium to dampen the negative information-leakage effect on profits. By contrast, when the seller's original data is inaccurate, the seller does not introduce any noise and makes the sold data very easy to digest in equilibrium.

We find that market variables—in particular, price informativeness (how much information is revealed in the asset price), the cost of capital (the expected difference between the asset cash flow and its price), and return volatility—often exhibit different properties in the two types of economies. This is because when the seller adds noise in equilibrium, which makes data clarity endogenous. Given that we are interested in the scenario in which the seller plays a strategic role, we focus on economies in which the seller's original data is accurate, and leave to the Online Appendix the analysis of economies in which the seller is

[4]For example, although the buy side has invested a lot in new data sets and hiring data scientists, not every fund involved can process the data well and reach the best insights fast (see Appendix B).

[5]Admati and Pfleiderer (1986) show that personalized allocations dominate a large class of allocations, and conjecture that the monopolist cannot do better than in the personalized allocation.

endowed with inaccurate data.

We further consider two cases, without and with skill-acquisition uncertainty (i.e., the gold rush feature of skill acquisition). In the case without skill-acquisition uncertainty, the seller will add noise in a way such that price informativeness, the cost of capital, and return volatility all remain constant, which completely neutralizes the effect of changing skill-acquisition costs on these variables.

However, in the case with skill-acquisition uncertainty, the seller is no longer able to fully undo the information leakage effect. This is because investors ending up with different skills will use data differently for the same level of data clarity. Specifically, now investors are faced with uncertainty in skill acquisition and may end up being a low-skilled type (associated with low utilities), thereby reducing their incentives to acquiring skills in the first place. As skill-acquisition costs increase, in order to encourage investors to acquire skills and thus purchase data, the seller sells clearer data. In consequence, quite surprisingly, price informativeness can increase with skill-acquisition costs and both the cost of capital and return volatility can decrease with skill-acquisition costs. We also conduct comparative statics with respect to the average level of skills (skill mean) and the volatility of skills (skill volatility). Again, counterintuitively, price informativeness can decrease with the skill mean, and the cost of capital and return volatility can instead increase with it. Skill volatility has a non-monotonic effect on these market variables.

Our model helps understand some puzzling asset pricing facts. For example, as the real-time, granular data become easier to collect, the seller owns relatively accurate information about firms. With the long history of mature firms, the investors face little uncertainty in acquiring relevant skills to process the purchased data and thus, the seller is able to navigate a constant price informativeness. This is consistent with the stagnant price informativeness in Panel A of Figure 1, as well as that in Figure 7 of Farboodi, Matray, and Veldkamp (2017). On the other hand, the decreasing price informativeness of the average listed firm (especially small firms) in Panel B of Figure 1 can be understood as our data-sales economy with skill-acquisition uncertainty: as the technology to crunch data improves over time (i.e., the skill-acquisition cost decreases, or the skill mean increases), price informativeness can instead decrease.

Additionally, the findings in Brandt, Brav, Graham, and Kumar (2009) suggest a U-shaped relation between institutional ownership and return volatility. For instance, by considering a comparative statics with respect to skill volatility, our model generates a U-shaped relation between the population of skilled investors and return volatility. To the extent that skilled investors correspond to institutions, our result offers a potential explanation for the findings in Brandt, Brav, Graham, and Kumar (2009).

Our analysis further investigates other important variables. We find that the equilibrium population of investors acquiring skills and the seller's profits tend to move together in response to changes in virtually all exogenous primitive parameters. This finding points to a basic tenet that the asset management industry and the data industry are fundamentally connected via the financial market, so that both industries foster each other's growth and development. In the context of alternative data, a practical view is that the fast growth of alternative data sets may shrink the funds industry in the future, because some asset management firms cannot adapt to the radically changing landscape.[6] Our analysis suggests that the fast growth of alternative data may be accompanied with an expanding funds industry that will rely more on alternative data to seek investment insights.

Finally, we show that the performance of skilled investors (relative to the unskilled investors) improves as the cost of acquiring skills increases, which follows from the fact that skilled investors have to be compensated for the incurred skill-acquisition costs. This result offers a potential explanation for the performance heterogeneity in the active asset management industry across different assets and markets (Gârleanu and Pedersen, 2018). For example, given that skill-acquisition costs are larger for international financial markets than domestic ones, and larger for private companies than public ones, the active funds focusing on the former tend to deliver better performance.

Our paper contributes to two strands of literature. The first is the information-sales literature, such as Admati and Pfleiderer (1986, 1988), Allen (1990), Fishman and Hagerty

[6]For instance, a recent opimas article wrote: "The explosion of alternative data will require hedge funds and other asset managers to make large investments to acquire the necessary skills and infrastructure to leverage these sources of information. We expect that alternative data will contribute significantly to a further shrinkage in the hedge fund population, as firms unable to exploit the information needed to compete effectively in the new world of intelligent investing will fall behind." ("AI and alternative data: Moving to trading's next model," Axel Pierron, July 24, 2017, http://www.opimas.com/research/267/detail/)

(1995), Naik (1997), Cespa (2008), García and Sangiorgi (2011), Chen and Wilhelm (2012), and Easley, O'Hara, and Yang (2016). In this literature, all investors have equal (full) capacity to interpret and trade on information. We instead relax this assumption by allowing investors to acquire skills to trade on the purchased information. Our paper is about "direct sales of information," where the data seller sells data to investors who in turn trade on the data. Admati and Pfleiderer (1990) and García and Vanden (2009) have considered "indirect sales of information," where information sellers set up funds and sell fund shares. Gârleanu and Pedersen (2018) explore how asset managers produce information and then interact with clients. Their paper focuses more on the buy-side information, i.e., the information produced by fund managers. By contrast, our paper is more on the sell-side information, i.e., the information produced by data vendors. In our setting, fund managers (skilled investors) do not acquire fundamental information directly; rather, they acquire the skills to analyze and trade on the data purchased from the sell side.

A second stream of related research considers the possibility that investors have different abilities for processing data, albeit in different settings. Indjejikian (1991) and Pagano and Volpin (2012) analyze models in which investors have different sophistication levels in understanding public disclosure—i.e., free data—made by firms or asset issuers. Vives and Yang (2017) analyze a setting in which investors need to spend resources to acquire skills in interpreting the asset price, which is again freely observable by investors. In our setting, investors have different skills in understanding the data sold by a data vendor, and there are important interactions between skill-acquisition activities of investors and information-sales activities of the seller. Myatt and Wallace (2012) consider a "beauty-contest" coordination game in which players choose how much costly attention to pay to various informative signals. In their setting, each signal has an underlying accuracy (how precisely it identifies the state) and a clarity (how easy it is understood). Our information structure closely follows Myatt and Wallace (2012), but in our setting the clarity is endogenously chosen by the data vendor and is affected by how investors acquire skills for interpreting the data.

2 A Model of Skill Acquisition and Data Sales

We study a data-sales setting with one data seller and a continuum of investors. We extend the classic information-sales literature, such as Admati and Pfleiderer (1986) and García and Sangiorgi (2011), in two dimensions. First, interpreting and trading on data require skills, and investors need to pay a cost to acquire these skills with uncertainty. Second, to compensate skilled investors for the costly skill acquisition, we introduce Nash bargaining to split the rents from trading on the purchased data between the seller and skilled investors. Our economy has three dates, $t = 0$, 1, and 2. The order of events is described in Figure 2. At date 0, the data market opens and investors make skill-acquisition decisions. At date 1, the asset market opens and investors trade. At date 2, the assets pay off and all agents consume. For ease of reference, the main model variables are tabulated and explained in Appendix A.

[FIGURE 2 ABOUT HERE]

2.1 Assets and Asset Markets

There are two tradable assets in the date-1 financial market: a risk-free asset and a risky asset. The payoff of the risk-free asset has a constant value of 1 for simplicity, and is in unlimited supply. The risky asset is traded at an endogenous price $\tilde{p}$ per unit in the date-1 market, and its total supply is normalized as one share. The risky asset pays an uncertain cash flow $\tilde{\theta}$ at date 2, where $\tilde{\theta} \sim N\left(0, \tau_\theta^{-1}\right)$ with $\tau_\theta \in (0, \infty)$. There is noisy demand $\tilde{u}$ for the risky asset, where $\tilde{u} \sim N\left(0, \tau_u^{-1}\right)$ with $\tau_u \in (0, \infty)$. Therefore, the effective supply of the risky asset is $1 - \tilde{u}$. Noisy trading $\tilde{u}$ is independent of all other random variables in the economy, and it serves the usual role of preventing private information from being perfectly revealed by the price.

2.2 Information Structure

We follow Admati and Pfleiderer (1986) and assume that there is one data seller. This data seller can be broadly interpreted as professional data vendors (e.g., Bloomberg L.P. and alternative data providers such as DataMinr and Interactive Data) or brokerage analysts who sell various reports and newsletters to buy-side clients. The seller is endowed with information about the asset payoff $\tilde{\theta}$, in the form of $\tilde{\theta}+\tilde{\eta}$, where $\tilde{\eta} \sim N\left(0, \tau_\eta^{-1}\right)$, $\tau_\eta \in (0,\infty)$, and $\tilde{\eta}$ is independent of $\tilde{\theta}$. Parameter τ_η measures the accuracy level of the seller's endowed information.

Our analysis allows sales of data with personalized noises, as considered by Admati and Pfleiderer (1986, 1988).[7] That is, when selling data, the seller can add idiosyncratic noises into her endowed signal $\tilde{\theta}+\tilde{\eta}$. One often-cited interpretation for personalized signals is that data vendors intentionally pass on the data to investors in a vague way, so that investors make independent interpretations of the data (see Admati and Pfleiderer (1986) and García and Sangiorgi (2011) for more discussions). The innovation of our approach is that investors have heterogeneous skills of interpreting the purchased data.

Specifically, the seller offers investor i of type k a signal $\tilde{s}_{k,i}$ of the form

$$\tilde{s}_{k,i} = \tilde{\theta} + \tilde{\eta} + \tilde{\varepsilon}_{k,i}, \text{ with } \tilde{\varepsilon}_{k,i} \sim N\left(0, \frac{1}{x z_k}\right), \tag{1}$$

where $(\tilde{\theta}, \tilde{\eta}, \{\tilde{\varepsilon}_{k,i}\}, \tilde{u})$ are mutually independent. Specification (1) follows closely the information structure described in Myatt and Wallace (2012): the term $\tilde{\eta}$ is "sender noise," while the personalized error term $\tilde{\varepsilon}_{k,i}$ represents "receiver noise." Variable $x \in [0,\infty]$ is a constant chosen by the seller, and it controls the "clarity" of the data passed to investors. We allow x to take values of 0 and ∞. If $x = 0$, the seller is selling a completely uninformative signal since she is adding an error with an infinite variance for any $z_k > 0$. Instead, if $x = \infty$, the seller does not introduce any personalized noise into the sold data, as the variance of $\tilde{\varepsilon}_{k,i}$ degenerates to zero.

Variable z_k represents investors' skills of interpreting the data. We assume that z_k can take three values: 0, Z_L, and Z_H; that is, $z_k \in \{0, Z_L, Z_H\}$, where $0 \leq Z_L \leq Z_H < \infty$. When $z_k = 0$, investors are unskilled, and they cannot process the data at all. These investors can represent retail investors or passive funds. When $z_k > 0$, investors are skilled

[7] We have also considered the possibility of the seller adding photocopied noises. Nonetheless, we find that in equilibrium the seller optimally chooses not to add photocopied noises into the sold data. This is because personalized signals always dominate photocopied ones in terms of the seller's profits (Admati and Pfleiderer, 1986).

and can interpret the data, and they can represent active funds.[8] Skilled investors admit two types, High ($z_k = Z_H$) or Low ($z_k = Z_L$), so that we allow heterogeneity among active fund managers, as documented in Berk and van Binsbergen (2015). This feature of data-interpretation skills is particularly relevant in the context of alternative data, because most alternative data sources are not in a format that lends itself directly to investing.

2.3 Investors

There exists a continuum of investors, indexed by $i \in [0,1]$. Investors acquire skills and purchase data at date 0, trade assets at date 1, and consume at date 2. All investors derive expected utility over their date-2 wealth according to a constant absolute risk aversion (CARA) utility with a common risk aversion coefficient γ.

At date 0, knowing the data clarity, investors first decide whether to acquire skills to process the data. Acquiring skills incurs costs. For instance, if we interpret investors as individuals, these costs can be resources spent on attending educational programs. If we interpret investors as institutions, these costs can represent costs associated with hiring data specialists and with building infrastructure to analyze data. We use λ to denote the fraction of investors who decide to acquire skills. Skill-acquisition costs are specified to be heterogeneous among investors. Let $C(i)$ denote investor i's skill-acquisition cost. Without loss of generality, we assume that $C(i)$ is increasing in i. In addition, we assume that $C(i)$ is continuous, $C(0) = 0$, and $C(1) = +\infty$. This ensures that the equilibrium mass λ^* of skilled investors always admits an interior solution, that is, $\lambda^* \in (0,1)$.

Skill acquisition is uncertain. After investor i decides to acquire skills, her skill level may end up with Z_L or Z_H with equal probabilities, where $0 \leq Z_L \leq Z_H < \infty$. We refer to investors with skill level Z_H as "high-type skilled investors," and to those with skill level Z_L as "low-type skilled investors." As a result, there are a mass $\frac{\lambda}{2}$ of high-type skilled investors (e.g., star active fund managers), a mass $\frac{\lambda}{2}$ of low-type skilled investors (e.g., mediocre active

[8] In reality, fund managers may not only process information provided by the sell side (such as data vendors or analysts), but also develop information on their own. This buy-side information can be simply incorporated into our setting by assuming that skilled investor i develops a signal $\tilde{s}_{b,i} = \tilde{\theta} + \tilde{\xi}_i$, where $\tilde{\xi}_i \sim N(0, \tau_\xi^{-1})$ with $\tau_\xi \in (0,\infty)$. Gârleanu and Pedersen (2018) have studied information production by funds.

fund managers), and a mass $1-\lambda$ of unskilled investors (e.g., retail investors or passive funds). We define the mean and the volatility of skill levels as

$$\bar{Z} \equiv \frac{Z_H + Z_L}{2} \text{ and } \Delta \equiv \frac{Z_H - Z_L}{2}, \tag{2}$$

respectively. We introduce uncertainty into the skill-acquisition process to reflect the fact that to decode data resembles "gold rushes," as mentioned in the Introduction.

After making skill-acquisition decisions, investors decide whether to purchase the data. Naturally, unskilled investors will not buy data since they lack the ability to process data. Skilled investors will buy data even though they may end up with low skills. We assume that skilled investors make data-purchase decisions before they uncover their types; that is, only after investors exploit some data, can they understand whether they can successfully decode and then trade on the data. This assumption is consistent with the fact that when purchasing alternative data, active fund managers are not sure if they can unlock the data successfully and timely. Similarly, in the context of machine learning, institutions can know whether their algorithms work only after exploring a large amount of data.

At date 1, both skilled and unskilled investors trade in the financial market. As standard in the rational-expectations equilibrium literature, all investors submit demand schedules and they can condition their trades on prices. Skilled investors also observe the signals purchased from the seller and uncover their skill levels. Formally, investor i chooses demand D_i for the risky asset to maximize $E\left[-e^{-\gamma D_i(\tilde{\theta}-\tilde{p})}\middle|\mathcal{F}_i\right]$, where $\mathcal{F}_i$ represents investor i's information set (including the asset price $\tilde{p}$).

2.4 Nash Bargaining and Data Price

Both data and skills are scarce resources and thus both factors should earn economic rents. We use Nash bargaining as a modeling device to implement profit sharing. Our approach is similar to Gârleanu and Pedersen (2018) who employ Nash bargaining to determine the fee charged by mutual funds, and is also consistent with Anand and Galetovic (2000) who argue that when the information seller has local monopoly power, information prices are often determined through bilateral bargaining between the buyer and the seller. In reality, the Nash bargaining captures the feature of modern markets for financial information; for example, the buy-side often hires teams of people to speak with data owners, and the two sides bargain over the price directly (see Appendix B for a detailed description on how data vendors and the buy-side negotiate the data price bilaterally). In our setting, each skilled investor and the seller form a pair to engage in Nash bargaining. The bargaining outcome yields the price q of the data sold by the seller.

We use $\beta \in [0, 1]$ to denote the bargaining power of the data seller. When $\beta = 1$, the seller has the full bargaining power and will extract all surplus resulting from the informed trading, which corresponds to the settings analyzed in the literature.[9] When $\beta < 1$, investors as data buyers have non-negligible bargaining power and can keep some surplus in the bargaining game. One may argue that $\beta = 1$ is more applicable in a setting with a single seller. But note that in our setting, the single-seller assumption is just a simplification, and being the only seller does not imply that she can extract all the rents. In fact, if the seller was able to extract all rents, then no investors would acquire skills in our model economy and thus the seller could not sell any data. In reality, the sell side typically consists of multiple sellers, and we can treat these sellers as the representative seller in our setting (by ignoring the competition effect). As Berk and van Binsbergen (2015) document, active fund managers do add value to their companies (about \$3.2 million per year) and this value is largely captured by asset management companies, which suggests that skills earn rents in reality.

2.5 The Seller's Problem

A mass λ of investors purchase data at price q. So, the profit of selling data is $\pi = \lambda q$. At the very beginning of the economy, the seller chooses data clarity x in data sales to maximize the profit π. That is, the seller's optimization problem is:

$$\max_{x} \lambda(x)\, q(x)\,. \tag{3}$$

In the above problem, the seller is forward looking in the sense that she anticipates how the choice of x affects the data demand λ and the data price q.

[9]When the skill-acquisition cost C is 0, our setting with $\beta = 1$ degenerates to Admati and Pfleiderer (1986).

3 Equilibrium

The economy is defined by a tuple of seven exogenous parameters and one exogenous function, $\mathcal{E} \equiv \{\tau_\theta, \tau_\eta, \tau_u, \gamma, \beta, Z_H, Z_L, C(\cdot)\}$. Note that parameters $\bar{Z}$ and Δ specified in (2) are equivalent to $\{Z_H, Z_L\}$ given that $Z_H = \bar{Z} + \Delta$ and $Z_L = \bar{Z} - \Delta$. For a given economy $\mathcal{E}$, we define an equilibrium in a subgame-perfect sense.

Definition 1 *An equilibrium is characterized by a clarity level x^* of the data sold by the seller, a skill-acquisition decision function $A(i;x)$ of investor i, a mass $\lambda(x)$ of skilled investors, a data price $q(x)$, a price function $p(\tilde{\theta} + \tilde{\eta}, \tilde{u})$, and a demand schedule $D(\mathcal{F}_i)$ of investor i, such that:*

(1) In the date-1 financial market, (i) demand schedule $D(\mathcal{F}_i)$ maximizes investor i's expected utility conditional on her information set $\mathcal{F}_i$; and (ii) the price function $p(\tilde{\theta} + \tilde{\eta}, \tilde{u})$ clears the asset market almost surely.

(2) At date 0, (i) the data price $q(x)$ is generated from Nash bargaining between the seller and a typical skilled investor; (ii) the skill-acquisition decision function $A(i;x)$ maximizes investor i's ex-ante expected utility (where $A(i;x) = 1$ or 0), and $\lambda(x) = \int_0^1 A(i;x)\,di$; and (iii) x^* maximizes the seller's profit $\lambda(x)\,q(x)$.

3.1 Financial Market Equilibrium

The equilibrium concept in the financial market is the standard noisy rational-expectations equilibrium (noisy-REE). Constructing a noisy-REE boils down to solving a price function that depends on skilled investors' private information $\tilde{s}_{k,i}$ and noise trading $\tilde{u}$. By the law of large numbers, the noise terms $\tilde{\varepsilon}_{k,i}$ in the private signals $\tilde{s}_{k,i}$ will wash out and thus the price $\tilde{p}$ becomes a function of $(\tilde{\theta} + \tilde{\eta}, \tilde{u})$. We follow the literature and consider a linear price function as follows:

$$\tilde{p} = a_0 + a_\theta(\tilde{\theta} + \tilde{\eta}) + a_u \tilde{u}, \tag{4}$$

where the a-coefficients will be endogenously determined in equilibrium.

For any investor, the information contained in the price is equivalent to the following signal $\tilde{s}_p$:

$$\tilde{s}_p \equiv \frac{\tilde{p} - a_0}{a_\theta} = \tilde{\theta} + \tilde{\eta} + \alpha^{-1}\tilde{u}, \text{ with } \alpha \equiv \frac{a_\theta}{a_u}, \tag{5}$$

where $\tilde{s}_p$ is normally distributed, with mean $\tilde{\theta} + \tilde{\eta}$ and precision $\alpha^2 \tau_u$. The CARA-normal setup implies that investor i's demand function is

$$D(\mathcal{F}_i) = \frac{E(\tilde{\theta}|\mathcal{F}_i) - \tilde{p}}{\gamma Var(\tilde{\theta}|\mathcal{F}_i)}, \tag{6}$$

where $\mathcal{F}_i$ is investor i's information set (including the asset price $\tilde{p}$).

Investors differ in their data-processing abilities and hence information sets. In the date-1 financial market, there are three types of investors depending on their skill levels of processing data: (i) high-type skilled investors, labeled by H, with data-processing ability Z_H; (ii) low-type skilled investors, labeled by L, with data-processing ability Z_L; and (iii) unskilled investors, labeled by U, without any data-processing ability. The masses of the three types of investors are $\frac{\lambda}{2}$, $\frac{\lambda}{2}$, and $1 - \lambda$, respectively.

An H-type skilled investor i has information set $\{\tilde{p}, \tilde{s}_{H,i}\}$, which is equivalent to $\{\tilde{s}_p, \tilde{s}_{H,i}\}$. Her demand function is $D_H(\tilde{p}, \tilde{s}_{H,i}) = \frac{E(\tilde{\theta}|\tilde{p}, \tilde{s}_{H,i}) - \tilde{p}}{\gamma Var(\tilde{\theta}|\tilde{p}, \tilde{s}_{H,i})}$. By equation (1), private signal $\tilde{s}_{H,i}$ has precision xZ_H in predicting $\tilde{\theta} + \tilde{\eta}$. Applying Bayes' rule, we can compute the conditional moments of an H-type skilled investor i as follows:

$$E(\tilde{\theta}|\tilde{p}, \tilde{s}_{H,i}) = \frac{\tau_\eta}{\tau_\theta + \tau_\eta} \frac{\alpha^2 \tau_u \tilde{s}_p + xZ_H \tilde{s}_{H,i}}{\frac{\tau_\theta \tau_\eta}{\tau_\theta + \tau_\eta} + \alpha^2 \tau_u + xZ_H}, \tag{7}$$

$$Var(\tilde{\theta}|\tilde{p}, \tilde{s}_{H,i}) = \frac{1}{\tau_\theta + \tau_\eta} + \left(\frac{\tau_\eta}{\tau_\theta + \tau_\eta}\right)^2 \frac{1}{\frac{\tau_\theta \tau_\eta}{\tau_\theta + \tau_\eta} + \alpha^2 \tau_u + xZ_H}. \tag{8}$$

An L-type skilled investor i has information set $\{\tilde{p}, \tilde{s}_{L,i}\}$, equivalent to $\{\tilde{s}_p, \tilde{s}_{L,i}\}$, where $\tilde{s}_{L,i}$ has lower precision xZ_L in predicting $\tilde{\theta} + \tilde{\eta}$. Her demand function is $D_L(\tilde{p}, \tilde{s}_{L,i}) = \frac{E(\tilde{\theta}|\tilde{p}, \tilde{s}_{L,i}) - \tilde{p}}{\gamma Var(\tilde{\theta}|\tilde{p}, \tilde{s}_{L,i})}$. We can compute the conditional moments as follows:

$$E(\tilde{\theta}|\tilde{p}, \tilde{s}_{L,i}) = \frac{\tau_\eta}{\tau_\theta + \tau_\eta} \frac{\alpha^2 \tau_u \tilde{s}_p + xZ_L \tilde{s}_{L,i}}{\frac{\tau_\theta \tau_\eta}{\tau_\theta + \tau_\eta} + \alpha^2 \tau_u + xZ_L}, \tag{9}$$

$$Var(\tilde{\theta}|\tilde{p}, \tilde{s}_{L,i}) = \frac{1}{\tau_\theta + \tau_\eta} + \left(\frac{\tau_\eta}{\tau_\theta + \tau_\eta}\right)^2 \frac{1}{\frac{\tau_\theta \tau_\eta}{\tau_\theta + \tau_\eta} + \alpha^2 \tau_u + xZ_L}. \tag{10}$$

An unskilled investor can forecast the asset fundamental only based on price $\tilde{p}$, and her demand function is $D_U(\tilde{p}) = \frac{E(\tilde{\theta}|\tilde{p})-\tilde{p}}{\gamma Var(\tilde{\theta}|\tilde{p})}$. Using Bayes' rule, we have

$$E(\tilde{\theta}|\tilde{p}) = \frac{\tau_\eta}{\tau_\theta+\tau_\eta}\frac{\alpha^2\tau_u\tilde{s}_p}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u}, \tag{11}$$

$$Var(\tilde{\theta}|\tilde{p}) = \frac{1}{\tau_\theta+\tau_\eta}+\left(\frac{\tau_\eta}{\tau_\theta+\tau_\eta}\right)^2\frac{1}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u}. \tag{12}$$

The market-clearing condition is

$$\int_0^{\lambda/2} D_H(\tilde{p},\tilde{s}_{H,i})\,di+\int_{\lambda/2}^{\lambda} D_L(\tilde{p},\tilde{s}_{L,i})\,di+(1-\lambda)\,D_U(\tilde{p})+\tilde{u}=1. \tag{13}$$

To derive the equilibrium price function, we insert the demand functions into the market-clearing condition to solve the price in terms of $\tilde{\theta}+\tilde{\eta}$ and $\tilde{u}$, and then compare with the conjectured price function in equation (4) to obtain and solve a system that characterizes the unknown a-coefficients.

Proposition 1 (Financial market equilibrium) *Given $(\lambda, x; \tau_\theta, \tau_\eta, \tau_u, \gamma, Z_H, Z_L)$, there exists a unique linear equilibrium price function, given by equation (4), where a_0, a_θ, and a_u are given in Appendix C. The equilibrium is characterized by the ratio $\alpha = \frac{a_\theta}{a_u}$, which is the unique real positive root of the following equation:*

$$\gamma\alpha\left(xZ_H+\tau_u\alpha^2+\tau_\eta\right)-\frac{1}{2}x\lambda\tau_\eta\left(Z_H+Z_L+\frac{xZ_L\left(Z_H-Z_L\right)}{xZ_L+\tau_u\alpha^2+\tau_\eta}\right)=0. \tag{14}$$

The endogenous variable α defined by (5) is positively related to price informativeness, which measures how much fundamental information is revealed by the price. Formally, we can measure price informativeness by $\frac{1}{Var(\tilde{\theta}|\tilde{p})}$ (e.g., Grossman and Stiglitz, 1980). According to equation (12), price informativeness $\frac{1}{Var(\tilde{\theta}|\tilde{p})}$ is determined by four variables: $\tau_\theta, \tau_\eta, \tau_u$, and α. To the extent that parameters τ_θ, τ_η, and τ_u are exogenous, the endogenous variable α indeed captures price informativeness.

Corollary 1 (Price informativeness) *Given $(x; \tau_\theta, \tau_\eta, \tau_u, \gamma, Z_H, Z_L)$, price informativeness $\frac{1}{Var(\tilde{\theta}|\tilde{p})}$ increases with the mass λ of skilled investors.*

3.2 Nash-Bargaining Equilibrium

At date 0, the data price is determined as an outcome of Nash bargaining between the seller and one skilled investor who has not uncovered her skill type yet. Recall that the bargaining equilibrium depends on agents' utilities in the events of agreement versus no agreement. For the data seller, if an agreement is achieved, her additional profit will increase by the data price q. For the investor, we follow Gârleanu and Pedersen (2018) and specify the investor's bargaining objective in terms of certainty-equivalent wealth. That is, the utility when agreeing on a data price q is $CE_S - q$, where CE_S is the ex-ante certainty equivalent of the skilled investor. If no agreement is reached, the investor's outside option is to invest as the unskilled, yielding a utility of CE_U, where CE_U is the ex-ante certainty equivalent of an unskilled investor. Hence, the investor's gain from agreement is $CE_S - CE_U - q$.

The difference $CE_S - CE_U$ captures the total gain from interpreting and trading on the data. We denote this gain by G, that is, $G \equiv CE_S - CE_U$. Note that this trading gain G has gross of the skill-acquisition cost C. That is, when we compute CE_S, we follow Gârleanu and Pedersen (2018) and do not subtract the skill-acquisition cost C, because at the bargaining stage, an investor has already invested to acquire skills and thus the cost C is sunk. In Appendix C, we compute the trading gain G as follows:

$$G(x,\alpha) = \frac{1}{\gamma} \ln \left[\frac{\sqrt{Var(\tilde{\theta}|\tilde{p})}}{\frac{1}{2}\sqrt{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})} + \frac{1}{2}\sqrt{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})}} \right], \tag{15}$$

where $Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})$, $Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})$, and $Var(\tilde{\theta}|\tilde{p})$ are given by equations (8), (10), and (12), respectively. In equation (15), we have explicitly expressed G as a function of the two endogenous variables (x,α).

In the Nash-bargaining game, the seller's gain from agreement is the data price q and the investor's gain from agreement is $G - q$. The seller has a bargaining power $\beta \in [0,1]$. The bargaining outcome maximizes the product of the utility gains from agreement:

$$\max_q q^{\beta} (G-q)^{1-\beta} \Rightarrow q^* = \beta G, \tag{16}$$

where q^* denotes the equilibrium data price.

In equilibrium, the total gain G generated from a skilled investor (who interprets and trades on data) is split between the seller and the investor as follows:

$$G = q + (G-q) = \underbrace{\beta G}_{\text{seller's rent}} + \underbrace{(1-\beta)\, G}_{\text{skilled investor's rent}}. \tag{17}$$

Intuitively, if we interpret a skilled investor as an active fund, then part of the fund's revenue goes to data providers (as compensation for data provision) and the remaining is left to the fund (as compensation for trading skills). The gain division is determined by the bargaining power $(1-\beta)$ of the fund manager in negotiating prices with data providers.

3.3 Skill-Acquisition Decisions

An investor trades off between the benefit and cost of acquiring skills. If investor i decides to acquire skills, she expects to receive a benefit of $(1-\beta)\,G$ resulting from the Nash-bargaining game, but acquiring skills incurs a cost of $C\,(i)$. If $(1-\beta)\,G > C\,(i)$, investor i decides to acquire skills; otherwise, she decides not. Thus,

$$A\,(i) = \begin{cases} 1, & \text{if } (1-\beta)\,G \geq C\,(i)\,, \\ 0, & \text{otherwise,} \end{cases} \tag{18}$$

where $A\,(i) = 1$ indicates that investor i chooses to acquire skills.

Since $C\,(i)$ monotonically increases from 0 to $+\infty$, there always exists an interior equilibrium mass $\lambda^* \in (0,1)$ of skilled investors. That is, $A\,(i) = 1$ as long as $i \in [0, \lambda^*]$. The equilibrium mass λ^* is determined by

$$(1-\beta)\,G\,(x,\alpha) = C\,(\lambda^*)\,, \tag{19}$$

where $G\,(x,\alpha)$ is given by equation (15).

3.4 The Seller's Decision and the Overall Equilibrium

The seller chooses data clarity x to maximize profits: $\max_x \lambda\,(x)\,q\,(x)$. In computing this maximization problem, we need to figure out functions $\lambda\,(x)$ and $q\,(x)$ in two steps. First, we use the financial market equilibrium condition (14) and the skill-acquisition equilibrium condition (19) to jointly solve α and λ as functions of x. In Appendix C, we show that these two functions are well defined: for a given x, there exists a unique pair (α, λ) that solves (14) and (19). Second, we use the Nash-bargaining equilibrium condition (16) to express q as a function of x: $q\,(x) = \beta G\,(x, \alpha\,(x))$.

Proposition 2 (Overall equilibrium) *There exists an overall equilibrium such that:*
(1) Clarity level $x^* \in [0, +\infty]$ *solves the seller's profit maximization problem (3), where*

the population of skilled investors $\lambda(x^*)$ and the data price $q(x^*)$ are jointly determined by equations (14), (16), and (19);

(2) Demand schedules are given by equation (6), and the price function is characterized by Proposition 1.

3.5 Two Types of Equilibria

Two types of data clarity x^* arise in equilibrium, depending on the accuracy level τ_η of the seller's original data. When data accuracy τ_η is high, the seller sells data in a way with limited clarity to control the information leakage via the asset price. On the other hand, when data accuracy τ_η is low, the seller sells data "as is."

The data seller needs to balance two competing forces in choosing clarity x to maximize profit $\pi(x) = \lambda(x) q(x)$. On one hand, the data price q is negatively related to the amount of information revealed by the price (information leakage) and thus, the seller has an incentive to decrease the amount of information in the price by adding noise to the sold data (i.e., $q(x)$ decreases with x). On the other hand, adding noise directly lowers the data value, reducing investors' willingness to pay and hence the population λ of skilled investors (i.e., $\lambda(x)$ increases with x). When the data accuracy is sufficiently high, the seller is more concerned with information leakage and thus adds noise to the sold data, leading to a finite level x^* of data clarity. By contrast, if the data accuracy is sufficiently low, the seller cannot afford to add noise and she simply sells the data "as is," i.e., the optimal clarity level x^* is ∞.

Proposition 3 (Data clarity x^*) *In equilibrium, there exist two types of data clarity x^*:*
(1) If $\lambda^*\tau_\eta > \alpha^*\gamma$, the seller sells data with added noise, i.e., $x^* < \infty$;
(2) If $\lambda^*\tau_\eta \leq \alpha^*\gamma$, the seller sells data "as is," i.e., $x^* = \infty$.

In the remaining of this paper, we focus on economies in which the seller's data is sufficiently accurate (i.e., high τ_η) for the following reasons. First, only when the seller owns accurate data does she engage in strategic sales of data; otherwise she just sells her data "as is." Given that we are interested in the effect of data markets on financial markets, we emphasize the case in which the seller plays an active role. Second, as will be shown below, many novel theoretical results often emerge in economies with a high data-accuracy level, in which the data seller can adjust her data clarity in response to changes in exogenous parameters. We leave the complete analysis of the case in which the seller owns inaccurate data to the Online Appendix.

4 Comparative Statics

In this section, we examine how skill-acquisition technology feeds back into data sales, financial markets, and the asset management industry, by conducting comparative statics with respect to parameters governing the skill-acquisition cost function $C(\cdot)$ and parameters governing skill levels ($\bar{Z}$ and Δ). We consider two economies. In the first economy, we shut down the "gold rush" feature of data interpretation, so that a skilled investor knows for sure her ability to interpret the purchased data (i.e., $\Delta = 0$, or equivalently, $Z_L = Z_H$). We find that when the accuracy level τ_η of the seller's original data is sufficiently high, the seller can fully undo the price's information leakage effect. In the second economy, a skilled investor faces uncertainty about her ability to read the purchased data (i.e., $\Delta > 0$, or equivalently, $Z_L < Z_H$). In the presence of skill uncertainty, the seller is no longer able to use data clarity to fully neutralize the information leakage effect even when she owns very accurate data. This leads to novel theoretical results such as that price informativeness can increase with skill-acquisition costs and can decrease with the average skill level of investors.

4.1 Skill Acquisition without Uncertainty

Assume $Z_L = Z_H$. So, by paying a cost, an investor obtains for sure a prespecified skill level of interpreting data. We conduct comparative statics with respect to skill-acquisition costs. To facilitate analysis, we specify that the skill-acquisition cost function $C(\cdot)$ is governed by parameter $c > 0$, so that an increase in c shifts the entire curve $C(\cdot)$ upward (i.e., $\frac{\partial C}{\partial c} > 0$). Skill-acquisition cost parameter c can be associated with market conditions, technology changes, or development of professional educational programs. For instance, it may be more costly for fund managers to develop skills in a more complex asset market. Also, it is more costly for active funds to hire researchers if the wage rates in labor markets become higher.

We use Figure 3 to graphically illustrate our results for $C(i) = \frac{ci}{1-i}$ with $c > 0$. The parameter values are as follows: $\tau_\theta = 1, \tau_u = 5, \tau_\eta = 1000, \gamma = 1, \beta = 0.3$, and $Z_L = Z_H = 10$.

[FIGURE 3 ABOUT HERE]

It is intuitive that as c increases, fewer investors choose to acquire skills (i.e., $\frac{\partial \lambda^*}{\partial c} < 0$). This feeds back to the supply side of data (data clarity x^* and data price q^*). With very accurate original data (i.e., high τ_η), the seller optimally adds noise into the sold data. Given that there are fewer skilled investors caused by the increase in skill-acquisition costs, the information leakage effect is less of a concern and thus, the seller chooses to add less noise and deliver cleaner data (i.e., $\frac{\partial x^*}{\partial c} > 0$). In equilibrium, the improvement in data clarity exactly cancels the adverse effect of the skill-acquisition cost on the mass of skilled investors, thereby making price informativeness unchanged.[10] This means that the seller completely undoes the effect of the exogenous skill-acquisition costs on price informativeness, and "brings the entire market to an information level that is only a function of her own information" (Admati and Pfleiderer, 1986, p.426). Recall that data price q^* is in proportion to the trading gain, which is in turn determined by the extra precision brought about interpreting the purchased data in addition to the information revealed by the price (see equations (15) and (16)). Hence, given the constant price informativeness, better data clarity leads to higher precision of interpreting the purchased data, which raises the equilibrium data price (i.e., $\frac{\partial q^*}{\partial c} > 0$).

In addition to price informativeness, we also examine two additional asset price variables: the cost of capital $E(\tilde{\theta} - \tilde{p})$ and return volatility $\sigma(\tilde{\theta} - \tilde{p})$. The cost of capital is the expected difference between the cash flow generated by the risky asset and its price. This difference arises from the compensation required to induce investors to hold the risky asset. The return on the risky asset is $\tilde{\theta} - \tilde{p}$, and thus its volatility can be measured by $\sigma(\tilde{\theta} - \tilde{p})$. Since the seller adds noise to navigate a fixed level of price informativeness, the uncertainty faced by investors does not change with the skill-acquisition cost, making the cost of capital and return volatility constant.

Finally, we report the seller's profit and skilled investors' performance. The seller's profit

[10] Specifically, the equilibrium price informativeness is a function of τ_η and τ_θ only, i.e., $\frac{1}{Var(\tilde{\theta}|\tilde{p})} = \frac{2\tau_\theta(\tau_\eta + \tau_\theta)}{\tau_\eta + 2\tau_\theta}$.

is $\pi^* = \lambda^* q^*$. As c increases, λ^* decreases but q^* increases. Nonetheless, the effect of λ^* dominates, so that the seller's profit π^* decreases with c. We follow Gârleanu and Pedersen (2018) and proxy the performance of skilled investors using the disparity between a skilled investor's certainty equivalent and an unskilled investor's certainty equivalent. Specifically, the performance of skilled investor i of type k is

$$Performance_{k,i} = \frac{1}{2\gamma} \ln \frac{Var(\tilde{\theta}|\tilde{p})}{Var(\tilde{\theta}|\tilde{p}, \tilde{s}_{k,i})}, \text{ where } i \in [0,1], \text{ and } k \in \{H, L\}, \tag{20}$$

which captures the additional expected utility (outperformance) of a high- or low-skilled investor relative to an unskilled (uninformed) one. As Gârleanu and Pedersen (2018) pointed out, an investor's expected utility is directly linked to her (squared) Sharpe ratio. We do not subtract data price q or skill-acquisition costs $C(i)$ from the performance because in practice fund performance is examined without consideration of the expenses assumed by the fund side. The performance of a typical skilled investor is the weighted average of the performance of a high- and a low-skilled investor:

$$Performance = \frac{1}{2} \left(\frac{1}{2\gamma} \ln \frac{Var(\tilde{\theta}|\tilde{p})}{Var(\tilde{\theta}|\tilde{p}, \tilde{s}_{H,i})} + \frac{1}{2\gamma} \ln \frac{Var(\tilde{\theta}|\tilde{p})}{Var(\tilde{\theta}|\tilde{p}, \tilde{s}_{L,i})} \right). \tag{21}$$

This roughly corresponds to the total revenue of an active fund in reality. Since trading performance compensates investors' skill-acquisition costs, performance is higher in financial markets with higher skill-acquisition costs. For example, to the extent the skill-acquisition costs are higher for international financial markets than domestic ones, and higher for private companies than public ones, our analysis helps to explain why active funds focusing on the former may deliver better performance.

The following proposition formally summarizes the aforementioned results.

Proposition 4 (Implications of skill-acquisition costs without skill-acquisition uncertainty) *Suppose that there is no uncertainty involved in skill acquisition and the seller owns sufficiently accurate data. When skill-acquisition cost parameter c increases (meaning that $C(\cdot)$ shifts upward), the population λ^* of skilled investors decreases, data clarity x^* increases, data price q^* increases, and skilled investors' performance increases, but price informativeness $\frac{1}{Var(\tilde{\theta}|\tilde{p})}$, the cost of capital $E(\tilde{\theta} - \tilde{p})$, and return volatility $\sigma(\tilde{\theta} - \tilde{p})$ remain unchanged. That is, if $\Delta = 0$, then for sufficiently high τ_η, $\frac{\partial \lambda^*}{\partial c} < 0$, $\frac{\partial x^*}{\partial c} > 0$, $\frac{\partial q^*}{\partial c} > 0$, $\frac{\partial Performance}{\partial c} >$*

$0, \frac{\partial}{\partial c} \frac{1}{Var(\tilde{\theta}|\tilde{p})} = 0, \frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c} = 0,$ *and* $\frac{\partial \sigma(\tilde{\theta}-\tilde{p})}{\partial c} = 0.$

4.2 Skill Acquisition with Uncertainty

Now we introduce the "gold rush" feature of skill acquisition by assuming skill acquisition is uncertain (i.e., $\Delta > 0$, or equivalently, $Z_L < Z_H$). For instance, in the context of machine learning, an institution needs to input data first to determine how useful its algorithm is. We conduct comparative statics with respect to three skill-acquisition parameters: the cost c of acquiring skills, the mean $\bar{Z}$ of skill levels, and the volatility Δ of skill levels.

4.2.1 Comparative Statics with Respect to c

With skill-acquisition uncertainty, price informativeness increases with skill-acquisition costs. This surprising result arises because the data seller loses her full control of information leakage in the presence of skill-acquisition uncertainty. When there is no uncertainty in the skill-acquisition process, the data seller knows the skill level of the data buyer and so she can adjust data clarity to completely neutralize the effect of skill acquisition cost on information leakage. When skill acquisition becomes uncertain, the data seller is no longer certain about the data buyer's skill levels in the data transaction. Investors ending up with different skill levels will use data differently for the same level of data clarity. Since the seller can only set one clarity level for all investors, her ability to control information leakage is impaired. To be specific, even after paying skill-acquisition costs, investors may end up with low skills (and thus low utilities), which reduces their incentives to acquire skills in the first place. To encourage investors to acquire skills and buy data, the data seller will provide clearer data than she does in the economy without skill-acquisition uncertainty. In consequence, as skill-acquisition costs increase, more information is leaked via the price, thereby increasing price informativeness. In addition, both the cost of capital and return volatility decrease as a result of the increased price informativeness and data clarity. The following proposition summarizes the results.

Proposition 5 (Skill-acquisition cost) Suppose that skill acquisition is uncertain and *the seller owns sufficiently accurate data. When skill-acquisition cost parameter c increases* *(meaning that $C(\cdot)$ shifts upward), the population λ^* of skilled investors decreases, data clarity x^* increases, data price q^* increases, price informativeness $\frac{1}{Var(\tilde{\theta}|\tilde{p})}$ increases, the cost of capital $E(\tilde{\theta}-\tilde{p})$ decreases, return volatility $\sigma(\tilde{\theta}-\tilde{p})$ decreases, and skilled investors' performance increases. That is, if $\Delta > 0$, then for sufficiently high τ_η, $\frac{\partial \lambda^*}{\partial c} < 0, \frac{\partial x^*}{\partial c} > 0, \frac{\partial q^*}{\partial c} > 0, \frac{\partial}{\partial c}\frac{1}{Var(\tilde{\theta}|\tilde{p})} > 0, \frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c} < 0, \frac{\partial \sigma(\tilde{\theta}-\tilde{p})}{\partial c} < 0$, and $\frac{\partial Performance}{\partial c} > 0$.*

Figure 4 graphically illustrates Proposition 5 for the same parameter values as those in Figure 3. We here also set skill volatility Δ at 5. The patterns of variables are consistent with Proposition 5.

[FIGURE 4 ABOUT HERE]

4.2.2 Comparative Statics with Respect to $\bar{Z}$

The mean $\bar{Z}$ of skill levels can be interpreted as the average ability of the active asset management industry. We plot the implications of skill mean in Figure 5. The parameter values are the same as those in Figure 4 with $c = 0.2$.

[FIGURE 5 ABOUT HERE]

As $\bar{Z}$ increases, both the population λ^* of skilled investors and data price q^* increase (i.e., $\frac{\partial \lambda^*}{\partial \bar{Z}} > 0$ and $\frac{\partial q^*}{\partial \bar{Z}} > 0$). The seller optimally adds more noise into the sold data to control information leakage. That is, data clarity x^* decreases with $\bar{Z}$ in Figure 5. As a result, price informativeness declines despite the fact that there are more skilled investors. The cost of capital and return volatility in turn increase with $\bar{Z}$ because investors face more uncertainty when trading the risky asset.

By equation (21), the performance of skilled investors is determined by three factors: data clarity, skill levels, and price informativeness. Data clarity and skill levels jointly reflect the edge that skilled investors have over unskilled investors, and hence these two factors positively affect the performance of skilled investors. Price informativeness represents the expected utility of unskilled investors, which constitutes the outside option for skilled investors and negatively influences the performance of skilled investors. As skill mean increases, both data clarity and price informativeness decrease, but they affect skilled investor performance in opposite ways. Overall, the price informativeness effect dominates so that the performance of skilled investors increases in skill mean.

Proposition 6 (Skill mean) Suppose that skill acquisition is uncertain and *the seller owns sufficiently accurate data. When skill mean $\bar{Z}$ increases, the population λ^* of skilled investors increases, data clarity x^* decreases, data price q^* increases, price informativeness $\frac{1}{Var(\tilde{\theta}|\tilde{p})}$ decreases, the cost of capital $E(\tilde{\theta} - \tilde{p})$ increases, return volatility $\sigma(\tilde{\theta} - \tilde{p})$ increases, and skilled investors' performance increases. That is, if $\Delta > 0$, then for sufficiently high τ_η, $\frac{\partial \lambda^*}{\partial \bar{Z}} > 0$, $\frac{\partial x^*}{\partial \bar{Z}} < 0$, $\frac{\partial q^*}{\partial \bar{Z}} > 0$, $\frac{\partial}{\partial \bar{Z}} \frac{1}{Var(\tilde{\theta}|\tilde{p})} < 0$, $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial \bar{Z}} > 0$, $\frac{\partial \sigma(\tilde{\theta}-\tilde{p})}{\partial \bar{Z}} > 0$, and $\frac{\partial Performance}{\partial \bar{Z}} > 0$.*

4.2.3 Comparative Statics with Respect to Δ

The volatility Δ of investor skills captures the uncertainty in the process of decoding the purchased data. Especially for alternative data, whether an institution can timely transfer the unstructured data to investment ideas remains uncertain, and many funds have to hire top notch data scientists in the hope of increasing their success odds and speed. The range of Δ is $(0, \bar{Z})$, which is implied by the condition $0 < Z_L < Z_H$. We use Figure 6 to examine the implications of Δ, where the parameter values are the same as those in Figure 5 with $\bar{Z} = 10$.

[FIGURE 6 ABOUT HERE]

As skill volatility increases, fewer investors acquire skills, and the seller adds less noise into the sold data, but data price decreases. Interestingly, price informativeness is hump-shaped in skill volatility. This is due to the competition between two effects. On one hand, as Δ increases, investors face larger uncertainty in acquiring skills and thus, fewer investors choose to acquire skills. This tends to decrease price informativeness. On the other hand, the data seller will sell clearer data since information leakage is of less a concern when fewer investors acquire skills. This tends to increase price informativeness. When Δ is small, the mild uncertainty in the skill-acquisition process does not deter investors from acquiring skills and the second positive effect dominates, so that price informativeness increases with Δ. The result is reversed when Δ is large. Accordingly, both the cost of capital and return volatility exhibit a U-shaped pattern.

Proposition 7 (Skill volatility) Suppose that skill acquisition is uncertain and *the seller owns sufficiently accurate data. When skill volatility Δ increases, the population λ^* of skilled investors increases, data clarity x^* increases, data price q^* decreases, and skilled investors' performance decreases. Moreover, price informativeness $\frac{1}{Var(\tilde{\theta}|\tilde{p})}$ first decreases and then increases; the cost of capital $E(\tilde{\theta}-\tilde{p})$ and return volatility $\sigma(\tilde{\theta}-\tilde{p})$ first decrease and then increase.*

5 Implications and Applications

5.1 Price Informativeness Over Time

As data-processing costs have plummeted and data availability has expanded, one natural and fundamental question to ask is whether financial market prices become more informative. Bai, Philippon, and Savov (2016) derive a theory-based measure of price informativeness and find that it has risen for firms in the S&P 500 since 1960. Farboodi, Matray, and Veldkamp (2017) show that most of this rise comes from a composition effect (i.e., the rise is mostly the result of a change in composition toward older and larger firms). In their Figure 7, Farboodi, Matray, and Veldkamp (2017) find that price informativeness is almost constant in a relatively composition-bias-free sample that includes only those firms that remain listed for at least 40 years, consistent with our Panel A of Figure 1. As Farboodi, Matray, and Veldkamp (2017, p.15) argued, "since older firms should have more informative prices, that makes the lack of a trend here all the more striking." Our result in Section 4.1 provides a potential explanation for this puzzling phenomenon by considering a strategic information sell side in the financial market. Given the long history of those firms, investors may have little uncertainty in developing costly skills of interpreting and trading on the traditional data about mature firms ($\Delta \approx 0$). According to Proposition 4 (Figure 3), data vendors will optimally adjust data clarity to target a fixed level of price informativeness in maximizing profits. This clarity adjustment leads to virtually constant price informativeness for the mature companies.

Meanwhile, as in our Panel B of Figure 1, Farboodi, Matray, and Veldkamp (2017) also document that the price informativeness of the average public firm (especially small firms) is deteriorating. This finding appears surprising given that over the past decades, more and more talents have been drawn to the financial sector (see discussions in Glode and Lowery (2016)). Our analysis offers a complementary explanation for this phenomenon. Although recent technological advancements have enabled data vendors to collect real-time, granular indicators of firms' fundamentals, small firms suggest potential uncertainty involved in securing and analyzing such data ($\Delta > 0$). As the knowledge and infrastructure to leverage these sources of data become more accessible (i.e., skill-acquisition costs decrease), the data vendor optimally reduces the clarity of data, leading to lower price informativeness (see Proposition 5 and Figure 4).

5.2 Institutional Investors and Return Volatility

Brandt, Brav, Graham, and Kumar (2009) find that among low-priced stocks, a higher level of institutional ownership predicts lower idiosyncratic volatility and that among high-priced stocks, the opposite is true. Since low-priced stocks are dominated by retail traders and high-priced stocks are dominated by institutional investors, the finding of Brandt, Brav, Graham, and Kumar (2009) suggests a U-shaped relation between return volatility and institutional ownership. The non-monotone effects of skill volatility on return volatility in our data-sales economy provide a potential explanation for this phenomenon by factoring a strategic data seller into the picture. Specifically, in Figure 6, as the population λ^* of skilled investors increases (caused by a decrease in skill volatility), return volatility decreases in λ^* when it is small and increases in λ^* when it is large. If we interpret skilled investors as institutions (and unskilled as retail investors), this model-implied U-shape between λ^* and return volatility is consistent with the empirical regularity on return volatility and institutional ownership.

5.3 Data Industry and Funds Industry

Our setting illustrates a generic point that the asset management industry and the data industry foster each other. We can interpret skilled investors as asset management firms and thus, the population λ^* of the skilled investors can be a measure for the size of the asset management industry. The data seller's profit π^* can be a proxy for the size of the data industry. Examining Figures 3–6, we find that λ^* and π^* tend to move in the same direction in response to changes in exogenous parameters, which suggests that the two industries tend to foster each other. As we mentioned in the Introduction, some practitioners argue that the rapidly growing alternative data will lead to a shrinkage in the funds industry. However, our analysis predicts that the abundance of alternative data will instead spur the rapid growth of active funds using this kind of data.

Our prediction is empirically consistent with the observation that both the asset management industry and the data industry often prosper or stagnate simultaneously. For instance, assets under management (AUM) of institutional investors (e.g., mutual funds, exchange-traded funds, and institutional funds) and the global spending on information/analysis have experienced an annual growth rate of 7.7% and 2.6% from 2011 to 2016, respectively (2017 Investment Company Fact Book; "Global market data demand," Burton-Taylor International Consulting, 2016). In addition, as the leader in the financial data industry, Bloomberg L.P. keeps enjoying yearly increases in its terminal subscriptions over the past decades except for two only drops, both of which were accompanied by the shrinkage of the asset management industry. Specifically, the first decline followed the financial crisis in 2009 when the company lost 20,000 terminals, while the second decline—a mild loss of 3,145 terminals from 2015—can be attributed to the cutbacks overall in financial institutions ("Bloomberg suffers rare drop in terminal numbers as banks cut back," *Financial Times*, March 28, 2017).

6 Conclusion

In modern financial markets, data providers (such as data vendors) sell valuable data to sophisticated traders (such as hedge funds or active mutual funds). Interpreting and trading on this type of data require skills, and developing such skills is costly and uncertain to investors. We provide a data-sales model to capture these features. We model that both data providers and skilled investors share the gains generated from trading on the data to reflect the fact that both data and trading skills are limited resources. We focus on the economy in which the seller's data is very accurate so that the seller adds personalized noise to the sold data. We find that when there is uncertainty in skill-acquisition processes, the seller loses her full control of price informativeness, and thus she can no longer fully undo the information-leakage effect as she does when there is no skill-acquisition uncertainty. As a result, price informativeness can deteriorate as skill-acquisition costs decrease or as the average level of investor skills increases. These novel theoretical results help understand certain regularities in real financial markets.

Appendix A: List of Variables

Variables	Description
Introduced in Section 2	
	Exogenous Variables
$\tilde{\theta}$	Risky asset value with prior precision τ_θ
$\tilde{u}$	Noise-trader demand with prior precision τ_u
$\tilde{\eta}$	Uncertainty of the seller's data, with precision of τ_η
z_k	Investor's skills of interpreting the data, $z_k \in \{0, Z_L, Z_H\}$
$\bar{Z}$	Skill mean, $\bar{Z} = \frac{Z_H+Z_L}{2}$
Δ	Skill volatility, $\Delta = \frac{Z_H-Z_L}{2}$
β	Data sellers' bargaining power, $\beta \in [0,1]$
γ	Investors' coefficient of risk aversion
	Exogenous Functions
$C(\cdot)$	Skill-acquisition cost function
	Endogenous Variables
$\tilde{p}$	Price of risky asset
x	Clarity of the data passed to investors
$\tilde{\varepsilon}_{k,i}$	Personalized noise added by the seller
$\tilde{s}_{k,i}$	Signal sold to investor i: $\tilde{s}_{k,i} = \tilde{\theta} + \tilde{\eta} + \tilde{\varepsilon}_{k,i}$
D_i	Investor i's demand for the risky asset
λ	Mass of skilled investors, $\lambda \in [0,1]$
π	Profit of selling data
q	Price of data
Introduced in Section 3	
a_0, a_θ, a_u	Coefficients of price function
α	$\alpha = a_\theta/a_u$
A_i	Investor i's decision to acquire skills
CE	Ex-ante certainty equivalent of an investor
G	The gain from trading on the purchased data

Appendix B: The Modern Markets for Financial Data

Here we briefly introduce the modern markets for financial data. We focus on the main real-world issues related to acquiring skills to analyze the purchased data, which is the main point we model theoretically. While the traditional data markets are featured with information in the form of newsletters, advising services, etc. (Admati and Pfleiderer, 1986), here we emphasize the alternative data that represents an accelerating trend of data markets.

There is no clear definition of alternative data, but it is generally considered to be anything outside of traditional data like economic statistics and corporate reports, such as satellite images, credit card sales, sentiment analysis, mobile geolocation data and website scraping. The landscape of alternative data keeps evolving, and they used to be stock prices and fundamental information decades ago. One natural consequence is that every dataset has a life cycle. That is, as the data set becomes more and more standard, its value will decay over time.

We next introduce the sell side and the buy side in information markets.

The Sell Side Any data owners can potentially sell their data. In addition to the professional data vendors such as Bloomberg and Thomson Reuters, many tech companies now generate data as a by-product of their core activity, and try to monetize these data. For example, Twitter evolved its original firehose business into a full-blown enterprise data platform, GNIP; Mastercard offers data indices and research products through MasterIntelligence ("The new gold rush? Wall Street wants your data," Matt Turck, January 17, 2017). Moreover, banks used to crunch data on company earnings, price targets, and other mundane metrics to inform investing and trading decisions, but now they are pulling data from unorthodox sources such as social-media sentiment and geospatial mapping, and increasingly making their data feeds available directly to clients, without surrounding research notes ("Wall Street analysts are now selling more data, less analysis," *Wall Street Journal*, November 7, 2018).

Selling data involves a lot of aspects. First, the value of a data set depends heavily on such details as accuracy, time series, release schedule, uniqueness, and compliance. The more granular the data, the longer history that can be backtested, the more frequent the release, and the more unique the data, the more valuable the data set is. For compliance issues, the sold data need to be strictly anonymized and avoid material, non-public information in violation of securities laws, and the Terms of Service (TOS) should enable the data owner to sell it.

Second, data owners should know their customers and their needs. Different types of institutional investors often have different requirements for data. For quant funds, the data should speak to a lot of companies and have a long time series. One example is a panel of consumer transactions touching many public companies that has a correlation with share prices. Alternatively, for fundamental investors, the data can be about specific public companies.

Third, to determine the prices of the data, since the buy-side organizations often employ teams of people looking to speak with data owners, data owners can speak to these people directly without working with a data broker or other intermediary, who can demand a revenue share of 50% or more ("The ultimate guide to selling data to hedge funds," AlternativeData.org, August 16, 2017).

The Buy Side Hedge funds are at the forefront of the trend to leverage the new data to gain an edge over their competitors (e.g., low-priced exchange traded funds) and generate alpha, through accurate predictions. Since each data set has its life cycle, as some data sets become widely available, hedge funds move on to other forms of data. However, the behavior has also spread to many more mainstream mutual funds and across the world ("Asia investors boost use of unorthodox data sources in battle to beat benchmarks," *Business News*, April 25, 2017). According to a survey of investors, the buy side has spent $373 million on alternate data sets and the number is expected to hit $1 billion by 2020 ("Asset managers double spending on new data in hunt for edge," *Financial Times*, May 8, 2018).

However, data alone is not enough. In February 2018, Battlefin organized a conference about alternative data, bringing together 107 asset managers, 94 data providers, and around 100 other industry professionals. One common theme throughout the conference is that access to certain data sources is no longer the main source of alpha, but rather the ability to process the data well and reach the best insights the fastest. As put by Barry Hurewitz, UBS's global head of Evidence Lab, "People say data is the new oil, but there is a refiner needed.”

In fact, integrating data analysts and engineers in the investment process is a major challenge for institutional investors looking to leverage alternative data. The estimated cost for a full data team (comprised of one Data Engineer, three Data Analysts, one Data Scientist, one Data Scout, and one Head of Data) would start at \$1.5m – \$2.5m at an entry level. With consideration for insurance, benefits, overhead, etc., it is likely that the true cost could be twice as much. Moreover, from the size of some teams and anecdotal research, several top funds are already spending over \$10m on alternative data teams (“Buy-side alternative data employee analysis,” AlternativeData.org, February 7, 2018).

According to a report by AlternativeData.org (“Buy-side alternative data employee analysis,” AlternativeData.org, February 7, 2018), the backgrounds of the alternative-data employees on the buy-side have relatively high concentration on science, technology, engineering, and math (STEM). Given the technical sophistication of the role, over 40% of Data Scientist positions hold a graduate degree. In addition, the buy side has substantially increased their hiring from tech companies, academia, and data providers, rather than the traditional sourcing from other funds or the sell side. In terms of work experience, the majority of the buy-side hire individuals with 11 and more years of experience.

With those investments in the technologies and talents, over time it has gotten easier and easier to process large amounts of unstructured data.

Appendix C: Proofs

Proof of Proposition 1

Inserting equations (7)–(12) into the demand functions for different types of investors, we can rearrange the market-clearing condition (13) and compute the price as follows:

$$\tilde{p} = \frac{A\tilde{s}_p + B(\tilde{\theta} + \tilde{\eta}) + \tilde{u} - 1}{\frac{\lambda}{2\gamma Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})} + \frac{\lambda}{2\gamma Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})} + \frac{1-\lambda}{\gamma Var(\tilde{\theta}|\tilde{p})}}, \tag{C1}$$

where

$$\begin{aligned} A &= \frac{\lambda}{2\gamma Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})}\frac{\tau_\eta}{\tau_\theta+\tau_\eta}\frac{\alpha^2\tau_u}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u+xZ_H} \\ &\quad + \frac{\lambda}{2\gamma Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})}\frac{\tau_\eta}{\tau_\theta+\tau_\eta}\frac{\alpha^2\tau_u}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u+xZ_L} \\ &\quad + \frac{1-\lambda}{\gamma Var(\tilde{\theta}|\tilde{p})}\frac{\tau_\eta}{\tau_\theta+\tau_\eta}\frac{\alpha^2\tau_u}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u}, \\ B &= \frac{\lambda}{2\gamma Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})}\frac{\tau_\eta}{\tau_\theta+\tau_\eta}\frac{xZ_H}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u+xZ_H} \\ &\quad + \frac{\lambda}{2\gamma Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})}\frac{\tau_\eta}{\tau_\theta+\tau_\eta}\frac{xZ_L}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u+xZ_L}. \end{aligned}$$

From equation (C1), we have $\alpha = B$. Using the expression of B above, we can rearrange equation $\alpha = B$ as (14) in Proposition 1. We use $F(\alpha)$ to denote the left-hand-side (LHS) of equation (14). It is clear that $F(\alpha)$ is negative when $\alpha = 0$, and $F(\alpha)$ is positive when α is large enough. Thus, by the intermediate value theorem, there exists a positive root to (14). In addition, $F(\alpha)$ is monotonically increasing in α and thus, the real root is unique.

After some algebra, we can compute the price coefficients as follows:

$$a_0 = -\frac{1}{\frac{\lambda}{2\gamma Var\left(\tilde{v}|\tilde{p},\tilde{s}_{H,i}\right)} + \frac{\lambda}{2\gamma Var\left(\tilde{v}|\tilde{p},\tilde{s}_{L,i}\right)} + \frac{1-\lambda}{\gamma Var(\tilde{v}|\tilde{p})}}, \tag{C2}$$

$$a_\theta = \frac{A+B}{\frac{\lambda}{2\gamma Var\left(\tilde{v}|\tilde{p},\tilde{s}_{H,i}\right)} + \frac{\lambda}{2\gamma Var\left(\tilde{v}|\tilde{p},\tilde{s}_{L,i}\right)} + \frac{1-\lambda}{\gamma Var(\tilde{v}|\tilde{p})}}, \tag{C3}$$

$$a_u = B^{-1}a_\theta. \tag{C4}$$

QED.

Proof of Corollary 1

It suffices to prove that α is increasing in λ, where α is determined by equation (14). Note that the LHS of equation (14) is increasing in α and decreasing in λ. Thus, when λ increases, α must increase to maintain equation (14). So, $\frac{\partial\alpha}{\partial\lambda} > 0$. Meanwhile, because $Var(\tilde{\theta}|\tilde{p})$ is decreasing with α (see equation (12)), we have $\frac{\partial}{\partial\lambda}\frac{1}{Var(\tilde{\theta}|\tilde{p})} > 0$. QED.

Proof of Equation (15)

Following Grossman and Stiglitz (1980), we can compute the ex-ante expected utility of a high-type and a low-type skilled investor respectively as follows:

$$\begin{aligned}
V_H &= -\sqrt{\frac{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})}{Var(\tilde{\theta}-\tilde{p})}}\exp\left(-\frac{\left[E(\tilde{\theta}-\tilde{p})\right]^2}{2Var(\tilde{\theta}-\tilde{p})}\right), \\
V_L &= -\sqrt{\frac{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})}{Var(\tilde{\theta}-\tilde{p})}}\exp\left(-\frac{\left[E(\tilde{\theta}-\tilde{p})\right]^2}{2Var(\tilde{\theta}-\tilde{p})}\right),
\end{aligned}$$

where we have ignored the skill-acquisition cost, since this cost is sunk at the stage of Nash bargaining as we have discussed in the text. Given that a skilled investor can be of a high type with probability $\frac{1}{2}$ and a low type with probability $\frac{1}{2}$, the ex-ante expected utility of a skilled investor is $V_S = \frac{1}{2}V_H + \frac{1}{2}V_L$. The certainty equivalent of a skilled investor is

$$\begin{aligned}
CE_S &\equiv -\frac{1}{\gamma}\log\left(-V_S\right) \\
&= -\frac{1}{\gamma}\ln\left[\frac{\frac{1}{2}\sqrt{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})}+\frac{1}{2}\sqrt{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})}}{\sqrt{Var(\tilde{\theta}-\tilde{p})}}\right]+\frac{\left[E(\tilde{\theta}-\tilde{p})\right]^2}{2\gamma Var(\tilde{\theta}-\tilde{p})}.
\end{aligned}$$

Similarly, we can compute the certainty equivalent of an unskilled investor as follows:

$$CE_U = -\frac{1}{\gamma}\ln\left[\frac{\sqrt{Var(\tilde{\theta}|\tilde{p})}}{\sqrt{Var(\tilde{\theta}-\tilde{p})}}\right]+\frac{\left[E(\tilde{\theta}-\tilde{p})\right]^2}{2\gamma Var(\tilde{\theta}-\tilde{p})}.$$

QED.

Proof of Proposition 2

To prove the existence of an overall equilibrium, it suffices to show that given the data seller's choice x, there is a unique solution (λ,α) to the system of equations (14) and (19). This implies that the data seller can use x to effectively control the demand for data and the price of data.

We use equation (14) to first determine λ as a function of α, which is in turn inserted into equation (19). We then characterize the property of equation (19) in terms of the single unknown α.

Specifically, equation (14) determines λ as a function of α, denoted by $\lambda = \lambda(\alpha)$:

$$\lambda(\alpha) = \frac{2\alpha\gamma\left(xZ_H + \tau_u\alpha^2 + \tau_\eta\right)\left(xZ_L + \tau_u\alpha^2 + \tau_\eta\right)}{x\tau_\eta\left(\left(\tau_u\alpha^2 + \tau_\eta\right)\left(Z_H + Z_L\right) + 2xZ_HZ_L\right)}.$$

From the proof of Corollary 1, we know that $\lambda(\alpha)$ is an increasing function of α. Thus, the right-hand-side (RHS) of equation (19), $C(\lambda(\alpha))$ (which is expressed as a function α), is increasing in α, since $C(\cdot)$ is an increasing function.

We next prove that the LHS of equation (19) is decreasing in α. By equation (15), we can express G as

$$G = -\frac{1}{\gamma}\ln\left[\frac{1}{2}\sqrt{\frac{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})}{Var(\tilde{\theta}|\tilde{p})}} + \frac{1}{2}\sqrt{\frac{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})}{Var(\tilde{\theta}|\tilde{p})}}\right].$$

Using the expressions of $Var(\tilde{\theta}|\tilde{p})$, $Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})$, and $Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})$, we have

$$\frac{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})}{Var(\tilde{\theta}|\tilde{p})} = \frac{\frac{1}{\tau_\theta+\tau_\eta} + \left(\frac{\tau_\eta}{\tau_\theta+\tau_\eta}\right)^2 \frac{1}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u+xZ_H}}{\frac{1}{\tau_\theta+\tau_\eta} + \left(\frac{\tau_\eta}{\tau_\theta+\tau_\eta}\right)^2 \frac{1}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u}},$$

$$\frac{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})}{Var(\tilde{\theta}|\tilde{p})} = \frac{\frac{1}{\tau_\theta+\tau_\eta} + \left(\frac{\tau_\eta}{\tau_\theta+\tau_\eta}\right)^2 \frac{1}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u+xZ_L}}{\frac{1}{\tau_\theta+\tau_\eta} + \left(\frac{\tau_\eta}{\tau_\theta+\tau_\eta}\right)^2 \frac{1}{\frac{\tau_\theta\tau_\eta}{\tau_\theta+\tau_\eta}+\alpha^2\tau_u}}.$$

Direct computation shows that both $\frac{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{H,i})}{Var(\tilde{\theta}|\tilde{p})}$ and $\frac{Var(\tilde{\theta}|\tilde{p},\tilde{s}_{L,i})}{Var(\tilde{\theta}|\tilde{p})}$ are increasing in α, and thus G is decreasing in α. That is, the LHS of equation (19) decreases with α. Given that the RHS increases with α, if there is a solution α to equation (19), it must be unique.

We establish the existence using the intermediate value theorem. Suppose $\alpha = 0$, the LHS minus RHS of equation (19) is positive given that the RHS $C(0) = 0$ but the LHS is still positive. By contrast, suppose that α takes the value which is determined by equation (14) with $\lambda = 1$. Given that $C(1) = +\infty$, the LHS minus RHS of equation (19) is negative given that the RHS is infinite and the LHS is finite. According to the intermediate value theory, there is a solution of the equation (19). In addition, the LHS minus RHS of equation (19) is decreasing with α, and thus the solution is unique.

Once $\alpha(x)$ is well defined, the fraction $\lambda(x)$ of active investors is determined by equation (14). By the Nash-bargaining equilibrium condition (16), $q(x) = \beta G(x, \alpha(x))$. Thus, the product of $\lambda(x)$ and $q(x)$ yields the total profit of the seller. Then the seller can simply choose x^* to maximize $\lambda(x) q(x)$. Note that we allow x to take the value of ∞ (i.e., the seller sells her data "as is") and $\lambda(x) q(x)$ is a continuous function of x, and hence there always exist solutions to the seller's maximization problem according to the Weierstrass Theorem. QED.

Proof of Proposition 3

We use the method of Lagrangian multipliers to solve for the seller's profit-maximization problem (3) subject to the constraints (14) and (19). By equation (19), for $\beta < 1$, the seller's net profit can be written as

$$\pi(x) = \frac{\beta\lambda C(\lambda)}{1-\beta}. \tag{C5}$$

We only focus on the case with $\beta < 1$ because it is more relevant. If $\beta = 1$, then the information market will break down since no investors will incur costs to acquire skills and purchase data.

The Lagrangian function is

$$L = \frac{\beta\lambda C(\lambda)}{1-\beta} - \upsilon_1 \left[\begin{array}{c} \gamma\alpha\left(xZ_H + \tau_u\alpha^2 + \tau_\eta\right)\left(xZ_L + \tau_u\alpha^2 + \tau_\eta\right) \\ -\frac{1}{2}x\lambda\tau_\eta\left(\left(\tau_u\alpha^2 + \tau_\eta\right)\left(Z_H + Z_L\right) + 2xZ_HZ_L\right) \end{array} \right] - \upsilon_2\left[(1-\beta)G - C(\lambda)\right],$$

where $\upsilon_1, \upsilon_2 \geq 0$ are the Lagrangian multipliers. The first-order conditions (FOC) with respect to x and υ_1 are respectively:

$$-\frac{1}{2}\upsilon_1 \left(\begin{array}{c} 4Z_LZ_H\left(\alpha\gamma - \lambda\tau_\eta\right)x \\ +\left(Z_H + Z_L\right)\left(\tau_\eta + \alpha^2\tau_u\right)\left(2\alpha\gamma - \lambda\tau_\eta\right) \end{array} \right) - \upsilon_2(1-\beta)\frac{\partial G}{\partial x} = 0, \tag{C6}$$

$$Z_HZ_L\left(\alpha\gamma - \lambda\tau_\eta\right)x^2 + \frac{1}{2}\left(Z_H + Z_L\right)\left(\tau_\eta + \alpha^2\tau_u\right)\left(2\alpha\gamma - \lambda\tau_\eta\right)x + \alpha\gamma\left(\tau_\eta + \alpha^2\tau_u\right)^2 = 0 \tag{C7}$$

Equation (C7) is a quadratic function of x. If $\lambda\tau_\eta > \alpha\gamma$, there is a unique positive solution:

$$x^* = \frac{\left(\tau_\eta + \alpha^2\tau_u\right)\left(\begin{array}{c} \left(Z_H + Z_L\right)\left(2\alpha\gamma - \lambda\tau_\eta\right) \\ +\sqrt{4\alpha^2\gamma^2\left(Z_H - Z_L\right)^2 - 4\alpha\gamma\lambda\tau_\eta\left(Z_H - Z_L\right)^2 + \lambda^2\tau_\eta^2\left(Z_H + Z_L\right)^2} \end{array} \right)}{4Z_HZ_L\left(\lambda\tau_\eta - \alpha\gamma\right)}. \tag{C8}$$

Note that the numerator is positive if $\lambda\tau_\eta > \alpha\gamma$ because $\left[4\alpha^2\gamma^2 (Z_H - Z_L)^2 - 4\alpha\gamma\lambda\tau_\eta (Z_H - Z_L)^2 + \lambda^2\tau_\eta^2 (Z_H + Z_L)^2\right] - \left[(Z_H + Z_L)(2\alpha\gamma - \lambda\tau_\eta)\right]^2 = 16\alpha\gamma Z_H Z_L (\lambda\tau_\eta - \alpha\gamma)$.

If $\lambda\tau_\eta \leq \alpha\gamma$, the LHS of equation (C7) is always positive, which yields $\upsilon_1 = 0$. Plugging $\upsilon_1 = 0$ into equation (C6), we have $\frac{\partial G}{\partial x} = 0$ (note that $\upsilon_2 \neq 0$ since (19) always holds). When taking derivative of G with respect to x, we obtain

$$\frac{\partial G}{\partial x} = \frac{\tau_\eta^2 \left(\frac{Z_H}{\left(\frac{\tau_\eta\tau_\theta}{\tau_\eta+\tau_\theta} + xZ_H + \alpha^2\tau_u\right)^2 \sqrt{\frac{\tau_\eta + xZ_H + \alpha^2\tau_u}{\tau_\eta\tau_\theta + xZ_H(\tau_\eta+\tau_\theta) + \alpha^2\tau_u(\tau_\eta+\tau_\theta)}}} + \frac{Z_L}{\left(\frac{\tau_\eta\tau_\theta}{\tau_\eta+\tau_\theta} + xZ_L + \alpha^2\tau_u\right)^2 \sqrt{\frac{\tau_\eta + xZ_L + \alpha^2\tau_u}{\tau_\eta\tau_\theta + xZ_L(\tau_\eta+\tau_\theta) + \alpha^2\tau_u(\tau_\eta+\tau_\theta)}}} \right)}{2\gamma (\tau_\eta + \tau_\theta)^2 \left(\sqrt{\frac{\tau_\eta + xZ_H + \alpha^2\tau_u}{\tau_\eta\tau_\theta + xZ_H(\tau_\eta+\tau_\theta) + \alpha^2\tau_u(\tau_\eta+\tau_\theta)}} + \sqrt{\frac{\tau_\eta + xZ_L + \alpha^2\tau_u}{\tau_\eta\tau_\theta + xZ_L(\tau_\eta+\tau_\theta) + \alpha^2\tau_u(\tau_\eta+\tau_\theta)}} \right)} > 0.$$

Thus, in equilibrium, the optimal choice $x^* = +\infty$ when $\lambda\tau_\eta \leq \alpha\gamma$. With an infinite data clarity x^*, we can solve α^* from equation (14): $\alpha^* = \frac{\lambda^*\tau_\eta}{\gamma}$. Taken together, if $\lambda^*\tau_\eta > \alpha^*\gamma$, $x^* < +\infty$, and if $\lambda^*\tau_\eta \leq \alpha^*\gamma$, $x^* = +\infty$. QED.

Proof of Proposition 4

Suppose $\tau_\eta \to +\infty$. When $Z_L = Z_H$, equation (14) can be simplified to

$(xZ_H + \tau_u\alpha^2 + \tau_\eta)\left[\gamma\alpha (xZ_H + \tau_u\alpha^2 + \tau_\eta) - x\lambda\tau_\eta Z_H\right] = 0$, which yields

$$x^* = \frac{\alpha\gamma (\tau_u\alpha^2 + \tau_\eta)}{Z_H (\lambda\tau_\eta - \alpha\gamma)}. \tag{C9}$$

Plugging x^* into equation (19) and imposing $Z_L = Z_H$ yields

$$\exp\left(\frac{\gamma C(\lambda)}{1-\beta}\right) \sqrt{1 - \frac{\alpha\gamma\tau_\eta}{\alpha^2\lambda\tau_u\tau_\theta + \tau_\eta (\alpha\gamma + \alpha^2\lambda\tau_u + \lambda\tau_\theta)}} = 1. \tag{C10}$$

So, the seller's problem is equivalent to choosing α to maximize profits given by (C5) subject to equation (C10). Because both equations (C5) and $C(\lambda)$ are increasing with λ, the data seller could choose α to minimize $\sqrt{1 - \frac{\alpha\gamma\tau_\eta}{\alpha^2\lambda\tau_u\tau_\theta + \tau_\eta(\alpha\gamma + \alpha^2\lambda\tau_u + \lambda\tau_\theta)}}$ to get the maximum profit. Taking derivative of $\sqrt{1 - \frac{\alpha\gamma\tau_\eta}{\alpha^2\lambda\tau_u\tau_\theta + \tau_\eta(\alpha\gamma + \alpha^2\lambda\tau_u + \lambda\tau_\theta)}}$ with respect to α and setting it to zero yields

$$\alpha^* = \sqrt{\frac{\tau_\eta \tau_\theta}{\tau_u (\tau_\eta + \tau_\theta)}}, \tag{C11}$$

which does not depend on skill-acquisition costs. Thus, $\frac{\partial \alpha^*}{\partial c} = 0$ and $\frac{\partial}{\partial c} \frac{1}{Var(\tilde{\theta}|\tilde{p})} = 0$. Plugging x^*, α^*, and $Z_L = Z_H$ into equation (19) yields

$$\exp\left(\frac{\gamma C(\lambda)}{1-\beta}\right) \frac{1}{\sqrt{1 + \frac{\gamma}{2\lambda}\sqrt{\frac{\tau_\eta}{\tau_u \tau_\theta (\tau_\eta + \tau_\theta)}}}} = 1.$$

The LHS of the above equation is increasing in λ and increasing in c. Thus, when c increases, λ should decrease to maintain the equation: $\frac{\partial \lambda^*}{\partial c} < 0$. Further, plugging α^* into x^* yields $x^* = \frac{\gamma \tau_\eta \left(\frac{\tau_\theta}{\tau_\eta + \tau_\theta} + 1\right)\sqrt{\frac{\tau_\eta \tau_\theta}{\tau_u (\tau_\eta + \tau_\theta)}}}{Z_H \left(\lambda \tau_\eta - \gamma \sqrt{\frac{\tau_\eta \tau_\theta}{\tau_u (\tau_\eta + \tau_\theta)}}\right)}$ and it is easy to get that $\frac{\partial x^*}{\partial c} = \frac{\partial x^*}{\partial \lambda}\frac{\partial \lambda}{\partial c} > 0$.

Further, as $\tau_\eta \to +\infty$, based on (C9) and (C11), we obtain that $x^* \to \frac{\alpha\gamma}{Z_H \lambda}$ and $\alpha^* \to \sqrt{\frac{\tau_\theta}{\tau_u}}$, which yields $G \to \frac{1}{\gamma} \ln \frac{\sqrt{2\lambda\tau_\theta + \gamma\sqrt{\frac{\tau_\theta}{\tau_u}}}}{\sqrt{2\lambda\tau_\theta}}$ when inserting the expressions of x^* and α^* into G. Thus $\frac{\partial G}{\partial \lambda} \to -\frac{\sqrt{\frac{\tau_\theta}{\tau_u}}}{4\lambda^2\tau_\theta + 2\lambda\gamma\sqrt{\frac{\tau_\theta}{\tau_u}}} < 0$. Based on equation (16), $\frac{\partial q^*}{\partial c} = \frac{\beta}{1-\beta}\frac{\partial G}{\partial \lambda}\frac{\partial \lambda}{\partial c} > 0$. Also, given finite λ^*, α^* and infinite τ_η, we know $\lambda^* \tau_\eta > \alpha^* \gamma$. So, based on Proposition 3, x^* takes a finite value, which is indeed the case.

Next, based on equation (4), the cost of capital $E(\tilde{\theta} - \tilde{p})$ can be simplified to

$$E(\tilde{\theta} - \tilde{p}) = -a_0. \tag{C12}$$

Inserting $Z_L = Z_H$ and x^* into equation (C12) yields

$$E(\tilde{\theta} - \tilde{p}) = \frac{\gamma (\tau_u \alpha^2 + \tau_\eta)}{\tau_\theta (\tau_u \alpha^2 + \tau_\eta) + \tau_\eta \alpha (\gamma + \tau_u \alpha)},$$

which depends only on the endogenous variable α. So, $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c} = \frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial \alpha}\frac{\partial \alpha}{\partial c} = 0$.

By (4), return volatility $\sigma(\tilde{\theta} - \tilde{p})$ can be simplified as follows:

$$\sigma(\tilde{\theta} - \tilde{p}) = \sqrt{(1-a_\theta)^2 \frac{1}{\tau_\theta} + a_\theta^2 \frac{1}{\tau_\eta} + a_u^2 \frac{1}{\tau_u}}. \tag{C13}$$

Plugging $Z_L = Z_H$ and x^* into equation (C13) yields

$$\sigma(\tilde{\theta} - \tilde{p}) = \sqrt{\frac{(\tau_\eta + \alpha^2 \tau_u)\left(\tau_\theta \tau_u (\tau_\eta + \alpha^2 \tau_u) + \tau_\eta (\gamma + \alpha \tau_u)^2\right)}{\tau_u \left(\alpha^2 \tau_\theta \tau_u + \tau_\eta (\tau_\theta + \alpha (\gamma + \alpha \tau_u))\right)^2}},$$

which depends only on the endogenous variable α. So, $\frac{\partial \sigma(\tilde{\theta}-\tilde{p})}{\partial c} = \frac{\partial \sigma(\tilde{\theta}-\tilde{p})}{\partial \alpha}\frac{\partial \alpha}{\partial c} = 0$.

Finally, as $\tau_\eta \to \infty$, with $Z_L = Z_H$, performance (21) can be simplified to *Performance* $\to \frac{1}{2\gamma} \ln \left(1 + \frac{\gamma}{2\lambda\sqrt{\tau_u \tau_\theta}}\right)$. So, $\frac{\partial Performance}{\partial c} > 0$. QED.

Proof of Proposition 5

Denote $\tau_\eta \equiv \frac{1}{w}$. Let us consider the process of $w \to 0$ (which is equivalent to the process of $\tau_\eta \to \infty$). The idea of the proof is to obtain $x(\alpha, \lambda; w)$ around $w = 0$ from equation (14), plug $x(\alpha, \lambda; w)$ into equation (19), solve $\alpha(\lambda; w)$ from its FOC, and finally insert $\alpha(\lambda; w)$ back into equation (19) to obtain the effect of the parameter change on λ^* and other endogenous variables.

When $w = 0$, we have $x(\alpha, \lambda; 0) = \frac{2\alpha\gamma}{\lambda(Z_H + Z_L)}$ by equation (14). Applying the implicit function theorem to equation (14) and imposing $w = 0$, we have $\frac{\partial x}{\partial w}(\alpha, \lambda; 0) = \frac{2\alpha^2\gamma\left(2\alpha\lambda Z_H Z_L \tau_u + \left(Z_H^2 + Z_L^2\right)(2\gamma + \alpha\lambda\tau_u)\right)}{\lambda^2 (Z_H + Z_L)^3}$. Now the Taylor series of $x(\alpha, \lambda; w)$ around $w = 0$ is

$$\begin{aligned} x(\alpha, \lambda; w) &= x(\alpha, \lambda; 0) + \frac{\partial x}{\partial w}(\alpha, \lambda; 0)(w - 0) \\ &= \frac{2\alpha\gamma \begin{pmatrix} 2\lambda Z_H Z_L (\alpha^2 w \tau_u + 1) \\ +Z_H^2 (\lambda + \alpha^2 \lambda w \tau_u + 2\alpha\gamma w) + Z_L^2 (\lambda + \alpha^2 \lambda w \tau_u + 2\alpha\gamma w) \end{pmatrix}}{\lambda^2 (Z_H + Z_L)^3} \end{aligned} \tag{C14}$$

Next when $\tau_\eta = +\infty$, equation (19) becomes

$$\frac{1}{2} e^{\frac{\gamma C}{1-\beta}} \left(\sqrt{\frac{\tau_\theta + \alpha^2 \tau_u}{\tau_\theta + x Z_H + \alpha^2 \tau_u}} + \sqrt{\frac{\tau_\theta + \alpha^2 \tau_u}{\tau_\theta + x Z_L + \alpha^2 \tau_u}}\right) = 1. \tag{C15}$$

Taking derivative of the LHS of equation (C15) with respect to α (note that x is a function of α) yields $\frac{1}{4}e^{\frac{\gamma C}{1-\beta}}\left(\begin{array}{c}-\frac{Z_H}{\sqrt{\frac{\tau_\theta+\alpha^2\tau_u}{\tau_\theta+xZ_H+\alpha^2\tau_u}}(\tau_\theta+xZ_H+\alpha^2\tau_u)^2}\\ -\frac{Z_L}{\sqrt{\frac{\tau_\theta+\alpha^2\tau_u}{\tau_\theta+xZ_L+\alpha^2\tau_u}}(\tau_\theta+xZ_L+\alpha^2\tau_u)^2}\end{array}\right)\Omega=0$, where $\Omega(\alpha,\lambda;w)=\alpha\tau_u\left(\alpha\frac{\partial x}{\partial\alpha}-2x\right)+\tau_\theta\frac{\partial x}{\partial\alpha}$. Plugging x and the expression of $\frac{\partial x}{\partial\alpha}$ in equation (C14) into Ω and the FOC of the LHS of equation (C15) is equivalent to

$$\Omega(\alpha,\lambda;w)=\frac{2\gamma\left(\begin{array}{c}2\lambda Z_HZ_L\left(\tau_\theta+\alpha^4w\tau_u^2+\alpha^2\tau_u\left(3w\tau_\theta-1\right)\right)\\ +\left(Z_H^2+Z_L^2\right)\left(\alpha^4\lambda w\tau_u^2+\alpha^2\lambda\tau_u\left(3w\tau_\theta-1\right)+\tau_\theta(\lambda+4\alpha\gamma w)\right)\end{array}\right)}{\lambda^2\left(Z_H+Z_L\right)^3}=0. \tag{C16}$$

When $w=0$, $\Omega(\alpha,\lambda;0)=\frac{2\gamma\left(\tau_\theta-\tau_u\alpha^2\right)}{\lambda(Z_H+Z_L)}$, and $\Omega(\alpha,\lambda;0)=0$ yields $\alpha(\lambda;0)=\sqrt{\frac{\tau_\theta}{\tau_u}}$. Applying the implicit function theorem to equation (C16) and imposing the extreme condition ($w=0$) delivers $\frac{\partial\alpha}{\partial w}(\lambda;0)=\frac{2\tau_\theta\left(2\lambda Z_HZ_L\sqrt{\tau_\theta\tau_u}+\left(Z_H^2+Z_L^2\right)\left(\gamma+\lambda\sqrt{\tau_\theta\tau_u}\right)\right)}{\lambda\tau_u(Z_H+Z_L)^2}$. Thus, the Taylor series of $\alpha(\lambda;w)$ around $w=0$ is

$$\alpha(\lambda;w)=\alpha(\lambda;0)+\frac{\partial\alpha}{\partial w}(\lambda;0)(w-0)=\frac{2\gamma w\tau_\theta\left(Z_H^2+Z_L^2\right)}{\lambda\tau_u\left(Z_H+Z_L\right)^2}+\frac{\sqrt{\tau_\theta}\left(2w\tau_\theta+1\right)}{\sqrt{\tau_u}}. \tag{C17}$$

Now plugging equations (C14) and (C17) into equation (C15) yields

$$\frac{1}{2}e^{\frac{\gamma C}{1-\beta}}\left(\sqrt{\frac{1}{\frac{\gamma Z_H}{\lambda(Z_H+Z_L)\sqrt{\tau_\theta\tau_u}}+1}}+\sqrt{\frac{1}{\frac{\gamma Z_L}{\lambda(Z_H+Z_L)\sqrt{\tau_\theta\tau_u}}+1}}\right)=1. \tag{C18}$$

The LHS of equation (C18) is increasing in λ and c. Thus, when c increases, λ should decrease to maintain the equation, which implies that $\frac{\partial\lambda^*}{\partial c}<0$. By equation (C17), we have $\frac{\partial\alpha^*}{\partial c}=\frac{\partial\alpha^*}{\partial\lambda}\frac{\partial\lambda}{\partial c}>0$ and thus, $\frac{\partial}{\partial c}\left(\frac{1}{Var\left(\tilde{\theta}|\tilde{p}\right)}\right)>0$. Based on $x(\alpha;0)=\frac{2\alpha\gamma}{\lambda(Z_H+Z_L)}$, $\frac{\partial x}{\partial c}=\frac{\partial x}{\partial\alpha}\frac{\partial\alpha}{\partial c}+\frac{\partial x}{\partial\lambda}\frac{\partial\lambda}{\partial c}>0$. Further, with $x(\alpha,\lambda;0)$ and $\alpha(\lambda;0)$ information gain G can be simplified to $G=\frac{1}{\gamma}\ln\frac{\frac{2}{\sqrt{\tau_\theta}}}{\frac{1}{\sqrt{\tau_\theta+\frac{\gamma Z_H\sqrt{\frac{\tau_\theta}{\tau_u}}}{\lambda(Z_H+Z_L)}}}+\frac{1}{\sqrt{\tau_\theta+\frac{\gamma Z_L\sqrt{\frac{\tau_\theta}{\tau_u}}}{\lambda(Z_H+Z_L)}}}}$. Thus,

$\frac{\partial G}{\partial \lambda} = -\frac{\sqrt{\frac{\tau_\theta}{\tau_u}}\left(Z_H\left(\frac{\lambda(Z_H+Z_L)}{\lambda Z_L\tau_\theta+Z_H\left(\lambda\tau_\theta+\gamma\sqrt{\frac{\tau_\theta}{\tau_u}}\right)}\right)^{\frac{3}{2}}+Z_L\left(\frac{\lambda(Z_H+Z_L)}{\lambda Z_H\tau_\theta+Z_L\left(\lambda\tau_\theta+\gamma\sqrt{\frac{\tau_\theta}{\tau_u}}\right)}\right)^{\frac{3}{2}}\right)}{2\lambda^2(Z_H+Z_L)\left(\frac{1}{\sqrt{\tau_\theta+\frac{\gamma Z_H\sqrt{\frac{\tau_\theta}{\tau_u}}}{\lambda(Z_H+Z_L)}}}+\frac{1}{\sqrt{\tau_\theta+\frac{\gamma Z_L\sqrt{\frac{\tau_\theta}{\tau_u}}}{\lambda(Z_H+Z_L)}}}\right)} < 0$. So, $\frac{\partial G^*}{\partial c} = \frac{\partial G^*}{\partial \lambda}\frac{\partial \lambda}{\partial c} > 0$. Thus, $\frac{\partial q^*}{\partial c} > 0$ according to equation (16).

Next, we examine market quality. Plugging $\tau_\eta = +\infty$ and $x(\alpha, \lambda; 0) = \frac{2\alpha\gamma}{\lambda(Z_H+Z_L)}$ into equation (C12), the cost of capital $E(\tilde{\theta} - \tilde{p})$ can be simplified as follows:

$$E(\tilde{\theta} - \tilde{p}) = \frac{\gamma}{\alpha\gamma + \tau_u\alpha^2 + \tau_\theta}.$$

Hence, $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c} = \frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial \alpha}\frac{\partial \alpha}{\partial c} < 0$. Further, with $\tau_\eta = +\infty$ and $x(\alpha, \lambda; 0) = \frac{2\alpha\gamma}{\lambda(Z_H+Z_L)}$, return variance $Var(\tilde{\theta} - \tilde{p})$ can be simplified as

$$Var(\tilde{\theta} - \tilde{p}) = \frac{(\gamma + \alpha\tau_u)^2 + \tau_\theta\tau_u}{\tau_u(\tau_\theta + \alpha(\gamma + \alpha\tau_u))^2}.$$

Taking derivative of it with respect to α yields

$$\frac{\partial Var(\tilde{\theta} - \tilde{p})}{\partial \alpha} = -\frac{2(\gamma^3 + \alpha^3\tau_u^3 + 3\alpha\gamma^2\tau_u + \alpha\tau_u^2(3\alpha\gamma + \tau_\theta))}{\tau_u(\alpha\gamma + \tau_\theta + \alpha^2\tau_u)^3} < 0.$$

So, $\frac{\partial \sigma(\tilde{\theta}-\tilde{p})}{\partial c} = \frac{\partial \sigma(\tilde{\theta}-\tilde{p})}{\partial \alpha}\frac{\partial \alpha}{\partial c} < 0$.

With $x(\alpha, \lambda; 0) = \frac{\alpha\gamma}{\lambda\bar{Z}}$ and $\alpha(\lambda; 0) = \sqrt{\frac{\tau_\theta}{\tau_u}}$, performance (21) can be simplified as follows

$$Performance = \frac{1}{4\gamma}\left(\ln\left(1 + \frac{\gamma(\bar{Z} - \Delta)}{2\bar{Z}\lambda\sqrt{\tau_u\tau_\theta}}\right) + \ln\left(1 + \frac{\gamma(\bar{Z} + \Delta)}{2\bar{Z}\lambda\sqrt{\tau_u\tau_\theta}}\right)\right).$$

Thus, $\frac{\partial Performance}{\partial c} = \frac{\partial Performance}{\partial \lambda}\frac{\partial \lambda}{\partial c} = -\frac{1}{4\gamma}\left(\frac{1}{\gamma+\frac{2\bar{Z}\lambda\sqrt{\tau_u\tau_\theta}}{\bar{Z}-\Delta}} + \frac{1}{\gamma+\frac{2\bar{Z}\lambda\sqrt{\tau_u\tau_\theta}}{\bar{Z}+\Delta}}\right)\frac{\partial \lambda}{\partial c} > 0$. QED.

Proof of Proposition 6

The proof follows the proof of Proposition 5. Replacing $Z_H = \bar{Z}+\Delta$ and $Z_L = \bar{Z}-\Delta$ in equation (C18) yields

$$\frac{1}{\sqrt{2}}e^{\frac{\gamma C}{1-\beta}}\left(\sqrt{\frac{\lambda\bar{Z}\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}-\Delta)+2\lambda\bar{Z}\sqrt{\tau_\theta\tau_u}}}+\sqrt{\frac{\lambda\bar{Z}\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}+\Delta)+2\lambda\bar{Z}\sqrt{\tau_\theta\tau_u}}}\right)=1. \tag{C19}$$

Taking derivative of $\sqrt{\frac{\lambda\bar{Z}\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}-\Delta)+2\lambda\bar{Z}\sqrt{\tau_\theta\tau_u}}}+\sqrt{\frac{\lambda\bar{Z}\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}+\Delta)+2\lambda\bar{Z}\sqrt{\tau_\theta\tau_u}}}$ with respect to λ yields

$$\frac{\gamma\left(\left(\bar{Z}-\Delta\right)\left(\frac{\lambda\bar{Z}\sqrt{\tau_\theta}\sqrt{\tau_u}}{\bar{Z}\left(\gamma+2\lambda\sqrt{\tau_\theta}\sqrt{\tau_u}\right)-\gamma\Delta}\right)^{3/2}+\left(\bar{Z}+\Delta\right)\left(\frac{\lambda\bar{Z}\sqrt{\tau_\theta}\sqrt{\tau_u}}{\bar{Z}\left(\gamma+2\lambda\sqrt{\tau_\theta}\sqrt{\tau_u}\right)+\gamma\Delta}\right)^{3/2}\right)}{2\lambda^2\bar{Z}\sqrt{\tau_\theta}\sqrt{\tau_u}}>0.$$

Together with $\frac{\partial C(\cdot)}{\partial\lambda}>0$, we know that the LHS of equation (C19) is increasing in λ. Taking derivative of the LHS of equation (C19) with respect to $\bar{Z}$ generates $\frac{\gamma\Delta e^{\frac{\gamma C}{1-\beta}}}{2\sqrt{2}\lambda\bar{Z}^2\sqrt{\tau_\theta}\sqrt{\tau_u}}\left(\left(\frac{\lambda\bar{Z}\sqrt{\tau_\theta}\sqrt{\tau_u}}{\bar{Z}\left(\gamma+2\lambda\sqrt{\tau_\theta}\sqrt{\tau_u}\right)+\gamma\Delta}\right)^{3/2}-\left(\frac{\lambda\bar{Z}\sqrt{\tau_\theta}\sqrt{\tau_u}}{\bar{Z}\left(\gamma+2\lambda\sqrt{\tau_\theta}\sqrt{\tau_u}\right)-\gamma\Delta}\right)^{3/2}\right)<0$, where the inequality follows because $\left(\frac{\lambda\bar{Z}\sqrt{\tau_\theta}\sqrt{\tau_u}}{\bar{Z}\left(\gamma+2\lambda\sqrt{\tau_\theta}\sqrt{\tau_u}\right)+\gamma\Delta}\right)-\left(\frac{\lambda\bar{Z}\sqrt{\tau_\theta}\sqrt{\tau_u}}{\bar{Z}\left(\gamma+2\lambda\sqrt{\tau_\theta}\sqrt{\tau_u}\right)-\gamma\Delta}\right)=-\frac{2\gamma\Delta\lambda\bar{Z}\sqrt{\tau_\theta}\sqrt{\tau_u}}{\bar{Z}^2\left(\gamma+2\lambda\sqrt{\tau_\theta}\sqrt{\tau_u}\right)^2-\gamma^2\Delta^2}<0$. So the LHS of equation (C19) is decreasing in $\bar{Z}$. Therefore, when $\bar{Z}$ increases, λ should increase to maintain the equation, which implies that $\frac{\partial\lambda^*}{\partial\bar{Z}}>0$. Next, replacing $Z_H = \bar{Z}+\Delta$ and $Z_L = \bar{Z}-\Delta$ in equation (C17) and taking derivative with respect to $\bar{Z}$ yield

$$\frac{\partial\alpha}{\partial\bar{Z}}(\lambda;0)=-\frac{\gamma w\tau_\theta\left(2\lambda\Delta^2+\bar{Z}\left(\bar{Z}^2+\Delta^2\right)\lambda'(\bar{Z})\right)}{\lambda^2\bar{Z}^3\tau_u}<0.$$

With $x(\alpha,\lambda;0)=\frac{\alpha\gamma}{\lambda\bar{Z}}$, we have $\frac{\partial x^*}{\partial\bar{Z}}=\frac{\partial x^*}{\partial\alpha}\frac{\partial\alpha}{\partial\bar{Z}}+\frac{\partial x^*}{\partial\lambda}\frac{\partial\lambda}{\partial\bar{Z}}+\frac{\partial x^*}{\partial\bar{Z}}<0$. Further, by (19), we have $\frac{\partial G^*}{\partial\bar{Z}}=\frac{1}{1-\beta}\frac{\partial C(\lambda)}{\partial\lambda}\frac{\partial\lambda}{\bar{Z}}>0$. Thus, $\frac{\partial q^*}{\partial\bar{Z}}>0$ according to equation (16). And similar to the proof of $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c}<0$ and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial c}<0$ in the proof of Proposition 4, we can show $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial\bar{Z}}=\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial\alpha}\frac{\partial\alpha}{\partial\bar{Z}}>0$ and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\bar{Z}}=\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\alpha}\frac{\partial\alpha}{\partial\bar{Z}}>0$.

For the proof of $\frac{\partial Performance}{\partial\bar{Z}}>0$, we rely on the indifference condition for skill-acquisition decision of the marginal investor. When $\tau_\eta=+\infty$, equation (19) can be simplified as

$$-\frac{1-\beta}{\gamma}\ln\left[\frac{1}{2}\left(\frac{1}{\sqrt{1+\frac{x\left(\bar{Z}+\Delta\right)}{\alpha^2\tau_u+\tau_\theta}}}+\frac{1}{\sqrt{1+\frac{x\left(\bar{Z}-\Delta\right)}{\alpha^2\tau_u+\tau_\theta}}}\right)\right]=C\left(\lambda^*\right). \tag{C20}$$

We have shown that when $\tau_\eta = +\infty$, $\frac{\partial \lambda^*}{\partial \bar{Z}} > 0$. Together with equation (C20), we know that

$$\frac{\partial \left(\frac{1}{\sqrt{1+\frac{x(\bar{Z}+\Delta)}{\alpha^2\tau_u+\tau_\theta}}} + \frac{1}{\sqrt{1+\frac{x(\bar{Z}-\Delta)}{\alpha^2\tau_u+\tau_\theta}}} \right)}{\partial \bar{Z}} = -\frac{1}{2} \frac{\frac{\partial}{\partial \bar{Z}} \frac{x(\bar{Z}+\Delta)}{\alpha^2\tau_u+\tau_\theta}}{\left(1+\frac{x(\bar{Z}+\Delta)}{\alpha^2\tau_u+\tau_\theta}\right)^{\frac{3}{2}}} - \frac{1}{2} \frac{\frac{\partial}{\partial \bar{Z}} \frac{x(\bar{Z}+\Delta)}{\alpha^2\tau_u+\tau_\theta}}{\left(1+\frac{x(\bar{Z}+\Delta)}{\alpha^2\tau_u+\tau_\theta}\right)^{\frac{3}{2}}} < 0. \quad \text{(C21)}$$

Further, when $\tau_\eta = +\infty$, performance (21) can be simplified as

$$Performance = \frac{1}{4\gamma} \left(\ln \left[1 + \frac{x\left(\bar{Z}+\Delta\right)}{\alpha^2\tau_u+\tau_\theta} \right] + \ln \left[1 + \frac{x\left(\bar{Z}-\Delta\right)}{\alpha^2\tau_u+\tau_\theta} \right] \right).$$

Thus,

$$\frac{\partial Performance}{\partial \bar{Z}} = \frac{1}{4\gamma} \left(\frac{\frac{\partial\left(\frac{x(\bar{Z}+\Delta)}{\alpha^2\tau_u+\tau_\theta}\right)}{\partial \bar{Z}}}{1+\frac{x(\bar{Z}+\Delta)}{\alpha^2\tau_u+\tau_\theta}} + \frac{\frac{\partial\left(\frac{x(\bar{Z}-\Delta)}{\alpha^2\tau_u+\tau_\theta}\right)}{\partial \bar{Z}}}{1+\frac{x(\bar{Z}-\Delta)}{\alpha^2\tau_u+\tau_\theta}} \right) > 0,$$

where the inequality is based on (C21). QED.

Proof of Proposition 7

Denote $\tau_\eta \equiv \frac{1}{w}$. Equation (14) is equivalent to

$$\begin{aligned} 0 &= 2\alpha\gamma \left(wxZ_H + \alpha^2 w\tau_u + 1\right)\left(wxZ_L + \alpha^2 w\tau_u + 1\right) \\ &\quad -\lambda x \left(Z_H\left(2wxZ_L + \alpha^2 w\tau_u + 1\right) + Z_L\left(\alpha^2 w\tau_u + 1\right)\right), \end{aligned} \quad \text{(C22)}$$

and equation (19) is equivalent to

$$\frac{1}{2} e^{\frac{\gamma C}{1-\beta}} \left(\begin{array}{c} \sqrt{\frac{(\tau_\theta+\alpha^2\tau_u(w\tau_\theta+1))(wxZ_H+\alpha^2 w\tau_u+1)}{(\alpha^2 w\tau_u+1)(\tau_\theta+Z_H(wx\tau_\theta+x)+\alpha^2\tau_u(w\tau_\theta+1))}} \\ +\sqrt{\frac{(\tau_\theta+\alpha^2\tau_u(w\tau_\theta+1))(wxZ_L+\alpha^2 w\tau_u+1)}{(\alpha^2 w\tau_u+1)(\tau_\theta+Z_L(wx\tau_\theta+x)+\alpha^2\tau_u(w\tau_\theta+1))}} \end{array} \right) = 1. \quad \text{(C23)}$$

Suppose $w \to 0$. Conditions $\Delta = \bar{Z}$ and $\Delta = 0$ are equivalent to $Z_L = 0$ and $Z_L = Z_H$, respectively. The idea of the proof is to obtain $x(\alpha, \lambda; w, Z_L)$ around $w = 0$ and $\Delta = \bar{Z}$ (or $w = 0$ and $\Delta = 0$) from (C22), plug $x(\alpha, \lambda; w, Z_L)$ into (C23) and solve $\alpha(\lambda; w, Z_L)$ from the FOC, and finally insert $\alpha(\lambda; w, Z_L)$ back into equation (C23) to obtain the effect of the parameter change on λ^* and other endogenous variables.

First, assume $\tau_\eta = +\infty$ (i.e., $w = 0$) and $\Delta = \bar{Z}$ (i.e., $Z_L = 0$). When $w = 0$ and $Z_L = 0$, $x(\alpha, \lambda; 0, 0) = \frac{2\alpha\gamma}{\lambda Z_H}$ by equation (C22). Applying the implicit function theorem to equation (C22) and imposing the extreme conditions ($w = 0$ and $Z_L = 0$), we have $\frac{\partial x}{\partial Z_L}(\alpha, \lambda; 0, 0) = \frac{x\left(2wxZ_H(\lambda-\alpha\gamma w)-\left(\alpha^2 w\tau_u+1\right)(2\alpha\gamma w-\lambda)\right)}{Z_H(4wxZ_L(\alpha\gamma w-\lambda)+(\alpha^2 w\tau_u+1)(2\alpha\gamma w-\lambda))+Z_L(\alpha^2 w\tau_u+1)(2\alpha\gamma w-\lambda)} = -\frac{2\alpha\gamma}{\lambda Z_H^2}$, where the second equation follows by inserting $w = 0$, $Z_L = 0$, and $x = \frac{2\alpha\gamma}{\lambda Z_H}$. Similarly, when $w = 0$ and $Z_L = 0$, $\frac{\partial x}{\partial w}(\alpha, \lambda; 0, 0) = \frac{2\alpha^2\gamma(2\gamma+\alpha\gamma\tau_u)}{\lambda^2 Z_H}$. Now the Taylor series of $x(\alpha, \lambda; w, Z_L)$ around $Z_L = 0$ and $w = 0$ is

$$\begin{aligned} x(\alpha; w, Z_L) &= x(\alpha; 0, 0) + \frac{\partial x}{\partial Z_L}(\alpha; 0, 0)(Z_L - 0) + \frac{\partial x}{\partial w}(\alpha; 0, 0)(w - 0) \\ &= \frac{2\alpha\gamma\left(Z_H\left(\lambda + \alpha^2\lambda w\tau_u + 2\alpha\gamma w\right) - \lambda Z_L\right)}{\lambda^2 Z_H^2}. \end{aligned} \tag{C24}$$

Next, we solve $\alpha(\lambda; w, Z_L)$. Inserting equation (C24) into equation (C23) and taking derivative of its LHS with respect to α (note x is a function of α) yield

$$\frac{1}{2}e^{\frac{\gamma C}{1-\beta}}\left(\begin{array}{c} -\frac{Z_H\left(\alpha\tau_u\left(-2wx^2Z_H(w\tau_\theta+1)+\alpha(2w\tau_\theta+1)\frac{\partial x}{\partial\alpha}-2x(2w\tau_\theta+1)\right)+\alpha^3 w\tau_u^2(w\tau_\theta+1)\left(\alpha\frac{\partial x}{\partial\alpha}-4x\right)+\tau_\theta\frac{\partial x}{\partial\alpha}\right)}{2(\alpha^2 w\tau_u+1)^2\sqrt{\frac{(\tau_\theta+\alpha^2\tau_u(w\tau_\theta+1))(wxZ_H+\alpha^2 w\tau_u+1)}{(\alpha^2 w\tau_u+1)(\tau_\theta(wxZ_H+1)+xZ_H+\alpha^2\tau_u(w\tau_\theta+1))}}(\tau_\theta(wxZ_H+1)+xZ_H+\alpha^2\tau_u(w\tau_\theta+1))^2} \\ -\frac{Z_L\left(\alpha\tau_u\left(-2wx^2Z_L(w\tau_\theta+1)+\alpha(2w\tau_\theta+1)\frac{\partial x}{\partial\alpha}-2x(2w\tau_\theta+1)\right)+\alpha^3 w\tau_u^2(w\tau_\theta+1)\left(\alpha\frac{\partial x}{\partial\alpha}-4x\right)+\tau_\theta\frac{\partial x}{\partial\alpha}\right)}{2(\alpha^2 w\tau_u+1)^2\sqrt{\frac{(\tau_\theta+\alpha^2\tau_u(w\tau_\theta+1))(wxZ_L+\alpha^2 w\tau_u+1)}{(\alpha^2 w\tau_u+1)(\tau_\theta(wxZ_L+1)+xZ_L+\alpha^2\tau_u(w\tau_\theta+1))}}(\tau_\theta(wxZ_L+1)+xZ_L+\alpha^2\tau_u(w\tau_\theta+1))^2} \end{array}\right) = 0. \tag{C25}$$

Denote the LHS of equation (C25) as $\Psi_1(\alpha, \lambda; Z_L, w)$. When $w = 0$ and $Z_L = 0$, $\Psi_1(\alpha, \lambda; 0, 0) = \frac{\gamma\lambda e^{\frac{\gamma C}{1-\beta}}\left(\alpha^2\tau_u-\tau_\theta\right)}{2\sqrt{\frac{\lambda(\tau_\theta+\alpha^2\tau_u)}{2\alpha\gamma+\lambda\tau_\theta+\alpha^2\lambda\tau_u}}(2\alpha\gamma+\lambda\tau_\theta+\alpha^2\lambda\tau_u)^2} = 0$, which yields $\alpha(\lambda; 0, 0) = \sqrt{\frac{\tau_\theta}{\tau_u}}$. Applying the implicit function theorem to equation (C25) and imposing extreme conditions ($w = 0$ and $Z_L = 0$), we know that $\frac{\partial\alpha}{\partial Z_L}(\lambda; 0, 0) = 0$, $\frac{\partial\alpha}{\partial w}(\lambda; 0, 0) = -\frac{\tau_\theta}{2}\sqrt{\frac{\tau_\theta}{\tau_u}}$, $\frac{\partial^2\alpha}{\partial Z_L^2}(\lambda; 0, 0) = 0$,
$\frac{\partial^2\alpha}{\partial w^2}(\lambda; 0, 0) = \frac{2\gamma^3\sqrt{\tau_\theta}\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}+6\gamma^2\lambda\tau_\theta\sqrt{\tau_u}\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}+2\lambda^3\tau_\theta^2\tau_u^{3/2}\left(\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}-1\right)+\gamma\lambda^2\tau_\theta^{3/2}\tau_u\left(6\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}+1\right)}{4\lambda^2 Z_H\tau_u^{3/2}\left(\gamma+\lambda\sqrt{\tau_\theta\tau_u}\right)}$,
and $\frac{\partial^2\alpha}{\partial w\partial Z_L}(\lambda; 0, 0) = \frac{2\lambda\tau_\theta^3+\frac{5\gamma\tau_\theta^{5/2}}{\sqrt{\tau_u}}}{8\gamma+8\lambda\sqrt{\tau_\theta\tau_u}}$. The Taylor series of $\alpha(\lambda; w, Z_L)$ around $w = 0$ and $Z_L = 0$

becomes

$$
\begin{aligned}
\alpha(\lambda; w, Z_L) &= \alpha(\lambda;0,0) + \frac{\partial \alpha}{\partial Z_L}(\lambda;0,0)(Z_L-0) + \frac{\partial \alpha}{\partial w}(\lambda;0,0)(w-0) \\
&\quad + \frac{1}{2}\frac{\partial^2 \alpha}{\partial Z_L^2}(\lambda;0,0)(Z_L-0)^2 + \frac{1}{2}\frac{\partial^2 \alpha}{\partial w^2}(\lambda;0,0)(w-0)^2 + \frac{\partial^2 \alpha}{\partial w \partial Z_L}(\lambda;0,0)(Z_L-0)(w-0) \\
&= \frac{\sqrt{\tau_\theta}\left(\begin{array}{c} \lambda^2 Z_H \tau_u \left(2\lambda\sqrt{\tau_\theta \tau_u}\left(w\tau_\theta\left(w\tau_\theta - 4\right)+8\right) + \gamma\left(w\tau_\theta\left(5w\tau_\theta - 8\right)+16\right)\right) \\ +4wZ_L \left(\begin{array}{c} 2\gamma^3\sqrt{1-\frac{\gamma}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}} + 6\gamma^2\lambda\sqrt{\tau_\theta\tau_u}\sqrt{1-\frac{\gamma}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}} \\ +2\lambda^3\tau_\theta^{3/2}\tau_u^{3/2}\left(\sqrt{1-\frac{\gamma}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}-1\right) \\ +\gamma\lambda^2\tau_\theta\tau_u\left(6\sqrt{1-\frac{\gamma}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}+1\right) \end{array}\right) \end{array}\right)}{16\lambda^2 Z_H \tau_u^{3/2}\left(\gamma+\lambda\sqrt{\tau_\theta\tau_u}\right)}. \qquad \text{(C26)}
\end{aligned}
$$

Inserting equations (C24) and (C26) into equation (C23) and imposing $w = 0$, we have

$$
\frac{1}{2}e^{\frac{\gamma C}{1-\beta}}\left(\sqrt{\frac{\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}{2\gamma\Delta+\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}} + \sqrt{\frac{\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}{2\gamma\Delta(\bar{Z}-\Delta)+\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}}\right) = 1. \qquad \text{(C27)}
$$

Taking derivative of $\sqrt{\frac{\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}{2\gamma\Delta+\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}} + \sqrt{\frac{\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}{2\gamma\Delta(\bar{Z}-\Delta)+\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}}$ with respect to λ yields $\frac{\gamma\left(\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}\right)^{3/2}}{2\lambda^2\sqrt{\tau_\theta\tau_u}} > 0$. Together with $\frac{\partial C(\cdot)}{\partial \lambda} > 0$, the LHS of equation (C27) is increasing in λ. Taking derivative of the LHS of equation (C27) with respect to Δ yields $\frac{\gamma e^{\frac{\gamma C}{1-\beta}}\left(2-2\left(\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}\right)^{3/2}\right)}{16\lambda\bar{Z}\sqrt{\tau_\theta\tau_u}} > 0$ and thus, the LHS of equation (C27) is increasing in Δ. Therefore, when Δ increases, λ should decrease to maintain the equation: $\frac{\partial \lambda^*}{\partial \Delta} < 0$.

Next, we solve $\frac{\partial \alpha^*}{\partial \Delta}$. In equation (C26), replacing $Z_H = \bar{Z}+\Delta$ and $Z_L = \bar{Z}-\Delta$, taking derivative with respect to Δ (note λ is a function of Δ), and imposing $\Delta = \bar{Z}$ ($Z_L = 0$), we can show

$$
\begin{aligned}
\frac{\partial\alpha}{\partial\Delta}(\lambda;w,0) &= -\frac{w\tau_\theta\left(\begin{array}{c}4\gamma^3+\gamma\lambda\tau_\theta\tau_u\left(3\bar{Z}w\tau_\theta\lambda'(\Delta)\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}-2\lambda\left(\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}-6\right)\right)\\+2\gamma^2\lambda\sqrt{\tau_\theta\tau_u}\left(\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}+6\right)-4\lambda^3\tau_\theta^{3/2}\tau_u^{3/2}\left(\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}-1\right)\end{array}\right)}{16\lambda\bar{Z}\tau_u\sqrt{1-\frac{\gamma}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}\left(\gamma+\lambda\sqrt{\tau_\theta\tau_u}\right)^2}\\
&\simeq -\frac{w\tau_\theta\left(\begin{array}{c}4\gamma^3+\gamma\lambda\tau_\theta\tau_u\left(-2\lambda\left(\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}-6\right)\right)\\+2\gamma^2\lambda\sqrt{\tau_\theta\tau_u}\left(\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}+6\right)-4\lambda^3\tau_\theta^{3/2}\tau_u^{3/2}\left(\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}-1\right)\end{array}\right)}{16\lambda\bar{Z}\tau_u\sqrt{1-\frac{\gamma}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}\left(\gamma+\lambda\sqrt{\tau_\theta\tau_u}\right)^2}\\
&= \frac{w\sqrt{\tau_\theta}\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}\left(\begin{array}{c}-2\gamma^2+2\lambda^2\tau_\theta\tau_u\left(\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}-1\right)\\-\gamma\lambda\sqrt{\tau_\theta\tau_u}\left(\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}+4\right)\end{array}\right)}{8\lambda^2\bar{Z}\tau_u^{3/2}}\\
&< 0.
\end{aligned}
$$

The approximation follows because by applying the implicit function theorem to equation (C27), we have $\lambda'(\Delta)=-\frac{2-2\left(\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}\right)^{3/2}}{8\bar{Z}\lambda\sqrt{\tau_\theta\tau_u}\left(\frac{\left(\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}\right)^{3/2}}{2\lambda^2\sqrt{\tau_\theta\tau_u}}+\frac{1+\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}{1-\beta}C'(\lambda)\right)}$ and thus, $3\bar{Z}w\tau_\theta\lambda'(\Delta)\sqrt{\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+\lambda\sqrt{\tau_\theta\tau_u}}}=0$ as $w=0$. Also, inserting $\alpha(\lambda;0,0)$ into $x(\alpha,\lambda;0)$, taking derivative with respect to Δ and imposing $\Delta=\bar{Z}$ and $w=0$ yield $\frac{\partial x^*}{\partial\Delta}=-\frac{\gamma}{\bar{Z}\lambda^2}\sqrt{\frac{\tau_\theta}{\tau_u}}\lambda'(\Delta)>0$. Further, by equation (19), we have $\frac{\partial G^*}{\partial\Delta}=\frac{1}{1-\beta}\frac{\partial C(\lambda)}{\partial\lambda}\frac{\partial\lambda}{\partial\Delta}<0$. Thus, $\frac{\partial q^*}{\partial\Delta}<0$ by equation (16). And similar to the proof of $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c}<0$ and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial c}<0$ for sufficiently high τ_η in the proof of Proposition 5, we can show $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial\Delta}=\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial\alpha}\frac{\partial\alpha}{\partial\Delta}>0$ and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\Delta}=\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\alpha}\frac{\partial\alpha}{\partial\Delta}>0$.

When $w=0$ and $Z_L=Z_H$, $x(\alpha,\lambda;0,Z_H)=\frac{\alpha\gamma}{\lambda Z_H}$ based on equation (C22). Applying the implicit function theorem to equation (C22) and imposing the extreme conditions ($w=0$ and $Z_L=Z_H$), we have $\frac{\partial x}{\partial Z_L}(\alpha;0,Z_H)=\frac{x\left(2wxZ_H(\lambda-\alpha\gamma w)-\left(\alpha^2w\tau_u+1\right)(2\alpha\gamma w-\lambda)\right)}{Z_H(4wxZ_L(\alpha\gamma w-\lambda)+(\alpha^2w\tau_u+1)(2\alpha\gamma w-\lambda))+Z_L(\alpha^2w\tau_u+1)(2\alpha\gamma w-\lambda)}=-\frac{\alpha\gamma}{2\lambda Z_H^2}$, where the second equation follows by inserting $w=0$, $Z_L=Z_H$ and $x=\frac{\alpha\gamma}{\lambda Z_H}$. Similarly, as $w=0$ and $Z_L=Z_H$, $\frac{\partial x}{\partial w}(\alpha;0,Z_H)=\frac{\alpha^2\gamma(\gamma+\alpha\gamma\tau_u)}{\lambda^2 Z_H}$. Now the Taylor series of $x(\alpha;w,Z_L)$ around $w=0$ and $Z_L=Z_H$ is

$$
\begin{aligned}
x(\alpha;w,Z_L) &= x(\alpha;0,Z_H)+\frac{\partial x}{\partial Z_L}(\alpha;0,Z_H)(Z_L-Z_H)+\frac{\partial x}{\partial w}(\alpha;0,Z_H)(w-0)\\
&= \frac{\alpha\gamma\left(Z_H\left(3\lambda+2\alpha^2\lambda w\tau_u+2\alpha\gamma w\right)-\lambda Z_L\right)}{2\lambda^2 Z_H^2}. \qquad \text{(C28)}
\end{aligned}
$$

Next, we solve $\alpha\left(\lambda; w, Z_L\right)$. Inserting equation (C28) into equation (C23), and taking derivative of its LHS with respect to α (note that x is a function of α), we have

$$\frac{e^{\frac{C\gamma}{1-\beta}}\gamma\left(\begin{array}{c}\frac{\left(\begin{array}{c} w\gamma\lambda^2 Z_L^2\tau_u\left(w\tau_\theta+1\right)\alpha^3 \\ -\lambda Z_H Z_L\left(\begin{array}{c} w\lambda(4w\alpha\gamma+3\lambda)\tau_u^2\left(w\tau_\theta+1\right)\alpha^4 \\ +\tau_u\left(4w^2\alpha^2\gamma^2+6w\alpha\lambda\gamma+\lambda^2+2w\left(2w^2\alpha^2\gamma^2+3w\alpha\lambda\gamma+\lambda^2\right)\tau_\theta\right)\alpha^2-\lambda^2\tau_\theta \end{array}\right) \\ +Z_H^2\left(\begin{array}{c} 2w^2\lambda^2(2w\alpha\gamma+\lambda)\tau_u^3\left(w\tau_\theta+1\right)\alpha^6+w\lambda\tau_u^2\left(\begin{array}{c} 8w^2\alpha^2\gamma^2+16w\alpha\lambda\gamma+7\lambda^2 \\ +w\left(8w^2\alpha^2\gamma^2+16w\alpha\lambda\gamma+5\lambda^2\right)\tau_\theta \end{array}\right)\alpha^4 \\ +\tau_u\left(4w^3\alpha^3\gamma^3+12w^2\alpha^2\lambda\gamma^2+9w\alpha\lambda^2\gamma+w^2\alpha(2w\alpha\gamma+3\lambda)^2\tau_\theta\gamma+3\lambda^3\right)\alpha^2 \\ -\lambda^2(4w\alpha\gamma+3\lambda)\tau_\theta \end{array}\right) \end{array}\right)}{\left(\alpha\gamma\lambda Z_L\left(w\tau_\theta+1\right)-Z_H\left(\begin{array}{c} 2\lambda(w\alpha\gamma+\lambda)\tau_u\left(w\tau_\theta+1\right)\alpha^2+\gamma(2w\alpha\gamma+3\lambda)\alpha \\ +\left(2w^2\alpha^2\gamma^2+3w\alpha\lambda\gamma+2\lambda^2\right)\tau_\theta \end{array}\right)\right)^2 \times\sqrt{\frac{\left(Z_H\left(2w^2\gamma^2\alpha^2+2w\lambda(w\alpha\gamma+\lambda)\tau_u\alpha^2+3w\gamma\lambda\alpha+2\lambda^2\right)-w\alpha\gamma\lambda Z_L\right)\left(\tau_u(w\tau_\theta+1)\alpha^2+\tau_\theta\right)}{\left(w\tau_u\alpha^2+1\right)\left(Z_H\left(2\lambda(w\alpha\gamma+\lambda)\tau_u(w\tau_\theta+1)\alpha^2+\gamma(2w\alpha\gamma+3\lambda)\alpha+\left(2w^2\alpha^2\gamma^2+3w\alpha\lambda\gamma+2\lambda^2\right)\tau_\theta\right)-\alpha\gamma\lambda Z_L(w\tau_\theta+1)\right)}}} \\ +\frac{Z_L\left(\begin{array}{c} w\gamma\lambda^2 Z_L^3\tau_u\left(w\tau_\theta+1\right)\alpha^3-2w\gamma\lambda Z_H Z_L^2\tau_u\left(2w\lambda\tau_u\alpha^2+2w\gamma\alpha+3\lambda\right)\left(w\tau_\theta+1\right)\alpha^3 \\ +\lambda^2 Z_H^3\left(2w^2\lambda\tau_u^3\left(w\tau_\theta+1\right)\alpha^6+w\tau_u^2\left(4w\alpha\gamma+7\lambda+w(4w\alpha\gamma+5\lambda)\tau_\theta\right)\alpha^4+3\lambda\tau_u\alpha^2-(4w\alpha\gamma+3\lambda)\tau_\theta\right) \\ +Z_H^2 Z_L\left(\begin{array}{c} 4w^3\gamma\lambda^2\tau_u^3\left(w\tau_\theta+1\right)\alpha^7+w\lambda\left(8w^2\alpha^2\gamma^2+12w\alpha\lambda\gamma-3\lambda^2\right)\tau_u^2\left(w\tau_\theta+1\right)\alpha^4 \\ +\tau_u\left(\begin{array}{c} 4w^3\alpha^3\gamma^3+12w^2\alpha^2\lambda\gamma^2+9w\alpha\lambda^2\gamma-\lambda^3 \\ +w\left(4w^3\alpha^3\gamma^3+12w^2\alpha^2\lambda\gamma^2+9w\alpha\lambda^2\gamma-2\lambda^3\right)\tau_\theta \end{array}\right)\alpha^2+\lambda^3\tau_\theta \end{array}\right) \end{array}\right)}{\sqrt{\frac{\left(2\lambda^2\left(w\tau_u\alpha^2+1\right)Z_H^2+w\alpha\gamma Z_L\left(2w\lambda\tau_u\alpha^2+2w\gamma\alpha+3\lambda\right)Z_H-w\alpha\gamma\lambda Z_L^2\right)\left(\tau_u(w\tau_\theta+1)\alpha^2+\tau_\theta\right)}{\left(w\tau_u\alpha^2+1\right)\left(2\lambda^2\left(\tau_u(w\tau_\theta+1)\alpha^2+\tau_\theta\right)Z_H^2+\alpha\gamma Z_L\left(2w\lambda\tau_u\alpha^2+2w\gamma\alpha+3\lambda\right)(w\tau_\theta+1)Z_H-\alpha\gamma\lambda Z_L^2(w\tau_\theta+1)\right)}} \times\left(2\lambda^2\left(\tau_u\left(w\tau_\theta+1\right)\alpha^2+\tau_\theta\right)Z_H^2+\alpha\gamma Z_L\left(2w\lambda\tau_u\alpha^2+2w\gamma\alpha+3\lambda\right)\left(w\tau_\theta+1\right)Z_H-\alpha\gamma\lambda Z_L^2\left(w\tau_\theta+1\right)\right)^2} \end{array}\right)}{2\left(w\tau_u\alpha^2+1\right)^2}=0. \tag{C29}$$

Denote the LHS of equation (C29) as $\Psi_2\left(\alpha, \lambda; Z_L, w\right)$. When $w=0$ and $Z_L=Z_H$, $\Psi_2\left(\alpha, \lambda; 0, Z_H\right)=\frac{\gamma\lambda e^{\frac{\gamma C}{1-\beta}}\left(\alpha^2\tau_u-\tau_\theta\right)}{2\sqrt{\frac{\lambda\left(\tau_\theta+\alpha^2\tau_u\right)}{\alpha\gamma+\lambda\tau_\theta+\alpha^2\lambda\tau_u}}\left(\alpha\gamma+\lambda\tau_\theta+\alpha^2\lambda\tau_u\right)^2}=0$, which yields $\alpha\left(\lambda; 0, Z_H\right)=\sqrt{\frac{\tau_\theta}{\tau_u}}$. Applying the implicit function theorem to equation (C29) and imposing extreme conditions ($w=0$ and $Z_L=Z_H$), $\frac{\partial\alpha}{\partial Z_L}\left(\lambda; 0, Z_H\right)=0$, $\frac{\partial\alpha}{\partial w}\left(\lambda; 0, Z_H\right)=-\frac{\tau_\theta}{2}\sqrt{\frac{\tau_\theta}{\tau_u}}$, $\frac{\partial^2\alpha}{\partial Z_L^2}\left(\lambda; 0, Z_H\right)=0$, $\frac{\partial^2\alpha}{\partial w^2}\left(\lambda; 0, Z_H\right)=\frac{20\lambda\tau_\theta^3+\frac{13\gamma\tau_\theta^{5/2}}{\sqrt{\tau_u}}}{8\gamma+16\lambda\sqrt{\tau_\theta}\sqrt{\tau_u}}$, and $\frac{\partial^2\alpha}{\partial w\partial Z_L}\left(\lambda; 0, 0\right)=0$. The Taylor series of $\alpha\left(\lambda; w, Z_L\right)$ around $w=0$ and $Z_L=Z_H$ is

$$\begin{aligned}\alpha\left(\lambda; w, Z_L\right) &= \alpha\left(\lambda; 0, Z_H\right)+\frac{\partial\alpha}{\partial Z_L}\left(\lambda; 0, Z_H\right)\left(Z_L-Z_H\right)+\frac{\partial\alpha}{\partial w}\left(\lambda; 0, Z_H\right)(w-0) \\ &\quad +\frac{1}{2}\frac{\partial^2\alpha}{\partial Z_L^2}\left(\lambda; 0, Z_H\right)\left(Z_L-Z_H\right)^2+\frac{1}{2}\frac{\partial^2\alpha}{\partial w^2}\left(\lambda; 0, Z_H\right)(w-0)^2 \\ &\quad +\frac{\partial^2\alpha}{\partial w\partial Z_L}\left(\lambda; 0, Z_H\right)\left(Z_L-Z_H\right)(w-0) \\ &= \sqrt{\frac{\tau_\theta}{\tau_u}}+\frac{w^2\left(20\lambda\tau_\theta^3+\frac{13\gamma\tau_\theta^{5/2}}{\sqrt{\tau_u}}\right)}{16\left(\gamma+2\lambda\sqrt{\tau_\theta\tau_u}\right)}-\frac{w\tau_\theta^{3/2}}{2\sqrt{\tau_u}}.\end{aligned} \tag{C30}$$

Now inserting equations (C28) and (C30) into equation (C23) and imposing $w = 0$, we have

$$\frac{1}{\sqrt{2}}e^{\frac{\gamma C}{1-\beta}}\left(\begin{array}{c}\sqrt{\frac{\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}+2\Delta)+2\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}} \\ +\sqrt{\frac{\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}^2+\bar{Z}\Delta-2\Delta^2)+2\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}}\end{array}\right) = 1. \tag{C31}$$

Taking derivative of $\sqrt{\frac{\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}+2\Delta)+2\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}}+\sqrt{\frac{\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}^2+\bar{Z}\Delta-2\Delta^2)+2\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}}$ with respect to λ and imposing $\Delta = 0$ yield $\frac{\gamma\left(\frac{\lambda\sqrt{\tau_\theta\tau_u}}{\gamma+2\lambda\sqrt{\tau_\theta\tau_u}}\right)^{3/2}}{\lambda^2\sqrt{2\tau_\theta\tau_u}} > 0$. Together with $\frac{\partial C(\cdot)}{\partial\lambda} > 0$, the LHS of equation (C31) is increasing in λ. Taking derivative of the LHS of equation (C31) with respect to Δ yields (note that λ is a function of Δ), we have

$$\begin{aligned}
& \frac{\gamma\bar{Z}e^{\frac{\gamma C}{1-\beta}}\left((\bar{Z}+5\Delta)\left(\frac{\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}^2+\bar{Z}\Delta-2\Delta^2)+2\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}\right)^{3/2}-(\bar{Z}+\Delta)\left(\frac{\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}+2\Delta)+2\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}\right)^{3/2}\right)}{2\lambda(\bar{Z}+\Delta)^3\sqrt{2\tau_\theta\tau_u}} \\
> & \frac{\gamma\bar{Z}e^{\frac{\gamma C}{1-\beta}}\left((\bar{Z}+\Delta)\left(\frac{\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}^2+\bar{Z}\Delta-2\Delta^2)+2\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}\right)^{3/2}-(\bar{Z}+\Delta)\left(\frac{\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}+2\Delta)+2\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}\right)^{3/2}\right)}{2\lambda(\bar{Z}+\Delta)^3\sqrt{2\tau_\theta\tau_u}} \\
> & \ 0,
\end{aligned}$$

where the last inequality follows because $\frac{\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}^2+\bar{Z}\Delta-2\Delta^2)+2\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}} - \frac{\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}{\gamma(\bar{Z}+2\Delta)+2\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}}$ $= \frac{2\gamma\lambda\Delta(\bar{Z}+\Delta)(\bar{Z}+2\Delta)\sqrt{\tau_\theta\tau_u}}{\left(\gamma(\bar{Z}+2\Delta)+2\lambda(\bar{Z}+\Delta)\sqrt{\tau_\theta\tau_u}\right)\left(\gamma(\bar{Z}^2+\bar{Z}\Delta-2\Delta^2)+2\lambda(\bar{Z}+\Delta)^2\sqrt{\tau_\theta\tau_u}\right)} > 0$. The LHS of equation (C31) is increasing in Δ. Therefore, when Δ increases, λ should decrease to maintain the equation. Hence, $\frac{\partial\lambda^*}{\partial\Delta} < 0$.

Finally, we compute $\frac{\partial\alpha^*}{\partial\Delta}$. In equation (C30), taking derivative with respect to Δ (note that λ is a function of Δ), and imposing $\Delta = 0$ ($Z_L = Z_H$), we have

$$\frac{\partial\alpha}{\partial\Delta}(\lambda; w, Z_H) = -\frac{3\gamma w^2\tau_\theta^3\lambda'(\Delta)}{8\left(\gamma+2\lambda\sqrt{\tau_\theta\tau_u}\right)^2} > 0.$$

Given that $x(\alpha, \lambda; 0) = \frac{\alpha\gamma}{\lambda\bar{Z}}$, we have $\frac{\partial x^*}{\partial\Delta} = \frac{\partial x}{\partial\alpha}\frac{\partial\alpha}{\partial\Delta} + \frac{\partial x}{\partial\lambda}\frac{\partial\lambda}{\partial\Delta} > 0$. Further, similar to the proof of $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c} < 0$ and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial c} < 0$ in the proof of Proposition 5, we can show $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial\Delta} = \frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial\alpha}\frac{\partial\alpha}{\partial\Delta} > 0$ and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\Delta} = \frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\alpha}\frac{\partial\alpha}{\partial\Delta} > 0$. QED.

References

Admati A.R., Pfleiderer, P., 1986. A monopolistic market for information. Journal of Economic Theory 39, 400-438.

Admati A.R., Pfleiderer, P., 1988. A theory of intraday patterns: Volume and price variability. Review of Financial Studies 1, 3-40.

Admati, A.R., Pfleiderer, P., 1990. Direct and indirect sale of information. Econometrica 58, 901-928.

Allen, F., 1990. The market for information and the origin of financial intermediation. Journal of Financial Intermediation 1, 3-30.

Anand, B.N., Galetovic, A., 2000. Information, nonexcludability, and financial market structure. Journal of Business 73, 357-402.

Bai, J., Philippon, T., Savov, A., 2016. Have financial markets become more informative? Journal of Financial Economics 122, 625-654.

Berk, J.B., van Binsbergen, J.H., 2015. Measuring skill in the mutual fund industry. Journal of Financial Economics 118, 1-20.

Berk, J.B., van Binsbergen, J.H., 2017. Active managers are skilled: On average, they add more than $3 million per year. Journal of Portfolio Management 42, 131-139.

Brandt, M.W., Brav, A., Graham, J.R., Kumar, A., 2009. The idiosyncratic volatility puzzle: Time trend or speculative episodes? Review of Financial Studies 23, 863-899.

Brown, B.R., 2010. Chasing the same signals: How black-box trading influences stock markets from Wall Street to Shanghai. John Wiley & Sons.

Cespa, G., 2008. Information sales and insider trading with long lived information. Journal of Finance 63, 639-672.

Chen, Z., Wilhelm, W.J., 2012. Sell-side information production in financial markets. Journal of Financial and Quantitative Analysis 47, 763-794.

Easley, D., O'Hara, M., Yang, L., 2016. Differential access to price information in financial markets. Journal of Financial and Quantitative Analysis 51, 1071-1110.

Farboodi, M., Matray, A., Veldkamp, L., 2017. Where has all the big data gone? Unpublished working paper.

Fishman, M.J., Hagerty, K.M., 1995. The incentive to sell financial market information.

Journal of Financial Intermediation 4, 95-115.

Hellwig, M.F., 1980. On the aggregation of information in competitive markets. Journal of Economic Theory 22, 477-498.

Gârleanu, N.B., Pedersen, L.H., 2018. Efficiently inefficient markets for assets and asset management. Journal of Finance 73, 1663-1712.

García, D., Sangiorgi, F., 2011. Information sales and strategic trading. Review of Financial Studies 24, 3069-3104.

García, D., Vanden, J.M., 2009. Information acquisition and mutual funds. Journal of Economic Theory 144, 1965-1995.

Glode, V., Lowery, R., 2016. Compensating financial experts. Journal of Finance 71, 2781-2808.

Grossman, S.J., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. American economic review 70, 393-408.

Indjejikian, R.J., 1991. The impact of costly information interpretation on firm disclosure decisions. Journal of Accounting Research 29, 277-301.

Myatt, D.P., Wallace, C., 2012. Endogenous information acquisition in coordination games. Review of Economic Studies 79, 340-374.

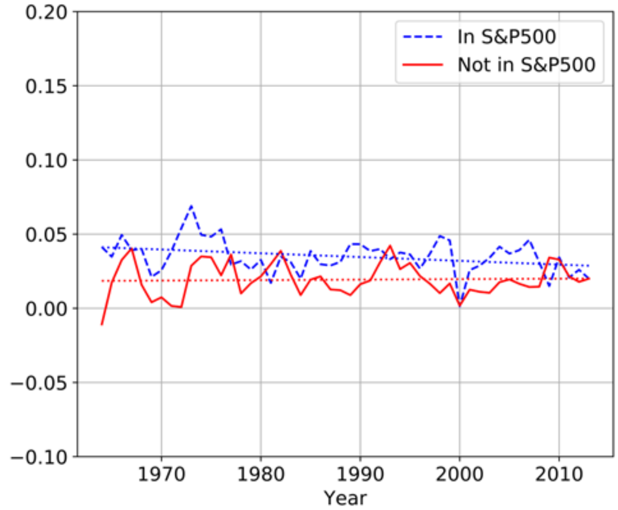
Naik, N.Y., 1997. Multi-period information markets. Journal of Economic Dynamics and Control 21, 1229-1258.

Pagano, M., Volpin, P., 2012. Securitization, disclosure and liquidity. Review of Financial Studies 25, 2417-2453.

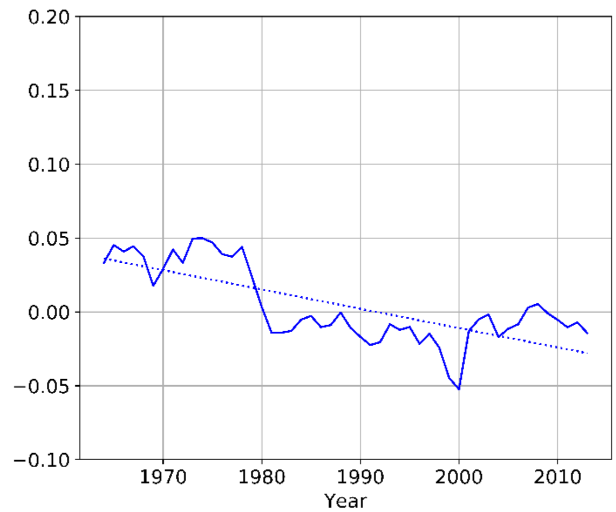
Vives, X., Yang, L., 2017. Costly interpretation of asset prices. Unpublished working paper.

Figure 1: Time trend of price informativeness

Panel A. Stagnant price informativeness for the long-lived firms



Panel B. Decreasing price informativeness for the average firms



This figure plots the time trend of price informativeness. Panel A plots price informativeness over time at a 3-year horizon for firms that exist for at least 45 year from 1964 to 2018: the blue dashed line indicates firms that were at some point in S&P500 and the red solid line indicates firms that were never in S&P500 firms. The blue and red dotted lines are linear trends that fit the blue and red time trends, respectively. Panel B plots price informativeness over time at a 3-year horizon for all listed firms from 1964 to 2018. Price informativeness is calculated based on Bai, Philippon, and Savov (2016).

Figure 2: Timeline

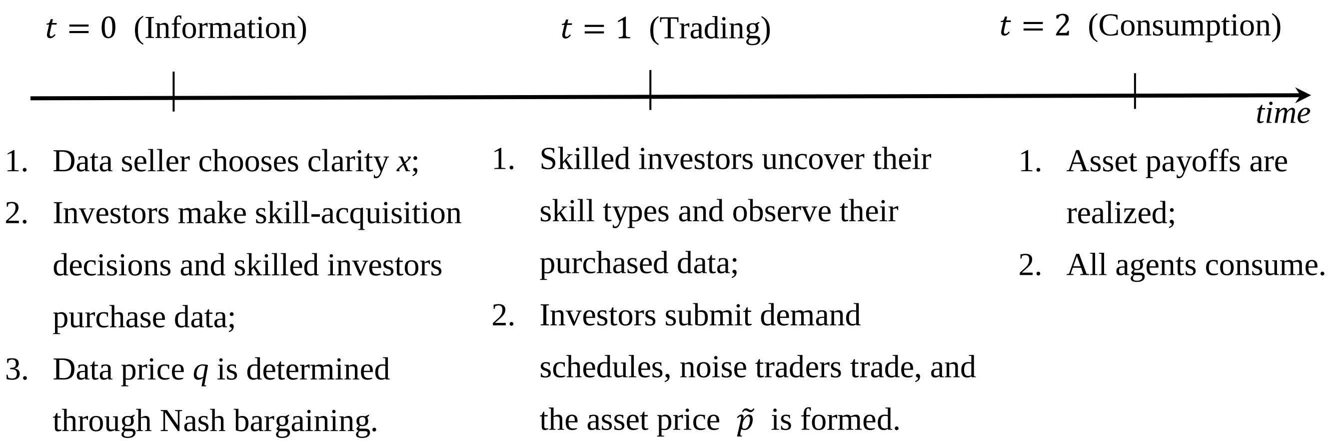
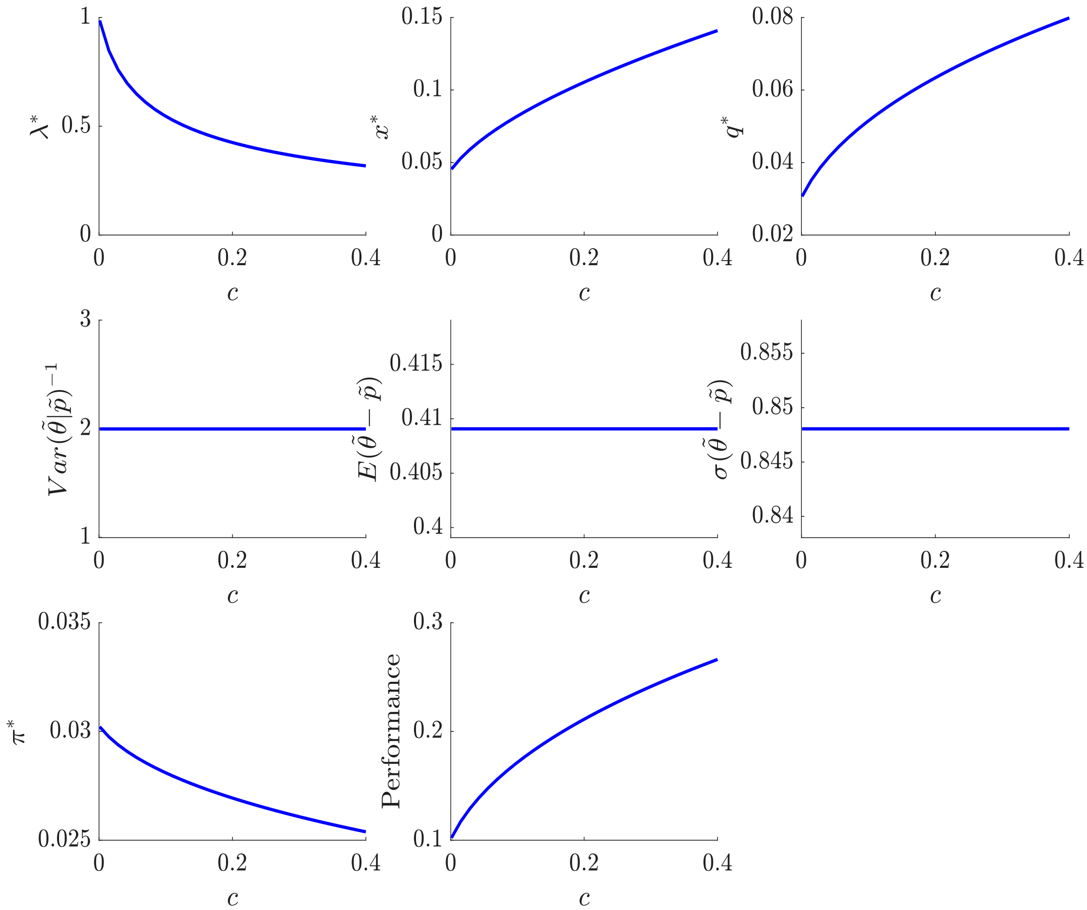
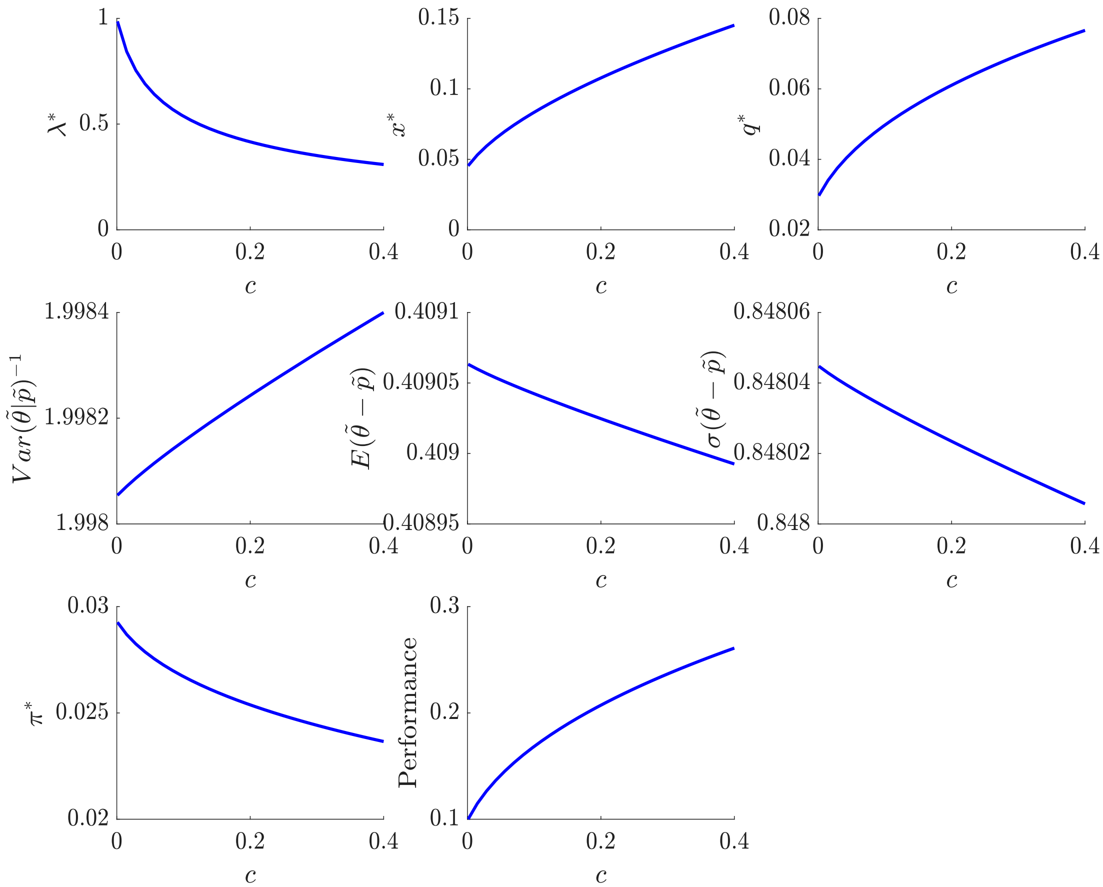


Figure 3: Implications of skill-acquisition costs without skill-acquisition uncertainty



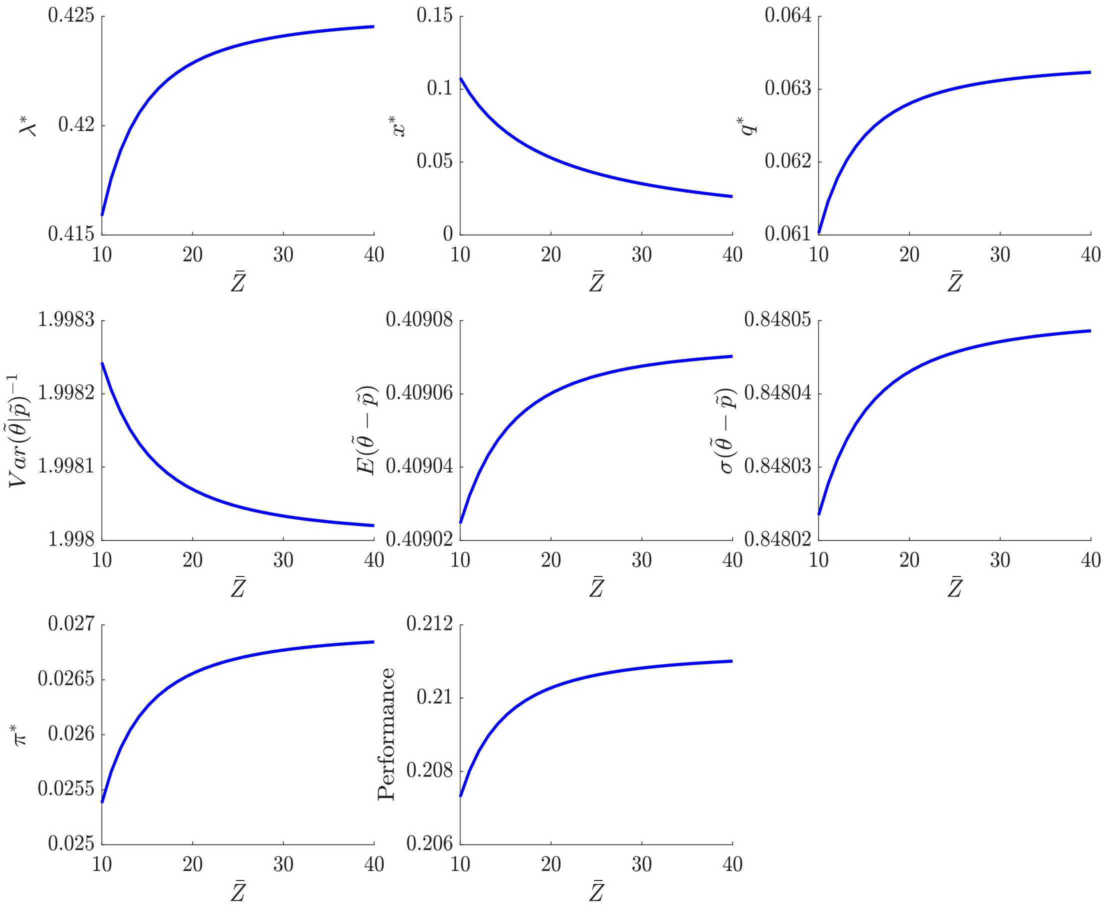
This figure plots the effects of skill-acquisition costs in the economy without skill-acquisition uncertainty. The skill-acquisition cost function is specified as $C(\mathrm{i}) = ci/(1-i)$ with $c > 0$. The other parameters are: $\tau_\theta = 1, \tau_u = 5, \tau_\eta = 1000, \beta = 0.3, \gamma = 1, \bar{Z} = 10,$ and $\Delta = 0$.

Figure 4: Implications of skill-acquisition costs with skill-acquisition uncertainty



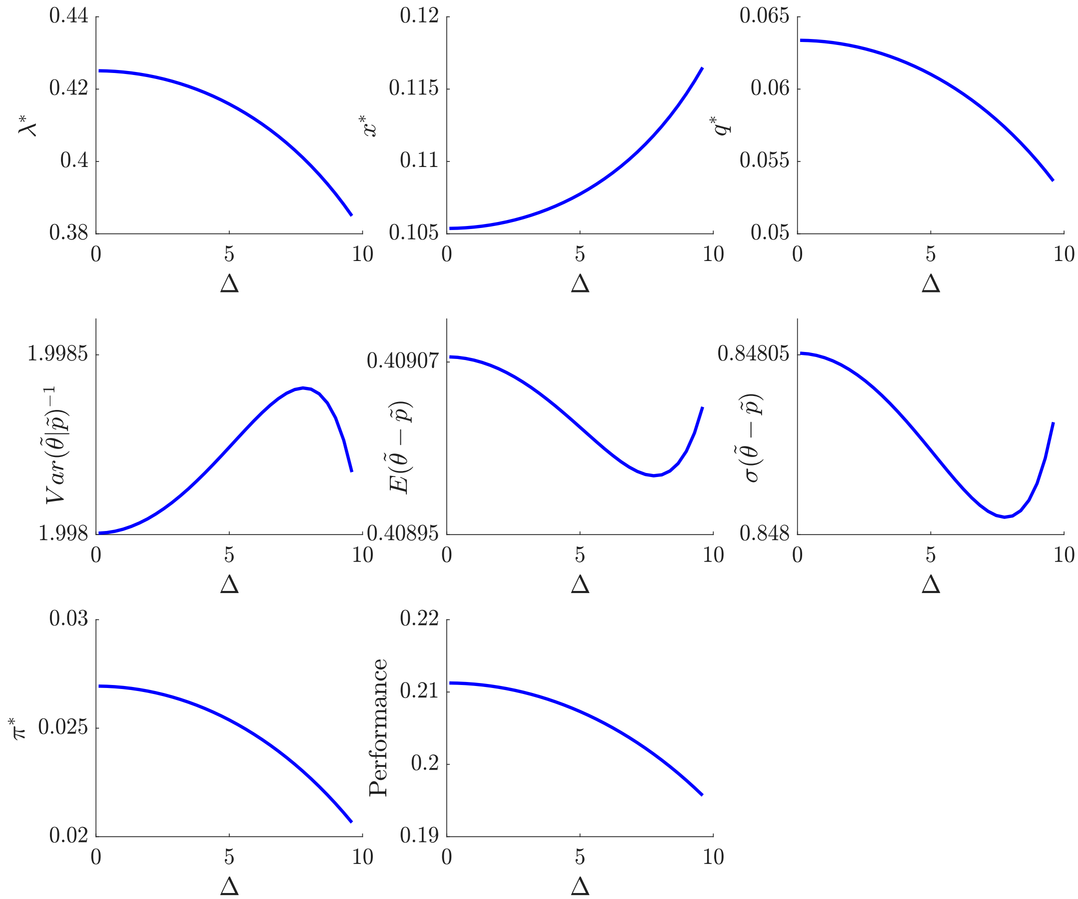
This figure plots the effects of skill-acquisition costs. The skill-acquisition cost function and the parameters are the same as those in Figure 3 with $\Delta = 5$.

Figure 5: Implications of skill mean ($\bar{Z}$) with skill-acquisition uncertainty



This figure plots the effects of skill mean $\bar{Z}$. The skill-acquisition cost function and the parameters are the same as those in Figure 4 with $c = 0.2$.

Figure 6: Implications of skill volatility (Δ) with skill-acquisition uncertainty



This figure plots the effects of skill volatility Δ. The skill-acquisition cost function and the parameters are the same as those in Figure 5 with $\bar{Z} = 10$.

Online Appendix to "Skill Acquisition and Data Sales"

In the main paper, we focus on the economy where the seller is endowed with accurate data. Now we study the economy in which the seller owns inaccurate data by examining how skill-acquisition technology affects information sales, financial markets, and the asset management industry. We find that market variables, such as price informativeness, the cost of capital, and return volatility, tend to exhibit different patterns in the two economies, which differ in the accuracy τ_η of the seller's original data.

First, as in Section 4.1 and Section 4.2.1, we study the effect of skill-acquisition cost. When the seller's data is relatively inaccurate (i.e., low τ_η) and there is no skill-acquisition uncertainty, the seller will choose not to add any noise into the sold data in order to preserve the information value (i.e., $x^* = \infty$). In this case, as skill-acquisition costs increase, still fewer investors acquire skills and hence purchase data (i.e., $\frac{\partial \lambda^*}{\partial c} < 0$), and less information is injected into the price. As a result, price informativeness declines (i.e., $\frac{\partial}{\partial c} \frac{1}{Var(\tilde{\theta}|\tilde{p})} < 0$). Since less information is leaked by the price, the data is valued more by investors, and its price q^* increases (i.e., $\frac{\partial q^*}{\partial c} > 0$). Hence, both for high and low values of τ_η, data price q^* increases with c, but for different reasons: for high value of τ_η, price informativeness is constant, and q^* increases because of the improved data clarity; for low value of τ_η, data clarity x^* is constant (at ∞), and q^* increases due to the worse price informativeness. In contrast, when τ_η is low, the seller sells data "as is" and both price informativeness and the mass of skilled investors decrease with c, which raises the uncertainty faced by investors and hence the cost of capital and return volatility.

The results for the case with skill-acquisition uncertainty are the same. The seller sells data "as is" ($x^* = \infty$). When skill-acquisition costs increase: fewer investors acquire skills and the price of data increases; price informativeness decreases, and both the cost of capital and return volatility increase; the seller's profit decreases and skilled investors' performance increases.

Proposition OA1 (Implications of skill-acquisition costs with low τ_η) *Assume the seller owns very inaccurate data (i.e., $\tau_\eta \to 0$). Whether there is uncertainty involved in skill acquisition or not, when skill-acquisition cost parameter c increases (meaning that $C(\cdot)$ shifts* *upward), when skill-acquisition cost parameter c increases (meaning that $C(\cdot)$ shifts upward), the fraction λ^* of skilled investors decreases, data price q^* increases, skilled investors' performance increases, data clarity x^* remains at ∞, price informativeness $\frac{1}{Var(\tilde{\theta}|\tilde{p})}$ decreases, the cost of capital $E(\tilde{\theta}-\tilde{p})$ increases, and return volatility $\sigma(\tilde{\theta}-\tilde{p})$ increases. That is, whether $\Delta > 0$ or not, for sufficiently low τ_η, $x^* = \infty$, $\frac{\partial \lambda^*}{\partial c} < 0$, $\frac{\partial q^*}{\partial c} > 0$, $\frac{\partial Performance}{\partial c} > 0$, $\frac{\partial}{\partial c}\frac{1}{Var(\tilde{\theta}|\tilde{p})} < 0$, $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c} > 0$, and $\frac{\partial \sigma(\tilde{\theta}-\tilde{p})}{\partial c} > 0$.*

Second, we examine the effect of the mean of skill levels in the economy with skill uncertainty (i.e., $\Delta > 0$, or $Z_L < Z_H$) as in Section 4.2.2. Again, if the seller's data is very inaccurate (i.e., if τ_η is low), she optimally sells the data "as is." That is, data clarity x^* is set at ∞. In consequence, price informativeness increases, and both the cost of capital and return volatility decrease, since more investors choose to acquire skills and become informed when $\bar{Z}$ becomes higher. Further, given the constant data clarity and increasing price informativeness, the performance of skilled investors declines with skill mean $\bar{Z}$.

Proposition OA2 (Implications of skill mean with low τ_η) A*ssume that the seller owns very inaccurate data (i.e., $\tau_\eta \to 0$). When skill mean $\bar{Z}$ increases, data clarity x^* remains at ∞, the fraction λ^* of skilled investors increases, data price q^* increases, price informativeness $\frac{1}{Var(\tilde{\theta}|\tilde{p})}$ increases, the cost of capital $E(\tilde{\theta}-\tilde{p})$ decreases, return volatility $\sigma(\tilde{\theta}-\tilde{p})$ decreases, and skilled investors' performance decreases. That is, for sufficiently low τ_η, $\frac{\partial \lambda^*}{\partial \bar{Z}} > 0$, $\frac{\partial q^*}{\partial \bar{Z}} > 0$, $\frac{\partial}{\partial \bar{Z}}\frac{1}{Var(\tilde{\theta}|\tilde{p})} > 0$, $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial \bar{Z}} < 0$, $\frac{\partial \sigma(\tilde{\theta}-\tilde{p})}{\partial \bar{Z}} < 0$, and $\frac{\partial Performance}{\partial \bar{Z}} < 0$.*

Last, unlike the results in Section 4.2.3, when the data seller owns very inaccurate data (low τ_η), as skill volatility Δ increases, price informativeness unambiguously decreases due to the presence of fewer skilled investors, and both the cost of capital and return volatility increase accordingly. This is because now the seller sells data "as is" so market quality is affected only by the change in the mass of skilled investors. Further, skilled investors' performance exhibits different patterns depending on the accuracy level τ_η: when τ_η is high, performance decreases with Δ; but when τ_η is low, performance increases with Δ.

Proposition OA3 (Implications of skill volatility with low τ_η) A*ssume that the seller owns very inaccurate data (i.e., $\tau_\eta \to 0$). When skill volatility Δ increases, data clarity x^**

remains at ∞, the fraction λ^* of skilled investors increases, data price q^* increases, price informativeness $\frac{1}{Var(\tilde{\theta}|\tilde{p})}$ decreases, the cost of capital $E(\tilde{\theta}-\tilde{p})$ increases, return volatility $\sigma(\tilde{\theta}-\tilde{p})$ increases, and skilled investors' performance increases.

Proofs of Proposition OA1-OA3

1. Proof of Proposition OA1 *1.1. The case without skill-acquisition uncertainty.*

When τ_η is sufficiently low and $Z_L = Z_H$, by Proposition 3, we have $x^* = +\infty$ and $\alpha^* = \frac{\lambda^* \tau_\eta}{\gamma}$. Now, inserting $x = +\infty$ and $\alpha = \frac{\lambda \tau_\eta}{\gamma}$ into equation (19) generates

$$\exp\left(\frac{\gamma C(\lambda)}{1-\beta}\right)\sqrt{1-\frac{\gamma^2\tau_\eta}{(\tau_\eta+\tau_\theta)\left(\gamma^2+\lambda^2\tau_u\tau_\eta\right)}}=1. \tag{OA1}$$

The LHS of equation (OA1) is increasing in λ and increasing in c. Thus, when c increases, λ should decrease to maintain the equation: $\frac{\partial \lambda^*}{\partial c} < 0$. Given $\alpha^* = \frac{\lambda^* \tau_\eta}{\gamma}$ and equation (12), we know $\frac{\partial \alpha^*}{\partial c} < 0$, and $\frac{\partial}{\partial c}\frac{1}{Var(\tilde{\theta}|\tilde{p})} < 0$.

With $x^* = +\infty$, the information gain G can be simplified as $G = \frac{1}{2\gamma}\log\left[\frac{\tau_\eta+\alpha^2\tau_u}{[\tau_\eta\tau_\theta+\alpha^2\tau_u(\tau_\eta+\tau_\theta)](\tau_\eta+\tau_\theta)}\right]$. Taking derivative of G with respect to α yields $\frac{\partial G}{\partial \alpha} = -\frac{\alpha\tau_u\tau_\eta}{\gamma(\tau_\eta+\alpha^2\tau_u)[\tau_\eta\tau_\theta+\alpha^2\tau_u(\tau_\eta+\tau_\theta)]} < 0$. So, $\frac{\partial G^*}{\partial c} = \frac{\partial G^*}{\partial \alpha}\frac{\partial \alpha}{\partial c} > 0$. Thus, $\frac{\partial q^*}{\partial c} > 0$ by equation (16).

Next, we examine market-quality variables. Inserting $x = +\infty$ and $\alpha = \frac{\lambda^* \tau_\eta}{\gamma}$ into equation (C12), the cost of capital $E(\tilde{\theta}-\tilde{p})$ can be simplified as follows:

$$E(\tilde{\theta}-\tilde{p}) = \frac{\gamma^3+\gamma\lambda^2\tau_\eta\tau_u}{\tau_\theta\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)+\lambda\tau_\eta\left(\gamma^2+\lambda\tau_\eta\tau_u\right)}, \tag{OA2}$$

which is a function of λ only. Thus, $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c} = \frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial \lambda}\frac{\partial \lambda}{\partial c} = -\frac{\gamma^3\tau_\eta\left((2-\lambda)\lambda\tau_\eta\tau_u+\gamma^2\right)}{\left(\tau_\theta\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)+\lambda\tau_\eta(\gamma^2+\lambda\tau_\eta\tau_u)\right)^2}\frac{\partial \lambda}{\partial c} > 0$.

With $x = +\infty$ and $\alpha = \frac{\lambda\tau_\eta}{\gamma}$, return variance $Var(\tilde{\theta}-\tilde{p})$ can be simplified to

$$Var(\tilde{\theta}-\tilde{p}) = \frac{\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)\left(\tau_\theta\tau_u\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)+\left(\gamma^2+\lambda\tau_\eta\tau_u\right)^2\right)}{\tau_u\left(\tau_\theta\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)+\lambda\tau_\eta\left(\gamma^2+\lambda\tau_\eta\tau_u\right)\right)^2},$$

which is a function of λ only. Taking derivative of it with respect to λ yields

$$\frac{\partial Var(\tilde{\theta}-\tilde{p})}{\partial \lambda} = -\frac{2\gamma^2\tau_\eta \left(\begin{array}{c} \gamma^6 + \gamma^4\lambda\tau_u\left(3\tau_\eta + \tau_\theta\right) \\ +\gamma^2\lambda\tau_\eta\tau_u^2\left(3\lambda\tau_\eta + \left(\lambda^2+1\right)\tau_\theta\right) + \lambda^3\tau_\eta^2\tau_u^3\left(\tau_\eta+\tau_\theta\right) \end{array} \right)}{\tau_u\left(\gamma^2\tau_\theta + \lambda\tau_\eta\left(\gamma^2+\lambda\tau_\theta\tau_u\right) + \lambda^2\tau_\eta^2\tau_u\right)^3} < 0.$$

Thus, $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial c} = \frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\lambda}\frac{\partial\lambda}{\partial c} > 0$.

Finally, for sufficiently low τ_η, performance (21) can be simplified as $Performance = \frac{1}{2\gamma}\log\left(1+\frac{\tau_\eta^2}{\tau_\eta\tau_\theta+\alpha^2\tau_u(\tau_\eta+\tau_\theta)}\right)$, which is a function of α only. Since we have shown that $\frac{\partial\alpha^*}{\partial c} < 0$, it must be the case that $\frac{\partial Performance}{\partial c} > 0$.

1.2. The case with skill-acquisition uncertainty.

When τ_η is sufficiently low, based on the above proof, $x^* = +\infty$ and $\lambda^*\tau_\eta = \alpha^*\gamma$. Now, inserting $x^* = +\infty$ and $\alpha^* = \frac{\lambda^*\tau_\eta}{\gamma}$ into equation (19) leads to (OA1). Note that with $x^* = +\infty$, the skill levels Z_H and Z_L play no role and thus, we obtain the same simplification of equation (19) as in the case with $x^* = +\infty$, $\alpha^* = \frac{\lambda^*\tau_\eta}{\gamma}$, and $Z_L = Z_H$. So, similar to the above proof, with $x^* = +\infty$ and $\alpha^* = \frac{\lambda^*\tau_\eta}{\gamma}$, we can show $\frac{\partial\lambda^*}{\partial c} < 0$, $\frac{\partial\alpha^*}{\partial c} < 0$, $\frac{\partial}{\partial c}\frac{1}{Var\left(\tilde{\theta}|\tilde{p}\right)} < 0$, $\frac{\partial q^*}{\partial c} > 0$, $\frac{\partial\pi^*}{\partial c} < 0$, $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c} > 0$, and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial c} > 0$.

For performance, with $x^* = +\infty$, performance (21) can be simplified as *Performance* $= \frac{1}{2\gamma}\log\left(1+\frac{\tau_\eta^2}{\tau_\eta\tau_\theta+\alpha^2\tau_u(\tau_\eta+\tau_\theta)}\right)$, which is a function of α only. Taking derivative of it with respect to α yields $\frac{\partial Performance}{\partial\alpha} = -\frac{\alpha\tau_\eta^2\tau_u}{\gamma(\tau_\eta+\alpha^2\tau_u)(\tau_\eta\tau_\theta+\alpha^2\tau_u(\tau_\eta+\tau_\theta))} < 0$. Therefore, for sufficiently low τ_η, $\frac{\partial Performance}{\partial c} = \frac{\partial Performance}{\partial\alpha}\frac{\partial\alpha}{\partial c} > 0$. QED.

2. Proof of Proposition OA2 By the proof of Proposition OA1, replacing $Z_H = \bar{Z}+\Delta$ and $Z_L = \bar{Z}-\Delta$ in (OA1) yields

$$\frac{1}{2}\exp\left(\frac{\gamma C(\lambda)}{1-\beta}\right)\left[\begin{array}{c} \sqrt{1-\frac{\gamma^4 x\tau_\eta(\bar{Z}-\Delta)}{\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)\left(\gamma^2\tau_\eta\left(\tau_\theta+x(\bar{Z}-\Delta)\right)+\gamma^2x\tau_\theta(\bar{Z}-\Delta)+\lambda^2\tau_\eta^2\tau_\theta\tau_u+\lambda^2\tau_\eta^3\tau_u\right)}} \\ +\sqrt{1-\frac{\gamma^4 x\tau_\eta(\bar{Z}+\Delta)}{\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)\left(\gamma^2\tau_\eta\left(\tau_\theta+x(\bar{Z}+\Delta)\right)+\gamma^2x\tau_\theta(\bar{Z}+\Delta)+\lambda^2\tau_\eta^2\tau_\theta\tau_u+\lambda^2\tau_\eta^3\tau_u\right)}} \end{array}\right] = 1. \tag{OA3}$$

The LHS of equation(OA3) is increasing in λ. Taking derivative of the LHS of equation (OA3) with respect to $\bar{Z}$ yields $-\frac{\gamma^4 x\tau_\eta^2 e^{\frac{\gamma C}{1-\beta}}\left(\gamma^2\tau_\theta+\lambda^2\tau_\eta\tau_u(\tau_\eta+\tau_\theta)\right)(H(Z_H)+H(Z_L))}{4\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)} < 0$, where

$$H(Z) = \frac{1}{\sqrt{\frac{(\gamma^2\tau_\theta+\lambda^2\tau_\eta\tau_u(\tau_\eta+\tau_\theta))(\gamma^2 xZ+\tau_\eta(\gamma^2+\lambda^2\tau_\eta\tau_u))}{(\gamma^2+\lambda^2\tau_\eta\tau_u)(\gamma^2 xZ(\tau_\eta+\tau_\theta)+\tau_\eta(\gamma^2\tau_\theta+\lambda^2\tau_\eta\tau_u(\tau_\eta+\tau_\theta)))}}\left(\gamma^2 xZ\left(\tau_\eta+\tau_\theta\right)+\tau_\eta\left(\gamma^2\tau_\theta+\lambda^2\tau_\eta\tau_u\left(\tau_\eta+\tau_\theta\right)\right)\right)^2}.$$

The LHS of equation (OA3) is decreasing in $\bar{Z}$. Thus, when $\bar{Z}$ increases, λ should increase to maintain the equation. Hence, $\frac{\partial\lambda^*}{\partial\bar{Z}} > 0$. Given $\alpha^* = \frac{\lambda^*\tau_\eta}{\gamma}$ and equation (12), $\frac{\partial\alpha^*}{\partial\bar{Z}} > 0$, and $\frac{\partial}{\partial\bar{Z}}\frac{1}{Var\left(\tilde{\theta}|\tilde{p}\right)} > 0$. The data price is $q = \beta G = \frac{\beta}{1-\beta}C(\lambda)$, where the second equation follows from equation (19). Thus, $\frac{\partial q^*}{\partial\bar{Z}} = \frac{\beta}{1-\beta}\frac{\partial C}{\partial\lambda}\frac{\partial\lambda}{\partial\bar{Z}} > 0$. Also, the seller's profit is $\pi = \frac{\beta\lambda}{1-\beta}C(\lambda)$ and $\frac{\partial\pi^*}{\partial\bar{Z}} = \frac{\beta}{1-\beta}\left(\frac{\partial\lambda}{\partial\bar{Z}}C(\lambda)+\lambda\frac{\partial C}{\partial\lambda}\frac{\partial\lambda}{\partial\bar{Z}}\right) > 0$. Finally, similar to the proof of $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c} > 0$ and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial c} > 0$ for sufficiently low τ_η in the proof of Proposition 4, with $x^* = +\infty$ and $\alpha^* = \frac{\lambda^*\tau_\eta}{\gamma}$, we can show $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial\bar{Z}} = \frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial\lambda}\frac{\partial\lambda}{\partial\bar{Z}} < 0$ and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\bar{Z}} = \frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\lambda}\frac{\partial\lambda}{\partial\bar{Z}} < 0$.

Finally, similar to the proof in part (2) of Proposition OA1, we can show that $\frac{\partial Performance}{\partial\bar{Z}} = \frac{\partial Performance}{\partial\alpha}\frac{\partial\alpha}{\partial\bar{Z}} < 0$. QED.

3. Proof of Proposition OA3 The proof is based on the proof of Proposition OA2. Taking derivative of the LHS of equation (OA3) with respect to Δ yields $\frac{\gamma^4 x\tau_\eta^2 e^{\frac{\gamma C}{1-\beta}}\left(\gamma^2\tau_\theta+\lambda^2\tau_\eta\tau_u(\tau_\eta+\tau_\theta)\right)(H(Z_L)-H(Z_H))}{4\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)} > 0$. The inequality follows because

$$\frac{\partial H}{\partial Z} = -\frac{\gamma^2 x\left(4\gamma^2 xZ(\tau_\eta+\tau_\theta)+\tau_\eta\left(4\gamma^2\tau_\theta+\tau_\eta\left(3\gamma^2+4\lambda^2\tau_\theta\tau_u\right)+4\lambda^2\tau_\eta^2\tau_u\right)\right)}{\begin{array}{c}2\left(\gamma^2 xZ+\tau_\eta\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)\right)\sqrt{\frac{\left(\gamma^2\tau_\theta+\lambda^2\tau_\eta\tau_u(\tau_\eta+\tau_\theta)\right)\left(\gamma^2 xZ+\tau_\eta\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)\right)}{\left(\gamma^2+\lambda^2\tau_\eta\tau_u\right)\left(\gamma^2 xZ(\tau_\eta+\tau_\theta)+\tau_\eta\left(\gamma^2\tau_\theta+\lambda^2\tau_\eta\tau_u(\tau_\eta+\tau_\theta)\right)\right)}} \\ \times\left(\gamma^2 xZ\left(\tau_\eta+\tau_\theta\right)+\tau_\eta\left(\gamma^2\tau_\theta+\lambda^2\tau_\eta\tau_u\left(\tau_\eta+\tau_\theta\right)\right)\right)^3\end{array}} < 0,$$

and thus $H(Z_L) > H(Z_H)$. The LHS of equation (OA3) is increasing in Δ and increasing in λ. So, when Δ increases, λ should decrease to maintain the equation. Hence, $\frac{\partial\lambda^*}{\partial\Delta} < 0$. Given $\alpha^* = \frac{\lambda^*\tau_\eta}{\gamma}$ and equation (12), $\frac{\partial\alpha^*}{\partial\Delta} < 0$, and $\frac{\partial}{\partial\Delta}\frac{1}{Var\left(\tilde{\theta}|\tilde{p}\right)} < 0$. The data price is $q = \beta G = \frac{\beta}{1-\beta}C(\lambda)$, where the second equation follows from equation (19). Thus, $\frac{\partial q^*}{\partial\Delta} = \frac{\beta}{1-\beta}\frac{\partial C}{\partial\lambda}\frac{\partial\lambda}{\partial\Delta} < 0$. Next, similar to the proof of $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial c} > 0$ and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial c} > 0$ for sufficiently low τ_η in the proof of Proposition 4, with $x^* \to +\infty$ and $\alpha^* = \frac{\lambda^*\tau_\eta}{\gamma}$, we can show $\frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial\Delta} = \frac{\partial E(\tilde{\theta}-\tilde{p})}{\partial\lambda}\frac{\partial\lambda}{\partial\Delta} > 0$ and $\frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\Delta} = \frac{\partial\sigma(\tilde{\theta}-\tilde{p})}{\partial\lambda}\frac{\partial\lambda}{\partial\Delta} > 0$.

Similar to the proof in part (2) of Proposition OA1, we can show that $\frac{\partial Performance}{\partial\Delta} = \frac{\partial Performance}{\partial\alpha}\frac{\partial\alpha}{\partial\Delta} > 0$. QED.