Rational Inattention and Social Value of Standardization*

Peicong (Keri) Hu Sang Wu†

October 24, 2025

Abstract

This paper proposes a theoretical framework to study the social value of various accounting method adoption choices in the presence of investor rational inattention, generalizing the implications of the trade-off between the precision effect and the attention effect on accounting standardization. While the locally preferred accounting method of each firm measures their underlying transactions more precisely, the adoption of a common method may improve investment efficiency and increase firm value by concentrating investors' limited attention on a single method, thereby facilitating the processing of financial reports by investors. We find that the socially optimal number of accounting methods changes non-monotonically in investor attention capacity. Moreover, in stark contrast to the adoption of distinct local methods, the adoption of a common method does not necessarily emerge as an equilibrium even when it is socially optimal. In particular, mandatory standardization is socially beneficial when investor attention capacity is moderately large.

JEL classification: D83, D91, G14, M41, M48

Keywords: rational inattention, standardization, network effect, reporting quality, endogenous information acquisition

*We thank Tim Baldenius, Jeremy Bertomeu, Anne Beyer, Kim Sau Chung, Wouter Dessein, Paul Fischer, Henry Friedman, Pingyang Gao, Jonathan Glover, Ilan Guttman, Zeqiong Huang, Xu Jiang, Bjorn Jorgensen, Chandra Kanodia, Benson Tsz Kin Leung, Pierre Jinghong Liang, Mario Milone, Lin Nan (discussant), Suil Pae, Joel Sobel, Phillip Stocken, Shyam Sunder, Alfred Wagenhofer, Liyan Yang, Ming Yang, Erina Ytsma, Gaoqing Zhang, seminar participants at Carnegie Mellon University and Yale University, and conference participants at 2025 Columbia Accounting Theory Conference, 2025 Accounting and Economics Society Annual Conference, and 2025 Minnesota-Chicago Accounting Theory Conference for helpful discussions and comments.

†Corresponding author.

Hu is from the University of Hong Kong (kerihu@hku.hk). Wu is from Columbia University (sw3724@columbia.edu).

1. Introduction

The standardization of accounting methods is a central concern for standard setters and accounting researchers (Merino and Coe, 1978; Beaver, 1998; Dye and Sunder, 2001). In the U.S., reducing diverse practices is the most frequently cited reason by the Financial Accounting Standards Board (FASB) to take on a project (Jiang, Wang and Wangerin, 2018). Internationally, there has been extensive adoption of International Financial Reporting Standards (IFRS) since the 1990s as well as calls for convergence between local GAAP and IFRS to harmonize accounting standards across countries (De George, Li and Shivakumar, 2016). While the adoption of a common accounting method may facilitate investors' processing of financial information by enhancing comparability (see, e.g., Leuz and Wysocki, 2016; Gao, Jiang and Zhang, 2019), there are significant compliance costs associated with standardization (Hail, Leuz and Wysocki, 2010; De George, Ferguson and Spear, 2013). Moreover, uniformity could undermine reporting precision by disregarding the unique characteristics of individual firms (Leftwich, 1980; Dye and Sridhar, 2008). Thus, it is not clear whether the adoption of a common method is desirable from a social perspective. Furthermore, given that firms may voluntarily adopt a common method (Barth, Landsman and Lang, 2008), why would regulators need to mandate standardization in the first place?

In this paper, we shed light on the above issues by studying heterogeneous firms' choices of accounting methods in a capital market setting where investors have limited attention to learn about accounting methods. Viewing the output of an accounting information system as a noisy signal about the firm's productivity, each firm has a distinct locally preferred method which maximizes its reporting precision. More generally, the noisiness of the measurement process is determined by the distance between the characteristics of the underlying firm and its chosen accounting method.[1] Upon observing the accounting method

[1] Specifically, even the same transaction may have different value implications on different firms due to their distinct operational models. For example, consider two airline companies, United Airlines and Southwest Airlines. Southwest Airlines primarily operates domestic flights, whereas United Airlines has more international flights. Thus, it might be more value relevant for Southwest to use accelerated depreciation and for United to adopt straight-line depreciation. In a similar vein, although both companies operate within the

choice made by each firm, the investors decide how to allocate their limited attention to learn about each accounting method and reduce post-reporting uncertainty about the firms. Importantly, we consider situations where the investors learn about accounting methods prior to the release of financial reports. The learning is at the method level and independent of the adopting firm.[2] The notion of learning about accounting methods aligns with the conceptual framework established by the Financial Accounting Standards Board (FASB, 2018, QC22) and the International Accounting Standards Board (IASB, 2018, 2.26) in that consistency, achieved through the adoption of a common method, enhances comparability by concentrating investors' attention on a single method, thereby facilitating their processing of financial reports.[3] We model the impact of attention as a reduction in the entropy-based measure of uncertainty, which exhibits decreasing returns to learning.

To focus on the externalities of firms' method choices in the presence of investor rational inattention, we consider a parsimonious model in which the firm's incentive is perfectly aligned with the investors' in terms of maximizing information in the capital market. Specifically, the investors' post-learning uncertainty about the firm's productivity is lower when (i) the firm chooses an accounting method that is closer to its characteristics and (ii) the investors devote more attention to the chosen accounting method. We refer to the former as the *precision effect* and the latter as the *attention effect*. In an economy with a single representative firm, investors optimally allocate all the attention to any chosen method. Thus, the firm will optimally adopt its locally preferred method to maximize reporting precision. However, in an economy with heterogeneous firms, inefficiency may arise from a

oil and gas industry, ExxonMobil employs the last-in first-out (LIFO) method for inventory costing, while BP utilizes the first-in first-out (FIFO) approach.

[2]There is substantial literature emphasizing the critical role of understanding accounting methods in interpreting financial reports, see, e.g., Dyckman (1964), Jensen (1966), Watts and Zimmerman (1986), Maines (1995), and Libby, Bloomfield and Nelson (2002).

[3]The FASB (2018, QC22) and the IASB (2018, 2.26) define consistency in their conceptual framework as "the use of of the same methods for the same items, either from period to period within a reporting entity or in a single period across entities." In our context, consistency can be used interchangeably with uniformity, which refers to the use of the same accounting methods across firms (Flynn, 1965).

firm's deviation to its own local method, as it inadvertently distracts attention away from the common method.

When acquiring information about the entire universe of accounting methods, investors optimally allocate their limited attention to equate the marginal return of attention across the methods. Intuitively, the investors will not waste attention on methods that are not adopted by any firm, but instead devote attention to the more important methods in the sense of being adopted by more firms and/or generating higher post-reporting uncertainty. Perhaps surprisingly, some methods may not necessarily receive any attention even if they are adopted by some firms. Indeed, when attention is scarce, the investors will only learn about the most important method despite decreasing returns to attention. This happens when the marginal return of learning more about the most important method exceeds that of all the other methods even if it receives all the attention. As their attention capacity increases, the investors are induced to spread their attention across more methods due to decreasing returns to learning.

The socially optimal standard trades off the precision effect and the attention effect of method adoption. While the magnitude of the precision effect is governed by the characteristics of the underlying accounting method, the attention effect is endogenously determined by investor rational inattention. Ideally, the total post-learning uncertainty across all firms is minimized by inducing high reporting precision as well as concentrating attention on fewer methods for more in-depth learning. The former drives the socially optimal standard towards allowing for distinct local methods, whereas the latter makes the adoption of a common method more desirable. The attention effect dominates only for intermediate values of attention capacity, rendering the adoption of a common method socially optimal. When attention is sufficiently scarce, the gain from concentrating attention on fewer methods is always dominated by the loss from lower reporting precision due to the adoption of non-local methods. Conversely, when attention is sufficiently abundant, the marginal value of attention is limited due to decreasing returns to learning, making it optimal for some firm(s) to adopt their local methods from a social perspective. In short, the trade-off between the precision effect and the attention effect leads the socially optimal number of methods to change non-monotonically, first decreasing and and then increasing, in investor attention capacity.

While the adoption of distinct local methods can always be sustained as an equilibrium whenever it is socially optimal, the adoption of a common method does not necessarily emerge as an equilibrium, notwithstanding its optimality. In particular, firms have an incentive to coordinate with each other in adopting a common method only when investor attention capacity is moderately small. Intuitively, the marginal value of attention devoted to a method is increasing in the number of firms adopting this method. As a method is adopted by more firms, the investors will optimally allocate more attention to this particular method, which in turn makes it more attractive to be adopted by the other firms. The complementarity between attention and method adoption is a manifestation of the *network effect* (see, e.g., Katz and Shapiro, 1985, 1986). On the other hand, the amount of attention that an individual firm attracts by deviating to its local method is limited under relatively scarce attention. The adoption of the socially optimal common method arises as an equilibrium outcome because it is in firms' best interest to coordinate on a common method so that the investors will concentrate all of their limited attention on the common method. Essentially, the investors' rational allocation of limited attention mitigates the externalities that one firm can potentially impose on the other firms through a deviation to a different method, which may distract attention away from the common method. Rational inattention suppresses firms' deviation incentives because of their fear of being the minority, thereby coordinating firms to "speak the same language."

A special case arises when there are only two firms, where mandatory standardization on a common method becomes unnecessary for any levels of attention capacity. This result is in stark contrast to the extant literature that advocates the mandate of common standards using a setting of two firms (see, e.g., Gao et al., 2019; Corona, Huang and Hwang, 2024).

The reason is that, in the presence of investor rational inattention, a firm necessarily attracts less than half of the attention when deviating to its local method due to the increased reporting precision. Indeed, the other firm's adoption of a non-local method makes the common method relatively more important to learn about. Thus, whenever the adoption of a common method results in less post-learning uncertainty than the adoption of distinct methods, where each firm receives exactly half of the attention, sticking with the common method strictly dominates deviation to the local method. On the other extreme where the number of firms is infinitely large, each firm receives minimal, if any, attention by deviating to its local method. Hence, there is again no need to mandate the adoption of the socially optimal common method.

Nevertheless, in a larger economy with finitely many firms, individual firms may be incentivized to inefficiently deviate to their own local methods as attention becomes more abundant. From a social perspective, when the total attention is moderately but not sufficiently large, a significant fraction of the measurement noise can be mitigated through learning. Hence, the social benefit of concentrating attention on a single method still renders the adoption of a common method desirable. However, the investors will optimally allocate some attention to the minority method due to decreasing returns to learning, limiting the firm's loss in attention of deviating from the common method. Since the firm does not fully internalize the externalities it imposes on the other firms arising from the diluted attention on the common method, the positive precision effect of deviating to the local method dominates the negative attention effect from the individual firm's perspective. Hence, mandatory adoption of a common method is socially beneficial under moderately large attention.

Last but not least, there exist some intermediate regions of investor attention capacity in which the adoption of neither a common method nor distinct methods is socially optimal. In these cases, symmetry among firms does not imply symmetry in the socially optimal method choice. That is, the socially optimal standard in those regions does not feature the division of firms into several equal-sized groups, each of which adopts a common method. Instead, due to the complementarity between attention and method adoption, it is more efficient to have only one large group adopt a common method, while the other firms each adopt their own local methods. Furthermore, whenever a method is adopted by multiple firms, the socially optimal common method is the one that maintains equal distances to all adopting firms, thereby effectively minimizing the total post-reporting uncertainty among them. In other words, the socially optimal common method equally accounts for the idiosyncrasies of all constituent firms.

Our model generates several empirical and policy implications. By studying the implications of investors' rational allocation of limited attention on firms' method choices, we provide an alternative explanation for the existence of only a few prevailing accounting methods even when "one size does not fit all." In addition, our findings highlight that mandatory standardization may not always be necessary. The adoption of distinct methods is socially optimal in terms of maximizing information in the market when the investors' information processing capacity is sufficiently small or sufficiently large. Even when the adoption of a common method is socially optimal, firms will voluntarily coordinate on the socially optimal common method under relatively scarce attention. A need to mandate the adoption of a common method arises only when attention is relatively abundant, implying that mandatory standardization may become more important as information processing capacity increases. We view this insight as consistent with the recent adoption of ASC 606, a new accounting rule that standardizes revenue recognition across different industries (Hinson, Pündrich and Zakota, 2024), despite the advancement in information technology and language processing tools. Meanwhile, firms may be inclined to make "pro forma" (i.e. non-GAAP) disclosures, which are found to be more informative in predicting future performance (Curtis, McVay and Whipple, 2014; Black, Christensen, Ciesielski and Whipple, 2018; Campbell, Gee and Wiebe, 2022). Furthermore, the regime shift in the socially optimal standard from allowing for diverse practices to restricting to a common method as attention increases also seems supported by the disclosure dynamics of firms' environmental, social, and governance (ESG) performances. In particular, ESG reporting has been transitioning in recent years from a discretionary to a more standardized approach guided by the framework proposed by the Sustainability Accounting Standards Board (SASB) (Khan, Serafeim and Yoon, 2016; SEC, 2024). More broadly, viewing accounting standards as a set of accounting methods for various transactions, our model also speaks to the equilibrium and the optimal standard adoption choices by each firm and harmonization of accounting standards across different countries.

2. Related Literature

This paper studies the value of standardization by investigating the potential discrepancy between the socially optimal and equilibrium choices of accounting methods through the lens of rational inattention. It is related to the large literature on the "economic consequence" of accounting standards (see, e.g., Holthausen and Leftwich, 1983; Watts and Zimmerman, 1990; Fields, Lys and Vincent, 2001) and the real effects of accounting measurement policies (Kanodia, 2007; Kanodia and Sapra, 2016). In particular, we contribute to the analytical work that examines the consequence of accounting uniformity and comparability by endogenizing the trade-off between the precision effect and the network effect. Barth, Clinch and Shibano (1999) are the first to examine the implication of international accounting harmonization on investors' cost of processing financial information. They show that harmonization may have deleterious effects on price informativeness as it crowds out private information. Dye and Sridhar (2008) find that investors and firms often have "orthogonal" attitudes towards uniformity. Subsequent studies examine information externalities of accounting comparability by measuring comparability as the correlation in measurement processes (Fang, Iselin and Zhang, 2022; Jiang, Tang and Zhang, 2023; Corona et al., 2024; Wu and Zheng, 2025). Most closely related to our paper, Ray (2018) and Gao et al. (2019) study the consequences of the adoption of common standards relative to two standards in resource allocation and capital market settings, respectively, where both the precision effect and the network effect are fixed exogenously. We extend this line of research by considering an arbitrary number firms and the entire universe of accounting methods without imposing any restriction on the method choices. In our setting, comparability is achieved through the adoption of a common method, i.e., uniformity, by concentrating investors' attention on a single method.

Our paper speaks to the debate over mandatory versus voluntary disclosures (see Dye, 1990; Bertomeu and Magee, 2015; Bertomeu, Vaysman and Xue, 2021). We reconcile existing findings on the voluntary adoption and the consequences of a common standard (Barth et al., 2008; De George et al., 2016) as well as the extensive accounting reporting discretion and large variety of accounting procedures (Beatty and Weber, 2006; Bushman and Williams, 2012). We also complement the analytical literature advocating the adoption of common standards (see, e.g., Ray, 2018; Gao et al., 2019; Corona et al., 2024) by investigating the conditions under which (i) the adoption of a common accounting method is socially optimal and (ii) whether the adoption of the socially optimal common method emerges as an equilibrium, shedding light on when it is socially beneficial to mandate a common method. Given that implementing a mandatory regulation involves high cost, our paper has the same spirit as the proposal of Beaver (1977) that we should provide explicit cost-benefit analyses before expanding regulations. By studying the socially optimal and equilibrium choices of accounting methods, we deepen our understanding about firms' incentives and, hence, the appropriate scope of government and policy resolutions (Bertomeu and Cheynel, 2013).

Our study broadly combines two threads of literature, rational inattention and strategic interaction among capital market participants (see Blankespoor, deHaan and Marinovic, 2020, for a review on disclosure processing costs). To the best of our knowledge, we are the first to deploy rational inattention in an accounting method choice setting. In our model, endogenous learning at the method rather than the firm level creates incentives for firms to coordinate on a common method. In contrast, Fishman and Hagerty (1989) show that firms may voluntarily disclose an excessive amount of information to compete for attention under the premise that investors can only observe one disclosure. Jiang and Yang (2017), Lu (2022), and Bertomeu, Hu and Liu (2023) examine the implications of investors' limited attention on financial reporting by considering a single representative firm. Meanwhile, the accounting literature has devoted considerable attention to coordination in various settings (Gao, 2008; Arya and Mittendorf, 2016; Chen, Lewis, Schipper and Zhang, 2017; Liang and Zhang, 2019; Jiang et al., 2023; Zhang and Zheng, 2024; Corona and Wu, 2024; Wu, 2024), but the effect of receivers' limited attention on senders' coordination decisions remains unexplored.

3. The Model

We consider a risk-neutral economy with $n \geq 2$ firms indexed by $i \in \mathcal{I} = \{1, \ldots, n\}$ and a continuum of investors $a \in [0, 1]$. Each firm is owned and managed by an entrepreneur. The investors buy the firms from the entrepreneurs after the release of financial reports, but before the realization of cash flows.

Firm i's cash flow $\tilde{v}_i$ is determined jointly by (i) its productivity $\tilde{\omega}_i$ and (ii) the aggregate investment I_i made by all investors in the economy (see, e.g., Gao and Jiang, 2020). More specifically, given the investment $I_i \equiv \int_0^1 I_{ia} \mathrm{d}a$,

$$\tilde{v}_i = 2\tilde{\omega}_i I_i - I_i^2. \tag{1}$$

The firm's productivity is unobservable but measured by an accounting system. Thus, we also call $\tilde{\omega}_i$ the economic substance of a representative transaction or simply the state (henceforth, the firm's *fundamental*).

The accounting system regulated by method j produces a noisy signal, $\tilde{r}_{ij}$, about the firm's fundamental. Since signal structures and belief distributions are equivalent formalisms (see, e.g., Kamenica and Gentzkow, 2011), the posterior variance induced by the accounting report (henceforth, *post-reporting* uncertainty), $\mathrm{Var}(\tilde{\omega}_i | r_{ij}) = \sigma_{ij}^2$, is a summary statistic of the outcome of the measurement process. The noisiness of the measurement process is determined by the Euclidean distance between the characteristics of firm i and its chosen accounting method j. Specifically, let $\boldsymbol{x_i} = (x_{i,k})_k \in \mathbb{R}^n$ and $\boldsymbol{y_j} = (y_{j,k})_k \in \mathbb{R}^n$ represent the characteristics of firm $i \in \mathcal{I}$ and the measurement process prescribed by method $j \in \mathcal{J}$, respectively. The post-reporting variance is then represented as

$$\sigma_{ij}^2 = \sigma_0^2 + \sum_{k=1}^{n} (x_{i,k} - y_{j,k})^2, \tag{2}$$

where $\sigma_0^2 := \tau_0^{-1}$ with $\tau_0 > 0$ being the highest reporting precision induced by the firm's local method i such that $y_{i,k} = x_{i,k}$ for any $k \in \{1, \ldots, n\}$. The further the reporting method is from the firm characteristics, the noisier the reporting process is. Moreover, given a set of heterogeneous firms, without loss of generality, we orthogonalize and normalize each dimension of firm characteristics that are relevant to the value implications of the underlying transaction, such that $x_{i,i} = 1$ and $x_{i,k} = 0$ for any $k \neq i$. Indeed, the trade-off between the precision effect and the attention effect exists only if the locally preferred method differs across firms; otherwise, the socially optimal method is unambiguously the common local method of all firms.

Upon observing the accounting method choice made by each firm, the investors decide how to allocate their limited attention, K, to learn about each accounting method. Specifically, by learning about accounting method j, the investors acquire an additional signal, $\tilde{s}_j$, about the method. Note that the learning process is method-specific and independent of the adopting firm. Put differently, only the systematic component of the measurement error arising from the adopted method is learnable, while the idiosyncratic error arising randomly in individual transactions is unlearnable even with expertise in the method. Together with the accounting report, knowledge about the adopted method induces a posterior variance $\mathrm{Var}(\tilde{\omega}_i | r_{ij}, s_j) = \hat{\sigma}_{ij}^2$ about firm i's fundamental (henceforth, *post-learning uncertainty*).

We assume only $\chi \in (0,1)$ fraction of the post-reporting uncertainty can be reduced by learning about the method (see, e.g., Van Nieuwerburgh and Veldkamp, 2009, 2010; De Marco, Macchiavelli and Valchev, 2022). Thus, for firm $i \in \mathcal{I}_j$, where $\mathcal{I}_j$ denotes the set of firms that adopt method j, knowledge about method j reduces the investors' post-reporting uncertainty arising from the systematic component of the measurement error from $\chi\sigma_{ij}^2$ to $\hat{\sigma}_{ij}^2 - (1-\chi)\sigma_{ij}^2$, whereas the uncertainty arising from the unlearnable idiosyncratic component of the measurement error remains at $(1-\chi)\sigma_{ij}^2$. Following Sims (2003), we model the amount of information transmitted as the reduction in entropy, the standard measure of information in information theory, achieved by conditioning on that additional information. Formally, suppose all random variables are normally distributed. Then the attention allocated to method j is represented by $\delta_j \equiv \log(\chi\sigma_{ij}^2) - \log(\hat{\sigma}_{ij}^2 - (1-\chi)\sigma_{ij}^2)$. Equivalently, the posterior uncertainty about firm i's fundamental can be expressed as

$$\hat{\sigma}_{ij}^2 = \left(1 - \chi + \chi e^{-\delta_j}\right)\sigma_{ij}^2 \tag{3}$$

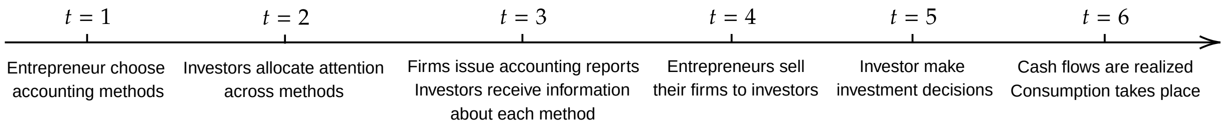
subject to the attention capacity constraint and the no-forgetting constraint:

$$\sum_{j\in\mathcal{J}} \delta_j \leq K, \tag{4}$$

$$\delta_j \geq 0. \tag{5}$$

The timeline of the model consists of six dates, as summarized in Figure 1. At date $t = 1$, the entrepreneur of each firm $i \in \mathcal{I}$ simultaneously chooses an accounting method $j \in \mathcal{J}$. At $t = 2$, observing the firms' method choices $\{\mathcal{I}_j\}_{j\in\mathcal{J}}$, the investors make attention allocation decisions, $\{\delta_j\}_{j\in\mathcal{J}}$, across methods, subject to the attention capacity constraint (4) and the no-forgetting constraint (5). At $t = 3$, each entrepreneur i prepares and issues a financial report, $\tilde{r}_{ij}$, about the its productivity, $\tilde{\omega}_i$, according to the chosen method j, and the investors receive additional information, $\{\tilde{s}_j\}_{j\in\mathcal{J}}$, about each method. At date $t = 4$, the

Figure 1: The Timeline of Events



entrepreneurs sell their firms to the investors, perhaps due to life cycle reasons, at a price P_i. At date $t = 5$, the investors make their investment decisions $\{I_{ia}\}_{i \in \mathcal{I}}$. At date $t = 6$, the entrepreneurs consume the proceeds from the sale, whereas the investors consume the cash flows generated by their portfolios.

All parameters are common knowledge unless specified otherwise.

Discussion of Assumptions Rational inattention considers that agents have information available but cannot process all due to cognitive limitation. Instead, agents choose which pieces of information to attend to and optimally ignore the less important. To focus on the attention allocation decisions made by the investors in our setting, we abstract away from the moral hazard problem of information acquisition by taking the quantity of information acquired, the attention capacity, as given. Relatedly, we assume a continuum of otherwise identical investors so that each atomistic investor is a price-taker at the trading stage and, hence, aims to minimize the total post-learning uncertainty about the fundamentals when making attention allocation decisions. In addition, perfect competition in the capital market allows the entrepreneurs to retain all the returns of investment. Consequently, in an economy with a single representative firm, the entrepreneur's incentive is perfectly aligned with the investors' in maximizing information in the capital market. Indeed, by assuming away the potential strategic interaction among investors, the investors' preferences are perfectly aligned with those of a benevolent social planner.[4] Inefficiency can only arise from the

[4] Myatt and Wallace (2012) find that investors tend to acquire the same pieces of information when their actions exhibit strategic complementarity. By contrast, Van Nieuwerburgh and Veldkamp (2010) show that the resulting information acquisition may reflect specialization in the case of strategic substitutability.

entrepreneurs' method choices as they fail to take into account the externalities on the other firms when competing for the investors' limited attention.

Our learning technology is based on the entropy of independently distributed normal random variables. Summarizing in a single number the uncertainty associated with a distribution, in our setting, entropy is solely determined by the posterior variance of the firm's fundamental (see, e.g., Wiederholt, 2010). We measure the cost of information as a function of the distance between belief distributions, or equivalently, the mutual information. Conceptually, this measure is tantamount to the additional amount of information acquired through learning and captures the difficulties of processing and understanding information (Shannon, 1948; Cover, 1999; Maćkowiak, Matějka and Wiederholt, 2023). Moreover, it is independent of the source of information and the information acquisition process, making the implications of our findings more general to a broader set of settings.

The unlearnable uncertainty has been introduced in prior literature to generate predictions consistent with empirical findings (see, e.g., Van Nieuwerburgh and Veldkamp, 2010; Nezafat, Schroder and Wang, 2017; De Marco et al., 2022; Huang, Lunawat and Wang, 2024; Dávila and Parlatore, forthcoming). In our setting, the existence of unlearnable uncertainty ensures that the precision effect still plays a role even when the attention capacity is large. Indeed, our main results go through as long as there is an arbitrarily small fraction of unlearnable uncertainty, i.e., $\chi < 1$. Otherwise, the adoption of a common method would be socially optimal for sufficiently large attention capacity, and the adoption of the socially optimal common method can always be sustained as an equilibrium. To keep the analysis tractable, we model the impact of attention as reducing the post-reporting uncertainty proportionately. More specifically, the fraction of post-reporting uncertainty that survives through the learning process is an exponential function of attention that has a decaying tail. Alternatively, the unlearnable component of measurement error can arise endogenously from noise trading in a Kyle (1985) setting. Our results also hold qualitatively for fat-tailed distributions, e.g., the Pareto distribution under the power law, even in the absence of unlearnable uncertainty as long as the marginal return to attention diminishes as attention goes to infinity (see Dessein, Galeotti and Santos, 2016). That is, the essential feature of our learning technology is decreasing returns to attention.[5]

4. Preliminary Analysis

The equilibrium consists of the method choices simultaneously made by the entrepreneurs, the attention allocation and investment decisions made by the investors, and the pricing rules applied by the market. We characterize a perfect Bayesian equilibrium in which all players make optimal decisions that maximize their utility given all their available information as well as their rational expectations regarding the strategic behavior of the other players, utilizing Bayes' rule to make inferences and update their beliefs.

We solve the equilibrium using backward induction. At date $t = 5$, each atomistic investor holds an identical proportion of the market portfolio. Since all investors are otherwise identical and there is no externality in their actions, we consider a representative investor's investment and attention allocation choices below. The representative investor chooses investment to maximize the expected cash flow of each firm:

$$I_i^* = \arg\max_{I_i} \ \mathbb{E}[\tilde{v}_i | r_{ij}, s_j] = \mathbb{E}[\tilde{\omega}_i | r_{ij}, s_j].$$

That is, the representative investor matches the level of investment with his posterior belief about the firm's fundamental.

[5]The benefits of the adoption of a common accounting method become evident when the learning technology exhibits increasing returns to attention. In that case, the investors will devote all the attention to a single method. The adoption of a common method is socially optimal for sufficiently large attention capacity, and the adoption of the socially optimal common method can always be sustained as an equilibrium. Indeed, increasing and decreasing returns to learning have been used to generate specialized and generalized learning, respectively (see Van Nieuwerburgh and Veldkamp, 2010).

With a perfectly competitive capital market populated by a continuum of investors, at date $t = 4$, the firm price is set equal to the investors' posterior belief of the cash flow:

$$P_i = \mathbb{E}[\tilde{v}_i(I_i^*)|r_{ij}, s_j] = (\mathbb{E}[\tilde{\omega}_i|r_{ij}, s_j])^2 = \mathbb{E}[\tilde{\omega}_i^2|r_{ij}, s_j] - \hat{\sigma}_{ij}^2. \tag{6}$$

Thus, the entrepreneur's ex-ante expected market price at date $t = 1$ is

$$\mathbb{E}[P_i] = \mathbb{E}[\tilde{\omega}_i^2] - \hat{\sigma}_{ij}^2 = \mathrm{Var}(\tilde{\omega}_i) + \mathbb{E}[\tilde{\omega}_i]^2 - \hat{\sigma}_{ij}^2. \tag{7}$$

While the first two terms in (7) are determined by the prior distribution of the firm's fundamental, the last term is determined endogenously by both the entrepreneur's choice of accounting method and the investors' attention allocation decision. In particular, as can be seen from equations (2) and (3), the representative investor's post-learning variance about the firm's fundamental is lower when (i) the entrepreneur chooses a method that is closer to its firm characteristics and (ii) the investor devotes more attention to the chosen method. We refer to the former as the *precision effect* and the latter as the *attention effect*.

Moreover, since each individual investor is a price-taker, they also aim at maximizing the ex-ante expected value of the market portfolio, which determines their final consumption, when making attention allocation decisions at date $t = 2$:

$$V = \sum_{i=1}^{n} \mathbb{E}\big[\mathbb{E}[\tilde{v}_i(I_i^*)|r_{ij}, s_j]\big] = \sum_{i=1}^{n} \left(\mathrm{Var}(\tilde{\omega}_i) + \mathbb{E}[\tilde{\omega}_i]^2 - \hat{\sigma}_{ij}^2\right). \tag{8}$$

Therefore, the incentives of an entrepreneur and the investors would be perfectly aligned in minimizing the investors' post-learning uncertainty, or equivalently, maximizing information in the capital market, about the firm's fundamental if there were only one representative firm in the economy. Consequently, the entrepreneur would choose its local method, and the investors would devote all attention to learning about the chosen method. However, in an economy with multiple firms of different characteristics, competition for the investors' limited attention might create some externality across firms and induce firms to deviate from the socially optimal method choice.

Proposition 1. *Given the firms' choices of accounting methods, $\delta_j = 0$ for any method such that $\mathcal{I}_j = \emptyset$. For the rest of methods such that $\mathcal{I}_j \neq \emptyset$, rank them according to $\sum_{i\in\mathcal{I}_j} \sigma_{ij}^2$ from the largest to the smallest with index ℓ. Then*

$$\delta_\ell = \begin{cases} \frac{1}{\ell^*}\left(K + \log \frac{(\sum_{i\in\mathcal{I}_\ell} \sigma_{i\ell}^2)^{\ell^*}}{\prod_{\ell'\leq\ell^*}\sum_{i\in\mathcal{I}_{\ell'}} \sigma_{i\ell'}^2}\right) & \text{if } \ell \leq \ell^*, \\ 0 & \text{if } \ell > \ell^*, \end{cases} \tag{9}$$

where ℓ^* is the largest ℓ such that $K + \log \frac{(\sum_{i\in\mathcal{I}_\ell} \sigma_{i\ell}^2)^{\ell}}{\prod_{\ell'\leq\ell}\sum_{i\in\mathcal{I}_{\ell'}} \sigma_{i\ell'}^2} > 0$.

Since learning about firms' chosen methods is beneficial in reducing the posterior uncertainty about the fundamental, thereby increasing the ex-ante expected firm value, the investors will not waste attention on methods that are not adopted by any firm. For the rest of methods that are adopted by some firm, the most important method in the sense of inducing the largest sum of post-reporting uncertainty, $\sum_{i\in\mathcal{I}_j} \sigma_{ij}^2$, will always receive some attention. In contrast, the method with the smallest sum of post-reporting uncertainty, even if it is adopted by some firms, may not be allocated any attention when attention is scarce.

In general, whether the rest of the methods other than the most important one will be devoted any attention depends on (i) investor attention capacity and (ii) the balancedness of firms' method choices. Note that the marginal return of allocating one additional unit of attention to a particular method j is expressed as

$$\frac{\partial V}{\partial \delta_j} = \chi e^{-\delta_j} \sum_{i\in\mathcal{I}_j} \sigma_{ij}^2. \tag{10}$$

The first component is the fraction of the learnable post-reporting uncertainty. The second component captures the fractional marginal decrease in the post-reporting uncertainty due to learning, which is decreasing in the amount of attention allocated. The last component represents the sum of post-reporting variances of all firms that adopt this method because learning is at the method level. When attention is scarce, despite decreasing returns to learning, the marginal return of learning more about the most important method $\ell = 1$ would exceed that of the second most important method $\ell = 2$ even if investors devote all their attention to the most important method; that is, $\chi e^{-K} \sum_{i \in \mathcal{I}_1} \sigma_{i1}^2 \geq \chi \sum_{i \in \mathcal{I}_2} \sigma_{i2}^2$. Thus, the investors will only learn about the most important method with the highest sum of post-reporting variances of adopting firms.

As the attention capacity increases, the investors will devote their attention to more methods due to decreasing returns to learning. They optimally allocate their limited attention to equate the marginal return of attention across those methods. As can be seen from equation (10), the balancedness of method choices does not just depend on the number of firms adopting a certain method, but also depends on the consequence of such adoption in terms of reporting precision. More specifically, firms' adoption of a non-local method makes the method more important to learn about relative to the adoption of a local method. The reason is that the reduction in uncertainty from learning is proportional to the post-reporting uncertainty and, hence, higher in the former case. Nevertheless, in cases where attention is abundant, the marginal value of learning about a particular method becomes minimal. Consequently, the investors have an incentive to spread attention across all adopted methods.

Corollary 1. *The post-learning uncertainty about firm i's fundamental is continuous and increasing in its post-reporting uncertainty, i.e.,* $\frac{\partial \hat{\sigma}_{ij}^2}{\partial \sigma_{ij}^2} > 0$.

Although the attention allocated to a method is increasing in the sum of post-reporting uncertainty of its adopting firms (see *Proposition 1*), it is not worthwhile to adopt a more distant method to attract more attention. In other words, as a firm searches the entire universe of methods, excluding those already adopted by other firms, the loss in post-reporting precision from adopting a non-local method always dominates the potential gain from increased attention. Even in the extreme case where the firm adopts a sufficiently remote method that it ends up attracting all the attention, the increase in the post-reporting uncertainty would be so high that the post-learning uncertainty will still be higher than the case where the firm sticks with its own local method. In short, the precision effect always dominates the attention effect.

Corollary 1 suggests that, when a firm is deviating from a method that is not its local method, the most profitable deviation is to adopt its local method. Conversely, if a firm is deviating from its local method, it will never deviate to a new method that is not adopted by any other firm. The only potentially profitable deviation is to adopt a method that is already in use by some other firm(s) so that the gain from increased attention may dominate the loss in post-reporting precision.

The adoption of distinct methods, however, may not be socially optimal due to the externalities that firms impose on the others when choosing reporting methods, which arise from investor rational inattention. In equilibrium, the entrepreneurs are paid the expected firm value whereas the investors break even in expectation in a perfectly competitive capital market. Thus, the social welfare is captured by the sum of firm values as in a pure exchange economy (see also Dye and Sridhar, 2004; Gao et al., 2019) and, from an ex-ante perspective, is expected to be

$$W = \sum_{i=1}^{n} \mathbb{E}\big[\mathbb{E}[\tilde{v}_i(I_i^*)|r_{ij}, s_j]\big] = \sum_{i=1}^{n} \left(\mathrm{Var}(\tilde{\omega}_i) + \mathbb{E}[\tilde{\omega}_i]^2 - \hat{\sigma}_{ij}^2\right). \tag{11}$$

Note that the ex-ante expected social welfare perfectly coincides with the investors' objective in equation (8), implying that inefficiency can only arise from the entrepreneurs' method choices when they fail to take into account the externalities on the other firms as a result of the investors' endogenous attention allocation decision.

Before characterizing the socially optimal standard in general in the following section, we first examine the socially optimal method choice for a given number of adopted methods. Perhaps surprisingly, instead of dividing firms into several equal-sized groups that adopt a common method, it is more efficient to have only one large group adopt a common method, while the other firms each adopt a distinct method.

Proposition 2. *For a given number of adopted accounting methods, m, the socially optimal method choice is such that $m-1$ firms adopt their own local methods, whereas the remaining $n-m+1$ firms adopt a common method characterized by their geometric center.*

With endogenously allocated attention prescribed in *Proposition 1*, the sum of post-learning uncertainty of all firms, S, can be written as follows:

$$S = (1-\chi)\sum_{\ell \leq \ell^*}\sum_{i \in \mathcal{I}_\ell} \sigma_{i\ell}^2 + \ell^* \chi e^{-\frac{K}{\ell^*}} \left(\prod_{\ell \leq \ell^*} \sum_{i \in \mathcal{I}_\ell} \sigma_{i\ell}^2 \right)^{\frac{1}{\ell^*}} + \sum_{\ell > \ell^*} \sum_{i \in \mathcal{I}_\ell} \sigma_{i\ell}^2, \tag{12}$$

where ℓ^* is the number of methods receiving attention in equilibrium. As can be seen from equation (12), from a social perspective, it is always optimal to reduce the sum of post-reporting uncertainty no matter whether a method receives attention or not. It follows that, if a method is adopted by multiple firms, it should be the geometric center of those firms, denoted as j^*, such that $y_{j^*,k} = \frac{1}{n_{j^*}}$ for any $k \in \mathcal{I}_{j^*}$ and $y_{j^*,k} = 0$ otherwise, yielding the minimum sum of post-reporting uncertainty:

$$\sum_{i \in \mathcal{I}_{j^*}} \sigma_{ij^*}^2 = n_{j^*}\sigma_0^2 + n_{j^*} - 1, \tag{13}$$

where $n_{j^*} \equiv |\mathcal{I}_{j^*}|$ is the number of firms adopting the common method. That is, the socially optimal common method equally accounts for the idiosyncrasies of all constituent firms. If a method is adopted by a single firm, then it should be the local method of that firm. Moreover, for the remaining $m - \ell^*$ firms whose underlying methods have not received any attention, they should optimally adopt their own local methods instead of forming a coalition to adopt a common method.

Furthermore, recall that the equilibrium attention allocation optimally equates the marginal return of attention across methods (see equation (10)). Consequently, for those receiving some attention, equation (12) suggests that the socially optimal standard also aims at minimizing their product of post-reporting variance due to the complementarity between attention and method adoption. Observe from equation (13) that the sum of post-reporting uncertainty of each group adopting a common method is linear in the number of firms in that group, i.e., the group size. Hence, the sum of post-reporting uncertainty across all ℓ^* groups, each of which adopts a common method, is determined by the total number of firms across all groups, $n - m + \ell^*$, and invariant of the distribution of the group size. The product of post-reporting variance is thus minimized by having only one large group adopt a common method, while the other $\ell^* - 1$ firms each adopt their own local methods.

Intuitively, since the marginal return of attention devoted to a method is increasing in the number of firms adopting this method (see equation (10)), the investors will optimally allocate more attention to a method as it is adopted by more firms. The increase in allocated attention in turn makes the method more attractive to be adopted by the other firms from a social perspective. The complementarity between attention and method adoption leads to a convexity in the value of attention and manifests as the *network effect*. As a result, fixing the number of adopted methods, the socially optimal method choice features only one large group adopting a common method, while the other remaining firms each adopt their own local methods. Such a method choice minimizes the sum of post-learning uncertainty, S, for any ℓ^* and hence the endogenously determined ℓ^* given by *Proposition 1*. By the symmetry among the $m - 1$ firms that adopt their own local methods, in equilibrium $\ell^* = 1$ if investor attention capacity is relatively small; otherwise, $\ell^* = m$.

5. Main Results

As discussed in the previous section, the socially optimal standard trades off the precision effect and the attention effect. While the magnitude of the precision effect is governed by the characteristics of the underlying accounting method, the attention effect is endogenously determined by investor rational inattention. Ideally, the total post-learning uncertainty is minimized by inducing high reporting precision as well as concentrating attention on fewer methods for more in-depth learning. The former drives the socially optimal standard towards allowing for diverse practices, whereas the latter makes the adoption of a common method more desirable.

To understand the trade-off between the two effects in the presence of investor rational inattention, in this section, we restrict our attention to the adoption of distinct methods, each of which is the firm's local method, and the adoption of a common method, which is characterized by the geometric center of all firms. We first characterize the conditions under which each of these two method choices is socially optimal, and then proceed to investigate whether the socially optimal method choice can be sustained in equilibrium.

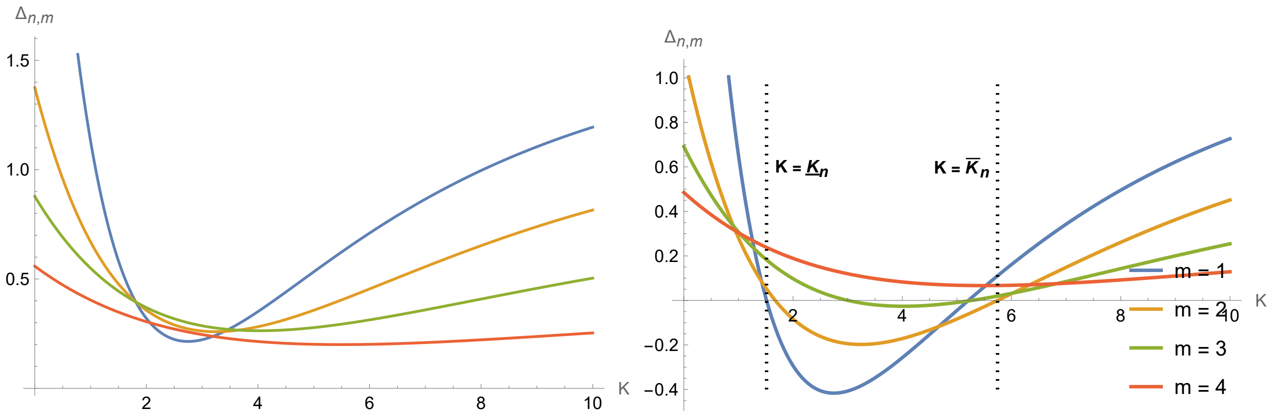
5.1. Distinct Methods

Proposition 3. *The adoption of distinct accounting methods is socially optimal for any level of attention capacity if the fraction of the learnable post-reporting uncertainty is relatively low, i.e.,* $\chi \leq \underline{\chi}$*; otherwise, it is socially optimal if and only if investor attention capacity is sufficiently small or sufficiently large, i.e.,* $K < \underline{K}_n$ *or* $K > \overline{K}_n$*, where* $\underline{K}_n < \overline{K}_n$ *and* $\underline{\chi} \in (0, 1)$ *are defined in the appendix.*

When attention is scarce, the gain from concentrating attention on a few common methods is always dominated by the loss from lower reporting precision. On the other hand, when attention is abundant, its marginal value is also limited due to decreasing returns to learning, making it optimal for some firm(s) to adopt their own local methods from a social perspective (see Figure 2(b)). Hence, the precision effect dominates for low and high values of attention capacity in general, rendering the adoption of distinct methods socially optimal.

Moreover, when the fraction of learnable uncertainty is sufficiently small, the attention effect will be dampened so much that the precision effect dominates even if all the attention is devoted to a common method. Intuitively, concentrated attention on fewer methods only

Figure 2: The Optimality of Distinct Methods



(a) $\chi \leq \underline{\chi}$ (b) $\chi > \underline{\chi}$

Note: Each curve represents the reduction in the total post-learning uncertainty of adopting distinct methods relative to the other number of adopted method(s). The parameter values are $n = 5$, $\tau_0 = 1$, and $\chi = 0.6$ for the left panel and $\chi = 0.7$ for the right panel.

plays a role on the learnable component of measurement errors. In the extreme case where no measurement error can be reduced by learning about reporting methods, i.e., $\chi = 0$, it is socially optimal for each firm to adopt its own local method to minimize post-reporting uncertainty and maximize firm value. Consequently, as shown in Figure 2(a), the adoption of distinct methods is always superior to any other method choice.

Corollary 2. *The adoption of distinct methods can be sustained as an equilibrium whenever it is socially optimal.*

Recall from *Corollary 1* that the only potentially profitable deviation for an individual firm from its local method is to adopt some other firm's local method so that there is a decrease in the total number of adopted methods and these two firms can jointly attract more attention from the investors. However, such a deviation is not desirable whenever the adoption of distinct methods is socially optimal. Suppose firm i deviates to the local method of firm i'. Since the adoption of distinct methods yields a higher social welfare than all other method choices, including the one where all firms adopt the local method of firm i', all firms other than firm i' must be strictly better off adopting their own local methods than adopting the local method of firm i':

$$\left(1-\chi+\chi e^{-\frac{K}{n}}\right)\sigma_0^2 < \left(1-\chi+\chi e^{-K}\right)(\sigma_0^2+2).$$

That is, even if receiving all the attention with a common method, the positive precision effect of adopting one's local method dominates the negative attention effect of distinct methods. As a result, deviating to some other firm's local method while all the other firms stick with their own local methods is not profitable because it induces the same level of post-reporting variance but attracts less attention than the common method, leading to a dampened negative attention effect.

In a nutshell, whenever the adoption of distinct methods is socially optimal, no firm will have an incentive to deviate to any other method. Put differently, the adoption of distinct methods is always sustained as an equilibrium whenever it is socially optimal, implying there is no need to mandate distinct methods.[6]

5.2. Common Method

Proposition 4. *The adoption of a common method is socially optimal if and only if*

(i) the fraction of the learnable post-reporting uncertainty is relatively high, i.e., $\chi > \underline{\chi}$*; and*

(ii) investor attention capacity is in an intermediate range, i.e., $\underline{K}_1 < K < \overline{K}_1$*,*

where $\underline{K}_1 < \overline{K}_1$ *are defined in the appendix.*

The net social benefit, in terms of the reduction in the total post-learning uncertainty across all firms, S (see equation (12)), of adopting a common method relative to the adoption of any other number of methods crucially depends on investor attention capacity. Recall that the learning technology exhibits decreasing returns to attention. When attention is

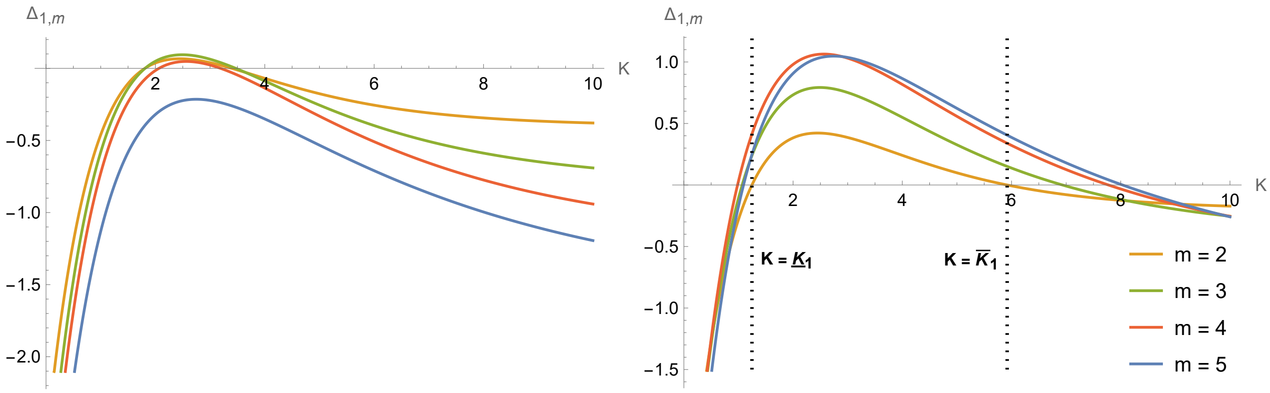
[6]We do not rule out the possibility of multiple equilibria, but instead focus on the Pareto-optimal equilibrium in our analysis (see, e.g., Harsanyi and Selten, 1988).

scarce, by attracting all attention to a single method, the adoption of a common method allows an increase in investor attention capacity to reduce the learnable uncertainty more significantly than any other method choice. In contrast, when attention is abundant, by spreading attention across methods, the adoption of different methods facilitates a greater reduction in the learnable uncertainty following an increase in investor attention capacity. Therefore, as depicted in Figure 3, the net benefit of adopting a common method is first increasing and then decreasing in the total attention.

Nevertheless, the adoption of a common method is socially optimal only for intermediate values of attention. The reason is that the extent to which learning about an accounting method reduces the post-learning uncertainty about the reporting firm is proportional to its post-reporting uncertainty. In the extreme case where the total attention is zero, adopting a common method makes everyone worse off by inducing higher post-reporting uncertainty. On the other extreme, when attention is unlimited, the adoption of a common method still provides no benefit because any adopted method will receive infinite amount of attention, effectively eliminating all learnable uncertainty. However, because the measurement errors cannot be perfectly identified even with infinite attention, the adoption of distinct methods is socially optimal by inducing the least post-learning uncertainty.

Lastly, to ensure that the benefit of a common method arising from concentrated attention ever exceeds its loss in reporting precision, the learnable fraction of post-reporting uncertainty has to be sufficiently high. Otherwise, as shown in *Proposition 3*, the adoption of a common method is always dominated by the adoption of distinct methods for any amount of the total attention (see Figures 3(a) and 4(a) for a graphical illustration). Taken together, *Proposition 3* and *Proposition 4* jointly suggest that the socially optimal number of adopted methods first decreases and then increases in investor attention capacity. In particular, the socially optimal standard allows for distinct local methods when attention is sufficiently scarce or sufficiently abundant, i.e., $K < \underline{K}_n$ and $K > \overline{K}_n$, but mandates a

Figure 3: The Optimality of A Common Method



(a) $\chi \leq \underline{\chi}$ (b) $\chi > \underline{\chi}$

Note: Each curve represents the reduction in the total post-learning uncertainty of adopting a common method relative to the other number of adopted methods. The parameter values are $n = 5$, $\tau_0 = 1$, and $\chi = 0.6$ for the left panel and $\chi = 0.8$ for the right panel.

common method when the total attention is in an intermediate range, i.e., $\underline{K}_1 < K < \overline{K}_1$, where $\underline{K}_n \leq \underline{K}_1 < \overline{K}_1 \leq \overline{K}_n$.

In stark contrast to the adoption of distinct methods, the adoption of a common method does not necessarily emerge as an equilibrium, notwithstanding its optimality. Before characterizing the conditions under which mandating the adoption of a common method is necessary, we investigate in the following proposition the scenarios in which the adoption of a common method can be sustained as an equilibrium whenever it is socially optimal.

Proposition 5. *When the adoption of a common method is socially optimal, it can be sustained as an equilibrium if investor attention capacity is relatively small, i.e.,* $K \leq K^* \equiv \frac{n}{n-2} \log \Big((n-1)\big(1 + \frac{n-1}{n}\tau_0\big)\Big)$.

Recall from *Corollary 1* that the most profitable deviation from the common method is to adopt one's own local method. Thus, it is incentive compatible for firms to stick with the common method if and only if

$$B_i(K) \geq C \equiv (1-\chi)\frac{n-1}{n}\tau_0, \tag{14}$$

where the left-hand side represents the individual benefit of lower learnable uncertainty due to concentrated attention devoted to the common method:

$$B_i(K) \equiv \begin{cases} \chi\Big(1 - \big(1 + \frac{n-1}{n}\tau_0\big)e^{-K}\Big) & \text{if } K < K_i, \\ \chi\left(\big((n-1)\big(1 + \frac{n-1}{n}\tau_0\big)\big)^{\frac{1}{2}} e^{-\frac{K}{2}} - \big(1 + \frac{n-1}{n}\tau_0\big)e^{-K}\right) & \text{if } K \geq K_i, \end{cases} \tag{15}$$

with $K_i \equiv \log\left((n-1)\big(1 + \frac{n-1}{n}\tau_0\big)\right)$, and the right-hand side represents the cost of higher unlearnable uncertainty arising from the increased post-reporting variance of the common method relative to the local method.

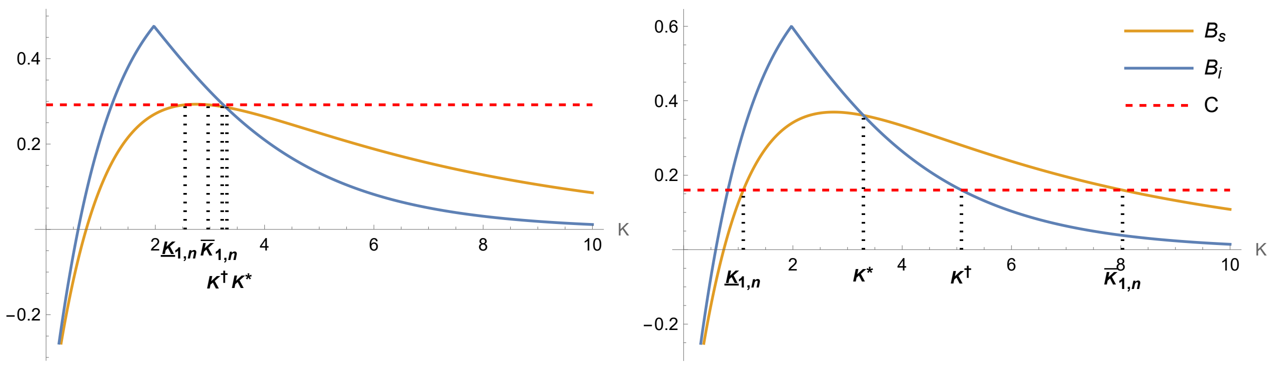
In addition, since the adoption of a common method is socially optimal, it must be more efficient than the adoption of distinct methods. That is,

$$B_s(K) \equiv \chi\Big(e^{-\frac{K}{n}} - \big(1 + \frac{n-1}{n}\tau_0\big)e^{-K}\Big) > C, \tag{16}$$

where the left-hand side represents the social benefit per firm of lower learnable uncertainty due to concentrated attention. Note that while the cost related to the unlearnable uncertainty is invariant to investor attention capacity, the benefit associated with the learnable uncertainty first increases and then decreases in the total attention at both the individual and the social level (see Figures 3 and 4).

The adoption of a common method can always be sustained as an equilibrium whenever it is socially optimal as long as the individual benefit exceeds the social benefit. As demonstrated in Figure 4, this happens for sufficiently small attention capacity, i.e., $K < K^*$. Intuitively, when attention is scarce, the amount of attention that the firm attracts by deviating to its local method is limited and, more importantly, less than the amount of attention it receives under the adoption of distinct methods. Put differently, firms efficiently coordinate on a common method under limited attention. In particular, when the investors' total attention is sufficiently small, i.e., $K < K_i$, they will not devote any attention to the local method following the firm's deviation. Therefore, the investors' small attention

Figure 4: Benefit and Cost of A Common Method Relative to Distinct Methods



(a) $\underline{\chi} < \chi \leq \overline{\chi}$

(b) $\chi > \overline{\chi}$

Note: The blue and yellow curves represent the individual benefit and the social benefit per firm of lower learnable uncertainty due to concentrated attention devoted to the common method relative to the adoption of local method(s), $B_i(K)$ and $B_s(K)$, defined in equations (15) and (16), respectively. The red dashed horizontal line represents the cost per firm of higher unlearnable uncertainty arising from the increased post-reporting variance of the common method relative to the local method, C, defined in equation (15). The parameter values are $n = 5$, $\tau_0 = 1$, and $\chi = 0.635$ for the left panel and $\chi = 0.8$ for the right panel.

capacity is a sufficient condition to sustain the adoption of the socially optimal common method as an equilibrium, suggesting no need for mandatory standardization.

Essentially, the investors' rational allocation of limited attention mitigates the externalities that one firm can potentially impose on the other firms through a deviation to a different method, which may distract attention away from the common method. Specifically, it suppresses firms' incentives to deviate to their own local methods for higher reporting precision. In other words, rational inattention coordinates firms to "speak the same language." The deviation incentive is deterred because of the fear of being the minority.

Corollary 3. *The threshold of investor attention capacity, K^*, below which the adoption of the socially optimal common method can always be sustained as an equilibrium*

(i) first decreases and then increases in the number of firms, n; and

(ii) increases in the reporting precision of local methods, τ_0.

The threshold of investor attention capacity below which mandatory standardization may not be necessary is determined by equating the individual and the social benefit per firm of adopting a common method relative to distinct methods. As the number of firms increases, the social benefit of adopting a common method becomes larger because each firm receives a smaller fraction of the total attention under the adoption of distinct methods. Meanwhile, the network effect of more firms adopting a common method enhances the individual firm's incentive to follow the crowd in adopting the same method. That is, the individual benefit of adopting the common method increases as well.

When there are only a few firms in the economy, the social benefit increases at a faster rate than the individual benefit. In the extreme case where there are only two firms, $\lim_{n\to 2} K^* = \infty$; that is, mandatory standardization is never necessary. Indeed, by deviating to its local method, the firm necessarily attracts less than half of the attention due to its lower post-reporting uncertainty relative to the other firm. Given that the adoption of a common method results in less post-learning uncertainty than the adoption of distinct methods, where each firm receives exactly half of the attention, sticking with the common method strictly dominates deviation to the local method. As a new firm enters the economy, the fractional loss in attention under the adoption of distinct methods is significant, which makes the adoption of a common method even more beneficial. By contrast, from an individual firm's perspective, the loss in attention is attenuated by its deviation. Indeed, the deviating firm does not internalize the negative externalities it imposes on the other firms by distracting attention away from the common method. Thus, the threshold of attention decreases, thereby mitigating the deviation incentive to the local method by making the deviation less profitable.

As the number of firms increases further, the fractional loss in attention under the adoption of distinct methods is minimal, limiting the desirability of the adoption of a common method. Meanwhile, the network effect makes deviation much less attractive to individual firms due to the complementarity between the attention allocated to the common method and the number of firms adopting the method. As a result, the threshold of attention increases, restoring the individual firm's deviation incentive. Indeed, the minority method will receive much less attention due to the complementarity between attention and the number of firms adopting the underlying method. The network effect suppresses individual firm's deviation incentive, making inefficient deviation less likely. In the extreme case where the number of firms is infinitely large, $\lim_{n\to\infty} K^* = \infty$. Each firm receives minimal, if any, attention by deviating to its local method. Hence, the individual benefit of adopting the common method always exceeds the social benefit, making mandatory standardization unnecessary.

As the reporting precision of local methods increases, perhaps surprisingly, there is less of a need to mandate the adoption of a common method as well. On the one hand, an individual firm is more incentivized to deviate to its local method due to the precision effect. On the other hand, investors with rational inattention optimally allocate more attention to the common method adopted by the other firms and less attention to the local method of the deviating firm. Taken together, the negative attention effect dominates the positive precision effect, undermining the firm's incentive to deviate from the common method.

5.3. Social Value of Standardization

Since the adoption of the socially optimal common method can always be sustained as an equilibrium in an economy with two firms, in this section, we focus on the case where $n \geq 3$ and $K > K^*$ to examine conditions under which there is a need to mandate the common method.

Proposition 6. *Firms have an incentive to deviate from the socially optimal common method if and only if*

(i) the fraction of the learnable post-reporting uncertainty is sufficiently high, i.e., $\chi > \overline{\chi}$; and

(ii) investor attention capacity is moderately large, i.e., $K^\dagger < K < \overline{K}_1$,

where $K^* < K^\dagger < \overline{K}_1$ and $\overline{\chi} > \underline{\chi}$ are defined in the appendix.

When investor attention capacity is relatively large, i.e., $K > K^*$, the social benefit of adopting a common method is greater than the individual benefit, i.e., $B_s(K) > B_i(K)$, implying a potential unilateral deviation from the socially optimal common method (see *Proposition 5*). However, as can be seen in Figure 4(b), when the total attention is moderately constrained, i.e., $K \leq K^{\dagger}$, the individual benefit of adopting the common method dominates the cost, i.e., $B_i(K) \geq C$, because the amount of attention attracted when deviating to the local method is limited under relatively scarce attention. Thus, firms have no incentive to deviate from the common method.

As attention becomes more abundant, i.e., $K > K^{\dagger}$, the gain in increased reporting precision from deviating to the local method outweighs the loss in attention. Due to decreasing returns to learning, the investors will optimally allocate some attention to the minority method, reducing its loss in attention of deviating from the common method. However, the individual firm does not fully internalize the externalities it imposes on the other firms arising from the diluted attention on the common method. From a social perspective, when the total attention is moderately but not sufficiently large, a significant fraction of measurement noises can be mitigated through learning. Therefore, the social benefit of concentrating attention on a single method still renders the adoption of a common method optimal despite decreasing returns to learning. Nevertheless, as the total attention increases further, i.e., $K \geq \overline{K}_{1,n}$, the cost of adopting a common method relative to the adoption of distinct methods exceeds its social benefit per firm, i.e., $B_s(K) \leq C$. Hence, the adoption of the common method is dominated by that of distinct methods and no longer socially optimal.

Furthermore, note that firms have no incentive to deviate from the socially optimal common method under relatively small total attention, i.e., $K \leq K^{*}$. When investor attention capacity is relatively large, however, the social benefit of the adoption of a common method is capped due to decreasing returns to attention. Thus, to ensure that the attention effect still plays an important role, we have a more stringent condition on the learnable fraction of post-reporting uncertainty.

Combining with *Proposition 4*, *Proposition 6* implies that mandatory standardization may be in place only when investor attention capacity is moderately large, i.e., $K^{\dagger} < K < \overline{K}_1$. Contrary to conventional wisdom, mandatory standardization may become more important as information processing capacity increases. We view this insight as consistent with the recent mandate of ASC 606 which aims at standardizing revenue recognition across different industries (Hinson et al., 2024) despite the advancement in information technology and language processing tools. It also provides an alternative explanation for the proliferation of "pro forma" (i.e. non-GAAP) earnings disclosures (Black et al., 2018) as well as the attempt by the U.S. Securities and Exchange Commission (SEC) to restrict such disclosures (Heflin and Hsu, 2008) despite their informativeness beyond GAAP earnings (Curtis et al., 2014; Campbell et al., 2022).

6. Conclusion

This paper proposes a theoretical framework to study the social value of various accounting method adoption choices, in particular, the mandatory adoption of a common method, by generalizing the implications of the trade-off between the precision effect and the attention effect of method adoption. Specifically, our theoretical framework considers an arbitrary number of firms and the entire universe of accounting methods without imposing any restriction on the method choices. The socially optimal standard trades off the precision effect and the attention effect. The former is governed by the characteristics of the underlying accounting method, driving the socially optimal standard towards allowing for diverse practices. The latter is endogenously determined by investor rational inattention, making the adoption of a common method more desirable.

We characterize the conditions under which each method choice is socially optimal and examine whether the socially optimal method choice can be sustained in equilibrium. In stark contrast to the adoption of distinct local methods, the adoption of a common method does not necessarily emerge as an equilibrium, notwithstanding its optimality. In particular, investor rational inattention induces firms to efficiently coordinate on a common method only when investor attention capacity is moderately small. Otherwise, due to decreasing returns to learning, it will lead to an inefficient deviation from the common method to the local method. The condition that the total attention is moderately but not sufficiently large for mandatory standardization to be welfare improving is left out by existing literature, yet provides meaningful policy implications. As stated by Anand (2006, p. 229), "an enabling governance regime coupled with mandatory disclosure of a firm's governance practices is likely to yield a high level of compliance at lower direct costs to the issuer than a wholly mandatory regime."

More broadly, reinterpreting the accounting method choice as the choice of selecting a proper language to communicate information, our theory has potential implications for related settings such as the choice of communication technologies to convey contents in media as well as technical languages adopted in multi-unit organizations (see, e.g., Cremer, Garicano and Prat, 2007; Myatt and Wallace, 2012; Dessein et al., 2016; Galperti and Trevino, 2020). To the extent that economic agents have to spare attention to different mediate outlets and organizational "codes" – the set of words used to describe characteristics of the problems (Arrow, 1974) – to facilitate communication in various contexts, we offer insights on when they may (in)efficiently coordinate on a common language and distinct languages, and when regulations might be necessary to ensure efficient communication.

Our study is subject to several caveats. First, we do not tackle the multiplicity of equilibria on accounting method choices. Instead, we implicitly assume that firms can self-enforce the Pareto-optimal equilibrium, but do not speak to how the socially optimal method adoption choice may evolve spontaneously over time. To answer this question, future research may deploy evolutionary game theory (see, e.g., Rubinstein, 2000) to study the evolution of generally accepted accounting principles as a social norm developed over time. Second, we consider a simultaneous move game among firms regarding their method adoption choices to abstract away any potential asymmetry among firms. Alternatively, one may consider the case in which firms move sequentially in choosing the accounting method. Anticipating the follower's best response under different levels of investor attention capacity, the leading firm may not necessarily stick with its own local method, despite the first-mover advantage in selecting its preferred equilibrium, to attract greater investor attention. This occurs when attention falls within an intermediate range, such that the network effect dominates the precision effect from the leading firm's perspective. Nevertheless, the socially optimal common method is never adopted in equilibrium almost surely: the leading firm can always improve its payoff by adopting a method that is closer to its local method yet still induces adoption by the follower. In addition, we abstract away from the potential correlation between firms' fundamentals and measurement processes. To the extent that comparability enhances learning about firm idiosyncrasies (Wu and Xue, 2023; Wu and Zheng, 2025), the adoption of a common method may be even more desirable beyond concentrating investors' limited attention to facilitate more in-depth learning. Lastly, we focus on externalities among firms, which arise endogenously as a result of investor rational inattention, and, hence, assume away potential strategic interactions among investors or information intermediaries. We leave the effect of the strategic interaction among investors on the optimality of a common accounting method to future research.

Appendix

Proof of Proposition 1. Inspecting equation (8), note that, given the firms' method choices, the representative investor allocates his attention to minimize the sum of post-learning uncertainty of each firm's fundamental subject to the attention capacity constraint (4) and the no-forgetting constraint (5). That is,

$$\min_{\{\delta_j\}_{j\in\mathcal{J}}} \sum_{j:\,\mathcal{I}_j\neq\emptyset}\sum_{i\in\mathcal{I}_j}\left(1-\chi+\chi e^{-\delta_j}\right)\sigma_{ij}^2 \tag{A1}$$

$$\text{s.t.} \sum_{j\in\mathcal{J}}\delta_j \leq K, \tag{A2}$$

$$\delta_j \geq 0. \tag{A3}$$

Note that the ex-ante expected firm value is strictly increasing in the attention allocated to method j adopted by firm i. Thus, the investor optimally allocates zero attention to methods that are not chosen by any firm, i.e., $\delta_j = 0$ if $\mathcal{I}_j = \emptyset$.

For $\mathcal{I}_j \neq \emptyset$, taking the first-order condition with respect to δ_j yields

$$\chi e^{-\delta_j}\sum_{i\in\mathcal{I}_j}\sigma_{ij}^2 - \lambda + \mu_j = 0, \tag{A4}$$

where λ and μ_j are the Lagrange multiplier on the attention capacity constraint and the no-forgetting constraint, respectively. If $\mu_j = 0$, then $\delta_j > 0$; otherwise, $\delta_j = 0$. It then follows that for any j, j' such that $\mathcal{I}_l, \mathcal{I}_{j'} \neq \emptyset$ and $\mu_j = \mu_{j'} = 0$, we have

$$\delta_j - \delta_{j'} = \log\frac{\sum_{i\in\mathcal{I}_j}\sigma_{ij}^2}{\sum_{i\in\mathcal{I}_{j'}}\sigma_{ij'}^2} \quad \text{and} \quad \sum_{j:\,\mathcal{I}_j\neq\emptyset}\delta_j = K. \tag{A5}$$

Therefore,

$$\delta_\ell = \begin{cases} \frac{1}{\ell^*}\left(K + \log\frac{\left(\sum_{i\in\mathcal{I}_\ell}\sigma_{i\ell}^2\right)^{\ell^*}}{\prod_{\ell'\leq\ell^*}\sum_{i\in\mathcal{I}_{\ell'}}\sigma_{i\ell'}^2}\right) & \text{if } \ell \leq \ell^*, \\ 0 & \text{if } \ell > \ell^*, \end{cases} \tag{A6}$$

where ℓ^* is the largest ℓ such that $K + \log \frac{(\sum_{i\in\mathcal{I}_\ell} \sigma_{i\ell}^2)^\ell}{\prod_{\ell'\leq \ell}\sum_{i\in\mathcal{I}_{\ell'}}\sigma_{i\ell'}^2} > 0$. □

Proof of Corollary 1. Suppose firm i adopts method j and is allocated some positive but not all the attention, i.e., $\delta_{ij} \in (0, K)$. Adapting to the notations in *Proposition 1*, we may re-index the method with index $\ell \leq \ell^*$, where $\ell^* \geq 2$. By *Proposition 1*, we have

$$\hat{\sigma}_{i\ell}^2 = \left(1 - \chi + \chi e^{-\frac{K}{\ell^*}} \left(\sum_{i\in\mathcal{I}_\ell} \sigma_{i\ell}^2\right)^{-1} \left(\prod_{\ell'\leq\ell^*}\sum_{i\in\mathcal{I}_{\ell'}} \sigma_{i\ell'}^2\right)^{\frac{1}{\ell^*}}\right)\sigma_{i\ell}^2$$
$$= \left(1 - \chi + \chi e^{-\frac{K}{\ell^*}} \left(\sum_{i\in\mathcal{I}_\ell} \sigma_{i\ell}^2\right)^{\frac{1}{\ell^*}-1} \left(\prod_{\substack{\ell'\neq\ell\\ \ell'\leq\ell^*}}\sum_{i\in\mathcal{I}_{\ell'}} \sigma_{i\ell'}^2\right)^{\frac{1}{\ell^*}}\right)\sigma_{i\ell}^2.$$

Hence, for a given number of methods, ℓ^*, that are allocated some attention in equilibrium, we have

$$\frac{\partial\hat{\sigma}_{i\ell}^2}{\partial\sigma_{i\ell}^2} = 1 - \chi + \chi e^{-\frac{K}{\ell^*}} \left(\sum_{i\in\mathcal{I}_\ell} \sigma_{i\ell}^2\right)^{-2} \left(\prod_{\ell'\leq\ell^*}\sum_{i\in\mathcal{I}_{\ell'}} \sigma_{i\ell'}^2\right)^{\frac{1}{\ell^*}} \left(\sum_{\substack{i'\neq i\\ i'\in\mathcal{I}_\ell}} \sigma_{i'\ell}^2 + \frac{1}{\ell^*}\sigma_{i\ell}^2\right) > 0.$$

Further note that an infinitesimal decrease in $\sigma_{i\ell}^2$ for $\ell < \ell^*$ such that $K = \log \frac{\prod_{\ell\leq\ell^*+1}\sum_{i\in\mathcal{I}_\ell}\sigma_{i\ell}^2}{\left(\sum_{i\in\mathcal{I}_{\ell^*+1}}\sigma_{i\ell^*+1}^2\right)^{\ell^*+1}}$ might lead to an increase in ℓ^*. However, such a change in ℓ^* does not result in any change in its post-learning uncertainty:

$$\begin{aligned}
&\hat{\sigma}_{i\ell}^2(\ell^*+1) - \hat{\sigma}_{i\ell}^2(\ell^*)\\
&= \chi\left(\sum_{i\in\mathcal{I}_\ell}\sigma_{i\ell}^2\right)^{-1}\left(e^{-\frac{K}{\ell^*+1}}\left(\prod_{\ell'\leq\ell^*+1}\sum_{i\in\mathcal{I}_{\ell'}}\sigma_{i\ell'}^2\right)^{\frac{1}{\ell^*+1}} - e^{-\frac{K}{\ell^*}}\left(\prod_{\ell'\leq\ell^*}\sum_{i\in\mathcal{I}_{\ell'}}\sigma_{i\ell'}^2\right)^{\frac{1}{\ell^*}}\right)\sigma_{i\ell}^2\\
&= 0.
\end{aligned}$$

Lastly, $\hat{\sigma}_{ij}^2 = \sigma_{ij}^2$ for $\delta_{ij} = 0$, and $\hat{\sigma}_{ij}^2 = \left(1 - \chi + \chi e^{-K}\right) \sigma_{ij}^2$ for $\delta_{ij} = K$, both of which increase in σ_{ij}^2. □

Proof of Proposition 2. Consider a choice of methods $\{\mathcal{I}_j\}_{j \in \mathcal{J}}$ under which m methods are adopted and ℓ^* methods are allocated some attention. For the effect of the sum of post-reporting variance of any adopted method on the sum of post-learning variance characterized in equation (12), we have

$$\frac{\partial S}{\partial \sum_{i \in \mathcal{I}_\ell} \sigma_{i\ell}^2} = \begin{cases} 1 - \chi + \chi e^{-\frac{K}{\ell^*}} \frac{\left(\prod_{\ell' \leq \ell^*} \sum_{i \in \mathcal{I}_{\ell'}} \sigma_{i\ell'}^2\right)^{\frac{1}{\ell^*}}}{\sum_{i \in \mathcal{I}_\ell} \sigma_{i\ell}^2} & \text{for } \ell \leq \ell^*, \\ 1 & \text{for } \ell > \ell^*. \end{cases}$$

Hence, for any given ℓ^*, in order to minimize the sum of post-learning uncertainty, the methods that are allocated some attention should be the geometric center of firms in each group that adopt a common method, and the methods that are not allocated any attention should be the local method of each firm.

Further note that an infinitesimal decrease in $\sum_{i \in \mathcal{I}_\ell} \sigma_{i\ell}^2$ for $\ell \leq \ell^*$ such that $K = \log \frac{\prod_{\ell \leq \ell^*} \sum_{i \in \mathcal{I}_\ell} \sigma_{i\ell}^2}{\left(\sum_{i \in \mathcal{I}_{\ell^*}} \sigma_{i\ell^*}^2\right)^{\ell^*}}$ might also lead to a change in ℓ^*. However, such a change in ℓ^* does not result in any change in the sum of post-learning uncertainty:

$$\begin{aligned} & S(\ell^*) - S(\ell^* - 1) \\ = & -\chi \sum_{i \in \mathcal{I}_{\ell^*}} \sigma_{i\ell^*}^2 + \ell^* \chi e^{-\frac{K}{\ell^*}} \left(\prod_{\ell \leq \ell^*} \sum_{i \in \mathcal{I}_\ell} \sigma_{i\ell}^2\right)^{\frac{1}{\ell^*}} - (\ell^* - 1) \chi e^{-\frac{K}{\ell^* - 1}} \left(\prod_{\ell \leq \ell^* - 1} \sum_{i \in \mathcal{I}_\ell} \sigma_{i\ell}^2\right)^{\frac{1}{\ell^* - 1}} \\ = & -\chi \sum_{i \in \mathcal{I}_{\ell^*}} \sigma_{i\ell^*}^2 + \ell^* \chi \sum_{i \in \mathcal{I}_{\ell^*}} \sigma_{i\ell^*}^2 - (\ell^* - 1) \chi \sum_{i \in \mathcal{I}_{\ell^*}} \sigma_{i\ell^*}^2 = 0. \end{aligned}$$

As a result, the sum of post-learning uncertainty, S, is continuous and increasing in the sum of of post-reporting uncertainty, $\sum_{i \in \mathcal{I}_\ell} \sigma_{i\ell}^2$, for any adopted method ℓ. Thus, for a given number of adopted methods, m, the optimal method choice is such that

$$\sum_{i\in\mathcal{I}_\ell}\sigma_{i\ell}^2=\begin{cases}n_\ell\sigma_0^2+n_\ell-1 & \text{for } \ell\le\ell^*,\\ \sigma_0^2 & \text{for } \ell>\ell^*,\end{cases}$$

where n_ℓ is the number of firms in each group such that $\sum_{\ell\le\ell^*}n_\ell=n-(m-\ell^*)$ and ℓ^* is endogenously determined according to *Proposition 1*. Hence, we have

$$\sum_{\ell\le\ell^*}\sum_{i\in\mathcal{I}_\ell}\sigma_{i\ell}^2=(n-m+\ell^*)\sigma_0^2+n-m,$$

$$\prod_{\ell\le\ell^*}\sum_{i\in\mathcal{I}_\ell}\sigma_{i\ell}^2\ge\left(\sigma_0^2\right)^{\ell^*-1}\left((n-m+1)\sigma_0^2+n-m\right).$$

In other words, the optimal method choice is such that $m-1$ firms adopt their own local methods, whereas the remaining $n-m+1$ firms adopt a common method characterized by their geometric center. That is, $S\ge\underline{S}(m)$, where

$$\underline{S}(m)=\begin{cases}S_L(m)=\Big((1-\chi)(m-1+A_m)+\chi(m-1+A_me^{-K})\Big)\tau_0^{-1} & \text{for } m\le\hat{m},\\ S_H(m)=\Big((1-\chi)(m-1+A_m)+\chi mA_m{}^{\frac{1}{m}}e^{-\frac{K}{m}}\Big)\tau_0^{-1} & \text{for } m>\hat{m},\end{cases}\tag{A7}$$

where $A_m\equiv n-m+1+(n-m)\tau_0$ and $\hat{m}\equiv n-\frac{e^K-1}{1+\tau_0}$. □

Proof of Proposition 3. Denote the net social benefit of adopting distinct methods relative to the adoption of m methods, where $1\le m\le n-1$, in terms of the reduction in the sum of post-learning uncertainty across all firms as

$$\Delta_{n,m}(K)\equiv\underline{S}(m;K)-\underline{S}(n;K)=\begin{cases}S_L(m;K)-\left(1-\chi+\chi e^{-\frac{K}{n}}\right)n\tau_0^{-1} & \text{if } K\le\hat{K}_m,\\ S_H(m;K)-\left(1-\chi+\chi e^{-\frac{K}{n}}\right)n\tau_0^{-1} & \text{if } K>\hat{K}_m,\end{cases}\tag{A8}$$

where $\hat{K}_m \equiv \log A_m$. Then

$$\begin{aligned}
\lim_{K\to 0} \Delta_{n,m}(K) &= n - m > 0, \\
\lim_{K\to\infty} \Delta_{n,m}(K) &= (1-\chi)(n-m) > 0, \\
\Delta'_{n,m}(K)\big|_{K<\hat{K}_m} &= \chi\left(e^{-\frac{K}{n}} - A_m e^{-K}\right)\tau_0^{-1} < 0, \\
\Delta'_{n,m}(K)\big|_{K>\hat{K}_m} &= \chi\left(e^{-\frac{K}{n}} - A_m{}^{\frac{1}{m}} e^{-\frac{K}{m}}\right)\tau_0^{-1} < 0 \iff K < \hat{K}_{n,m} \equiv \frac{n}{n-m}\log A_m.
\end{aligned}$$

Thus, $\Delta_{n,m}(K)$ is first decreasing on $(0, \hat{K}_{n,m})$ and then increasing on $(\hat{K}_{n,m}, \infty)$. Further note that

$$\Delta_{n,m}(\hat{K}_{n,m}) = (n-m)\left(1 - \chi - \chi e^{-\frac{\hat{K}_{n,m}}{n}}\tau_0^{-1}\right) \geq 0 \iff \chi \leq \frac{\tau_0}{e^{-\frac{\hat{K}_{n,m}}{n}} + \tau_0},$$

and

$$\begin{aligned}
\frac{\partial \hat{K}_{n,m}}{\partial m} &= \frac{n}{(n-m)^2}\kappa_n(m), \text{ where } \kappa_n(m) = \frac{1}{A_m} + \log A_m - 1, \\
\kappa'_n(m) &= -(n-m)\left(\frac{1+\tau_0}{A_m}\right)^2 < 0 \implies \kappa_n(m) > \kappa_n(n) = 0,
\end{aligned}$$

implying that $\hat{K}_{n,m}$ is increasing in m.

Therefore, whether the adoption of distinct methods will always be socially optimal depends on its comparison with the adoption of a common method. In particular, the adoption of distinct methods is always socially optimal if

$$\chi \leq \underline{\chi} \equiv \frac{\tau_0}{e^{-\frac{\hat{K}_{n,1}}{n}} + \tau_0} = \frac{\tau_0}{\big(n + (n-1)\tau_0\big)^{-\frac{1}{n-1}} + \tau_0}. \tag{A9}$$

Otherwise, it is socially optimal for $K \in (0, \underline{K}_n) \cup (\overline{K}_n, \infty)$, where

$$\underline{K}_n \equiv \min_{1\leq m\leq n-1} \underline{K}_{n,m} < \hat{K}_{n,1} = \frac{n}{n-1}\log\big(n + (n-1)\tau_0\big), \tag{A10}$$

$$\overline{K}_n \equiv \max_{1\le m\le n-1} \overline{K}_{n,m} > \hat{K}_{n,n-1} = n\log\left(2+\tau_0\right), \tag{A11}$$

with $\underline{K}_{n,m} < \hat{K}_{n,m} < \overline{K}_{n,m}$ being the two solutions to $\Delta_{n,m}(K) = 0$, if exist. □

Proof of Corollary 2. By *Corollary 1*, any deviation from the local method to a method that is not adopted by any firm makes the firm worse off. Now suppose firm i deviates to some other firm's local method i' and receives attention $\delta_{i'}$. Given that the adoption of distinct methods is socially optimal, it must yield a higher social welfare than all other method choices, including the one where all firms adopt the local method of firm i'. Since firm i' is strictly worse off when all firms each adopt their own local methods than when its local method is adopted as the common method, the other firms must be strictly better off. Therefore, we have

$$\left(1-\chi+\chi e^{-\frac{K}{n}}\right)\sigma_0^2 < \left(1-\chi+\chi e^{-K}\right)\left(\sigma_0^2+2\right) \le \left(1-\chi+\chi e^{-\delta_{i'}}\right)\left(\sigma_0^2+2\right),$$

implying that the firm will never deviate to method i'. □

Proof of Proposition 4. The proof follows a similar procedure as that of *Corollary 2*. The net social benefit of adopting a common method relative to the adoption of m methods, where $2 \le m \le n$, is

$$\Delta_{1,m}(K) \equiv \underline{S}(m;K) - \underline{S}(1;K) = \begin{cases} S_L(m;K) - \left(1-\chi+\chi e^{-K}\right)A_1\tau_0^{-1} & \text{if } K \le \hat{K}_m, \\ S_H(m;K) - \left(1-\chi+\chi e^{-K}\right)A_1\tau_0^{-1} & \text{if } K > \hat{K}_m. \end{cases} \tag{A12}$$

Then

$$\lim_{K\to 0} \Delta_{1,m}(K) = 1-m < 0,$$

$$
\begin{aligned}
&\lim_{K\to\infty} \Delta_{1,m}(K) = (1-\chi)(1-m) < 0,\\
&\Delta'_{1,m}(K)\big|_{K<\hat{K}_m} = \chi e^{-K}(m-1)(1+\tau_0)\tau_0^{-1} > 0,\\
&\Delta'_{1,m}(K)\big|_{K>\hat{K}_m} = \chi\Big(A_1 e^{-K} - A_m{}^{\frac{1}{m}} e^{-\frac{K}{m}}\Big)\tau_0^{-1} > 0 \iff K < \hat{K}_{1,m} \equiv \frac{1}{m-1}\log\frac{A_1{}^m}{A_m}.
\end{aligned}
$$

Thus, $\Delta_{1,m}(K)$ is first increasing on $(0, \hat{K}_{1,m})$ and then decreasing on $(\hat{K}_{1,m}, \infty)$. Further note that

$$
\Delta_{1,m}(\hat{K}_{1,m}) = (m-1)\Big(\chi A_1 e^{-\hat{K}_{1,m}} - (1-\chi)\tau_0\Big)\tau_0^{-1} > 0 \iff \chi > \frac{\tau_0}{A_1 e^{-\hat{K}_{1,m}} + \tau_0},
$$

and

$$
\begin{aligned}
&\frac{\partial \hat{K}_{1,m}}{\partial m} = \frac{1}{(n-1)^2}\kappa_1(m), \text{ where } \kappa_1(m) = \frac{A_1 - A_m}{A_m} + \log A_m - \log A_1,\\
&\kappa_1'(m) = (m-1)\left(\frac{1+\tau_0}{A_m}\right)^2 > 0 \implies \kappa_1(m) > \kappa_1(1) = 0,
\end{aligned}
$$

implying that $\hat{K}_{1,m}$ is increasing in m.

Therefore, whether the adoption of a common method will ever be socially optimal depends on its comparison with the adoption of distinct methods. In particular, the adoption of a common method is socially optimal if and only if

$$
\chi > \frac{\tau_0}{A_1 e^{-\hat{K}_{1,n}} + \tau_0} = \frac{\tau_0}{e^{-\frac{\hat{K}_{n,1}}{n}} + \tau_0} = \underline{\chi}, \tag{A13}
$$

where $\underline{\chi}$ is defined in the proof of *Proposition 3*, and $K \in (\underline{K}_1, \overline{K}_1)$, where

$$
\underline{K}_1 \equiv \max_{2\le m\le n} \underline{K}_{1,m} < \hat{K}_{1,n} = \frac{n}{n-1}\log\big(n + (n-1)\tau_0\big), \tag{A14}
$$

$$
\overline{K}_1 \equiv \min_{2\le m\le n} \overline{K}_{1,m} > \hat{K}_{1,2} = \log\frac{\big(n+(n-1)\tau_0\big)^2}{n-1+(n-2)\tau_0}, \tag{A15}
$$

with $\underline{K}_{1,m} < \hat{K}_{1,m} < \overline{K}_{1,m}$ being the two solutions to $\Delta_{1,m}(K) = 0$. Moreover, by the symmetry of function $\Delta_{m,m'}$ such that $\Delta_{1,n}(K) + \Delta_{n,1}(K) = 0$ and the definitions of $\underline{K}_n$ and $\overline{K}_n$ in equations (A10) – (A11), we have

$$\underline{K}_n \leq \underline{K}_{n,1} = \underline{K}_{1,n} \leq \underline{K}_1 < \overline{K}_1 \leq \overline{K}_{1,n} = \overline{K}_{n,1} \leq \overline{K}_n. \qquad \square$$

Proof of Proposition 5. When the adoption of a common method is socially optimal, it must be more efficient than the adoption of distinct methods; that is,

$$n\left(1-\chi+\chi e^{-K}\right)\left(\sigma_0^2+\frac{n-1}{n}\right) < n\left(1-\chi+\chi e^{-\frac{K}{n}}\right)\sigma_0^2 \tag{A16}$$

$$\iff B_s(K) \equiv \chi\Big(e^{-\frac{K}{n}} - \big(1+\frac{n-1}{n}\tau_0\big)e^{-K}\Big) > (1-\chi)\frac{n-1}{n}\tau_0. \tag{A17}$$

By the proof of *Proposition 4*, $B_s(K)$ is first increasing on $(0, \hat{K}_{1,n})$ and then decreasing on $(\hat{K}_{1,n}, \infty)$.

Consider firm i's deviation from the common method. By *Corollary 1*, if a firm ever finds it profitable to deviate from the common method, it will deviate to its local method which induces a post-learning variance

$$\hat{\sigma}_{ii}^2 = \begin{cases} \sigma_0^2 & \text{if } K \leq K_i, \\ \left(1-\chi+\chi e^{-\frac{K}{2}}\Big((n-1)\big(1+\frac{n-1}{n}\tau_0\big)\Big)^{\frac{1}{2}}\right)\sigma_0^2 & \text{if } K > K_i, \end{cases}$$

where $K_i = \log\Big((n-1)\big(1+\frac{n-1}{n}\tau_0\big)\Big)$.

When $K \leq K_i$, the firm will never deviate from the common method since

$$\sigma_0^2 > \left(1-\chi+\chi e^{-\frac{K}{n}}\right)\sigma_0^2 \geq \left(1-\chi+\chi e^{-K}\right)\left(1+\frac{n-1}{n}\tau_0\right)\sigma_0^2$$

by inequality (A16). When $K > K_i$, the firm does not have an incentive to deviate from the common method if

$$
\begin{aligned}
&\left(1-\chi+\chi e^{-\frac{K}{2}}\left((n-1)\big(1+\frac{n-1}{n}\tau_0\big)\right)^{\frac{1}{2}}\right)\sigma_0^2 \geq \left(1-\chi+\chi e^{-K}\right)\left(1+\frac{n-1}{n}\tau_0\right)\sigma_0^2 \\
\iff\ & 1-\chi+\chi e^{-\frac{K}{2}}\left((n-1)\big(1+\frac{n-1}{n}\tau_0\big)\right)^{\frac{1}{2}} \geq 1-\chi+\chi e^{-\frac{K}{n}} \\
\iff\ & K \leq K^* \equiv \frac{n}{n-2}\log\left((n-1)\big(1+\frac{n-1}{n}\tau_0\big)\right). \qquad \square
\end{aligned}
$$

Proof of Corollary 3. First, note that $\lim_{n\to 2} K^* = \infty > \lim_{n\to 3} K^*$. For $n \geq 3$, we have $\frac{\partial K^*}{\partial n} = \frac{q(n)}{(n-2)^2}$ where

$$
\begin{aligned}
q(n) &= n(n-2)\frac{1+\frac{(n-1)(n+1)}{n^2}\tau_0}{(n-1)\big(1+\frac{n-1}{n}\tau_0\big)} - 2\log\left((n-1)\big(1+\frac{n-1}{n}\tau_0\big)\right), \\
q'(n) &= \frac{n-2}{n}\frac{n(n-2)+2(n-1)(n-2)\tau_0+(n^2-2n-1)\big(\frac{n-1}{n}\big)^2\tau_0^2}{\left((n-1)\big(1+\frac{n-1}{n}\tau_0\big)\right)^2} > 0, \\
\frac{\partial q(n)}{\partial \tau_0} &= -\frac{1+2\big(\frac{n-1}{n}\big)^2\tau_0}{\big(1+\frac{n-1}{n}\tau_0\big)^2} < 0, \\
\lim_{\tau_0\to 0} q(n) &= \frac{n(n-2)}{n-1} - 2\log(n-1) > \lim_{\tau_0\to 0} q(2) = 0, \\
\lim_{\tau_0\to\infty} q(n) &= -\infty < 0.
\end{aligned}
$$

Thus, when τ_0 is sufficiently low, K^* always increases for $n \geq 3$. Otherwise, K^* first decreases and then increases. Combining this with the observation that K^* decreases from $n = 2$ to $n = 3$, we conclude that K^* first decreases and then increases in n in either case.

The second part of the corollary follows directly by inspecting the definition of K^*. $\square$

Proof of Proposition 6. When $K > K^*$, we have $B_s'(K) < 0$ since

$$
K^* - \hat{K}_{1,n} = n\left(\frac{\log(n-1)}{n-2} - \frac{\log n}{n-1} + \frac{\log\big(1+\frac{n-1}{n}\tau_0\big)}{(n-1)(n-2)}\right) > 0.
$$

To see this, consider $H(x) = \frac{\log(x+1)}{x}$ where $x > 0$. Then $H'(x) = \frac{h(x)}{x^2}$, where $h(x) = \frac{x}{x+1} - \log(x+1)$. Since $h'(x) = -\frac{x}{(x+1)^2} < 0$, we have $h(x) < h(0) = 0$ and hence $H'(x) < 0$.

When $K > K_i$, we have

$$\begin{aligned} B_i'(K) &= \chi e^{-\frac{K}{2}} \left(\big(1 + \frac{n-1}{n}\tau_0\big)e^{-\frac{K}{2}} - \frac{1}{2}\big((n-1)\big(1+\frac{n-1}{n}\tau_0\big)\big)^{\frac{1}{2}} \right) \\ &\sim \big(1 + \frac{n-1}{n}\tau_0\big)e^{-K} - \frac{1}{4}(n-1) \\ &< -\frac{(n-3)(n+1)}{4(n-1)} \leq 0. \end{aligned}$$

Recall from the proof of *Proposition 5* that $K_i < \hat{K}_{1,n} < K^*$. Hence, $B_s(K)$ is also decreasing on (K^*, ∞). In addition, note that

$$\begin{aligned} \lim_{K\to\infty} B_s(K) = \lim_{K\to\infty} B_i(K) = 0 < (1-\chi)\frac{n-1}{n}\tau_0, \\ B_s(K^*) = B_i(K^*) > (1-\chi)\frac{n-1}{n}\tau_0 \\ \iff \chi > \overline{\chi} \equiv \frac{\frac{n-1}{n}\tau_0}{e^{-\frac{K^*}{n}} - \left(1 + \frac{n-1}{n}\tau_0\right)e^{-K^*} + \frac{n-1}{n}\tau_0}, \end{aligned} \tag{A18}$$

where $\underline{\chi} < \overline{\chi} < 1$ because $B_s(\hat{K}_{1,n}) > B_s(K^*)$ and $K^* > \frac{n}{n-1}\log\left(1 + \frac{n-1}{n}\tau_0\right)$.

Therefore, by *Proposition 5*, a firm has a profitable deviation to its local method, despite the adoption of a common method being socially optimal, i.e., $B_i(K) < C < B_s(K)$, if and only if $K^\dagger < K < \overline{K}_{1,n}$, where $K^\dagger$ is the unique root of $B_i(K^\dagger) = C$ on (K^*, ∞) by the intermediate value theorem. □

References

Anand, Anita Indira (2006) 'An analysis of enabling vs. mandatory corporate governance structures post-Sarbanes-Oxley.' *Delaware Journal of Corporate Law* 31, 229

Arrow, Kenneth J (1974) *The Limits of Organization* (WW Norton & Company)

Arya, Anil, and Brian Mittendorf (2016) 'On the synergy between disclosure and investment beauty contests.' *Journal of Accounting and Economics* 61(2–3), 255–273

Barth, Mary E, Greg Clinch, and Toshi Shibano (1999) 'International accounting harmonization and global equity markets.' *Journal of Accounting and Economics* 26(1–3), 201–235

Barth, Mary E, Wayne R Landsman, and Mark H Lang (2008) 'International accounting standards and accounting quality.' *Journal of Accounting Research* 46(3), 467–498

Beatty, Anne, and Joseph Weber (2006) 'Accounting discretion in fair value estimates: An examination of sfas 142 goodwill impairments.' *Journal of Accounting Research* 44(2), 257–288

Beaver, William H (1977) 'The nature of mandated disclosure.' *Report of the SEC Advisory Committee on Corporate Disclosure* U. S. Government Printing Office (November), 618–656

_ (1998) *Financial Reporting: An Accounting Revolution* (Prentice Hall)

Bertomeu, Jeremy, and Edwige Cheynel (2013) 'Toward a positive theory of disclosure regulation: In search of institutional foundations.' *The Accounting Review* 88(3), 789–824

Bertomeu, Jeremy, and Robert P Magee (2015) 'Mandatory disclosure and asymmetry in financial reporting.' *Journal of Accounting and Economics* 59(2–3), 284–299

Bertomeu, Jeremy, Igor Vaysman, and Wenjie Xue (2021) ‘Voluntary versus mandatory disclosure.’ *Review of Accounting Studies* 26(2), 658–692

Bertomeu, Jeremy, Peicong Hu, and Yibin Liu (2023) ‘Disclosure and investor inattention: Theory and evidence.’ *The Accounting Review* 98(6), 1–36

Black, Dirk E, Theodore E Christensen, Jack T Ciesielski, and Benjamin C Whipple (2018) ‘Non-GAAP reporting: Evidence from academia and current practice.’ *Journal of Business Finance & Accounting* 45(3-4), 259–294

Blankespoor, Elizabeth, Ed deHaan, and Ivan Marinovic (2020) ‘Disclosure processing costs, investors’ information choice, and equity market outcomes: A review.’ *Journal of Accounting and Economics* 70(2–3), 101344

Bushman, Robert M, and Christopher D Williams (2012) ‘Accounting discretion, loan loss provisioning, and discipline of banks’ risk-taking.’ *Journal of Accounting and Economics* 54(1), 1–18

Campbell, John L, Kurt H Gee, and Zac Wiebe (2022) ‘The determinants and informativeness of non-GAAP revenue disclosures.’ *The Accounting Review* 97(7), 23–48

Chen, Qi, Tracy R Lewis, Katherine Schipper, and Yun Zhang (2017) ‘Uniform versus discretionary regimes in reporting information with unverifiable precision and a coordination role.’ *Journal of Accounting Research* 55(1), 153–196

Corona, Carlos, and Sang Wu (2024) ‘Heterogeneous interpretations of public disclosure and price efficiency.’ *Working Paper.* Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3388309

Corona, Carlos, Zeqiong Huang, and Hyun Hwang (2024) ‘Accounting uniformity, comparability, and resource allocation efficiency.’ *The Accounting Review* 99(1), 139–161

Cover, Thomas M (1999) *Elements of Information Theory* (John Wiley & Sons)

Cremer, Jacques, Luis Garicano, and Andrea Prat (2007) 'Language and the theory of the firm.' *The Quarterly Journal of Economics* 122(1), 373–407

Curtis, Asher B, Sarah E McVay, and Benjamin C Whipple (2014) 'The disclosure of non-GAAP earnings information in the presence of transitory gains.' *The Accounting Review* 89(3), 933–958

Dávila, Eduardo, and Cecilia Parlatore (forthcoming) 'Identifying price informativeness.' *Review of Financial Studies*

De George, Emmanuel T, Colin B Ferguson, and Nasser A Spear (2013) 'How much does IFRS cost? IFRS adoption and audit fees.' *The Accounting Review* 88(2), 429–462

De George, Emmanuel T, Xi Li, and Lakshmanan Shivakumar (2016) 'A review of the IFRS adoption literature.' *Review of Accounting Studies* 21(3), 898–1004

De Marco, Filippo, Marco Macchiavelli, and Rosen Valchev (2022) 'Beyond home bias: International portfolio holdings and information heterogeneity.' *Review of Financial Studies* 35(9), 4387–4422

Dessein, Wouter, Andrea Galeotti, and Tano Santos (2016) 'Rational inattention and organizational focus.' *American Economic Review* 106(6), 1522–36

Dyckman, Thomas R (1964) 'On the investment decision.' *The Accounting Review* 39(2), 285

Dye, Ronald A (1990) 'Mandatory versus voluntary disclosures: The cases of financial and real externalities.' *The Accounting Review* pp. 1–24

Dye, Ronald A, and Shyam Sunder (2001) 'Why not allow FASB and IASB standards to compete in the U.S.?' *Accounting Horizons* 15(3), 257–271

Dye, Ronald A, and Sri S Sridhar (2004) 'Reliability-relevance trade-offs and the efficiency of aggregation.' *Journal of Accounting Research* 42(1), 51–88

— (2008) 'A positive theory of flexibility in accounting standards.' *Journal of Accounting and Economics* 46(2-3), 312–333

Fang, Vivian W, Michael Iselin, and Gaoqing Zhang (2022) 'Consistency as a means to comparability: Theory and evidence.' *Management Science* 68(6), 4279–4300

FASB (2018) *Conceptual Framework for Financial Reporting* (Financial Accounting Foundation)

Fields, Thomas D, Thomas Z Lys, and Linda Vincent (2001) 'Empirical research on accounting choice.' *Journal of Accounting and Economics* 31(1-3), 255–307

Fishman, Michael J, and Kathleen M Hagerty (1989) 'Disclosure decisions by firms and the competition for price efficiency.' *The Journal of Finance* 44(3), 633–646

Flynn, Thomas D (1965) 'Uniformity in financial accounting: A progress report.' *Law and Contemporary Problems* 30, 623

Galperti, Simone, and Isabel Trevino (2020) 'Coordination motives and competition for attention in information markets.' *Journal of Economic Theory* 188, 105039

Gao, Pingyang (2008) 'Keynesian beauty contest, accounting disclosure, and market efficiency.' *Journal of Accounting Research* 46(4), 785–807

Gao, Pingyang, and Xu Jiang (2020) 'The economic consequences of discrete recognition and continuous measurement.' *Journal of Accounting and Economics* 69(1), 101250

Gao, Pingyang, Xu Jiang, and Gaoqing Zhang (2019) 'Firm value and market liquidity around the adoption of common accounting standards.' *Journal of Accounting and Economics* 68(1), 101220

Hail, Luzi, Christian Leuz, and Peter Wysocki (2010) 'Global accounting convergence and the potential adoption of IFRS by the US (Part I): Conceptual underpinnings and economic analysis.' *Accounting Horizons* 24(3), 355–394

Harsanyi, John C, and Reinhard Selten (1988) *A General Theory of Equilibrium Selection in Games* (The MIT Press)

Heflin, Frank, and Charles Hsu (2008) 'The impact of the SEC's regulation of non-GAAP disclosures.' *Journal of Accounting and Economics* 46(2-3), 349–365

Hinson, Lisa A, Gabriel P Pündrich, and Mark Zakota (2024) 'The decision-usefulness of asc 606 revenue disaggregation.' *The Accounting Review* 99(3), 225–258

Holthausen, Robert W, and Richard W Leftwich (1983) 'The economic consequences of accounting choice implications of costly contracting and monitoring.' *Journal of Accounting and Economics* 5, 77–117

Huang, Chong, Radhika Lunawat, and Qiguang Wang (2024) 'Disagreement about public information quality and informational price efficiency.' *Journal of Financial Economics* 152, 103762

IASB (2018) *Conceptual Framework for Financial Reporting* (IFRS Foundation)

Jensen, Robert E (1966) 'An experimental design for study of effects of accounting variations in decision making.' *Journal of Accounting Research* pp. 224–238

Jiang, John Xuefeng, Isabel Yanyan Wang, and Daniel D Wangerin (2018) 'How does the FASB make decisions? A descriptive study of agenda-setting and the role of individual board members.' *Accounting, Organizations and Society* 71, 30–46

Jiang, Xu, and Ming Yang (2017) 'Properties of optimal accounting rules in a signaling game.' *Journal of Accounting and Economics* 63(2–3), 499–512

Jiang, Xu, Chao Tang, and Gaoqing Zhang (2023) 'Harmonized accounting standards and investment beauty contests.' *The Accounting Review* 98(7), 377–404

Kamenica, Emir, and Matthew Gentzkow (2011) 'Bayesian persuasion.' *American Economic Review* 101(6), 2590–2615

Kanodia, Chandra (2007) 'Accounting disclosure and real effects.' *Foundations and Trends in Accounting* 1(3), 167–258

Kanodia, Chandra, and Haresh Sapra (2016) 'A real effects perspective to accounting measurement and disclosure: Implications and insights for future research.' *Journal of Accounting Research* 54(2), 623–676

Katz, Michael L, and Carl Shapiro (1985) 'Network externalities, competition, and compatibility.' *American Economic Review* 75(3), 424–440

_ (1986) 'Technology adoption in the presence of network externalities.' *Journal of Political Economy* 94(4), 822–841

Khan, Mozaffar, George Serafeim, and Aaron Yoon (2016) 'Corporate sustainability: First evidence on materiality.' *The Accounting Review* 91(6), 1697–1724

Kyle, Albert S (1985) 'Continuous auctions and insider trading.' *Econometrica* pp. 1315–1335

Leftwich, Richard (1980) 'Market failure fallacies and accounting information.' *Journal of Accounting and Economics* 2(3), 193–211

Leuz, Christian, and Peter D Wysocki (2016) 'The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research.' *Journal of Accounting Research* 54(2), 525–622

Liang, Pierre Jinghong, and Gaoqing Zhang (2019) 'On the social value of accounting objectivity in financial stability.' *The Accounting Review* 94(1), 229–248

Libby, Robert, Robert Bloomfield, and Mark W Nelson (2002) 'Experimental research in financial accounting.' *Accounting, Organizations and Society* 27(8), 775–810

Lu, Jinzhi (2022) 'Limited attention: Implications for financial reporting.' *Journal of Accounting Research* 60(5), 1991–2027

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt (2023) 'Rational inattention: A review.' *Journal of Economic Literature* 61(1), 226–273

Maines, Laureen A (1995) 'Judgment and decision-making research in financial accounting: A review and analysis.' In *Judgment and Decision-Making Research in Accounting and Auditing,* ed. Robert H. Ashton and Alison Hubbard Ashton (Cambridge University Press) pp. 76–101

Merino, Barbara D, and Teddy L Coe (1978) 'Uniformity in accounting: A historical perspective.' *Journal of Accountancy*

Myatt, David P, and Chris Wallace (2012) 'Endogenous information acquisition in coordination games.' *Review of Economic Studies* 79(1), 340–374

Nezafat, Mahdi, Mark Schroder, and Qinghai Wang (2017) 'Short-sale constraints, information acquisition, and asset prices.' *Journal of Economic Theory* 172, 273–312

Ray, Korok (2018) 'One size fits all? Costs and benefits of uniform accounting standards.' *Journal of International Accounting Research* 17(1), 1–23

Rubinstein, Ariel (2000) *Economics and Language: Five Essays* (Cambridge University Press)

SEC (2024) 'The Enhancement and Standardization of Climate-Related Disclosures for Investors: Release Nos. 33-11275; 34-99678.' Available at https://www.sec.gov/files/rules/final/2024/33-11275.pdf

Shannon, Claude Elwood (1948) 'A mathematical theory of communication.' *The Bell System Technical Journal* 27(3), 379–423

Sims, Christopher A (2003) 'Implications of rational inattention.' *Journal of Monetary Economics* 50(3), 665–690

Van Nieuwerburgh, Stijn, and Laura Veldkamp (2009) 'Information immobility and the home bias puzzle.' *The Journal of Finance* 64(3), 1187–1215

— (2010) 'Information acquisition and under-diversification.' *Review of Economic Studies* 77(2), 779–805

Watts, Ross L, and Jerold L Zimmerman (1986) *Positive Accounting Theory* (Prentice-Hall Inc)

— (1990) 'Positive accounting theory: A ten year perspective.' *The Accounting Review* pp. 131–156

Wiederholt, Mirko (2010) 'Rational inattention.' *The New Palgrave Dictionary of Economics (Online Edition ed.)* pp. 1–8

Wu, Sang (2024) 'Strategic communication of opposing opinions.' *Working Paper.* Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4013015

Wu, Sang, and Ronghuo Zheng (2025) 'Hierarchical standard setting and accounting comparability.' *Working Paper*

Wu, Sang, and Wenjie Xue (2023) 'Accounting comparability and relative performance evaluation by capital markets.' *Journal of Accounting and Economics* 75(1), 101535

Zhang, Gaoqing, and Ronghuo Zheng (2024) 'Optimal reporting systems in bank runs.' *The Accounting Review* 99(2), 457–481