ACCESS FLEXIBILITY WITH ESCALATION AND AUDIT[1]

Xia Zhao and M. Eric Johnson

Center for Digital Strategies
Tuck School of Business
Dartmouth College, Hanover NH 03755

X_Zhao3@uncg.edu; M.Eric.Johnson@darmouth.edu

Full paper: 5570 Words

Abstract

Managing information access in highly dynamic business environments is increasingly challenging. With thousands of employees accessing thousands of applications and data sources, managers strive to ensure the employees can access the information they need to create value while protecting information from misuse. We propose an access governance structure with escalation options, ensuring both flexibility and security of information systems. Using a game-theoretic approach, we show that properly coupling information access, audit, violation penalties and rewards can enable self-interested employees to access information in a timely manner, seizing business opportunities for the firm while managing security risks.

[1] This research was supported through the Institute for Security Technology Studies at Dartmouth College, under award Number 2006-CS-001-000001 from the U.S. Department of Homeland Security (NCSD). The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the Department of Homeland Security.

***Keywords**: Information security, access control, flexibility, audit, escalation*

1. Introduction

Pervasive and timely access to information is a source of competitive advantage for many firms such as investment banks, research laboratories, and hospitals. Technology has made information more available throughout and between organizations, enabling collaboration and fueling innovation. The literature on innovation has long discussed the benefits of free-flowing information, linking it to innovation productivity (e.g., Baker and Freeland 1972; Tsai, 2001; von Hippel 1994). Likewise, the services and supply chain literature have also extolled the benefits of increased information availability (e.g., Lee et al. 2000; Rathnam et al. 1995). With web-based tools linked to vast enterprise data sources, firms today have made much data and applications readily available to thousands of employees, business partners, and customers at very low cost. Thus, in environments where information can result in significant profits or is critical to outcome quality, firms are driven to invest in technologies that increase information availability.

Unfettered information access, however, can create significant security concerns, driving managers to constrict the availability of information. Such efforts become indispensible with the recent enforcement of many government regulations, such as Sarbanes-Oxley (SOX), Payment Card Industry Data Security Standard (PCI DSS), Health Insurance Portability and Accountability Act (HIPAA), Gramm-Leach-Bliley Act (GLBA), Personal Information Protection and Electronic Documents Act (PIPEDA), and the European Union Directive on Data Privacy (EU Directive), which all include language requiring firms to maintain some level of

access control. Driven by fears of data breaches, intellectual property losses, and compliance violations, firms are working to reduce information accesses through better controls and governance. Therefore, the role of access governance has become increasingly important in balancing security and availability.

Current practice of access governance focuses on the technical implementation of privileges and entitlement[2]. For example, access controls dictate user privileges to view a file, execute an application, share data with other agents, and so on. Users can only use data when they have the corresponding entitlements. By far, the firm's most important guideline of implementing access governance is to prevent misuse of data - either intentionally (such as using the data to make illegal stock trades) or unintentional (such as storing the data on device that is vulnerable to a security breach). One important criterion of access governance is known as "the rule of least privilege", i.e., each user is provided with the minimum entitlements needed to perform her/his task (Aveksa 2007). To ensure the rule of least privilege, an access control system must be customized and dynamically managed including five components—request, approve, administer, enforce and monitor. Specifically a user requests an entitlement; the owner (typically the business owner of the data) examines the request and then approves or rejects it; the administrator modifies the user's entitlements; the user accesses the resource and the system logs

[2] An entitlement is a resource that a person is authorized to access in a certain way; for example, "opening case files" might be an entitlement for application X. In practice, entitlement, privilege and permission are used interchangeably.

the user's activities; and the auditor examines the logs and evaluates users' activities. Figure 1 shows the access governance system with the rule of least privilege.

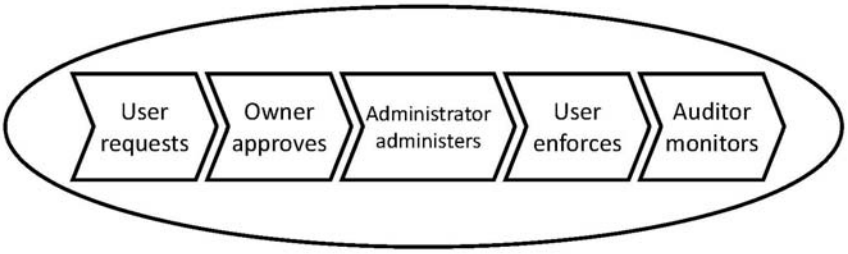

Figure 1. Access Governance System with the Rule of Least Privilege

To enforce the rule of least privilege, employees' accesses must be continually updated and audited to remain in synchronization with the changing organization. In large organizations with thousands of users interacting with thousands of different applications and data sources, each having many levels of privilege, the assignment and maintenance of access are daunting. The rule of least access is also limiting in many situations where it is difficult to foresee all information needs in advance. For example, in a hospital setting, emergencies arise where attending physicians may find themselves caring for another doctor's patient. In the increasingly dynamic environment, organizations frequently face unanticipated situations and have to adjust their organizational structures and personnel to adapt the consumers' needs. Rigid access control delays an organization's response to the changing markets, resulting in missed opportunities or degraded service quality.

In current practice, flexibility of access governance is sometimes achieved by overentitlement. In a field study of an investment bank, we found that 50-90% of employees are overentitled. This

outcome is rationalized by the argument that long-term employees are valuable and need quick access to information to create value for the firm. But, as the employees are permanently overentitled, they become larger security risks to the organization because their accesses could be used maliciously or accidentally. While the malicious insiders make the headlines (Jolly 2008), in many cases, benign overentitled employees pose a much larger risk to themselves and the organization because of secondary vulnerabilities like the loss of a laptop with sensitive data or because a malicious hacker could gain access to substantial firm information through a single overentitled account.

In an increasingly dynamic world, information governance must be flexible, yet secure. In this paper, we define access governance as an integrated system including policies, controls, incentives, and processes that manage user access to information resources. The goal of such access governance is to ensure the information systems to deliver the right information to the right people at the right time, but also protect the information from misuse, including security and privacy violations.

To achieve flexibility, we consider a different approach where employees are allowed to escalate into controlled data and applications when needed. This allows one-time access without any time-delaying approval process. In fact, we have witnessed cases where escalation is used to solve a failure of traditional access control system. For example, the investment banking sector refers to such an approach as “override” (Rissanen et al. 2004), and the health care sector refers to it as “break glass” (Ferreira et al. 2006). Escalation potentially breeds significant security risks since employees may abuse their ability to access information. For example, accessing

information not for business reasons but rather for personal benefit. To mitigate the associated security risks, the escalation activities are later audited, and employees found to be abusing their accesses are penalized. Auditing (or monitoring) with violation penalties have been implemented by firms seeking to drive desired behavior from employees or partners with respect to financial reporting, contract and regulation compliance. For example, Intel issues "speeding tickets" to employees that violate information security policies. In addition to penalties, we also consider the possibility that the firm uses rewards to motivate employees.

In this paper, we design an access governance policy with escalation options which couples escalation accesses with rewards, audit and violation penalties. We use a game-theoretic model to analyze the employees' incentives and the firm's policy design problem. The results show that a properly designed governance policy could provide the desired access flexibility with a significant level of control. Figure 2 shows the information governance system with escalation.

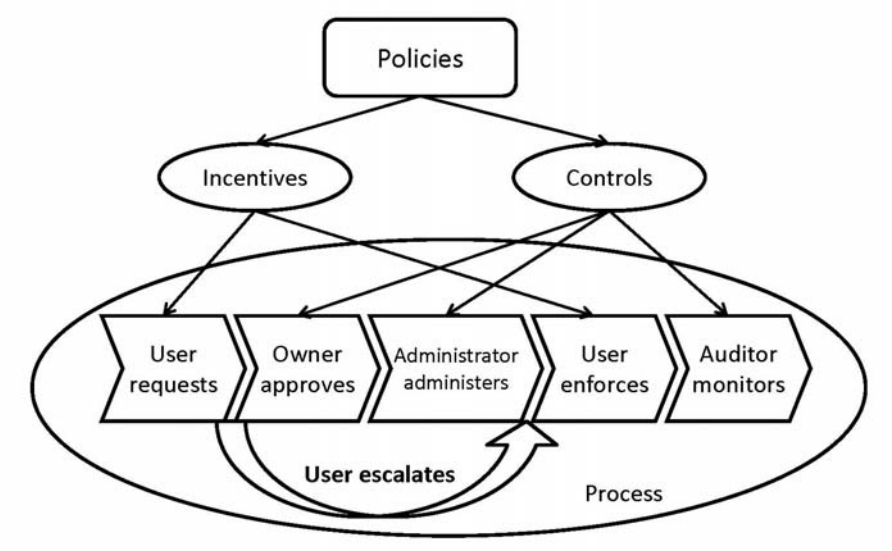

Figure 2. Access Governance System with Escalation

Of course, escalation must be confined to cases where the risk of failure or the cost of recovery is relatively low compared to the cost of not granting access (e.g., the potential value created through escalation). It may not be suited to some financial or trading systems where there is significant risk of massive fraud. Rather it is useful in cases where there are small risks or where the potential value of business opportunities is very high. For example, escalation is very effective in situations where emergency access may save someone's life, or in a time-critical system where the person with the necessary privileges may be unavailable (Povey 2000).

The paper is organized as follows. In Section 2, we review the related literature. In Section 3 and Section 4, we outline the model and analyze the game. We capture the important characteristics of the optimal access governance policy with escalation options. Finally we conclude with implementation guidance in Section 5.

2. Related Literature

The technological aspect of implementing escalation in access control has been studied in computer science literature. Povey (2000) broadly discussed an optimistic access control scheme with escalation and developed a formal model to ensure the integrity of computer systems including accountability, auditability and recoverability. Rissanen et al. (2004) emphasized the importance of audit and manual recovery in providing overriding of access control. Ferreira et al. (2006) described the design and initial implementation of a "Break-The-Glass" policy in a virtual Electronic Medical Record system. Our paper focuses on the economic aspect of the access

governance with escalation and uses a principal and agent setting to study the policy design problem.

Principle and agent models have been examined in a variety of contexts (e.g. Antle and Eppen 1985; Arrow 1985; Baiman 1990; Harris and Raviv 1979; Harris et al. 1982; Holmstrom 1979; Shavell 1979, etc.). Our paper closely relates to a large stream of literature which studies the audit policy in the principal and agent framework (Baron and Besanko 1984; Dye 1986; Harris and Raviv 1996; Kim and Suh 1992; Townsend 1979). Townsend (1979) was one of the first models to examine the costly verification. Dye (1986) showed that optimal monitoring policies are deterministic and lower-tailed. Kim and Suh (1992) also focused on the deterministic monitoring policy in which the optimal investment in audit technology is endogenously determined. They found the lower-tailed policy is one of the special cases. Baron (1984) investigated the random audit policy in a regulatory pricing problem. Firms are privately informed about their cost functions and required to report them to the regulator. Baron (1984) found that the optimal audit policy includes terms that firms may be penalized even though they report their best knowledge because of ex post uncertainty. And Harris and Raviv (1996) explored the random audit policy in the capital budgeting process and identified cases of overinvestment as well as underinvestment. In our paper, we characterize the optimal audit scheme which helps the firm achieve a significant level of flexibility at some expense of security risks.

3. Model

We consider the case where users gain access to data and applications through a system employing access control. We focus on the firm's optimal strategy in cases where there are only a few, discrete situations where employees may need more access – for example, when their boss is on vacation. In those situations, firms may allow employees to escalate access but then audit their actions (at a cost) afterward and penalize employees for misuse or reward for value generation.

We model the collection of applications and data as measured on a continuous scale of information, with each privilege weighted to reflect the amount and sensitivity of the data. Based on value generated by an employee and the associated information risk, the firm assigns the employee a regular access level to perform routine tasks.

Periodically employees may face an opportunity to create more value by accessing information beyond her/his regular access. We assume that with probability π_h (or π_l), an employee will observe such an opportunity with high (or low) revenue potential; with probability $\pi_0 = 1 - \pi_h - \pi_l$, s/he does not observe any opportunity. We refer to these situations as the high state, denoted as θ_h, low state θ_l and regular state θ_0. We assume $\theta_h > \theta_l > \theta_0 = 0$. We use a to denote the access level. The firm allows employees to escalate their access levels temporarily to seize the business opportunities. The net revenue from a business opportunity is determined by $\theta_i\,(i = 0, l, h)$ and the employee's escalated access level *a*, i.e., $U(\theta_i, a)$. Access control, while

providing a measure of security, restricts employees' flexibility to monetize the business opportunities. Therefore the more access rights an employee has, the more likely that s/he creates value for the firm. We assume that $U(\theta_i, a)$ is an increasing and concave function of *a*. $U'_a > 0$ and $U''_a \leq 0$. This is a reasonable assumption as increased availability of information can increase revenue generating potential, but is eventually limited by the skill and knowledge of the employee. The impact of flexibility on firm revenue is more significant when the firm observes a higher revenue potential than when it observes a lower revenue potential. Therefore, we assume that the marginal revenue of the information access in a higher state is larger than that in a lower state, $U''_{\theta a} > 0$. Figure 3 shows an example of the firm's revenue functions from emergent opportunities in three states.

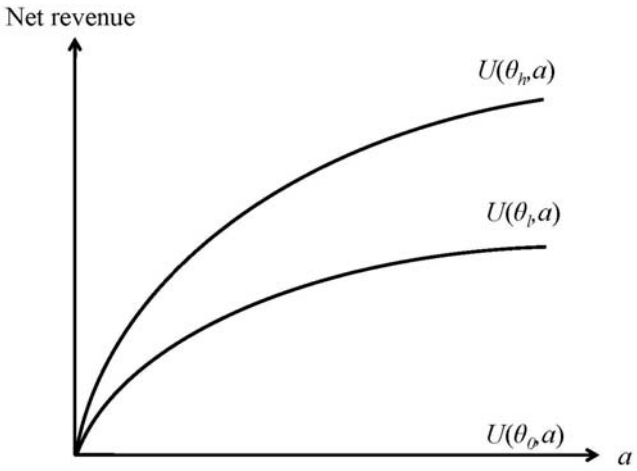

Figure 3. Firm's Revenue Functions in Three States

The firm bears costs associated with the escalation access level of $C(a)$ including additional security risks and routine technical support required to prudently maintain that access. $C(a)$ is

an increasing and convex function. $C_a' > 0$ and $C_a'' > 0$. This well-models the case where providing far too much access can eventually result in severe consequences (risks and cost to mitigate risk). To mitigate risk of unnecessary escalation, the firm controls the escalation flexibility and audits each instance of escalation. In particular, the firm offers three escalation options, $\{a_0, a_l, a_h\}$, corresponding to the states $\{\theta_0, \theta_l, \theta_h\}$ and motivates employees to choose the escalation access level a_i when the state θ_i arises.

Employees derive some private benefit by accessing information and data and prefer higher access levels to lower ones. Such "snooping" value is not uncommon - we have witnessed cases in health care, providers may examine the records of a patient for her/his own benefit. The employee's private benefit from escalation is $u(a)$. $u(a)$ is an increasing and concave function. $u_a' > 0$ and $u_a'' \le 0$. Since some employees may take advantage of the flexibility and not choose the right escalation options (choosing a_i in the state θ_j, $j \ne i$), the firm audits the instances of escalation at a cost and penalizes the escalation misuse. It is assumed that the firm can detect misuse with probability p by investing $D(p)$ in the audit capability. The audit spending includes hiring auditors, tracking escalation instances, and verifying the business opportunities by communicating with the manager or coworkers of the employees. We refer to *p* as the audit precision. $D(p)$ is an increasing and convex function. $D_p' > 0$ and $D_p'' > 0$. The employee will be penalized at the level of *F* if s/he is detected to misrepresent the state that s/he observes. We assume that the maximal violation penalty is $\overline{F}$. Without loss of generality, we assume that if an employee choice is consistent with the state (choosing a_i in the state θ_i), there is no audit error,

i.e., $p = 0$. In addition to audit and penalties, the firm may reward employees for choosing the right escalation options. w_i is used to denote the reward based on the escalation options the employee chooses. The audit precision, penalty and reward can be contingent on employees' choices. The firm maximizes its expected profit by designing an access governance policy with escalation options $\{(a_i, w_i, p_i, F_i) \mid i = 0, l, h\}$.

The sequence of events is showed in Figure 4. We use one employee as an example. At stage 1, the firm announces its access governance policy with escalation options; At stage 2, an employee observes the state and then chooses an escalation option; Finally, the firm audits the escalation instance, rewarding or penalizing the employee according to the announced access governance policy**.**

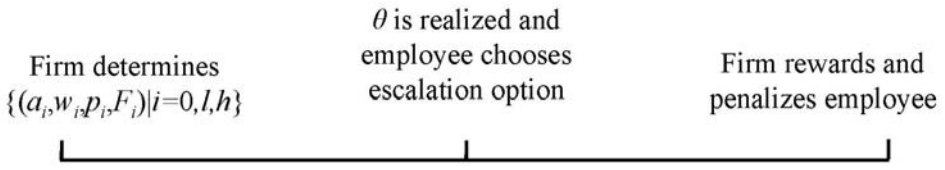

Figure 4. The Sequence of Events

The employee's expected payoff, denoted by $\Pi_{employee}$, can be represented by

$$\Pi_{employee} = \begin{cases} w_i + u(a_i) & \text{if s/he chooses } a_i \text{ when } \theta=\theta_i \\ -p_i F_i + (1-p_i) w_i + u(a_i) & \text{if s/he chooses } a_i \text{ when } \theta=\theta_j,\ j \neq i \end{cases} \qquad i, j = 0, l, h$$

The firm's expected profit is $E\left[U(\theta_i, a_i) - C(a_i) - D(p_i) - w_i\right]$. Let Π_{firm} be the maximum expected profit obtained by the following optimization problem.

$$\begin{aligned} \Pi_{firm} = & \max_{a_i, w_i, F_i, p_i} E\left[U(\theta_i, a_i) - C(a_i) - D(p_i) - w_i\right] \\ s.t.\ \ & w_i + u(a_i) \geq -p_j F_j + (1-p_j) w_j + u(a_j),\ j \neq i, \text{ if } \theta = \theta_i \qquad \text{(IC)} \\ & w_i + u(a_i) \geq 0 \qquad \text{(IR)} \\ & w_i \geq 0, a_i \geq 0, 0 \leq p_i \leq 1, 0 \leq F_i \leq \bar{F}, i = 0, l, h \end{aligned}$$

where (IC) are the employee's incentive constraints and (IR) are the employee's individual rationality constraints.

4. Analysis and Results

To gain managerial insight, we analyze the following (tractable) functional forms. We assume that the firm's revenue function is linear, $U(\theta_i, a) = \theta_i a,\ (i = 0, l, h)$ where θ_i represents the firm's marginal revenue of information access. The employee's private benefit function is also linear, $u(a) = ba$ where b is the employee's marginal private benefit of information access. The assumption of linear revenue and private benefit functions does not result in any loss of generality because the firm can always redefine the map between the collection of applications and data and the continuous scale of information, and transform the relationship between the benefit and information access to a linear one. We assume the cost functions are quadratic, $C(a) = \frac{1}{2} s a^2, s > 0$. Besides the frequent use of convex cost functions in the literature (e.g.,

Kannan and Telang 2005; Krishnan and Zhu 2006; Motta 1993), quadratic cost functions nicely capture the higher security risks associated with higher access as well as the cost of additional IT resources for maintaining access. Similarly, the audit cost function is $D(p) = \frac{1}{2}tp^2, t > 0$, which reflects the increasing difficulty of improving the audit precision.

4.1 Benchmark Case

We first consider a benchmark case where there is no information asymmetry between the firm and employees. The firm can directly observe the states (i.e. an opportunity with high revenue potential, an opportunity with low revenue potential, or no business opportunity) and assign the access levels to employees. In this case the firm does not need to implement any incentive scheme (neither reward nor penalty). The firm's optimization problem can be represented by

$$\Pi_{optimal} = \max_{a_i, i=0,l,h} E\left[\theta_i a_i - \frac{1}{2} s a_i^2\right]$$

The optimal access level is given by $a_i = \frac{1}{s}\theta_i$ $(i = 0, l, h)$. When the firm observes a business opportunity with high revenue potential, it will assign $\frac{1}{s}\theta_h$ to the employee; when it observes an opportunity with low revenue potential, it will assign $\frac{1}{s}\theta_l$ to the employee; otherwise, it will not assign any additional access to the employee. The firm's optimal profit is

$$\Pi_{optimal} = \frac{1}{2s}\left(\pi_h \theta_h^2 + \pi_l \theta_l^2\right).$$

4.2 Asymmetric Information

When there is information asymmetry between the firm and employees, the firm will design the escalation options in a way that the employee will choose the right option in each state, i.e. the employee will choose a_i if the state is θ_i. Therefore, the employee's incentive constraints are as follows.

High state: $\begin{cases} w_h + ba_h \geq -p_l F_l + (1-p_l) w_l + ba_l & (IC-HL) \\ w_h + ba_h \geq -p_0 F_0 + (1-p_0) w_0 + ba_0 & (IC-H0) \end{cases}$

Low state: $\begin{cases} w_l + ba_l \geq -p_h F_h + (1-p_h) w_h + ba_h & (IC-LH) \\ w_l + ba_l \geq -p_0 F_0 + (1-p_0) w_0 + ba_0 & (IC-L0) \end{cases}$

Regular state: $\begin{cases} w_0 + ba_0 \geq -p_h F_h + (1-p_h) w_h + ba_h & (IC-0H) \\ w_0 + ba_0 \geq -p_l F_l + (1-p_l) w_l + ba_l & (IC-0L) \end{cases}$

The first (or second) group of incentive constraints is for employees who observe business opportunities with high (or low) revenue potential. The third group of incentive constraints is for employees who do not observe any business opportunity. Presumably, if employees do not observe any business opportunity, the firm should not allow them to escalate, i.e. $a_0 = 0$. We do not impose this constraint in order to identify a better solution which gives the firm a higher profit.

Properly designed escalation options will induce employees to choose the right escalation levels and hence disclose their observations. We focus on the case where $p_0 = 0$. That is, if an employee chooses the option a_0 and s/he claims that s/he does not observe any business

opportunity, the firm does not audit such instances. Given that escalation is used to handle unusual situations, it is reasonable that the firm does not investigate the regular states[3]. We can substitute $p_0 = 0$ into (IC-H0) and (IC-L0) and obtain the following inequalities.

$$w_h + ba_h \geq w_0 + ba_0 \qquad \left(IC - H0^*\right)$$

$$w_l + ba_l \geq w_0 + ba_0 \qquad \left(IC - L0^*\right)$$

It is easy to find that (IC-HL), (IC-LH) and (IR) are not binding. For example (IC-HL) is implied by (IC-H0*) and (IC-0L).The firm's optimization problem can be simplified as

$$\begin{aligned}
\Pi_{firm} = \max_{\substack{a_h, w_h, p_h, F_h \\ a_l, w_l, p_l, F_l \\ a_0, w_0}} & \pi_h\left(\theta_h a_h - \tfrac{1}{2} s a_h^2 - \tfrac{1}{2} t p_h^2 - w_h\right) + \pi_l\left(\theta_l a_l - \tfrac{1}{2} s a_l^2 - \tfrac{1}{2} t p_l^2 - w_l\right) + \pi_0\left(-\tfrac{1}{2} s a_0^2 - w_0\right) \\
s.t. \quad & w_h + ba_h - w_0 - ba_0 \geq 0 && \text{(IC-H0}^*\text{)} \\
& w_l + ba_l - w_0 - ba_0 \geq 0 && \text{(IC-L0}^*\text{)} \\
& w_0 + ba_0 + p_h F_h - (1 - p_h) w_h - ba_h \geq 0 && \text{(IC-0H)} \\
& w_0 + ba_0 + p_l F_l - (1 - p_l) w_l - ba_l \geq 0 && \text{(IC-0L)} \\
& w_0, w_l, w_h \geq 0, a_0, a_l, a_h \geq 0, 0 \leq p_l, p_h \leq 1, 0 \leq F_l, F_h \leq \overline{F}
\end{aligned} \qquad (1)$$

Proposition 1*: If the firm detects that an employee misrepresents her/his observation, the firm will penalize her/him to the maximal level. i.e.,* $F_h = F_l = \overline{F}$.

(See Appendix for all proofs.)

---

[3] In practice, periodical entitlement reviews may be conducted to examine employees' regular access rights and en sure that employees have the adequate access rights to accomplish their tasks.

The audit and associated violation penalties deter employees from mirepresenting the business opportunities they observe. Since the firm does not incur any cost by penalizing employees after it detects misuse, it always penalizes them to the maximal level to reduce the audit spending. Proposition 2 characterizes the escalation options when the penalty can be extremely harsh, i.e. $\bar{F} \to \infty$.

Proposition 2: *If* $\bar{F}$ *approaches infinity, the firm only offers two options,* $\left\{\left(a_i, w_i, p_i, \bar{F}\right) | \, i = l, h\right\}$. *In particular,* $a_h = \frac{1}{s}\theta_h$, $a_l = \frac{1}{s}\theta_l$, $w_h = w_l = 0$, $p_h = p_l = \varepsilon \to 0$. *And the firm can achieve the optimal profit,* $\frac{1}{2s}\left(\pi_h \theta_h^2 + \pi_l \theta_l^2\right)$.

If the firm can render extreme penalties, for detected misuse, employees have no incentive to misrepresent their observations even though there is only a slight chance of being detected. The firm does not need to offer any additional information access to employees who do not observe any business opportunity. The firm can design the escalation options with a very low audit precision and no reward.

However, an infinite penalty is impossible to implement, e.g., the firm cannot take an employee's life. Next we consider the situation that there is an upper bound for the penalty. To avoid trivial cases, we assume that the difference between θ_h and θ_l is greater than $\frac{b}{\pi_h}$, i.e. $b < \pi_h\left(\theta_h - \theta_l\right)$, and that the audit is so costly that it is always not optimal for the firm to invest to achieve audit precision of $p = 1$. Proposition 3 characterizes the escalation options.

Proposition 3: *The optimal solutions of the optimization problem (1) are* $\{(a_i, w_i, p_i, F_i) \mid i = 0, l, h\}$ *where* $a_h = \frac{1}{s}\left(\theta_h - \frac{1-\pi_h}{\pi_h} b\right)$, $w_h = 0$, $p_h = \frac{1-\pi_h}{t\pi_h}\overline{F}$, $F_h = \overline{F}$, $a_l = \frac{1}{s}(\theta_l + b)$, $p_l = 0$, $w_l = -\frac{1}{s}b^2 + \frac{b}{s}(\theta_h - \theta_l) - \frac{1-\pi_h}{st\pi_h}\left(s\overline{F}^2 + tb^2\right)$, $a_0 = \frac{1}{s}b$, $w_0 = -\frac{1}{s}b^2 + \frac{b}{s}\theta_h - \frac{1-\pi_h}{st\pi_h}\left(s\overline{F}^2 + tb^2\right)$, $p_0 = 0$, $F_l, F_0 \in \left[0, \overline{F}\right]$[4].

(1) Information access: the access level for the business opportunity with high revenue potential is lower than the optimal level in the benchmark case (underentitlement); and the access levels for the low revenue potential and no business opportunity are higher than the optimal ones (overentitlement).

(2) Audit: the firm audits the escalation instances with higher precision if the employees choose a higher escalation level than it does if the employees choose a lower escalation level.

(3) Reward: The firm does not reward employees who choose the highest escalation level but rewards employees who choose the other two escalation levels.

The game has separating equilibria in which the firm offers the escalation options as proposition 3 presents and employees choose different escalation levels for different states.

The access levels, rewards, audit and violation penalties together motivate employees to escalate information access when necessary without the long-term security risks of overentitlement. The

---

[4] Since F_l and F_0 can be any value in the range of $\left[0, \overline{F}\right]$, the problem has infinite optimal solutions. However, F_l and F_0 do not matter because of $p_l = p_0 = 0$. We can regard this problem has a unique optimal solution.

access levels in the escalation options in the asymmetric information case deviate from the optimal ones in the benchmark case. The firm designs the escalation options in this way to save spending in audit capability and employee rewards. Consequently, it forgoes some revenue through underentitlement and voluntarily bears extra costs through overentitlement. It is counterintuitive that the firm maximizes its profit by allowing employees who do not observe any business opportunity to access extra information. It is worth remarking that designing escalation options with no escalation in the regular state is feasible (by solving the optimization problem (1) with an additional constraint that $a_0 = 0$). However, such a policy results in a lower profit.

It must be recognized that zero audit precision does not mean that the firm never audits escalation instances at all. The audit precision captures the level of additional time and effort by the firm in investigating the escalation instances compared to regular information access services. The firm should pay additional attention to the instances of high escalation level and handle other escalation instances as the regular services such as granting regular information access.

From the employees' perspective, the employees who observe high revenue potential obtain a high information access and generate high private benefit without the risk of being penalized. They will not choose other escalation options even though they will not be rewarded by the firm. The employees who observe lower revenue potential or no business opportunity are deterred

from over-claiming their observations by the audit possibility and potential penalties. They are also compensated by the firm through rewards for disclosing their observations.

The firm makes a positive profit by offering the escalation options in the asymmetric information case, $\Pi_{firm} = \frac{1}{2s}\left(\pi_h\theta_h^2 + \pi_l\theta_l^2 + (1-\pi_h)b^2 + 2b\pi_l\theta_l - 2(1-\pi_h)b\theta_h + \frac{(1-\pi_h)^2}{t\pi_h}\left(s\bar{F}^2 + tb^2\right)\right) > 0$, which justifies the provision of the escalation options. However, the firm's profit in the asymmetric information case is lower than the optimal profit in the benchmark case for three reasons. First, the access levels in the escalation options deviate from the optimal access levels in the benchmark case (overentitlement or underentitlement); second, the firm has to invest in audit capability; finally the firm shares its profit with the employees through rewards. The profit difference between the benchmark case and asymmetric information case is the value of information, i.e. how much the firm is willing to pay to observe the business opportunities ex ante.

$$\Delta\Pi = \Pi_{optimal} - \Pi_{firm} = \frac{1}{2s}\left(-(1-\pi_h)b^2 - 2\pi_l b\theta_l + 2(1-\pi_h)b\theta_h - \frac{(1-\pi_h)^2}{t\pi_h}\left(s\bar{F}^2 + tb^2\right)\right) > 0.$$

Proposition 4 summarizes some comparative statics.

Proposition 4: *The value of information is increasing in* θ_h*, b and t and decreasing in* θ_l *and s.*

The value of information is higher when it is more costly for the firm to motivate employees to disclose their observations. When the marginal revenue of information access in the high state is higher, the difference between the escalation access for the high state and that for the other two states is larger. Employees observing low business opportunities or no business opportunity are

more likely to cheat. The firm needs to reward more and/or audit with higher precision to prevent such behavior. When the marginal revenue of information access in the low state is higher, the escalation access for the low state is closer to that for the high state. Employees have less incentives to pretend to have observed a high business opportunity. Consequently, the value of information is lower.

The higher marginal private benefit of information access is, the more the employees have incentives to cheat. The firm has to distort the escalation access more and offer higher reward to drive employees to report truth, resulting in higher costs. Therefore, the value of information increases.

On the cost side, higher audit cost reduces the firm's capability to detect cheating behavior, which makes the information more valuable. The cost of security risks, on the other hand, reduces the value of information. The increase of security risks associated with additional information access lowers the firm's willingness to offer higher escalation accesses. Therefore, the differences between the information accesses for different states are lowers, which reduces employees' incentives to cheat and makes it easy for the firm to motivate employees.

5. Conclusion

Using game-theoretic analysis, we have shown how the firm can encourage value creation through flexible access governance, while controlling information misuse. By properly designing the access governance with escalation options, the firm seizes every business opportunity

without bearing significant security risks. Escalation levels, rewards, audit and violation penalties together provide employees with incentives to escalate their information accesses to the appropriate levels. Our analysis provides many interesting insights into the implementation challenges of access governance with escalation.

1. The firm should consider providing employees with more information access in escalation options than strictly needed because of information asymmetry. Such a strategy is optimal in that the firm can take advantage of the employees' private benefit to save audit expenditure and rewards. The proposed scheme does not imply that the firm should offer three escalation options, with employees escalating no matter whether there is a business opportunity or not. The firm can set two options instead of three, assigning the escalation level with additional access in place of the regular level (and thus freeing them from escalating from time to time when there is no business opportunity).
2. Controls are critical for the successful implementation of the escalation scheme. Escalation must be done within the allowable zone dictated by regulatory requirements. Some data or applications cannot be made available through an escalation scheme. By providing options with predefined access levels, the firm controls the limit for escalation.
3. Audit quality is an important element of our governance scheme. Without the ability to catch cheaters (i.e. the audit cost is extremely high), firms are better-off moving towards a more traditional rigid role-based access approach. Escalation must be done in a way that provides an audit trail, including records of who requested it, when, what data was accessed, and what value was created (e.g., the type of transaction being performed) (Rissanen et al., 2004). Nevertheless, perfect monitoring is technologically challenging or financially undesirable in

most cases. This study provides managers guidance on balancing the audit expenditure and the security risks.

4. Penalty instruments need not be monetary or be directly levied against the employees. For example, operational penalties could be very effective, such as mandatory attendance at compliance training for violators or requiring employees to file reports for the illegitimate escalation. We have also observed cases where the security fines were levied against the employees' manager, highlighting the manager's responsibility for training.

5. The firm needs to know employees' private benefit to properly design the escalation options. It is important for the firm to learn employees' characteristics over time or through other approaches, and only grant escalation flexibility to known employees.

6. The value of the access governance system with escalation options also includes the possibility that the firm learns the dynamics of the business environment from employees. Sometime the firm is unaware of potential business opportunities simply because employees forwent them. The escalation scheme creates an implicit communicate channel between the firm and employees. It is also possible for the firm to spot trends that could identify a potentially malicious insider. Finally, it can be very helpful in establishing regular access levels and understanding how employees' roles change over time (sometimes referred to as role drift). By observing employees' needs over time, the firm can adjust their regular accesses accordingly.

Appendix

Proof of Proposition 1

Proof: Since larger F_h and F_l make (IC-0H) and (IC-0L) easier to hold and F_h and F_l do not appear in the firm's expected profit function, the firm maximizes its profit by imposing the maximal level of penalty.

Proof of Proposition 2

Proof: If $\bar{F} \to \infty$, (IC-0H) and (IC-0L) are not binding if $p_h = p_l > 0$. The optimization problem can be simplified as

$$\Pi_{firm} = \max_{\substack{a_h, w_h, p_h \\ a_l, w_l, p_l \\ a_0, w_0}} \pi_h\left(\theta_h a_h - \tfrac{1}{2} s a_h^2 - \tfrac{1}{2} t p_h^2 - w_h\right) + \pi_l\left(\theta_l a_l - \tfrac{1}{2} s a_l^2 - \tfrac{1}{2} t p_l^2 - w_l\right) + \pi_0\left(-\tfrac{1}{2} s a_0^2 - w_0\right)$$

$$\begin{aligned} s.t. \quad & w_h + b a_h - w_0 - b a_0 \geq 0 && \text{(IC-H0}^*\text{)} \\ & w_l + b a_l - w_0 - b a_0 \geq 0 && \text{(IC-L0}^*\text{)} \\ & w_0, w_l, w_h \geq 0, a_0, a_l, a_h \geq 0, 0 < p_l, p_h \leq 1 \end{aligned}$$

Smaller a_0 and w_0 make the (IC-H0*) and (IC-L0*) easy to hold and increase the firm's expected profit, the firm will set $a_0 = w_0 = 0$. (IC-H0*) and (IC-L0*) are not binding. The firm's optimization problem can be further simplified as

$$\Pi_{firm} = \max_{\substack{a_h, w_h, p_h \\ a_l, w_l, p_l}} \pi_h\left(\theta_h a_h - \tfrac{1}{2} s a_h^2 - \tfrac{1}{2} t p_h^2 - w_h\right) + \pi_l\left(\theta_l a_l - \tfrac{1}{2} s a_l^2 - \tfrac{1}{2} t p_l^2 - w_l\right)$$

$$s.t. \quad w_l, w_h \geq 0, a_l, a_h \geq 0, 0 < p_l, p_h \leq 1$$

We can obtain $a_h = \frac{1}{s}\theta_h$, $a_l = \frac{1}{s}\theta_l$, $w_h = w_l = 0$. Since the penalty is effective only if misuse can be detected, the firm has to audit escalation instances. $p_h = p_l = \varepsilon \to 0$. The firm's profit approaches $\frac{1}{2s}\left(\pi_h\theta_h^2 + \pi_l\theta_l^2\right)$.

Proof of Proposition 3

Proof: The Lagrangian of the firm's optimization problem can be represented as

$$\begin{aligned} L = {} & \pi_h\left(\theta_h a_h - \tfrac{1}{2}sa_h^2 - \tfrac{1}{2}tp_h^2 - w_h\right) + \pi_l\left(\theta_l a_l - \tfrac{1}{2}sa_l^2 - \tfrac{1}{2}tp_l^2 - w_l\right) + \pi_0\left(-\tfrac{1}{2}sa_0^2 - w_0\right) \\ & + \lambda_1\left(w_h + ba_h - w_0 - ba_0\right) + \lambda_2\left(w_l + ba_l - w_0 - ba_0\right) \\ & + \lambda_3\left(w_0 + ba_0 + p_h\overline{F} - \left(1 - p_h\right)w_h - ba_h\right) + \lambda_4\left(w_0 + ba_0 + p_l\overline{F} - \left(1 - p_l\right)w_l - ba_l\right) \end{aligned}$$

FOC w.r.t. a_h : $\pi_h\left(\theta_h - sa_h\right) - \lambda_1 b - \lambda_3 b = 0$

FOC w.r.t. p_h : $\pi_h\left(-tp_h\right) + \lambda_3\left(\overline{F} + w_h\right) = 0$

FOC w.r.t. w_h : $\pi_h\left(-1\right) + \lambda_1 - \lambda_3\left(1 - p_h\right) = 0$

FOC w.r.t. a_l : $\pi_l\left(\theta_l - sa_l\right) + \lambda_2 b - \lambda_4 b = 0$

FOC w.r.t. p_l : $\pi_l\left(-tp_l\right) + \lambda_4\left(\overline{F} + w_l\right) = 0$

FOC w.r.t. w_l : $\pi_l\left(-1\right) + \lambda_2 - \lambda_4\left(1 - p_l\right) = 0$

FOC w.r.t. a_0 : $\pi_0\left(-sa_0\right) - \lambda_1 b - \lambda_2 b + \lambda_3 b + \lambda_4 b = 0$

FOC w.r.t. w_0 : $\pi_0\left(-1\right) - \lambda_1 - \lambda_2 + \lambda_3 + \lambda_4 = 0$

$\lambda_1\left(w_h + ba_h - w_0 - ba_0\right) = 0, \lambda_1 \geq 0, w_h + ba_h - w_0 - ba_0 \geq 0$

$\lambda_2\left(w_l+ba_l-w_0-ba_0\right)=0,\lambda_2\geq 0,w_l+ba_l-w_0-ba_0\geq 0$

$\lambda_3\left(w_0+ba_0+p_h\overline{F}-\left(1-p_h\right)w_h-ba_h\right)=0,\lambda_3\geq 0,w_0+ba_0+p_h\overline{F}-\left(1-p_h\right)w_h-ba_h\geq 0$

$\lambda_4\left(w_0+ba_0+p_l\overline{F}-\left(1-p_l\right)w_l-ba_l\right)=0,\lambda_4\geq 0,w_0+ba_0+p_l\overline{F}-\left(1-p_l\right)w_l-ba_l\geq 0$

$w_0,w_l,w_h\geq 0,a_0,a_l,a_h\geq 0,0\leq p_l,p_h\leq 1$

We obtain $a_h=\frac{1}{s}\left(\theta_h-\frac{1-\pi_h}{\pi_h}b\right)<\frac{1}{s}\theta_h$, $w_h=0$, $p_h=\frac{1-\pi_h}{t\pi_h}\overline{F}$, $F_h=\overline{F}$, $a_l=\frac{1}{s}\left(b+\theta_l\right)>\frac{1}{s}\theta_l$, $p_l=0$, $w_l=-\frac{1}{s}b^2+\frac{b}{s}\left(\theta_h-\theta_l\right)-\frac{1-\pi_h}{st\pi_h}\left(s\overline{F}^2+tb^2\right)$, $a_0=\frac{1}{s}b$, $w_0=-\frac{1}{s}b^2+\frac{b}{s}\theta_h-\frac{1-\pi_h}{st\pi_h}\left(s\overline{F}^2+tb^2\right)$. F_l and F_0 do not matter.

Proof of Proposition 4

Proof: $\Delta\Pi'_{\theta_h}=\frac{1}{s}\left(1-\pi_h\right)b>0$.

$\Delta\Pi'_t=\frac{\left(1-\pi_h\right)^2}{2t^2\pi_h}\overline{F}^2>0$.

$\Delta\Pi'_b=\frac{1}{s}\left(\left(1-\pi_h\right)\theta_h-\left(1-\pi_h\right)b-\pi_l\theta_l-\frac{\left(1-\pi_h\right)^2}{\pi_h}b\right)=\frac{1}{s}\left(\left(1-\pi_h\right)\left(\theta_h-\frac{1}{\pi_h}b\right)-\pi_l\theta_l\right)$.

Since $1-\pi_h=\pi_0+\pi_l>\pi_l$, $b<\pi_h\left(\theta_h-\frac{\pi_l}{1-\pi_h}\theta_l\right)$ given the assumption $b<\pi_h\left(\theta_h-\theta_l\right)$. Therefore, $\Delta\Pi'_b>0$.

$\Delta\Pi'_{\theta_l}=-\frac{1}{s}\pi_l b<0$.

$\Delta\Pi'_s=-\frac{1}{2s^2}\left(2\left(1-\pi_h\right)b\theta_h-\left(1-\pi_h\right)b^2-2b\pi_l\theta_l-\frac{\left(1-\pi_h\right)^2}{\pi_h}b^2\right)<0$

References

Antle, R. and Eppen, G. D. “Capital Rationing and Organizational Slack in Capital Budgeting,” *Management Science* (31:2), 1985, pp.163–174.

Arrow, K. J. “The Economics of Agency,” in *Principals and Agents: The Structure of Business,* Pratt, J.E., Zeckhauser, R.J and Arrow, K.J. (eds.) Harvard Business School Press, Boston, MA. 1985, pp. 37–53.

Aveksa. “Enterprise Roles-based Access Governance,” *Technical Report*, White Paper, 2007.

Baiman, S. “Agency Research in Managerial Accounting: A Second Look,” *Accounting Organizations and Society* (15:4), 1990, pp. 341–371.

Baker, N. R. and Freeland, J. R. “Structuring Information Flow to Enhance Innovation,” *Management Science* (19:1) *Theory Series*, 1972, pp. 105–116.

Baron, D. P. and Besanko, D. “Regulation, Asymmetric Information, and Auditing,” *The RAND Journal of Economics* (15:4), 1984, pp. 447–470.

Dye, R. A. “Optimal Monitoring Policies in Agencies,” *The RAND Journal of Economics* (17:3), 1986, pp. 339–350.

Ferreira, A., Cruz-Correia, R., Antunes, L., Farinha, P., Oliveira-Palhares, E., Chadwick, D., and Costa-Pereira, A. “How to Break Access Control in a Controlled Manner,” in *Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06)*, 2006, pp. 847-854.

Harris, M., Kriebel, C., and Raviv, A. “Asymmetric Information, Incentives and Intrafirm Resource Allocation,” *Management Science* (28:6), 1986, pp. 604–620.

Harris, M. and Raviv, A. “Optimal Incentive Contracts with Imperfect Information,” *Journal of Economic Theory* (20), 1979, pp. 231-259.

Harris, M. and Raviv, A. “The Capital Budgeting Process: Incentives and Information,” *Journal of Finance* (51:4), 1996, pp. 1139–1174.

Holmstrom, B. “Moral Hazard and Observability,” *Bell Journal of Economics* (10:1), 1979, pp. 74-91.

Jolly, D. “Fraud Costs French Bank $7.1 Billion,” *New York Times*, 2008.

Kannan, K. and Telang, R. “Market for Software Vulnerabilities? Think Again,” *Management Science* (51:5), 2005, pp. 726–740.

Kim, S. K. and Suh, Y. S. “Conditional Monitoring Policy Under Moral Hazard,” *Management Science* (38:8), 1992, pp. 1106–1120.

Krishnan, V. and Zhu, W. “Designing a Family of Development Intensive Products,” *Management Science* (52:6), 2006, pp. 813–825.

Lee, H. L., So, K. C., and Tang, C. S. “The Value of Information Sharing in a Two-level Supply Chain,” *Management Science* (46:5), 2000, pp. 626–643.

Motta, M. “Endogenous Quality Choice: Price vs. Quantity Competition,” *Journal of Industry Economics* (41:2), 1993, pp. 113–131.

Povey, D. “Optimistic Security: a New Access Control Paradigm,” In *Proceedings of the 1999 Workshop on New Security Paradigms*, ACM Press, 2000, pp. 40–45.

Rathnam, S., Mahajan, V., and Whinston, A. B. “Facilitating Coordination in Customer Support Teams: A Framework and Its Implications for the Design of Information Technology,” *Management Science* (41:12), 1995, pp. 1900–1922.

Rissanen, E., Firozabadi, S. B., and Sergot, M. “Towards a Mechanism for Discretionary Overriding of Access Control,” In *Proceedings of the 12th International Workshop on Security Protocols*, Cambridge, 2004.

Shavell, S. “Risk Sharing and Incentives in the Principal and Agent Relationship,” *Bell Journal of Economics* (10:1), pp. 55-73.

Townsend, R. M. “Optimal Contracts and Competitive Markets with Costly State Verification,” *Journal of Economy Theory* (21:2), 1979, pp. 265-293

Tsai, W. “Knowledge Transfer in Intraorganizational Networks: Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance,” *The Academy of Management Journal* (44:5), 2001, pp. 996–1004.

von Hippel, E. “Sticky Information and the Locus of Problem Solving: Implications for Innovation,” *Management Science* (40:4), pp. 429–439.