SMU ECONOMICS & STATISTICS
WORKING PAPER SERIES

SMU
SINGAPORE MANAGEMENT
UNIVERSITY

To Trust or to Monitor: A Dynamic Analysis

Fali Huang
September 2007

Paper No. 11-2007

To Trust or to Monitor: A Dynamic Analysis

Fali Huang

Abstract In a principal–agent framework, principals can mitigate moral hazard problems not only through extrinsic incentives such as monitoring, but also through agents' intrinsic trustworthiness. Their relative usage, however, changes over time and varies across societies. This paper attempts to explain this phenomenon by endogenizing agent trustworthiness as a response to potential returns. When monitoring becomes relatively cheaper over time, agents acquire lower trustworthiness, which may actually drive up the overall governance cost in society. Across societies, those giving employees lower weights in choosing governance methods tend to have higher monitoring intensities and lower trust. These results are consistent with the empirical evidence.

Keywords Monitoring · Trustworthiness · Trust · Screening · Economic Governance

JEL Classification Numbers D2 · J5 · L2 · M5 · Z13.

F. Huang

School of Economics, Singapore Management University,

90 Stamford Road, Singapore 178903, Singapore

Email: flhuang@smu.edu.sg

1 Introduction

All societies have to deal with moral hazard problems. But each society resolves such problems in different ways; some rely more on trust, while others depend on heavy use of governance and monitoring rules. In the late medieval period, for example, agency relations in the Maghribi traders were characterized by the prevalence of trust: "Despite the many opportunities for agents to cheat, only a handful of documents contain allegations of misconduct ..." This is, however, not the case in Italy, "where allegations of misconduct are well-reflected in the historical records" (Greif 1993). In current times, the labor-management relations in Japan depend on a high level of trust, while "[t]he twentieth-century American system of industrial labor relations, with its periodic massive layoffs, book-length contracts, and bureaucratic, rule-bound personal interactions, would seem the very model of low-trust social relations" (Fukuyama 1995, p. 218). In a sample of fifteen developed economies, the supervision intensities in UK, US and Canada are the highest, with an average over two times as high as that of the rest in 1980 (Gordon 1994); in most of these countries the monitoring intensity followed an upward trend from 1970s to 1990s in the manufacturing sector (Vernon 2001). Why do societies differ in their usage of trust and monitoring? How does it change over time? These questions are explored in this paper.

From a society's point of view, the interdependence between intrinsic trustworthiness and extrinsic governance in mitigating moral hazard problems is quite obvious. If the technology is so advanced that it costs very little to achieve perfect monitoring, the society may deem unnecessary to invest in individual trustworthiness. In contrast, if there is a sufficiently large supply of trustworthy individuals and there are easy ways to recognize them, the society may not invest to improve on monitoring. Since both monitoring technologies and cultivating trustworthy people are costly, most societies fall somewhere between these two extremes, and their exact locations depend on the relative costs of inculcation, screening, and monitoring.[1] These costs are not only affected by exogenous technical features, but more importantly, also by the incentive structure that shapes the relevant resource allocation decisions; both are crucial for understanding the relative usage of trust and monitoring over time and across societies.

This paper formalizes these insights using a principal-agent model, where the distribu-

---

[1]Considerable resources are involved in setting up schools, religious institutions, not to mention the time and resources spent by parents, to inculcate moral values in its people (Shavell 2002). Meanwhile, monitoring is also costly: "more than 70,000 U.S. companies spent more than $500 million on surveillance software between 1990 and 1992, and that by 1990 more than 10 million workers were under electronic surveillance" (Kipnis 1996). As a result, both intrinsic and extrinsic incentives are commonly used; see for example Baron and Kreps (1999) and Nagin et al. (2002).

tion of agent trustworthiness is endogenized through agents' skill investment choices. The main findings are as follows. Trustworthy agents need less monitoring and hence receive higher incomes than non-trustworthy ones in a competitive labor market. Changes in the relative cost of monitoring techniques affect the returns of being trustworthy, where the average trustworthiness declines as monitoring becomes cheaper because parents anticipate that trustworthiness will be less rewarded. The overall governance cost, however, may be driven up by cheaper monitoring technologies, which crowd out too much trustworthiness and hence force the society to rely excessively on extrinsic incentives.[2]

The relative costs of using trust and monitoring, however, are not dictated by natural laws. Instead, they are likely to be endogenously determined by resources spent in reducing them. As it turns out in equilibrium, principals and agents have conflicting interests in such matters: Agents are better off with higher trustworthiness and higher monitoring costs; in contrast, principals always prefer cheaper monitoring methods, and they do not necessarily benefit from hiring more trustworthy agents. A natural implication is that principals have strong incentives to reduce monitoring costs but much less to invest in agent trustworthiness, while the opposite is true for the agents.

The result may seem puzzling at first sight. If hiring trustworthy agents reduces governance costs, then principals should necessarily benefit from it; this is also the general impression one gets from discussions on social trust. But there are two problems in such an argument: it ignores the competition among principals and the endogeneity of agent trustworthiness. As a standard result of competition, the rent from hiring more trustworthy agents disappears in the equilibrium;[3] principals gain only when the bottom agents are more trustworthy, but then these agents have no incentives to invest in trustworthiness only to bring free windfall to principals. In other words, any rent captured by a principal is at risk of being bid away by ex post labor market competition and by ex ante investment of

---

[2]The experiences of American firms seem to be consistent with this result. The intensively monitored workplace and the "conflict-ridden state of labor-management relations in many American industries" are held partly to blame for "the low productivity and poor quality of American work" (Mills 1994). Faced with intense competition pressure from foreign firms, various high performance work practices relying more on cooperation efforts from employees started to be adopted from 1980s (Appelbaum and Batt 1994, Cappelli 1995, Cappelli and Neumark 2001). But such a transforming process is slow and difficult to sustain due to "insufficient trust" (Commission on the Future of Worker-Management Relations 1994). Actually the consequences of low trust have motivated lively discussions among the public and social scientists; see, among others, Putnam (1995), Cook (2001), James (2002), and Durlauf and Fafchamps (2005).

[3]The logic is similar in spirit to Becker's (1962) insight on firms' reluctance to invest in the general training of employees. In the light of this comparison, firms' willingness to invest in corporate culture (Rob and Zemsky 2002, Kreps 1997) and employee identity (Akerlof and Kranton 2005) follows a similar logic as their willingness to invest in firm-specific human capital.

agents. The existence of labor market frictions may enable principals to capture some rent; such a rent, however, is not only limited in value since it is bounded above by the degree of frictions, but also temporary in possession because it is rooted in the shifting sand of endogenous agent trustworthiness. In sharp contrast, the saving on governance costs due to cheaper monitoring methods is immune to both hazards.

Since principals have quite different incentives from agents in the choice of governance modes, which side has more weight in resource allocation becomes very important in shaping the relative costs and hence the actual usage of trust and monitoring. This yields the following cross-sectional variation: societies giving lower weights to the welfare of workers when choosing governance modes rely more on extrinsic governance and less on trust.[4] It is indeed supported by the empirical evidence (Esping-Anderson 1990, Gordon 1994, Rubery and Grimshaw 2003): The collective labor power is negatively correlated with the supervision intensity among developed economies; the US, UK, and Canada have the lowest labor powers and the highest supervision intensities, while the opposite is true in Denmark, Japan and Germany.[5]

These results provide a plausible explanation of how intrinsic trustworthiness and monitoring intensities evolve over time and across societies, which is the main contribution of this paper. The relative merits and costs of different governance modes are also investigated in related studies. Shavell (2002) analyzes the relative effectiveness of morality and law as means of controlling conduct; Sobel (2006) and Greif (1997) compare informal relational contracts with formal legal institutions in supporting cooperation. The emphasis of these work is more on the static comparison among competing modes than their dynamic interactions. Another strand of related literature shows that intrinsic motivation may be crowded out by high power incentive schemes (Rob and Zemsky 2002, Kreps 1997), by public policies (Bar-Gill and Fershtman 2005), by legal institutions (Bohnet, Bruno and Huck 2001, Huck 1998), and by explicit monitoring (Frey 1993). The current paper contributes to this literature by endogenizing both the monitoring intensities and intrinsic motivation so that their dynamic interactions, in addition to the usual one-way crowding out effects, can be

---

[4]A discrete version of this result is that an individualist society tends to rely more on monitoring than a group-oriented one. This is consistent with the distinct management styles in the United States versus Japan, and in the two trader groups mentioned above: In a survey cited by then Secretary of Labor Ann McLaughlin in 1988, "only 9 percent of American workers felt they would benefit from their companies' increased productivity compared to 93 percent of Japanese workers interviewed in a similar survey" (Mills 1994). The "social structure of the Maghribi traders' group was 'horizontal,' as traders functioned as agents and merchants at the same time," while agency relations were organized 'vertically' among the Italian traders in that "merchants and agents constitute two distinct subgroups" (Greif 1993).

[5]More details are in the working paper version (Huang 2006).

studied.

In this paper, trustworthiness is essentially a trait or skill that enables the owner to resist short-run opportunistic temptations and to maximize social welfare. This paper is thus naturally linked with a growing stream of studies of social preferences and ethnic behaviors in the economic literature. It is consistent with the extensive experimental evidence produced over the past four decades on human behavior in social dilemmas, which demonstrates that "internalized trust is a common phenomenon; that it is at least in part learned rather than innate; and that different individuals vary in their inclinations toward trust." (Stout and Blair 2001). Results in this paper regarding the negative relationship between screening and monitoring intensities and the positive relationship between screening and employee compensation are empirically tested and supported using data from a representative national employers survey in the US (Huang and Cappelli 2006). The theoretical side of the literature is too vast to be surveyed here; in the following only recent and more closely related studies are briefly discussed. In a principal-agent framework with adverse selection, the effects of different specifications of honest behaviors on the optimal contracts are studied in Alger and Renault (2006, 2007); Sliwka (2007) analyzes the interactions between the prevalence of ethic behaviors in society and an individual's choice of whether to conform to the ethic norm; these papers take ethic types as given and hence do not endogenize their distributions as in the current paper. Another strand of literature, in contrast, aims to study why people acquire social preferences in the first place. Frank (1987) explores whether an individual wants to choose his own utility function that allows others-regarding elements; Güth and Ockenfels (2005) study the endogeneity of moral preferences using the indirect evolutionary approach, which combines individual rational decision making with the evolutionary approach of preference determination; and finally, Kaplow and Shavell (2007) examine how a social planner would inculcate guilt and virtue in individuals to foster social welfare.

This paper proceeds as follows. In Section 2, a simple principal–agent model is introduced and then extended to costly screening for agent trustworthiness. The intergenerational dynamics are analyzed in Section 3, where individual trustworthiness is endogenized through parental investment. The final section presents conclusions.

2 The Principal-Agent Model with Monitoring and Screening

A principal hires an agent to complete a project. The outcome is stochastic: If the agent makes the appropriate effort e, he produces $h > 0$ with probability $p \in [0, 1]$ and 0 with probability $1 - p$; if the agent shirks, the probability of getting h is $q \in [0, 1)$, where $q < p$.

The cost of making effort e is $c(e)$, while shirking involves no cost. We assume $hp - c(e) > hq$ so that making effort e is the social optimal choice.

There is a continuum of agents indexed by $i \in [0, 1]$, who are heterogeneous in their predisposition to cooperate. Agent i has a degree of *trustworthiness* $\alpha_i \geq 0$ that measures the amount of guilt he feels if he shirks, whether caught or not by the principal.[6] The cumulative distribution function of trustworthiness among agents is $F(\cdot)$ with support $[0, A]$. It characterizes the quality of the workforce in this economy.

Principals are identical with unit mass. The reservation utility of agents and the alternative return for principals are normalized to zero. To reduce shirking, a principal may screen potential job candidates before hiring and/or monitor the agent on the job. The principal can find out an agent's trustworthiness α_i at some cost $s \geq 0$. We first analyze the case with $s = 0$ where an agent's trustworthiness is publicly observed, then proceed to the more general case with $s > 0$.

A principal chooses monitoring intensity $m_i \in [0, 1]$ for an agent with α_i; with probability m_i an agent who shirks is caught by the principal and paid nothing. The total monitoring cost is $m_i k$, where k measures the unit cost of using monitoring technologies such as video cameras in the workplace.

The payment to an agent with α_i has two components: one is the incentive pay w_i that will be forsaken if shirking is detected by the principal, and the other is the basic wage b_i that is independent of the agent's performance.[7] The agent's utility is $u(w_i) + u(b_i) - c(e)$ when he makes effort, and $(1 - m_i)u(w_i) + u(b_i) - \alpha_i$ if he shirks, where $u(0) = 0$, $u' > 0$ and $u'' \leq 0$. We define the governance cost M_i for an agent i as the sum of the monitoring cost $m_i k$ and the incentive pay w_i, since both of which are meant to reduce shirking:

$$M_i \equiv m_i k + w_i.$$

2.1 The Basic Model: Zero Screening Cost

The time line of the game with zero screening cost is as follows. Principals announce their incentive packages containing the incentive pay w_i and the monitoring intensity m_i as well as the basic wage b_i as functions of the agent's perceived trustworthiness α_i. Agents then match with principals. After the matching is finished, agents get the basic wage b_i

---

[6]This specific modelling of cooperative predisposition follows from the experimental evidence in Palfrey and Prisbrey (1997) and is widely adopted in the relevant literature (see, e.g. Huck 1998, Rob and Zemsky 2002). Modeling α_i as an intrinsic benefit of cooperation such as warm glow does not affect the results.

[7]A state-contingent wage scheme is less optimal here than a constant wage across states, since the monitoring intensity will be chosen to deter shirking completely so that the usual role of state-dependent wages for risk-averse agents is eliminated.

and choose whether to make effort or shirk. Principals then monitor agents, pay w_i if no shirking is found, and pay nothing if otherwise. The competitive equilibrium is reached in this game when there is no further changing of partners, and once in a match, nobody wants to deviate from their decisions. We solve the game backwards.[8]

Lemma 1 *In any given match, the optimal incentive package (w_i^*, m_i^*) is determined by (3) and (4) below for $\alpha_i \in [0, c(e))$, and is $(0,0)$ for $\alpha_i \geq c(e)$. Given this incentive scheme, all agents make effort. Both w_i^* and $m_i^* k$ decrease in α_i and increase in k, so does the governance cost M_i^*.*

Proof. Given the incentive package (w_i, m_i) and the basic wage b_i, an agent does not shirk if $u(w_i) + u(b_i) - c(e) \geq (1 - m_i)u(w_i) + u(b_i) - \alpha_i$. This is simplified to

$$\underset{\text{extrinsic incentive}}{m_i u(w_i)} + \underset{\text{intrinsic incentive}}{\alpha_i} \geq \underset{\text{cost of effort}}{c(e)}, \tag{1}$$

where an agent won't shirk if the sum of extrinsic and intrinsic incentives is larger than the cost of effort. For agents with $\alpha_i \geq c(e)$, their intrinsic incentive α_i alone is high enough to prevent shirking, so the principal would set both the monitoring intensity and the incentive pay at zero.

For agents with $\alpha_i < c(e)$, the minimum monitoring level required to deter shirking is

$$\underline{m_i} \equiv \frac{c(e) - \alpha_i}{u(w_i)}. \tag{2}$$

So if a positive monitoring level is ever chosen, the principal would get a profit $Q_i \equiv hp - w_i - \underline{m_i}k - b_i$, and her objective function is

$$\max_{w_i} hp - w_i - \frac{c(e) - \alpha_i}{u(w_i)}k - b_i.$$

The FOC w.r.t. w_i for an interior solution is

$$u'(w_i^*)(c(e) - \alpha_i)k - u(w_i^*)^2 = 0. \tag{3}$$

The SOC $SOC \equiv u''(w_i^*)(c(e) - \alpha_i)k - 2u(w_i^*)u'(w_i^*) < 0$ holds since $u' > 0$ and $u'' \leq 0$. Plugging w_i^* into (2), we get

$$m_i^* = \frac{c(e) - \alpha_i}{u(w_i^*)}. \tag{4}$$

---

[8]The qualitative results would not change if alternative combinations of schemes were used. For example, whatever repeated interactions can do to mitigate the moral hazard problem is either type-revealing or imposing extra extrinsic incentives, both of which are already represented by screening and monitoring in a one-period relationship.

The proof of comparative statics is relegated to the Appendix. ■

Since agents with higher α_i require lower governance expenditures, all principals prefer to hire them; but then competition among principals would bid up the basic wage b_i for these agents. Given that principals are identical with a mass equal to that of agents, all would end up earning the same profit in equilibrium. Let

$$\underline{\alpha} \equiv \min_{i \in [0,1]} \{\alpha_i\}$$

denote the lowest trustworthiness among agents. An agent with $\underline{\alpha}$ gets zero basic wage since there is no competitive bidding for him. So his principal earns a profit

$$Q_{\underline{\alpha}}^* = hp - M^*(\underline{\alpha}),$$

where $M^*(\underline{\alpha})$ is the corresponding governance cost for the bottom agent; $Q_{\underline{\alpha}}^*$ increases in $\underline{\alpha}$ and decreases in k by Lemma 1.

A principal hiring an agent with $\alpha_i > \underline{\alpha}$ gets a profit $Q_i^* = hp - M_i^* - b_i$, which must be the same as $Q_{\underline{\alpha}}^*$ in equilibrium. That is, she must pay a basic wage b_i^* such that $Q_i^* = Q_{\underline{\alpha}}^*$, which yields

$$b_i^* = M^*(\underline{\alpha}) - M_i^*. \tag{5}$$

Note the basic wage is essentially a *rent* generated by the agent's trustworthiness α_i, which reduces the governance cost from $M^*(\underline{\alpha})$ to M_i^*. The total compensation for an agent is

$$I_i^* \equiv w_i^* + b_i^* = M^*(\underline{\alpha}) - m_i^* k, \tag{6}$$

which increases in both α_i and k based on Lemma 1.

When $\underline{\alpha} = 0$, principals have to incur the highest governance cost

$$M_0^* = w_0^* + m_0^* k,$$

where $w_0^* \equiv w_i^*|_{\alpha_i=0}$ and $m_0^* \equiv m_i^*|_{\alpha_i=0}$ are respectively the maximum levels of optimal incentive pay and monitoring intensity. The corresponding profit is

$$Q_0^* \equiv hp - M_0^*; \tag{7}$$

in this case principals do not benefit from agent trustworthiness at all. In contrast, when $\underline{\alpha} > 0$, principals can capture a partial rent $Q_{\underline{\alpha}}^* - Q_0^* = M_0^* - M^*(\underline{\alpha}) > 0$.

Once all principals earn an identical profit $Q_{\underline{\alpha}}^*$, nobody wants to change agents anymore.[9] Similarly, agents do not gain from changing principals either. The distributions of agent income I_i^*, principal profit Q_0^*, and the governance cost M_i^* are illustrated in Figure 1. The relevant results are summarized in the following proposition.

---

[9]The profit of a principal is actually $\max\{Q_{\underline{\alpha}}^*, hq\}$, where hq arises when the principal pays an agent zero wage and incurs zero monitoring cost. So monitoring is chosen if and only if $Q_{\underline{\alpha}}^* \geq hq$; since $Q_{\underline{\alpha}}^* \geq Q_0^*$, this condition is true when $Q_0^* \geq hq$, which is assumed.
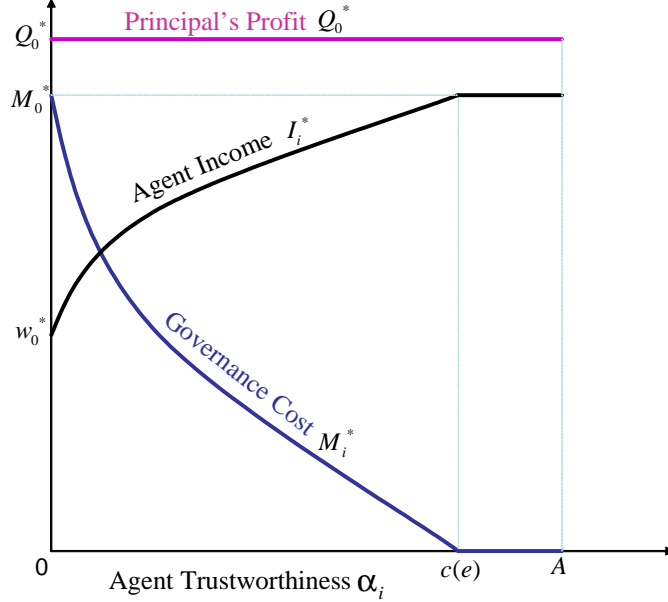
Figure 1: Principals' profit, agent incomes, and governance costs

Proposition 1 *In the competitive equilibrium of the basic model, each principal chooses the optimal basic wage b_i^* in (5) and the incentive package (w_i^*, m_i^*) derived in Lemma 1 for any agent with α_i; all agents make effort. The agent's income I_i^* increases in both k and α_i, whereas all principals make an identical profit $Q_{\underline{\alpha}}^*$, which is decreasing in k, increasing in $\underline{\alpha}$, but independent of α_i.*

Given an agent's trustworthiness, principals adjust their monitoring intensities and incentive pays accordingly to save governance costs. Because of perfect competition between principals, the cost saved is transferred to agents as a basic wage, leaving principals with a profit they would have made when hiring the least trustworthy agents in the population. In other words, principals do not benefit from hiring agents with higher trustworthiness in equilibrium, though they do gain when the bottom agents are more trustworthy since $Q_{\underline{\alpha}}^*$ increases in $\underline{\alpha}$. In this sense, the bottom level of trustworthiness $\underline{\alpha}$ serves as a public good for all principals.

2.2 The Model with Costly Screening

An extension to the basic model is studied in this subsection, where the principal has to pay a positive screening cost s to observe an agent's trustworthiness. The fixed screening cost implies two alternatives for a principal: pay s and observe an agent's trustworthiness

accurately, or pay nothing and get no information. Though screening is costly, it enables principals to reduce governance costs by using more suitable incentive packages. This may generate positive sorting in equilibrium, where screening principals are matched with more trustworthy agents.[10]

The time line of this screening game is similar as in the basic model, with some adjustment in the matching process. Principals first decide whether to screen or not. Those who choose to screen announce their selection criterion $\widehat{\alpha}$,[11] the incentive package (w_i^s, m_i^s) and a basic wage b_i^s for $\alpha_i \geq \widehat{\alpha}$. Those who do not screen would adopt a single incentive package (w_r, m_r) for any agent hired, since all look the same to them. Agents then decide where to apply for jobs. A screening principal screens job candidates, hires the first one with $\alpha_i \geq \widehat{\alpha}$ and rejects others. A non-screening principal hires whoever comes first. After matching is finished, agents get the basic wage if any and then choose whether to make effort or shirk. Principals monitor agents, pay nothing if shirking is detected and pay the incentive wages if otherwise. The competitive equilibrium is reached when all principals stick to their screening choices, there is no partner-changing, and once in a match, nobody wants to deviate from their decisions.

Again we solve the game backwards. When the labor market clears, the proportion of agents working for non-screening principals must be equal to that rejected by screening ones, which is $F(\widehat{\alpha})$; this is also the proportion of principals who choose not to screen. Since screening incurs a fixed lump-sum cost, the optimization problem of screening principals is exactly the same as in the basic model, and hence the same optimal incentive package (w_i^*, m_i^*) applies to an agent with $\alpha_i \in [\widehat{\alpha}, A]$. Competition among screening principals equalizes their profits, which in equilibrium would also be the same as the optimal profit Q_r^* made by all non-screening principals due to competition pressure from them. That is, just as in the basic model, all principals earn the same expected profit Q_r^* in equilibrium regardless of their screening choices.

An agent with $\alpha_i \geq \widehat{\alpha}$ works for a screening principal and gets a total income

$$I_i^s = b_i^s + w_i^*,$$

where w_i^* is in (3) and b_i^s is determined by

$$b_i^s \equiv hp - M_i^* - s - Q_r^*; \tag{8}$$

_____

[10]A more general screening technology yields similar results (Huang and Cappelli 2006). Since the maximum reduction in the governance cost is M_0^*, a necessary condition for a positive mass of screened agents is $s < M_0^*$.

[11]Since the screening cost s is the same regardless of the actual α_i while the potential profit strictly increases in α_i, there must exist a threshold $\widehat{\alpha}$ such that only agents with $\alpha_i \geq \widehat{\alpha}$ are hired after being screened.

the agent income I_i^s increases in α_i as before. Other agents will work for non-screening principals and get w_r^*. To make the marginal agent $\widehat{\alpha}$ indifferent between being screened or pooled with others,

$$I^s(\widehat{\alpha}) = w_r^* \tag{9}$$

must hold. Since I_i^s increases in α_i while w_r^* is fixed, agents with $\alpha_i \geq \widehat{\alpha}$ indeed prefer to work for screening principals while others not.

Given the incentive package (w_r, m_r) offered by a non-screening principal, agents with $\alpha_i \geq \alpha_r$ make effort while others shirk, where

$$\alpha_r \equiv c(e) - m_r u(w_r)$$

is determined by the non-shirking condition (1). Thus, from the principal's perspective, a non-screened agent shirks with probability $\frac{F(\alpha_r)}{F(\widehat{\alpha})}$; he is caught and loses w_r with probability m_r. So a non-screening principal's expected profit is

$$Q_r = (1 - \frac{F(\alpha_r)}{F(\widehat{\alpha})})(hp - w_r) + \frac{F(\alpha_r)}{F(\widehat{\alpha})}(hq - (1 - m_r)w_r) - m_r k. \tag{10}$$

Proposition 2 *In the competitive equilibrium of the screening game, there exists a unique level of trustworthiness $\widehat{\alpha}$ such that agents with $\alpha_i < \widehat{\alpha}$ choose to work for non-screening principals, whereas those with $\alpha_i \geq \widehat{\alpha}$ work for screening principals; $\frac{\partial \widehat{\alpha}}{\partial s} > 0$ and $\frac{\partial \widehat{\alpha}}{\partial k} < 0$ hold. A proportion $1 - F(\widehat{\alpha})$ of principals screen agents, hire only those with $\alpha_i > \widehat{\alpha}$, and offer (w_i^*, m_i^*) plus a basic wage b_i^s in (8); others do not screen agents and offer zero basic wage plus a uniform incentive package (w_r^*, m_r^*), which maximizes problem (10) and satisfies (9). All principals make the same profit Q_r^*, where $\frac{\partial Q_r^*}{\partial s} > 0$ and $Q_r^* > Q_{\underline{\alpha}}^*$.*

Proof. In the Appendix. ∎

Note that the basic wage of an agent with $\alpha_i \geq \widehat{\alpha}$, who works for a screening principal, can be rewritten as

$$b_i^s = b_i^* - (Q_r^* - Q_{\underline{\alpha}}^*) - s,$$

where b_i^* is the basic wage he gets in the basic model with zero screening cost; $b_i^s < b_i^*$ holds not only because it is the agent that ultimately pays the screening cost s, but also because principals earn a higher profit Q_r^* than before. So the positive screening cost increases the profit of principals but reduces agent incomes. The intuition is that the positive screening cost, a form of market friction, reduces competition among principals and hence enables them to capture some rent generated by agents. Again, screening principals do not gain from hiring agents with higher trustworthiness once $\alpha_i \geq \widehat{\alpha}$, while all principals do gain when bottom agents are more trustworthy.

11

3   Inter-Generational Dynamics: Endogenous Trustworthi-
ness

In this part we endogenize the distribution of α_i in society. Suppose principals and agents live for one period, each raising a child to replace his role when he dies. The underlying technologies remain the same over generations. All agent children are endowed with the identical productive ability as their parents. Their intrinsic trustworthiness is zero at birth, which can be improved by parental investment during childhood to maximize a child's lifetime income minus the investment cost.

The sequence of events is as follows. In the beginning of generation $n = 1, ..., \infty$, the distribution of α_{ni} is realized. Then the stage game with costly screening is played, where the competitive equilibrium derived in Proposition 2 prevails. At the same time, the agent α_{ni} inculcates trustworthiness $\alpha_{n+1,i}$ in his child, expecting him to get an equilibrium income $I^*(\alpha_{n+1,i})$ when the child becomes adult. The inculcation cost is $C(\alpha_{n+1,i}; \alpha_{ni})$ where $C(0; \alpha_{ni}) = 0$, $C_1 > 0$, $C_{11} > 0$, $C_2 < 0$, $C_{12} < 0$; here the parental trustworthiness α_{ni} indicates the efficacy of the inculcation process, and it costs nothing to retain the initial zero trustworthiness. Then generation $n+1$ replaces the old one, and the sequence of events goes on.

Since the pooling wage w_r^* is independent of an individual's trustworthiness α_i while the inculcation cost increases with α_i, there is no gain for a pooling agent to acquire any $\alpha_i > 0$. Hence, from the second generation onwards, any agents not screened must have zero trustworthiness. An obvious implication is that the lowest level of trustworthiness among future agents would never be positive, that is, $\underline{\alpha}_n = 0$ in generation $n \geq 2$.[12] Accordingly, non-screening principals will offer an incentive package (w_0^*, m_0^*) and make a profit Q_0^* in (7). Then screening principals also make Q_0^*, and the extra gain $Q_r^* - Q_0^*$ captured in the above static model is transferred back to agents. So principals do not benefit from agent trustworthiness once it becomes endogenous. Thus we have proved the following proposition.

Proposition 3 *For any generation $n \geq 2$, any non-screened agent has zero trustworthiness and gets income w_0^*, while a screened agent gets a basic wage $b^*(\alpha_{ni}) - s$ and income $I^*(\alpha_{ni}) - s$, where $b^*(\alpha_{ni}) = M_0^* - M^*(\alpha_{ni})$ and $I^*(\alpha_{ni}) = b^*(\alpha_{ni}) + w^*(\alpha_{ni})$ following (5) and (6) respectively. The principals' profit is Q_0^* as in the basic model.*

---

[12]Similarly, no agent would have $\alpha_{ni} > c(e)$, since doing so yields the same income as having $\alpha_{ni} = c(e)$ but incurs larger investment costs. We ignore the perverse case where α can be negative; even when it is allowed, there exists a lower bound for α, below which principals would choose zero monitoring and zero wage.

The objective function of a parent in generation $n = 1, 2, \ldots$ is $\max\{R(\alpha_{ni}), w_0^*\}$, where

$$R(\alpha_{ni}) \equiv \max_{\alpha_{n+1,i}} I^*(\alpha_{n+1,i}) - c(\alpha_{n+1,i}; \alpha_{ni}) - s \qquad (11)$$

is the return if he ever invests in $\alpha_{n+1,i}$ and his child is to be hired by a screening principal, and w_0^* is the return of his best alternative, which is no investment and his child with $\alpha_{n+1,i} = 0$ is to work for a non-screening principal. Intuitively, if a parent α_{ni} chooses to invest in his child, the resulting $\alpha_{n+1,i}$ must be high enough to enable the child to be hired after being screened and earn an income of at least $C(\alpha_{n+1,i}; \alpha_{ni}) + s + w_0^*$ to justify the investment. Since $R(\alpha_{ni})$ strictly increases in α_{ni}, there exists a unique threshold trustworthiness $\widetilde{\alpha}$ for every generation n such that

$$R(\widetilde{\alpha}) = w_0^*. \qquad (12)$$

So descendants of families with $\alpha_{1i} < \widetilde{\alpha}$ would have $\alpha_{n+1,i}^* = 0$ for all generations $n+1 \geq 2$ and work for non-screening principals. Only those with $\alpha_{ni} \geq \widetilde{\alpha}$ would ever invest in their children, whose optimal choices are stated in the following proposition.

Proposition 4 (*i*) *There exists a unique optimal solution* $\alpha_{n+1,i}^* \equiv g(\alpha_{ni})$ *to problem* (11). *The transition function* $g(\alpha_{ni})$ *is either equal to* $c(e)$, *or strictly increasing in* α_{ni}.

(*ii*) *In the steady state, there exists a unique trustworthiness level* $\alpha_c = g(\alpha_c) \in (0, c(e)]$ *for all agents from families with* $\alpha_{1i} \geq \widetilde{\alpha}$ *if* $C_{122} \geq 0$ *and* $C_{112} \geq 0$; *others have zero trustworthiness. Hence, in the steady state, a proportion* $\pi \equiv 1 - F(\widetilde{\alpha})$ *of agents have* α_c, *where* α_c *increase in* k *if* $-m_\alpha + w_{\alpha k} > 0$, *and* π *increases in* k *but decreases in* s.

(*iii*) *When the elasticity of* α_c *over* k *is high enough, the steady-state governance cost* $M^*(\alpha_c)$ *goes up when* k *is lower, contrary to the static result in Lemma 1.*

Proof. In the Appendix. ∎

This proposition suggests that when trustworthiness is endogenously determined, cheaper monitoring technologies may *increase* the governance cost, which is in stark contrast to the short run view in Lemma 1. The intuition is as follows. If monitoring is cheaper in the next generation, the lifetime return of trustworthiness is lower, so fewer agents will invest in it and those who do so would invest less in it; but when the overall level of trustworthiness is lower, principals have to monitor agents more intensively. When the effect of resulting higher monitoring intensity outweighs that of a lower unit cost of monitoring, the total governance cost goes up; this happens when the elasticity of trustworthiness over k is large enough.[13] A specific case is provided by the following example.

---

[13] The mechanism is in some sense similar to the familiar phenomenon that, when the demand is elastic, a lower price often leads to a much higher demand and thus a higher total expenditure.

Example. Suppose the utility function is $u(w) = w^z$ and the cost function is

$$c(\alpha_{n+1,i}; \alpha_{ni}) = \frac{c(e)^a - (c(e) - \alpha_{n+1,i})^a}{(1 + \alpha_{ni})^b},$$

where $z \in (0, 1]$, $a \in (0, 1)$ and $b \in R^+$.

When $a < \frac{1}{z+1}$, the unique optimal choice is

$$\alpha_{n+1,i}^* \equiv g(\alpha_{ni}) = c(e) - k^{\frac{-1}{1-a(z+1)}}(1 + \alpha_{ni})^{\frac{-b(z+1)}{1-a(z+1)}}\gamma, \tag{13}$$

where $\gamma \equiv (z(z+1)a)^{\frac{1}{1-a(z+1)}}$.[14] It is easy to check that $g(\alpha_{ni})$ is increasing and concave in α_{ni} and $\frac{\partial \alpha_{n+1,i}^*}{\partial k} > 0$. In any generation $n + 1$, the governance cost is

$$M_{n+1,i}^* \equiv m_{n+1,i}^* k + w_{n+1,i}^*$$

$$= \eta(k(c(e) - \alpha_{n+1,i}))^{\frac{1}{z+1}} \tag{14}$$

$$\underset{\text{plug in (13)}}{=} \eta\gamma^{\frac{1}{z+1}}(k^{-a}(1 + \alpha_{ni})^{-b})^{\frac{1}{1-a(z+1)}}, \tag{15}$$

where $\eta \equiv (1 + z^{-1})z^{\frac{1}{z+1}}$. It is straightforward to see that $\frac{\partial M_{n+1,i}^*}{\partial k}|_{\alpha_{n+1,i}} > 0$ holds in (14) when $\alpha_{n+1,i}$ is exogenous in the short run, while $\frac{\partial M_{n+1,i}^*}{\partial k} < 0$ is true in (15) when $\alpha_{n+1,i}$ is endogenously determined by k in the long run.

A social planner, when deciding on how to allocate resources in reducing monitoring, screening, and inculcation costs, would take into consideration the dynamic crowding-out and crowding-in effects by the monitoring technology on agent trustworthiness. Individual principals, however, do not necessarily internalize the negative externalities they impose on agent incomes when allocating resources to reduce monitoring costs. In other words, they tend to over-invest in monitoring technologies. The reason is that the profit of principals Q_0^* increases when monitoring is cheaper, but it does not change when inculcation or screening costs are lower; in contrast, the incomes of agents drop in the former case, but increase in the latter. So principals gain but agents lose when the unit cost of monitoring is lower; agents benefit from lower inculcation and screening costs, whereas principals are indifferent. This conflict of interests between principals and agents seems fundamental in determining the basic incentive structure of a society's resource allocation choices between reducing monitoring costs versus reducing inculcation and screening costs, and hence may shape the long term trends and cross-sectional variations of trust and monitoring intensity.

---

[14]When $a \geq \frac{1}{z+1}$, the marginal benefit of investing in trustworthiness is ever increasing before $c(e)$ is reached so that $\alpha_{n+1,i}^* = c(e)$.

4   Conclusions

This paper analyzes the dynamic relationship between trust and monitoring in reducing moral hazard problems in a principal-agent setting. Agent trustworthiness and monitoring intensities are both determined by fundamental forces in society such as the costs of monitoring and screening agents and the cost of inculcating trustworthiness; their long-term trends and cross-sectional variation are thus shaped by how these relevant costs change. While acknowledging the influence of exogenous technical features on the cost reduction process, we argue that an important role is also played by the inherent conflict of interests between principals and agents in equilibrium: principals benefit from lower monitoring costs but *not necessarily* from lower screening and inculcation costs, whereas the opposite is true for agents. When monitoring becomes relatively cheaper, trust declines and the monitoring intensity increases over time; they may do so at faster rates in societies where the interests of agents are weighed less than those of principals in the choice of monitoring schemes. The overall governance costs, however, may be driven up by cheaper monitoring technologies, which crowd out intrinsic incentives and induce society to rely too much on extrinsic ones. These results are indeed consistent with preliminary empirical evidence, though more rigorous tests are needed in future research.

The main insights of this paper also carry on to situations with general market frictions that give principals certain monopsony powers. The following results can be readily obtained with similar arguments as in the text. Principals may capture a rent from agent trustworthiness when there exist labor market frictions; the rent, however, is limited by the degree of frictions, and more importantly, it again disappears once trustworthiness becomes endogenous. Since labor market frictions increase principals' profits but reduce agent incomes, principals have less incentives to eliminate them, while the opposite is again true for agents.

This paper can be extended in various directions to get a more thorough understanding of relevant issues. For example, the resource allocation decisions on improving various governance modes can be explicitly modeled in a bargaining or political economy environment. The screening process can be fleshed out and repeated interactions between principals and agents may be added to better address potential problems associated with screening in a diverse and mobile society. The identical producing ability of agents assumed in this paper can also be relaxed to study the trade-off or complementarity between investment in cognitive and non-cognitive skills from the perspective of aggregate welfare. For instance, if the difficulty in monitoring arises when higher cognitive abilities are involved, then the thoroughness of screening and the combination of governance modes should vary across jobs

in some systematic way.

Appendix

Proof of comparative statics in Lemma 1 Based on (3) we get

$$
\begin{aligned}
w_\alpha &\equiv \frac{\partial w_i^*}{\partial \alpha_i} = \frac{-u'(w_i^*)k}{-SOC} < 0, \\
m_\alpha &\equiv \frac{\partial m_i^*}{\partial \alpha_i} = \frac{-1}{u(w_i^*)} - \frac{u'(w_i^*)(c(e) - \alpha_i)}{u(w_i^*)^2} w_\alpha \underset{\text{by (3)}}{=} \frac{-1}{u(w_i^*)} - \frac{1}{k} w_\alpha \\
&= \frac{u''(w_i^*)(c(e) - \alpha_i)k - u(w_i^*)u'(w_i^*)}{-u(w_i^*)SOC} < 0.
\end{aligned}
$$

Thus w_i^* and m_i^* decrease in α_i, and so does the governance cost $M_i^* \equiv m_i^* k + w_i^*$.

Both w_i^* and $m_i^* k$ increase in k while m_i^* decreases in k since

$$
\begin{aligned}
w_k &\equiv \frac{\partial w_i^*}{\partial k} = \frac{u'(w_i^*)(c(e) - \alpha_i)}{-SOC} > 0, \\
m_k &\equiv \frac{\partial m_i^*}{\partial k} = -\frac{u'(w_i^*)(c(e) - \alpha_i)}{u(w_i^*)^2} w_k \underset{\text{by (3)}}{=} -\frac{w_k}{k} < 0, \\
\frac{\partial m_i^* k}{\partial k} &= m_i^* + m_k k = m_i^* - w_k \\
&= (c(e) - \alpha_i)\left[\frac{1}{u(w_i^*)} + \frac{1}{k} w_\alpha\right] = (c(e) - \alpha_i)(-m_\alpha) > 0.
\end{aligned}
$$

So M_i^* also increases in k, where $\frac{\partial M_i^*}{\partial k} = m_i^* + m_k k + w_k = m_i^* > 0$.

Proof of Proposition 2 After re-arranging terms, the profit for non-screening principals is

$$
Q_r = hp - w_r - m_r k - \frac{F(\alpha_r)}{F(\widehat{\alpha})}(h(p - q) - m_r w_r),
$$

where $\alpha_r = c(e) - m_r u(w_r)$. Since $\widehat{\alpha}$ is taken as given by an individual non-screening principal, so is w_r^* that is determined by (9) in equilibrium. The first order condition with respect to m_r is

$$
\frac{F(\alpha_r)}{F(\widehat{\alpha})} w_r + (h(p - q) - m_r w_r)\frac{f(\alpha_r)}{F(\widehat{\alpha})} u(w_r) - k = 0.
$$

The second order condition

$$
-\frac{f(\alpha_r)}{F(\widehat{\alpha})} w_r u(w_r) - \frac{(h(p - q) - m_r w_r)}{F(\widehat{\alpha})} f'(\alpha_r) u(w_r)^2 < 0
$$

holds given $f'(\alpha_r) \geq 0$ and $h(p - q) - m_r w_r > 0$, which is true by (7) and footnote 10.

16

Not surprisingly, the principal's profit increases with $\widehat{\alpha}$:

$$\frac{\partial Q_r^*}{\partial \widehat{\alpha}} = \frac{f(\widehat{\alpha})F(\alpha_r)}{F^2(\widehat{\alpha})}(h(p-q) - w_r^* m_r^*) > 0.$$

Since $Q_r^*(\widehat{\alpha} = \underline{\alpha}) = Q_{\underline{\alpha}}^*$, it is obvious that $Q_r^* > Q_{\underline{\alpha}}^*$.

Now we prove the uniqueness of $\widehat{\alpha}$ and its relevant comparative statics. By definition of $\widehat{\alpha}$ we have $I(\widehat{\alpha}) = w_r^*(\widehat{\alpha})$, which can be written as

$$Q_i^*(\widehat{\alpha}) + w_i^*(\widehat{\alpha}) - Q_r^*(\widehat{\alpha}) - s - w_r^*(\widehat{\alpha}) = 0 \tag{16}$$

by plugging in (8). The first derivative of LHS with respect to $\widehat{\alpha}$ is

$$\frac{\partial LHS}{\partial \widehat{\alpha}} = \frac{\partial [Q_i^*(\widehat{\alpha}) + w_i^*(\widehat{\alpha})]}{\partial \widehat{\alpha}} - \frac{\partial [Q_r^*(\widehat{\alpha}) + w_r^*(\widehat{\alpha})]}{\partial \widehat{\alpha}} > 0$$

since the marginal gain of trustworthiness in an unconstrained maximization (measured by the first item) is larger than that in a constrained one (measured by the second item). Intuitively, the non-screening principal, because of imperfect information, adopts a higher-than-necessary wage and monitoring level for agent with $\widehat{\alpha}$; this is a waste of resources compared to the complete information case for the screening principal. Since $LHS(\widehat{\alpha} = \underline{\alpha}) = -s < 0$ while $\frac{\partial LHS}{\partial \widehat{\alpha}} > 0$, there exists a unique threshold $\widehat{\alpha}$ such that (16) holds.

For the same reason as $\frac{\partial LHS}{\partial \widehat{\alpha}} > 0$, we get

$$\frac{\partial LHS}{\partial k} = \frac{\partial [Q_i^*(\widehat{\alpha}) + w_i^*(\widehat{\alpha})]}{\partial k} - \frac{\partial [Q_r^*(\widehat{\alpha}) + w_r^*(\widehat{\alpha})]}{\partial k} > 0.$$

From equation (16) we get

$$\frac{\partial \widehat{\alpha}}{\partial s} = -\frac{-1}{\partial LHS/\partial \widehat{\alpha}} > 0,$$

$$\frac{\partial \widehat{\alpha}}{\partial k} = -\frac{\partial LHS}{\partial k} \Big/ \frac{\partial LHS}{\partial \widehat{\alpha}} < 0.$$

Together with $\frac{\partial Q_r^*}{\partial \widehat{\alpha}} > 0$, it is easy to see a principal's profit Q_r^* increases with s.

Proof of Proposition 4 The objective function is

$$\max_{\alpha_{n+1,i}} M_0^* - \frac{c(e) - \alpha_{n+1,i}}{u(w_{n+1,i}^*)} k - C(\alpha_{n+1,i}, \alpha_{ni}) - s, \tag{17}$$

where $w_{n+1,i}^*$ is specified in (3). The first order condition for an interior solution is

$$\frac{u(w_{n+1,i}^*) + u'(w_{n+1,i}^*)(c(e) - \alpha_{n+1,i})w_\alpha}{u(w_{n+1,i}^*)^2} k - C_1(\alpha_{n+1,i}, \alpha_{ni}) = 0, \tag{18}$$

17

where $w_\alpha \equiv \frac{\partial w^*_{n+1,i}}{\partial \alpha_{n+1,i}}$. Plugging (3) into (18) we get

$$\frac{k}{u(w^*_{n+1,i})} + w_\alpha - C_1(\alpha_{n+1,i}, \alpha_{ni}) = 0, \tag{19}$$

which yields the unique optimal choice $\alpha^*_{n+1,i} \equiv g(\alpha_{ni}) \in [0, c(e)]$ in each generation n when the second order condition

$$SOC_2 \equiv \frac{-u'kw_\alpha}{u^2} + w_{\alpha\alpha} - C_{11} < 0$$

holds; when $SOC_2 > 0$, then the marginal benefit of investing $\alpha^*_{n+1,i}$ is increasing so that $\alpha^*_{n+1,i} = c(e)$.

$\alpha^*_{n+1,i}$ is increasing in α_{ni} because

$$\frac{\partial \alpha^*_{n+1,i}}{\partial \alpha_{ni}} = \frac{-C_{12}(\alpha_{n+1,i}, \alpha_{ni})}{-SOC_2} > 0.$$

It's also increasing in k when $-m_\alpha + w_{\alpha k} > 0$ since

$$\frac{\partial \alpha^*_{n+1,i}}{\partial k} = \frac{\frac{\partial m^*_{n+1,i}k}{\partial k}\frac{1}{c(e)-\alpha_{n+1,i}} + w_{\alpha k}}{-SOC_2} = \frac{-m_\alpha + w_{\alpha k}}{-SOC_2}.$$

When the trustworthiness level $\alpha_{n+1,i}$ is endogenous, the governance cost

$$M_{n+1,i} = w^*_{n+1,i} + \frac{c(e) - \alpha_{n+1,i}}{u(w^*_{n+1,i})}k$$

may actually increase when k is lower, in contrary to the short run result. This happens when the elasticity of $\alpha_{n+1,i}$ over k is high enough. After cancelling terms and plugging in condition (3), we get

$$\frac{\partial M_{n+1,i}}{\partial k} = \underbrace{\left(\frac{\partial \alpha_{n+1,i}}{\partial k}\frac{k}{\alpha_{n+1,i}}\right)\frac{-\alpha_{n+1,i}}{u(w^*_{n+1,i})}}_{\substack{k\text{'s effect on} \\ \text{agent trustworthiness}}} + \underbrace{\frac{c(e) - \alpha_{n+1,i}}{u(w^*_{n+1,i})}}_{\substack{k\text{'s effect on} \\ \text{short-run governance cost}}}.$$

It is negative when the elasticity is large enough: $\frac{\partial \alpha_{n+1,i}}{\partial k}\frac{k}{\alpha_{n+1,i}} > \frac{c(e)}{\alpha_{n+1,i}} - 1$. At the steady state this condition becomes $\frac{\partial \alpha_c}{\partial k}\frac{k}{\alpha_c} > \frac{c(e) - \alpha_c}{\alpha_c}$.

Furthermore, $g(\alpha_{ni})$ is concave if

$$\frac{\partial^2 \alpha^*_{n+1,i}}{\partial \alpha^2_{ni}} = \frac{SOC_2 C_{122}(\alpha_{n+1,i}, \alpha_{ni}) + C_{112}C_{12}(\alpha_{n+1,i}, \alpha_{ni})}{(SOC_2)^2} \leq 0$$

holds, which is true when $C_{122} \geq 0$ and $C_{112} \geq 0$. Since the transition function $\alpha^*_{n+1,i} \equiv g(\alpha_{ni})$ is increasing and concave in $\alpha_{ni} \in [0, A]$, $g(0) > 0$ and $g(A) \leq c(e)$, there exists a

18

unique steady state trustworthiness level α_c such that $\alpha_c = g(\alpha_c)$. Note that $\frac{\partial \alpha_c}{\partial k} > 0$ since a higher k shifts up the transition function due to $\frac{\partial \alpha_{n+1,i}^*}{\partial k} > 0$.

Based on $\pi = 1 - F(\widetilde{\alpha})$ and (12) we get

$$\frac{\partial \pi}{\partial s} = -f(\widetilde{\alpha})\frac{\partial \widetilde{\alpha}}{\partial s} = f(\widetilde{\alpha})\frac{R'(s)}{R'(\widetilde{\alpha})} = f(\widetilde{\alpha})\frac{-1}{-C_2} < 0,$$

$$\frac{\partial \pi}{\partial k} = -f(\widetilde{\alpha})\frac{\partial \widetilde{\alpha}}{\partial k} = f(\widetilde{\alpha})\frac{R'(k)}{R'(\widetilde{\alpha})} = f(\widetilde{\alpha})\frac{m_0^* - m_i^* + \partial w_0^*/\partial k}{-C_2} > 0.$$

So the proportion of agents with positive trustworthiness decreases in screening cost s while increases in monitoring cost k.

References

Alger, I., Renault, R.: Screening ethics when honest agents care about fairness. Int Econ Rev **47**, 59–85 (2006)

Alger I., Renault, R.: Screening ethics when honest agents keep their word. Econ Theory **30**, 291-311 (2007)

Akerlof G., Kranton R.: Identity and the economics of organizations. J of Econ Perspect **19**, 9-32 (2005)

Appelbaum, E., and Batt R.: The new American workplace: transforming work systems in the U.S. Ithaca, New York: Cornell ILR Press 1994

Bar-Gill O., Fershtman C.: Public policy with endogenous preferences. J Public Econ Theory **7**, 841-857 (2005)

Baron J., Kreps D.: Strategic human resources. New York: John Wiley & Sons 1999

Becker G.: Investment in human capital: a theoretical analysis. J Polit Econ **70**, Part 2: Investment in Human Beings, 9-49 (1962)

Bohnet I., Frey B., Huck S.: More Order with Less Law: On Contract Enforcement, Trust, and Crowding. Am Polit Sci Rev **95**, 131-144 (2001)

Cappelli P.: Is the 'skills gap' really about attitudes? Cal Manage Rev Reprint Series **37**, 108-124 (1995)

Cappelli P., Neumark D.: Do 'high performance' work practices improve establishment-level outcomes? Indust Labor Relations Rev **54**, 737-775 (2001)

Commission on the Future of Worker-Management Relations, appointed by U.S. Secretaries of Labor and Commerce, Fact Finding Report (1994)

Cook K.: (Eds.) Trust in society. New York: Russell Sage Foundation 2001

Durlauf S., Fafchamps M.: Social capital. In: Aghion P., Durlauf S. (Eds.), Handbook of economic growth, Amsterdam: North Holland 2005

Esping-Anderson G.: The three worlds of welfare capitalism. Oxford: Polity Press 1990

Frank, R.: If homo economicus could choose his own utility function, would he want one with a conscience? Am Econ Rev **77**, 593-604 (1987)

Frey B.: Does monitoring increase work effort? the rivalry with trust and loyalty, Econ Inquiry **31**, 663-70 (1993)

Fukuyama F.: Trust: the social virtues and the creation of prosperity. New York, Simon & Schuster 1995

Gordon D.: Bosses of different stripes: a cross-national perspective on monitoring and supervision. Am Econ Rev **84**, 375-379 (1994)

Greif A.: Contract enforceability and economic institutions in early trade: the Maghribi traders' coalition. Am Econ Rev **83**, 525–48 (1993)

Greif A.: Contracting, enforcement, and efficiency: economics beyond the law. In: Bruno M., Pleskovic B. (Eds.), Annual World Bank Conference on Development Economics. Washington, D.C: The World Bank 1997

Güth, W., Ockenfels A.: The coevolution of morality and legal institutions: an indirect evolutionary approach. J of Institut Econ **1**, 155-174 (2005)

Huang F., Cappelli P.: Employee screening: theory and evidence. NBER working paper 12071 (2006)

Huck, S.: Trust, treason, and trials: an example of how the evolution of preferences can be driven by legal institutions. J Law Econ Organ **14**, 44-60 (1998)

James H.: The trust paradox: a survey of economic inquiries into the nature of trust and trustworthiness. J Econ Behavior and Organ **47**, 291-307 (2002)

Kaplow L., Shavell S.: Moral rules, the moral sentiments, and behavior: toward a theory of an optimal moral system. J Polit Econ **115**, 494–514 (2007)

Kipnis D.: Trust and technology. In: Kramer R.M., Tyler T.R. (Eds.), Trust in organizations: frontiers of theory and research. Thousand Oaks, California: Sage Publications 1996

Kreps D.: Intrinsic motivation and extrinsic incentives. Am Econ Rev **87**, 359-364 (1997)

Mills D.: Labor-management relations. New York : McGraw-Hill, 5th ed. 1994

Nagin D., Rebitzer J., Sanders S., Taylor L.: Monitoring, motivation, and management: the determinants of opportunistic behavior in a field experiment. Am Econ Rev **92**, 850–873 (2002)

Palfrey, T., Prisbrey, J.: Anomalous behavior in public goods experiments: how much and why? Am Econ Rev **87**, 829-846 (1997)

Putnam R.: Bowling alone: America's declining social capital. J Democracy **6**, 65–78 (1995)

Rob R., Zemsky P.: Social capital, corporate culture and the incentive intensity. RAND J Econ **33**, 243-257 (2002)

Rubery J., Grimshaw D.: The organization of employment: an international perspective. Palgrave MacMillan 2003

Shavell S.: Law versus morality as regulators of conduct. Am Law and Econ Rev **4**, 227–257 (2002)

Sliwka, D.: Trust as a signal of a social norm and the hidden costs of incentive schemes. Am Econ Rev **97**, 999–1012 (2007)

Sobel J.: For better or forever: formal versus informal enforcement. J Labor Econ **24**, 271–297 (2006)

Stout L., Blair M.: Trust, trustworthiness, and the behavioral foundations of corporate law. Univ Penn Law Rev **149**, 1735-1810 (2001)

Vernon G.: Work organization and comparative historical statistics on the extent of the managerial hierarchy. SKOPE Research Paper No. 18, University of Oxford (2001)