

TI 2013-198/III
Tinbergen Institute Discussion Paper



Endogenous Effort Norms in Hierarchical Firms

Jan Tichem

Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute, The Netherlands.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Endogenous Effort Norms in Hierarchical Firms*

Jan Tichem[†]

December 12, 2013

Abstract

This paper studies how a three-layer hierarchical firm (principal-supervisor-agent) optimally creates effort norms for its employees. The key assumption is that effort norms are affected by the example of superiors. In equilibrium, norms are eroded as one moves down the hierarchy. The reason is that, because exerting effort is costly, the supervisor only partially complies with the principal's example, and thereby transmits a lower norm to the agent. The principal optimally responds to norm erosion by setting a higher example to begin with. In equilibrium, norm erosion gives rise to three inefficiencies: the principal works too hard, the supervisor's norm is too high, and the agent's norm is too low. To reduce these inefficiencies, firms should keep the extent of hierarchy to a minimum, promote employees with the strongest sensitivity to social norms, and distort managerial spans of control.

Keywords: delayering, hierarchy, leading by example, norms, promotion, span of control

JEL Classification: D23, M50, M51

*I gratefully acknowledge helpful comments and suggestions by Robert Dur, Michiel Souverijn, and Robin Zoutenbier.

[†]Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute. E-mail: tichem@ese.eur.nl.

1 Introduction

Empirical research suggests that employees have a preference to conform to social norms within their firm, even if this requires taking costly actions.¹ Given that employees conform to social norms, firms have an incentive to affect these norms. After all, a firm only benefits from employee conformism to the extent that the existing norms prescribe to act in the firm's interest. From a practical point of view, an important question is therefore how firms can increase profits by creating efficient norms. From a theoretical point of view, an important question is whether and how firms' desire to create efficient norms can explain firm behavior and organization. This paper develops and analyzes a theoretical model that yields answers to both questions.

The novelty of the present analysis is its focus on a particular determinant of norms for employees, namely the actual behavior of superiors. Apart from being intuitive, there is empirical evidence suggesting that the example of superiors has normative implications for subordinates. Treviño et al. (1999) find a strong negative correlation between employees' perceptions of, on the one hand, the quality of ethical leadership of executives and supervisors within their firm, and on the other hand, unethical behavior within their firm (also see Treviño and Weaver 2003, chapter 9, where similar evidence is reported). In line with this, Posner and Schmidt (1992) document that 92% of American managers agree with the statement that "the behavior of those in charge is the *principle* determinant of the 'ethical tone' of my firm". This view is confirmed by J. Irwin Miller, a successful CEO himself, who claims that "all of the corporate standards of ethics don't mean anything unless the persons in the corporation perceive the top people to abide by them when the going is really tough" (quoted from Murphy and Enderle 1995).

The formal model studied below considers a hierarchical firm that consists of three layers: a residual claimant principal, a supervisor, and an agent. All players exert non-verifiable effort which yields valuable production for the principal's firm. Both employees (the supervisor and the agent) incur a psychological cost if their effort falls short of their norm for effort, as in e.g. Fischer and Huddart (2008). To reflect the evidence cited above, employees consider the example of their superior to be the norm for their own effort. Hence, the principal's effort constitutes the norm for the supervisor, and the supervisor's effort constitutes the norm for the agent. The remaining part of the Introduction describes the results of the paper.

¹Important studies documenting this phenomenon are Ichino and Maggi (2000), Falk and Ichino (2006), Bradley et al. (2007), Mas and Moretti (2009), and Bradler et al. (2013). See the next section for a detailed discussion of these papers.

The key result from the analysis is that, in equilibrium, norms are lower as one moves down the hierarchy. I call this phenomenon ‘norm erosion’. Norm erosion arises because when exerting effort, employees take into account the example of their superior, but never fully comply with it. The reason is simple: employees find it privately optimal to trade off some costs of performing below the norm against the benefit of avoiding some effort costs. It follows that each employee sets an example for his subordinate that is lower than the example he got from his superior. The upshot from this result is that effort norms within firms are not only determined by the example set at the top of the hierarchy, but also by how this example is transmitted to lower organizational layers. Since firms do not control this process, they incur a number of costs when creating optimal effort norms for employees, as explained next.

To understand why norm erosion is costly to firms, one needs to know that there is a unique first-best effort norm for employees. This holds because a higher norm induces more valuable effort from the employee, but also imposes more effort costs and norm violation costs upon the employee for which he must be compensated through his salary. The first-best norm optimally trades off these marginal costs and benefits. Next, when choosing how much effort to exert, the principal has two objectives in mind. On the one hand, she wants to set the first-best norm for her employees. On the other hand, she wants to maximize her own contribution to firm profits. The effort level that maximizes the principal’s own contribution to firm profits is referred to as the principal’s first-best effort level. Under a mild assumption that I use, it holds that the principal’s first-best effort level is identical to the first-best norm for both employees. It follows that if there were no norm erosion, the principal would maximize total profits by exerting her first-best effort level. However, knowing that the supervisor will erode her example, the principal optimally sets a higher example to begin with. In this way the principal makes sure that the agent’s equilibrium norm is closer to its first-best level. The cost of this strategy is two-fold: the principal works inefficiently hard, and the supervisor faces an inefficiently high norm. Compared to the first-best case, hierarchical firms are thus confronted with three inefficiencies in norm creation: the principal’s effort is too high, the supervisor’s norm is too high, and the agent’s norm is too low. The remaining results show that, to minimize these inefficiencies, hierarchical firms optimally adjust the way they are organized.

First, hierarchical firms optimally promote employees with the strongest sensitivity to social norms. A supervisor who is more sensitive to norms will conform to the principal’s example to a larger extent. As a result, norms are eroded to a lower extent, which is always valuable to the principal. The reason is that establishing a

given effort norm for the agent now requires a lower effort level from the principal. This is profitable since in equilibrium the principal works inefficiently hard and the supervisor's norm is inefficiently high. Alternatively, for a given effort level of the principal and norm for the supervisor, the agent will be faced with a higher norm. This is also profitable because in equilibrium the agent's norm is inefficiently low.

Second, hierarchical firms optimally distort the supervisor's span of control, that is, the number of agents heading under one supervisor. The analysis shows that firms may set the supervisor's span of control both above and below its first-best level. The mechanism that produces these results is the following. In the model I assume that for supervision technology reasons, there is some first-best supervisor span of control. When choosing how many supervisors and agents to hire, the principal thus has an incentive to stick as closely as possible to this exogenously given span of control. However, the firm's profits also depend on which kind of employee faces the most efficient effort norm in equilibrium. Recall that the equilibrium norm for supervisors is always above the first-best level, whereas the equilibrium norm for agents is always below the first-best level. Yet, the extent to which equilibrium norms diverge from the first-best level may be different for agents and supervisors. In fact, the principal optimally makes sure that the kind of employee that is relatively abundant faces the more efficient norm. The principal does this by raising her own effort level in the number of agents, and by decreasing her effort in the number of supervisors. Therefore, given that the supervision technology is such that it is attractive to hire relatively many agents (supervisors), the principal optimally makes sure that the effort norm for agents (supervisors) is more efficient than the effort norm for supervisors (agents), which in turn makes it attractive to hire even more agents (supervisors). As a result, the principal optimally distorts the supervisor's span of control away from its first-best level.

Last, hierarchical firms suffer from norm erosion precisely because they are hierarchical. A straightforward prediction following from the model is therefore that firms optimally keep the extent of hierarchy to a minimum.

The paper proceeds as follows. The next section describes the related literature. Section 3 develops the formal model. Section 4 solves the model, while section 5 explores implications for organizational design. Section 6 finishes with concluding remarks.

2 Related literature

Empirical research suggests that employees have a preference to conform to social norms within their firm. Ichino and Maggi (2000) show that employees engage less

in misconduct and absent themselves less often if their co-workers are less likely to do so (also see Bradley et al. 2007). Falk and Ichino (2006) report that, if employees work in pairs rather than individually, the standard deviation of output is smaller and output is higher. Moreover, the authors find that low-productivity employees respond strongest to working in pairs. Mas and Moretti (2009) find that supermarket cashiers improve their performance if a high-productivity cashier enters their shift. As in Falk and Ichino (2006), this effect is far more pronounced for low-productivity workers than for high-productivity workers. Mas and Moretti (2009) provide a social pressure interpretation for their results, based on the finding that cashiers only improve their performance if this can be observed by the new cashier. Finally, in a controlled work environment, Bradler et al. (2013) show that the provision of public recognition to employees improves their performance. Similar to Falk and Ichino (2006) and Mas and Moretti (2009), the authors find that their results are mainly driven by those employees who did not receive recognition, and thus learned that they performed worse than others. For this reason, Bradler et al. (2013) explain their results partly in terms of employees' preference to conform to a group norm.

An early theoretical paper studying social norms is Akerlof (1980). Akerlof's (1980) model shows that an existing norm may stay in place if deviating from the norm leads to a loss in reputation (also see Bernheim 1994, who assumes that people care for others' perceptions of one's preferences). Another early contribution is the paper by Kandell and Lazear (1992), who study peer pressure as a mechanism that may reduce free-riding problems in team production.

More recently, Sliwka (2007) develops a model where agents may prefer to conform their effort to the social norm, but are uncertain about what the social norm prescribes. This opens the door for the principal to affect the agent's perception of the social norm through her own actions. Specifically, Sliwka (2007) shows that offering incentives may be a credible signal that the social norm is to act selfishly, whereas offering no incentives may be a credible signal that the social norm is to act fairly, that is, to exert at least some effort. Fischer and Huddart (2008) study firms where norms exist for desirable actions (like exerting effort) and undesirable actions (like earnings management). The authors show that social norms multiply the impact of individual incentives on agents' behavior. The reason is that individual incentives change an individual agent's behavior, but this also affects the social norm to which all agents want to conform. Fischer and Huddart (2008) also show that it may be optimal to split firms in parts. The benefit is that having separate departments allows for the cultivation of different social norms for different tasks.

Stevens and Thevaranjan (2010) study optimal efficiency wage contracts for morally sensitive agents. The authors assume that the principal can declare what

the effort norm is at the contracting stage. Stevens and Thevaranjan (2010) show that, when hiring a morally sensitive agent, the principal may achieve first-best profits even when she does not use incentives. Huck et al. (2012) define social norms in terms of Pareto efficiency. Given this definition, the authors demonstrate that a social norm makes team-incentives more effective. The reason is that team-incentives create positive externalities among employees, which leads to a higher social norm for effort. The opposite reasoning applies to using relative incentives like tournaments. Huck et al. (2012) further show that, if team incentives are used, social norms may give rise to multiple equilibria, some of which may lead to inefficiently high social norms.

Akerlof and Kranton (2005) introduce the notion that a person's utility may depend on his or her social identity. In their model, a social identity exogenously implies some norm for behavior. Akerlof and Kranton (2005) show that employees identifying with the firm accept lower wages and require lower incentives. The authors consequently claim that it is valuable for firms to invest in changing their employees' identity (also see Heinle et al. 2012, who derive optimal contracts when agents may identify with the firm in a multi-task setting). Carlin and Gervais (2009) model morality as a self-imposed restriction to exert high effort. In their analysis, norms are thus a purely personal trait that cannot be affected by the firm in any way. The authors derive a number of predictions as to how the presence of virtuous agents in the labor market affect firms' optimal contract design, project choice, and extent of bureaucracy.

In contrast to all papers above, the starting point of the analysis here is that the actual example of superiors determines effort norms for employees. Norms are thus not exogenously given (as in Akerlof and Kranton 2005, Carlin and Gervais 2009, and Heinle et al. 2012), but depend on the behavior of superiors. Also, norms cannot be costlessly declared (as in Stevens and Thevaranjan 2010), but creating norms requires setting a costly example. Finally, the principal does not only indirectly influence norms through making institutional choices (as in Sliwka 2007, Fischer and Huddart 2008, and Huck et al. 2012), but also directly through setting her own example.

More distantly related to the present paper is Hermalin's (1998) analysis of leadership. Hermalin (1998) studies how a leader can credibly communicate information to his team-members about the marginal productivity of their effort. One of the mechanisms Hermalin (1998) considers is 'leading by example'. That is, if the leader exerts a high effort level and thereby incurs high effort costs, this credibly signals to the other team-members that exerting effort is valuable. The parallel between Hermalin (1998) and the present analysis is the importance of the leader's example

for the behavior of others. However, different problems are studied. In Hermalin's (1998) analysis the credible transmission of information is at stake, whereas here the principal sets her example so as to create optimal effort norms for her employees.

Finally, another important difference with the existing literature is that the present analysis considers hierarchical firms that consist of at least three layers. Hence, I do not only study how the example at the top of the hierarchy matters for effort norms, but also the process of transmitting this example to lower layers of the hierarchy. In fact, the key result of the paper is to show that this process is inefficient from the firm's perspective.

3 A model of effort norms in hierarchical firms

Consider a hierarchical firm owned by one principal. The principal's organizational layer is denoted by $l = 0$. In the most basic specification of the model, the principal hires one supervisor ($l = 1$) and one agent ($l = 2$) (subsection 5.2 extends the model to the case where the principal hires S supervisors and A agents). All players exert unverifiable effort e_l , which yields valuable production for the principal's firm. The marginal product of effort is given by p_l . I assume that $p_l \geq p_{l+1}$, implying that a player in a higher layer of the organization has a weakly higher marginal productivity. The costs of exerting effort are given by $\frac{1}{2}\theta e_l^2$.

Employee utility from the job is given by:

$$U_l = w_l - \frac{1}{2}\theta e_l^2 - \frac{1}{2}\gamma (n_l - e_l)^2, \quad (1)$$

where w_l denotes a fixed wage (in the Appendix, I show that as long as employees are risk averse, the results of the paper are insensitive to the use of performance pay). The last term in (1) describes the employee's preference for conforming to social norms, which consists of two parts. The term $n_l - e_l$ describes how much the employee's effort differs from the norm for effort, n_l . The parameter γ reflects the strength of the employee's sensitivity to norms. In the basic model, γ is assumed to be identical for all employees (section 5.1 extends the model to the case where employees differ in norm sensitivity). Throughout the paper, I assume that the principal can observe γ . The outside option utility is assumed to be identical for all employees, and equal to $u = 0$.

The novelty of the model is how social norms for effort are determined. Formally, the effort norm for an employee in organizational layer l is given by:

$$n_l = e_{l-1} \frac{p_l}{p_{l-1}}. \quad (2)$$

The effort norm consists of two parts. First, as laid out in the Introduction, the actual behavior of superiors plays an important role in shaping norms within organizations. Hence, I assume that the supervisor’s norm is affected by the principal’s effort, and the agent’s norm is affected by the supervisor’s effort.² Second, I assume that employees are aware of the fact that their superior’s effort may be more valuable than their own ($p_{l-1} > p_l$). If this is the case, from an efficiency point of view, the superior should work harder than the employee. For this reason, I assume that employees do not consider their superior’s effort to be normative to the extent that the superior’s marginal productivity is higher. This behavioral assumption implies that, qualitatively, the results of the paper do not depend on differences in marginal productivities p_l . For convenience, I therefore solve the model assuming that marginal productivities are identical across organizational layers, that is, $p_l = p$ for all l . It follows that norms are given by:

$$n_l = e_{l-1}. \tag{3}$$

The order of the game is as follows. First, the principal exerts an effort level, and thereby sets an example. Second, the principal hires at least one supervisor and one agent, and offers all employees a compensation contract. Third, the employees decide whether or not to accept the contract. If one of them rejects the contract, the game ends. Last, if all employees accepted the contract, they exert effort and payoffs realize. In the next section, I solve for the subgame perfect equilibrium of the game by using backwards induction.

4 Analysis

4.1 Contractible effort

In this subsection, I derive the first-best benchmark where the employees’ effort is verifiable and thus contractible. First note that, to satisfy the employees’ participation constraint, they must be paid a minimum salary equal to:

$$w_l = \frac{1}{2}\theta e_l^2 + \frac{1}{2}\gamma(n_l - e_l)^2. \tag{4}$$

²The agent’s norm may also directly be affected by the example of the principal. However, this would not affect the results of the paper qualitatively, as long as the agent attaches a non-negative weight to the example of the supervisor. This is not a strong assumption, since supervisors typically represent the firm and its policies toward their subordinates.

Given the lowest possible salary (4), and taking into account that the norm for layer l is given by the effort level exerted in layer $l - 1$, firm profits can be written as:

$$\Pi = \underbrace{pe_0 - \frac{1}{2}\theta e_0^2}_{\text{principal } (l=0)} + \underbrace{pe_1 - \frac{1}{2}\theta e_1^2 - \frac{1}{2}\gamma(e_0 - e_1)^2}_{\text{supervisor } (l=1)} + \underbrace{pe_2 - \frac{1}{2}\theta e_2^2 - \frac{1}{2}\gamma(e_1 - e_2)^2}_{\text{agent } (l=2)}. \quad (5)$$

Maximizing firm profits (5) over all players' effort levels yields that all players optimally exert the effort level $e_l^{FB} = \frac{p}{\theta}$. From this it follows directly that the first-best effort norm for employees is given by $n_l^{FB} = \frac{p}{\theta}$. First-best profits become:

$$\Pi^{FB} = \underbrace{\frac{1}{2} \frac{p^2}{\theta}}_{\text{principal } (l=0)} + \underbrace{\frac{1}{2} \frac{p^2}{\theta}}_{\text{supervisor } (l=1)} + \underbrace{\frac{1}{2} \frac{p^2}{\theta}}_{\text{agent } (l=2)}. \quad (6)$$

Lemma 1 *In the first-best case where effort is contractible, all players exert the effort level $e_l^{FB} = \frac{p}{\theta}$. The first-best effort norm for both employees therefore equals $n_l^{FB} = \frac{p}{\theta}$.*

Note that, when choosing her effort, the principal has two objectives in mind. On the one hand, she wants to maximize her individual production (first term of (5)). On the other hand, she wants to set the most efficient norm for the supervisor (second term of (5)). However, the principal does not face a trade-off between these two objectives. That is, the first-best effort level e_0^{FB} both maximizes the principal's individual production, and sets the most efficient norm for the supervisor. The reason is that the principal and the supervisor have the same marginal productivity of effort p . Moreover, even if the supervisor had a lower marginal productivity than the principal ($p_1 < p_0$), the principal would not face a trade-off between the objectives mentioned above. This holds because of the assumption that, to the extent that the principal's marginal productivity is higher, the supervisor does not consider the principal's effort level to be normative (see equation (2)). Also note that the exact same considerations apply when the principal mandates the supervisor's effort level. In the next subsection, I show that when effort is not contractible, the principal does face a trade-off between maximizing the value of her own production and setting efficient norms for her employees.

4.2 Non-contractible effort

In case effort is not contractible, the first step in solving the model is to derive the employees' effort level for any given effort norm and compensation contract. Since

the employee's wage w_l is fixed, the effort level that maximizes employee utility (1) equals:

$$e_l^* = \mu n_l, \quad (7)$$

where $\mu \equiv \frac{\gamma}{\gamma+\theta}$, and it holds that $0 < \mu < 1$. It can easily be verified that employee effort increases in μ , and therefore in the employee's norm sensitivity γ . The intuition is that deviating from the norm is more costly for an employee who possesses stronger norm sensitivity. Next, employee effort also increases in the norm n_l , as a higher norm implies higher costs of norm violation for a given level of effort. Finally, since it holds that $\mu < 1$, an employee always exerts less effort than his norm prescribes. The intuition is that employees trade off some costs of performing below the norm against the benefit of shirking. This simple fact has an important implication for hierarchical firms, namely that each employee transmits a norm towards the next layer of the organization that is lower than the norm he faces himself. Therefore, as one moves down the hierarchy, effort norms decrease. I call this phenomenon 'norm erosion'. The existence of norm erosion is the first result of the analysis, and all remaining results follow from this one.

Proposition 1 *A feature of hierarchical firms is norm erosion, that is, $n_l^* > n_{l+1}^*$ for all l .*

The second step in solving the model is to derive the optimal wage w_l . The optimal wage is the lowest possible wage that induces the employee to accept the job, conditional on the effort level the employee exerts (e_l^*), and the norm he faces (n_l). This constraint reads $U_l(w_l, e_l^*, n_l) \geq 0$. Rewriting this condition to w_l yields:

$$w_l^* = \frac{1}{2}\theta\mu n_l^2. \quad (8)$$

The employee's wage increases in the norm he faces. The reason is two-fold. First, a higher norm induces the agent to exert more effort. Second, given his effort choice, a higher norm implies that the agent will incur higher costs of norm violation. The optimal wage also increases in μ , and therefore in employee norm sensitivity γ . Norm sensitivity has two effects on the wage. First, the employee incurs higher costs of effort if norm sensitivity is stronger, implying the wage must increase. Second, norm sensitivity has two effects on the employee's costs of norm violation. On the one hand, given his effort level, stronger norm sensitivity implies that the employee incurs greater psychological costs from performing below the norm. On the other hand, higher norm sensitivity implies that the employee exerts a higher effort level, which reduces the violation of the norm. However, the net effect of norm sensitivity on the wage is always positive, as made apparent by equation (8).

The final step in solving the model is to derive the principal's effort level. Conditional on the definition of effort norms (3), the employees' effort levels (7) and wages (8), total firm profits can be written as:

$$\Pi = \underbrace{pe_0 - \frac{1}{2}\theta e_0^2}_{\text{principal } (l=0)} + \underbrace{p\mu e_0 - \frac{1}{2}\theta\mu e_0^2}_{\text{supervisor } (l=1)} + \underbrace{p\mu^2 e_0 - \frac{1}{2}\theta\mu^3 e_0^2}_{\text{agent } (l=2)}. \quad (9)$$

Clearly, the principal's effort not only affects the value of her individual production (first term), but, by setting an example, also affects the value of hiring the supervisor (second term) and the agent (third term). This is also reflected in the effort level the principal optimally exerts. Maximizing firm profits over e_0 yields that this is equal to:

$$e_0^* = \frac{p(1 + \mu + \mu^2)}{\theta(1 + \mu + \mu^3)}. \quad (10)$$

It can be easily checked that the principal's effort exceeds the first-best level derived in Lemma 1. This is the second result of the analysis:

Proposition 2 *The principal exerts an effort level that is higher than her first-best effort level ($e_0^* > e_0^{FB}$).*

To understand the result of Proposition 2, recall from Lemma 1 that the first-best norm for the supervisor and agent is given by $n_l^{FB} = \frac{p}{\theta}$. Hence, if the principal would exert the effort level $e_0^{FB} = \frac{p}{\theta}$, she would maximize the value of her individual production, and set the first-best norm for the supervisor. However, as derived in Proposition 1, the supervisor will erode the principal's example. The implication is that, if $e_0 = \frac{p}{\theta}$, the equilibrium norm for the agent will be inefficiently low. In response to this, the principal optimally sets a higher example to begin with, and thus raises her effort above e_0^{FB} .

Note that the result of Proposition 1 directly implies that the supervisor's norm is above the first-best level (recall that $n_l^{FB} = \frac{p}{\theta}$). In addition to this, it holds that the agent's norm is below the first-best level.³ Intuitively, it can never be optimal that the agent's norm exceeds the first-best level. After all, driving the agent's norm closer to the first-best level imposes two costs on the principal: the principal's effort is inefficiently high (Proposition 2), and as a result, the supervisor's norm is inefficiently high. It follows that it may be optimal to raise the agent's norm up to the first-best level, but not higher than that. Hence, norm erosion ultimately implies the following Corollary.

³To see this formally, note that the inequality $n_2^* < n_2^{FB}$ reduces to $\frac{1-\mu^2}{1+\mu+\mu^3} > 0$, which always holds because $0 < \mu < 1$.

Corollary 1 *Compared to the first-best benchmark, the effort norm for the supervisor is inefficiently high ($n_1^* > n_1^{FB}$), whereas the effort norm for the agent is inefficiently low ($n_2^* < n_2^{FB}$).*

Another result that norm erosion gives rise to is described in Corollary 2.

Corollary 2 *Firms optimally keep the number of organizational layers to a minimum.*

The intuition behind Corollary 2 is straightforward. So far, the analysis has shown that hierarchical firms are confronted with three inefficiencies in creating optimal norms for employees: the principal works too hard, the supervisor's norm is too high, and the agent's norm is too low. Hierarchical firms incur these costs because of norm erosion, which arises if intermediate layers of management are in place. Moreover, for any organizational layer added to the firm, the inefficiencies from norm erosion would be magnified. It follows that firms optimally keep the number of organizational layers to a minimum. In the remainder of the paper, I assume that the principal requires supervisors to let agents do their job. A motivation for this assumption is that the principal simply cannot supervise all the agents she hires, and hence needs to delegate this task. Also see subsection 5.2, where I study how many supervisors and agents the principal optimally hires.

5 Implications for organizational design

This section offers two extensions of the basic model analyzed above. Both extensions yield a result as to how hierarchical firms can be better designed to reduce the inefficiencies stemming from norm erosion. In following order, I discuss promotion decisions and managerial spans of control.

5.1 Promotion decisions

In contrast to what has been assumed in the basic model, employees are likely to differ in their sensitivity to social norms. Given this heterogeneity, a natural question is which kind of employee the principal should optimally promote to the position of supervisor. To study this question, denote by γ_q the norm sensitivity of employee q , and by γ_r the norm sensitivity of employee r . Using the definition $\mu \equiv \frac{\gamma}{\theta + \gamma}$, one can

also define μ_q and μ_r . Firm profits can now be written as:

$$\Pi = \underbrace{pe_0 - \frac{1}{2}\theta e_0^2}_{\text{principal } (l=0)} + \underbrace{p\mu_q e_0 - \frac{1}{2}\theta\mu_q e_0^2}_{\text{supervisor } (l=1)} + \underbrace{p\mu_q\mu_r e_0 - \frac{1}{2}\theta\mu_q^2\mu_r e_0^2}_{\text{agent } (l=2)}, \quad (11)$$

where it has been assumed that employee q is appointed supervisor.

Next, the principal optimally exerts the level of effort that maximizes firm profits. This effort level can be shown to be equal to:

$$e_0^* = \frac{p}{\theta} \frac{1 + \mu_q + \mu_q\mu_r}{1 + \mu_q + \mu_q^2\mu_r}, \quad (12)$$

which is greater than the first-best level, as in Proposition 2. Equilibrium profits become:

$$\Pi^* = \frac{1}{2} \frac{p^2 (1 + \mu_q + \mu_q\mu_r)^2}{\theta (1 + \mu_q + \mu_q^2\mu_r)}. \quad (13)$$

Note that in case employee r is promoted to the position of supervisor, equilibrium profits are simply given by interchanging the subscripts q and r in (13).

To determine whether the principal optimally promotes employee q or r , one can compute which case yields higher profits. Promoting employee q yields higher profits than promoting employee r if:

$$\frac{1}{2} \frac{p^2 (1 + \mu_q + \mu_q\mu_r)^2}{\theta (1 + \mu_q + \mu_q^2\mu_r)} > \frac{1}{2} \frac{p^2 (1 + \mu_r + \mu_r\mu_q)^2}{\theta (1 + \mu_r + \mu_r^2\mu_q)}. \quad (14)$$

After some rewriting, this inequality reduces to $\mu_q > \mu_r$. In words, it is most profitable to promote the employee who has the strongest sensitivity to norms.

Proposition 3 *Hierarchical firms optimally promote employees with the strongest sensitivity to social norms.*

The intuition behind this result can be explained as follows. If the supervisor possesses stronger norm sensitivity, norms are eroded to a lower extent. Less norm erosion is always valuable to the principal. On the one hand, for any given effort choice of the principal, the agent will face a higher norm in equilibrium. Since the agent's norm is always inefficiently low (Corollary 1), this is profit-increasing. On the other hand, the principal can reduce her effort such that the agent's norm remains the same. This is also profit-increasing, since the principal's effort and the supervisor's norm are always inefficiently high (Proposition 2 and Corollary 1).

5.2 Managerial spans of control

The basic model assumes that the principal hires only one supervisor and one agent. More realistic is that the firm hires a number of S supervisors and A agents. Moreover, firms tend to maintain some relationship between the number of hired supervisors and agents. In the literature, the number of agents heading under one supervisor is referred to as the supervisor's span of control. Analogously, the number of supervisors heading under the principal is referred to as the principal's span of control. In this extension, I study how managerial spans of control are affected by the principal's desire to create efficient norms for her employees.

In case the firm hires S supervisors and A agents, firm profits are given by:

$$\Pi = \underbrace{pe_0 - \frac{1}{2}\theta e_0^2}_{\text{principal } (l=0)} + S \underbrace{(pe_1 - w_1) - \frac{1}{2}k_1 S^2}_{\text{supervisors } (l=1)} + A \underbrace{(pe_2 - w_2) - \frac{1}{2}k_2 A^2}_{\text{agents } (l=2)}, \quad (15)$$

where w_1 and w_2 must satisfy the employees' participation constraint (4). The profit function above is different from the one used in the basic model in two respects. First, to obtain an interior solution, I assume that the marginal profitability from hiring a supervisor or agent is decreasing. This is reflected by the terms $-\frac{1}{2}k_1 S^2$ and $-\frac{1}{2}k_2 A^2$. Second, the parameters k_1 and k_2 together describe the supervision technology of the firm. As will be shown below, for low values of k_1 and high values of k_2 , the principal needs relatively many supervisors to supervise the agents, whereas for high values of k_1 and low values of k_2 the reverse holds.

Before I derive the optimal managerial spans of control, it is instructive to consider the first-best benchmark where effort is contractible. Using the results from Lemma 1, it is straightforward to derive the following Lemma.

Lemma 2 *In the first-best case where effort is contractible, the principal hires $S^{FB} = \frac{1}{2} \frac{p^2}{\theta k_1}$ supervisors and $A^{FB} = \frac{1}{2} \frac{p^2}{\theta k_2}$ agents. The first-best supervisor span of control is equal to $\left(\frac{A}{S}\right)^{FB} = \frac{k_1}{k_2}$.*

In the first-best case, there is no moral hazard problem in effort provision. Since supervisors and agents have the same marginal productivity of effort, it follows that differences in the marginal profitability of hiring a supervisor or agents can only arise because of differences in the technology parameters k_1 and k_2 . As a result, in the first-best case, the optimal supervisor span of control is uniquely determined by k_1 and k_2 .

In the remainder of this subsection, I study the optimal managerial spans of control in case effort is not contractible. Given the definition of effort norms (3),

employees' effort choice (7), and the employees' salary (8), firm profits can be written as:

$$\Pi = \underbrace{pe_0 - \frac{1}{2}\theta e_0^2}_{\text{principal } (l=0)} + S \underbrace{\left(p\mu e_0 - \frac{1}{2}\theta\mu e_0^2\right) - \frac{1}{2}k_1S^2}_{\text{supervisors } (l=1)} + A \underbrace{\left(p\mu^2 e_0 - \frac{1}{2}\theta\mu^3 e_0^2\right) - \frac{1}{2}k_2A^2}_{\text{agents } (l=2)}. \quad (16)$$

The principal will exert the effort level that maximizes (16), which is given by:

$$e_0^* = \frac{p}{\theta} \frac{1 + S\mu + A\mu^2}{1 + S\mu + A\mu^3}. \quad (17)$$

It holds that $e_0^* > e_0^{FB}$ for any $A > 0$, as in Proposition 2. It can easily be checked that the principal's effort decreases in S , but increases in A .⁴ To understand these results, remember that the norm for supervisors is always above its first-best level, whereas the norm for agents is always below its first-best level (Corollary 1). Further, if the number of supervisors increases, their organizational layer makes up a larger share of firm profits. Therefore it becomes more important for the principal to have supervisors face an efficient norm. It follows that the principal optimally reduces her effort if S increases. The reverse intuition holds for the number of agents hired. The principal thus adjusts her effort in the direction that improves the effort norm for the kind of employee that becomes more abundant.

Proposition 4 *The principal's effort decreases in the number of supervisors hired, $\frac{\partial e_0^*}{\partial S} < 0$, and increases in the number of agents hired, $\frac{\partial e_0^*}{\partial A} > 0$.*

The next step is to determine the optimal values of S and A . Given the principal's effort choice (17), firm profits can be written as:

$$\Pi = \frac{1}{2} \frac{p^2}{\theta} \frac{(1 + S\mu + A\mu^2)^2}{1 + S\mu + A\mu^3} - \frac{1}{2}k_1S^2 - \frac{1}{2}k_2A^2. \quad (18)$$

⁴To see this, note that:

$$\begin{aligned} \frac{\partial e_0^*}{\partial S} &= -\frac{p}{\theta} \frac{\mu^3 A (1 - \mu)}{(1 + S\mu + A\mu^3)^2} < 0, \text{ and} \\ \frac{\partial e_0^*}{\partial A} &= \frac{p}{\theta} \frac{\mu^2 (1 - \mu) (1 + S\mu)}{(1 + S\mu + A\mu^3)^2} > 0. \end{aligned}$$

The first-order conditions to S and A are given by:

$$\frac{1}{2} \frac{p^2}{\theta} \mu \frac{1 + A\mu^4 (2S - A + 2A\mu) + 2A\mu^3 + S^2\mu^2 + 2S\mu}{(1 + S\mu + A\mu^3)^2} = k_1 S, \text{ and} \quad (19)$$

$$\frac{1}{2} \frac{p^2}{\theta} \mu^2 \frac{(2 - \mu) (1 + 2S\mu + S^2\mu^2) + A\mu^2 (2 + 2S\mu + A\mu^3)}{(1 + S\mu + A\mu^3)^2} = k_2 A, \quad (20)$$

respectively. Although the first-order conditions are difficult to solve algebraically, they are instructive in two ways.

First, given the first-best solutions S^{FB} and A^{FB} derived in Lemma 2, both (19) and (20) are negative.⁵ Hence, the first-best solutions cannot be optimal in case effort is not contractible. Moreover, since k_1 and k_2 are assumed to be sufficiently large such that the second-order conditions are negative, in the optimum it must hold that $S^* < S^{FB}$ and $A^* < A^{FB}$.⁶ The intuition behind this result is straightforward. In case effort is not contractible, a moral hazard problem in effort provision arises. This problem is partly mitigated because employees are sensitive to social norms. However, as shown in Corollary 1, equilibrium norms for supervisors are always above first-best, whereas equilibrium norms for agents are always below first-best. As a result, it is less attractive to hire supervisors and agents. The implication is that the principal's span of control is distorted downwards relative to the first-best case.

Second, the marginal benefit of hiring an agent (LHS of (20)) may be higher than the marginal benefit of hiring a supervisor (LHS of (19)). Rewriting this inequality

⁵One can verify that, given the first-best solutions from Lemma 2, (19) and (20) reduce to:

$$\begin{aligned} \frac{\partial \Pi}{\partial S} &= -\frac{\frac{1}{2} \frac{p^2}{\theta} (1 - \mu)}{(k_1 p^2 \mu^3 + k_2 p^2 \mu + 2\theta k_1 k_2)^2} \left(\frac{p^4 \mu^5 k_1^2 + 2p^4 \mu^4 k_1 k_2 + p^4 \mu^2 k_2^2 + 4p^2 \theta \mu^3 k_1^2 k_2 + 4p^2 \theta \mu k_1 k_2^2 + 4\theta^2 k_1^2 k_2^2}{(k_1 p^2 \mu^3 + k_2 p^2 \mu + 2\theta k_1 k_2)^2} \right) < 0, \text{ and} \\ \frac{\partial \Pi}{\partial A} &= -\frac{\frac{1}{2} \frac{p^2}{\theta} (1 - \mu)}{(k_1 p^2 \mu^3 + k_2 p^2 \mu + 2\theta k_1 k_2)^2} \left[\frac{p^2 \mu^3 k_1 (2\mu p^2 k_2 + \mu^3 p^2 k_1 + 4\theta k_1 k_2) + 4\theta \mu (p^2 k_1 k_2^2 + \theta k_1^2 k_2^2) (1 + \mu - \mu^2) + p^4 \mu^2 k_2^2 (1 - \mu^2 + \mu^3)}{(k_1 p^2 \mu^3 + k_2 p^2 \mu + 2\theta k_1 k_2)^2} \right] < 0, \end{aligned}$$

where the signs follows from $0 < \mu < 1$, $p > 0$, $\theta > 0$, $k_1 > 0$, and $k_2 > 0$.

⁶The second-order conditions are given by:

$$\begin{aligned} \frac{\partial^2 \Pi}{\partial S^2} &= \frac{1}{2} \frac{p^2}{\theta} 2\mu^6 (1 - \mu)^2 \frac{A^2}{(A\mu^3 + S\mu + 1)^3} - k_1, \text{ and} \\ \frac{\partial^2 \Pi}{\partial A^2} &= \frac{1}{2} \frac{p^2}{\theta} 2\mu^4 (1 - \mu)^2 \frac{(S\mu + 1)^2}{(A\mu^3 + S\mu + 1)^3} - k_2, \end{aligned}$$

which are always negative if k_1 and k_2 are sufficiently large.

yields:

$$A > \frac{1 + S\mu}{\mu^2}. \quad (21)$$

Hence, given that agents are relatively abundant, the marginal benefit of hiring an agent exceeds the marginal benefit from hiring a supervisor. The intuition behind this result stems directly from Proposition 4. Recall that this Proposition essentially states that the principal sets her effort level such that the most abundant kind of employee faces the more efficient norm in equilibrium. Therefore, given that agents are relatively abundant, the principal ensures that the equilibrium norm for agents is closer to its first-best level than the equilibrium norm for supervisors.⁷ It follows that hiring agents yields a higher marginal benefit. Of course, the reverse may also hold: given that supervisors are relatively abundant, their marginal benefit is higher.

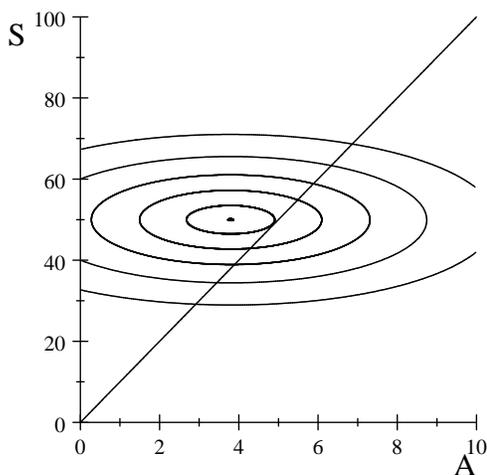


Figure 1. Firm profits as a function of S and A . $\frac{v^2}{\theta} = 20$, $\mu = 0.5$, $k_1 = 0.1$, and $k_2 = 1$.

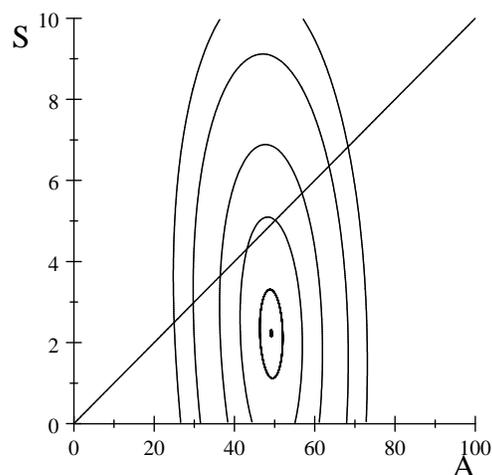


Figure 2. Firm profits as a function of S and A . $\frac{v^2}{\theta} = 20$, $\mu = 0.5$, $k_1 = 1$, and $k_2 = 0.1$.

Finally, each of the mechanisms described above may be at work in equilibrium, depending on the values of k_1 and k_2 . Figures 1 and 2 provide contour plots of the profit function (18) for a certain parameterization of the model. The inner contours represent a higher profit level than the outer contours, implying that the

⁷In fact, the norm for agents is closer to the first-best level than the norm for supervisors when it holds that $|n_2^* - n_2^{FB}| < |n_1^* - n_1^{FB}|$. Taking into account that $n_1^* = e_0^*$, $n_2^* = \mu e_0^*$, and $n_1^{FB} = n_2^{FB} = \frac{v}{\theta}$, one can show that this inequality is identical to condition (21).

maximum of the profit function is represented by the dots in the inner contours. In both figures, the 45-degrees line depicts the first-best supervisor span of control, $(\frac{A}{S})^{FB} = \frac{k_1}{k_2}$. Figure 1 depicts a case where the first-best supervisor span of control is low ($\frac{k_1}{k_2} = 0.1$). Next, as can be seen in the plot, the maximum of firm profits (18) lies above the 45-degrees line, meaning that the principal distorts the supervisor's span of control downwards relative to the first-best case. The reason is that the supervision technology is such that it is attractive to hire relatively many supervisors. However, as described above, given that supervisors are relatively abundant, the principal optimally makes sure that supervisors' effort norms are closer to the first-best level than agents' effort norms. As a result, supervisors become even more attractive to hire, implying that the principal optimally distorts the supervisor's span of control downwards. Figure 2 depicts a case where the first-best supervisor span of control is high ($\frac{k_1}{k_2} = 10$). By the same logic as above, the principal then optimally distorts the supervisor's span of control upwards. The final Proposition summarizes the results of this subsection.

Proposition 5 *Relative to the first-best case, hierarchical firms optimally distort the principal's span of control downwards. Relative to the first-best case, the supervisor's span of control is optimally distorted downwards if $\frac{k_1}{k_2}$ is low, and optimally distorted upwards if $\frac{k_1}{k_2}$ is high.*

6 Concluding remarks

This paper sets out to study the problem of creating effort norms within hierarchical firms. The analysis presented above rests on one key premise, namely that the actual example of superiors determines effort norms for subordinates. Norms are therefore shaped by two factors: the ultimate example set by those at the top of the hierarchy, like the CEO, and the process of transmitting this example to lower levels of the firm. Importantly, firms cannot control the process of norm transmission. After all, each hierarchical layer sets her own example for the next one. Consequently, norm transmission is prone to a moral hazard problem: no employee fully conforms with the norm he faces himself, as conforming requires him to exert costly effort. The implication is that, in equilibrium, norms erode as one moves down the hierarchy. To counteract the norm erosion effect, it has been shown that top managers optimally set a higher example to begin with. This is well in line with the fact that top managers work exceptionally long and hard. In addition to this, norm erosion gives rise to two comparative static results on the effort of top managers. Their effort decreases

in the number of middle-managers hired, and increases in the number of lower-level employees hired.

The analysis also yields three implications for organizational design. First, the inefficiencies due to norm erosion imply that firms optimally keep the extent of hierarchy to a minimum. Second, to reduce norm erosion, hierarchical firms optimally promote employees with the strongest sensitivity to social norms. Last, hierarchical firms optimally distort supervisor spans of control. The mechanism driving this result is that the relative profitability of hiring middle-managers and lower-level employees depends on whose equilibrium norm is more efficient.

Appendix: Allowing for Performance Pay

In this Appendix, I show that the results of the paper are not sensitive to the introduction of performance pay, as long as using performance pay generates some kind of agency costs.⁸ Suppose that employee effort yields a verifiable signal $y_l = pe_l + \varepsilon$, where ε reflects noise in the signal. The expected value of ε equals $E[\varepsilon] = 0$, and ε has a variance equal to $\sigma_\varepsilon^2 > 0$. Denote by s_l the base salary and by b_l the incentive intensity of the employee's contract. I assume that using performance pay is costly because employees are risk-averse. To model this assumption, employee utility is given by:

$$U_l = -\exp^{-r[s_l + b_l(pe_l + \varepsilon) - \frac{1}{2}\theta e_l^2 - \frac{1}{2}\gamma(n_l - e_l)^2]}, \quad (22)$$

where $r > 0$ denotes the intensity of risk-aversion. The certainty equivalent of (22) is given by:

$$E[U_l] = s_l + b_l p e_l - \frac{1}{2}\theta e_l^2 - \frac{1}{2}\gamma(n_l - e_l)^2 - \frac{1}{2}r\sigma_\varepsilon^2 b_l^2. \quad (23)$$

This is the same utility function as in (1), but with two terms added. The term $b_l p e_l$ describes the incentive portion of the employee's compensation. The term $\frac{1}{2}r\sigma_\varepsilon^2 b_l^2$ describes the risk-related disutility which performance pay imposes upon the employee.

The model can now be solved, following the same three steps as in subsection 4.2. First, given the contract and effort norm, employees choose the effort level that maximizes their utility. This effort level can be shown to be equal to:

$$e_l^* = \lambda b_l + \mu n_l, \quad (24)$$

⁸As will be shown below, if there are no agency costs of using incentives, the principal can achieve the first-best outcome derived in section 4.1.

where $\lambda \equiv \frac{p}{\theta+\gamma}$ and $\mu \equiv \frac{\gamma}{\theta+\gamma}$. Equation (24) shows that employee effort not only increases in the norm, but also in the incentive intensity of his contract.

Second, given employee effort (24), the lowest possible base salary that ensures participation can be written as:

$$s_i^* = \frac{1}{2}\theta\mu n_i^2 + \frac{1}{2}p\lambda b_i^2 + \frac{1}{2}r\sigma_\varepsilon^2 b_i^2 - b_i p e_i^*. \quad (25)$$

Last, given employee effort (24) and the base salary (25), the principal chooses her effort and the incentive intensities such that firm profits are maximized. This problem can be written as:

$$\begin{aligned} \max_{e_0, b_1, b_2} \Pi &= \underbrace{p e_0 - \frac{1}{2}\theta e_0^2}_{\text{principal } (l=0)} + \underbrace{p(\lambda b_1 + \mu e_0) - \frac{1}{2}\theta\mu e_0^2 - \frac{1}{2}p\lambda b_1^2 - \frac{1}{2}r\sigma_\varepsilon^2 b_1^2}_{\text{supervisor } (l=1)} + \quad (26) \\ &\underbrace{p[\lambda b_2 + \mu(\lambda b_1 + \mu e_0)] - \frac{1}{2}\theta\mu(\lambda b_1 + \mu e_0)^2 - \frac{1}{2}p\lambda b_2^2 - \frac{1}{2}r\sigma_\varepsilon^2 b_2^2}_{\text{agent } (l=2)}. \end{aligned}$$

The solutions to the maximization problem are given by:

$$e_0^* = \frac{p(1 + \mu + \mu^2 - \theta\mu^2\lambda^2 \frac{1+\mu}{r\sigma_\varepsilon^2 + p\lambda + \theta\lambda^2\mu})}{\theta(1 + \mu + \mu^3 - \theta\mu^2\lambda^2 \frac{\mu^2}{r\sigma_\varepsilon^2 + p\lambda + \theta\lambda^2\mu})}, \quad (27)$$

$$b_1^* = \frac{p\lambda + 2p\lambda\mu}{\theta\lambda^2\mu(1 + \mu) + (p\lambda + r\sigma_\varepsilon^2)(1 + \mu + \mu^3)}, \text{ and} \quad (28)$$

$$b_2^* = \frac{p\lambda}{p\lambda + r\sigma_\varepsilon^2}. \quad (29)$$

It is easy to check that, if using incentives is costly ($r > 0$ and $\sigma_\varepsilon^2 > 0$), it holds that $1 > b_1^* > b_2^*$. Hence, both employees receive an incentive share below 1, implying that employees will not fully internalize the benefits of their effort to the principal. As a result, norms will be eroded, as in Proposition 1. Since all the other results are driven by the existence of norm erosion, it holds that the results are insensitive to the introduction of performance pay. However, note that in case using incentives is costless ($r = 0$ or $\sigma_\varepsilon^2 = 0$), norm erosion can be avoided and all players exert the first-best effort level, $e_i^* = e_i^{FB} = \frac{p}{\theta}$. The reason is that the principal optimally provides full-powered incentives to the employees, that is, $b_1^* = b_2^* = 1$.

Finally, if using incentives is costly, it is found that the supervisor receives stronger incentives than the agent. The intuition behind this result is that providing incentives to the supervisor has the additional benefit of raising the norm for

the agent. Interestingly, this result is in line with Aggarwal and Samwick (2003), who document that CEOs receive stronger incentives than executives with oversight authority, who in turn receive stronger incentives than managers with divisional responsibility.

References

- [1] Aggarwal, R.K. and A.A. Samwick (2003), "Performance Incentives Within Firms: The Effect of Managerial Responsibility." *Journal of Finance*, 58(4): 1613-1649.
- [2] Akerlof, G.A. (1980), "A Theory of Social Custom, of Which Unemployment May be One Consequence." *Quarterly Journal of Economics*, 94(4): 749-775.
- [3] Akerlof, G.A. and R.E. Kranton (2005), "Identity and the Economics of Organizations." *Journal of Economic Perspectives*, 19(1): 9-32.
- [4] Bernheim, D.B. (1994), "A Theory of Conformity." *Journal of Political Economy*, 102(5): 841-877.
- [5] Bradler, C., R. Dur, S. Neckermann, and A. Non (2013), "Employee Recognition and Performance: A Field Experiment." Tinbergen Institute Discussion Paper, TI 2013-038/VII.
- [6] Bradley, S., C. Green, and G. Leeves (2007), "Worker Absence and Shirking: Evidence from Matched Teacher-School Data." *Labour Economics*, 14(3): 319-334.
- [7] Carlin, B.I. and S. Gervais (2009), "Work Ethic, Employment Contracts, and Firm Value." *Journal of Finance*, 64(2): 758-821.
- [8] Fischer, P. and S. Huddart (2008), "Optimal Contracting with Endogenous Social Norms." *American Economic Review*, 98(4): 1459-1475.
- [9] Heinle, M.S., C. Hofmann, and A.H. Kunz (2012), "Identity, Incentives, and the Value of Information." *Accounting Review*, 87(4): 1309-1334.
- [10] Hermalin, B.E. (1998), "An Economic Theory of Leadership." *American Economic Review*, 88(5): 1188-1206.

- [11] Huck, S., D. Kübler, and J. Weibull (2012), "Social Norms and Economic Incentives in Firms." *Journal of Economic Behavior and Organization*, 83(2): 173-185.
- [12] Ichino, A. and G. Maggi (2000), "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm." *Quarterly Journal of Economics*, 115(3): 1057-1090.
- [13] Kandel, E. and E.P. Lazear (1992), "Peer Pressure and Partnerships." *Journal of Political Economy*, 100(4): 801-817.
- [14] Mas, A. and E. Moretti (2009), "Peers at Work." *American Economic Review*, 99(1): 112-145.
- [15] Murphy, P.E. and G. Enderle (1995), "Managerial Ethical Leadership: Examples do Matter." *Business Ethics Quarterly*, 5(1): 117-128.
- [16] Posner, B.Z. and W.H. Schmidt (1992), "Values and the American Manager: An Update Updated." *California Management Review*, 34(3): 80-94.
- [17] Sliwka, D. (2007), "Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes." *American Economic Review*, 97(3): 999-1012.
- [18] Stevens, D.E. and A. Thevaranjan (2010), "A Moral Solution to the Moral Hazard Problem." *Accounting, Organizations and Society*, 35(1): 125-139.
- [19] Treviño, L.K., G.R. Weaver, D.G. Gibson, and B.L. Toffler (1999), "Managing Ethics and Legal Compliance: What Works and What Hurts." *California Management Review*, 41(2): 131-151.
- [20] Treviño, L.K. and G.R. Weaver (2003), "Ethics and the Broader Organizational Context: Ethical Climate and Ethical Culture." In: Treviño, L.K. and G.R. Weaver (Eds.), *Managing Ethics in Business Organizations*, Stanford, CA: Stanford University Press.