Erasmus School of Economics 2013-2014

Master Thesis

Social Relations and Monitoring in the Workplace: Implications in a Stylized Work Environment*

Jan Lampe
385678

July 17, 2014

Abstract

Social relations influence managerial decisions. This thesis develops a principal-supervisor-agent model of moral hazard. Agent and supervisor have social relations with each other. The principal does not adequately account for the real incentive distortions that the social relation can entail. Agents can be of high or low type. The supervisor can misrepresent the type of a worker to the principal in order to retain a worker he likes or lay off a worker he dislikes. When a high-type worker is fired or a low-type worker is retained due to the evaluation of the supervisor, this is has negative consequences for the principal. In extensions of the basic model, the principal can randomly learn the truthfulness of the supervisor's evaluation and a third type of agent, an agent who can steal from the firm but has to collude with the supervisor in order to steal successfully, is introduced. The principal's random learning of type makes it more likely that the supervisor's evaluation is truthful. The employee's collusion constraint is easiest to satisfy with positive social relations and is harder to satisfy with neutral or negative social relations. This suggests that, in certain situations, it may be beneficial to the principal when his employees dislike each other.

JEL-Codes: D03, D64, J53, M5
Keywords: social relations, co-worker relations, fraud, altruism, spite

*Comments, suggestions and support from Josse Delfgaauw, Suzanne Bijkerk, Joeri Sol, Xiaojian Zhao, Jan Stoop, Carolin Schaar, Mieke du Plessis, Sarah Mohr and Timo Krall are gratefully acknowledged. All errors are the author's.

Experience without theory is blind,
but theory without experience is
mere intellectual play.

Immanuel Kant

Disclaimer

In this thesis, female pronouns are used to refer to individuals in the model development, its extensions and the analysis thereof. Male pronouns are used elsewhere. No discrimination of any sort is meant by this distinction.

Contents

1 Introduction

Humans are social animals that care not only about themselves. They care also about other people's well-being. History books as well as contemporary and classic literature give innumerable examples of the importance of social interactions. Of the 36 main dramatic situations in literature as established by Polti (1921), at least thirteen directly deal with social or familial relations as a main theme.[1] More than a third of dramatic structures humans have employed to tell stories are based on social interactions. It is then clearly myopic to think that social interactions do not play a significant role when humans go about their daily lives rather than tell stories. The economics literature on socially influenced interactions and behavior has been growing since Becker (1974), yet much remains to be understood in detail. One environment where social relations are likely to play an important role is the workplace. With the advent of social networks, social ties have become easier to maintain and more salient. As social relations between employees are easier to nurture even outside of normal working hours —where they cannot be monitored— they likely will grow in importance. This thesis therefore focuses on the consequences of social relations in a stylized workplace environment.

At their core, possible effects of vertical, i.e. inter-hierarchy, social relations[2] of workers on the performance of a firm are dichotomous. On the one hand, it is compelling to argue that good social relations between workers have positive effects on worker motivation benefiting the firm, especially if there are synergies to be gained from worker cooperation. On the other hand, good inter-worker relations may facilitate collusion between them or lead to time spent on social interaction rather than work, leading to lower productivity. Moral hazard that arises from social relations can also lead to decisions that are not optimal from the firm's perspective.

In practice, team building exercises and excursions are commonly employed even in work environments where no large synergies exist (Salas et al., 1999). Such an observation is surprising, given potentially detrimental effects of social relations. This thesis focuses on potentially negative effects of social relations. It investigates the moral hazard aspect in an environment lacking synergies that benefit the principal to be gained from teamwork. A supervisor is tasked with telling the principal the agent's

[1] These dramatic situations, i.e. main plots, include, but may not be limited to (in the order established by Polti, 1921): Crime pursued by vengeance, vengeance taken for kindred upon kindred, revolt, enmity of kinsmen, rivalry of kinsmen, self-sacrifice for kindred, necessity of sacrificing loved ones, rivalry of superior and inferior, crimes of love, discovery of the dishonor of a loved one, an enemy loved, recovery of a lost one and loss of loved ones.

[2] The terms "social relation" and "altruism" are used interchangably in this thesis. The formal consequences of having a social relation with an individual that is either negative, neutral, or positive are the same as having a specific altruism parameter $\alpha \in (-1, 1)$ for that particular persons' utility in one's own utility function. Negative altruism (a negative value for α), relatedly, is used interchangably with "spite" or "enmity." In the following, "employees" is used to refer to all of a principal's employees where no distinction between them is warranted. Similarly, "monitor" and "supervisor" are used interchangably to refer to the same individual.

type. The principal prefers to have good rather than bad workers in his employ. Social relations influence the supervisor's decision about which type to convey to the principal. A moral hazard-style model is developed in order to study the effects of such social relations between a supervisor and an agent on a principal that is unaware of the incentive effects of social relations.[3] This is done in a stylized work environment where agents can make mistakes that are costly to the principal.

The contribution of this thesis to the literature is its description of the consequences of social relations between a monitor and an agent in a previously unstudied variant of a principal-supervisor-agent framework. This description allows it to derive testable predictions and implications that are meant to inform practice. It does so by making some simple assumptions about the work environment. First, it is assumed that a firm with a principal and supervisor already exists. Second, it is assumed that the principal *has to delegate* monitoring. Third, it is assumed that the agent and the principal cannot communicate directly. Fourth, it is assumed that the principal does not anticipate the emergence social relation between the supervisor and the agent and the incentive distortion it entails.

These assumptions are without doubt restrictive but not unwarranted. They are also not unrealistic. Firms often employ middlemen, whether they be foremen, supervisors or department heads, for instance because direct supervision of workers is unfeasible. It may be unfeasible for a variety of reasons: opportunity costs and the principal's lack of in-depth knowledge about the tasks performed by the agent or physical distance between them. Such restrictions make direct communication between the principal and the agent difficult and ensure that neither good nor bad agents can convince the principal they are good workers independently of their supervisor's evaluation. The assumption that the principal does not adequately anticipate a social relation between his employees is made to allow the moral hazard problem to arise. The principal's trust in the monitor's evaluation of the agent is implicit when the monitor can expect a bonus for making decisions in the interest of the principal.

These assumptions restrict the action set of the principal, which would certainly not be optimal from the principal's own perspective, if he was aware of the incentive distortion caused through the social relation. The principal thinks that it is optimal for him to trust the monitor. When incentives are distorted through the social relation, this is no longer true, although the principal does not know this. Hiring decisions in situations such as the one described undoubtedly occur often in practice. The stylized work environment consists of a principal, a supervisor and an agent. The agent and the supervisor engage in a two-period production process. The agents can be of either low

[3]This type of principal is constrained in his "vision" of the firm not unlike carriage horses are still today kept from seeing irrelevant or distracting parts of the environment through the use of blinkers. He is therefore referred to as being "socially blinkered." A definition follows in the model section.

or high type. It is equiprobable that an agent is of either type. Low type agents have a higher propensity to make mistakes in a period that are costly to the firm. Types are ex ante private information. It is possible for the supervisor but not the principal to observe the agent's type during period one. Given an evaluation of the agent by the supervisor, the principal can fire an agent he thinks to be of low type in period one. This gives the principal a chance to hire another agent who may then be of high type to produce in period two. Period two is purely production and entails no choices by any individual.

The agent and his supervisor have exogenously given social relations with each other. They are not aware of this relation before working together. The principal is completely unaware of the consequences of their social relation and does not take them into account when designing the incentive contracts for his employees. The social relation may give the supervisor an incentive to lie to the principal about the agent's true type. If the supervisor lies, bad agents may be retained while good agents may be fired. Therein lies the basic problem. When social relations are pronounced enough, they can distort incentives and lead to choices that are not desired by the principal. When the supervisor's goals of working with a friend or not working with a foe stand in contrast to the principal's goal of employing only good workers and the principal is unaware of this misalignment, he finances his employees' social relations at no benefit of his own.

Most closely related to the topic investigated here is the work of Lee & Persson (2011). Their model similarly is a principal-supervisor-agent model in which the agent and the supervisor are friends. In their model, the principal hires a supervisor to monitor an agent, while in this model, the agent is hired into an existing firm framework consisting of principal and supervisor. Lee & Persson then investigate whether a corporate governance strategy, built on either the principal's authority or employees' loyalty to the principal, is preferable by the principal. Their main finding is that the presence of social relations weakens the authority form of governance while strengthening the loyalty-based one. The Lee & Persson paper approaches the implications of positive social relations from an optimal governance perspective. They ask which governance form is better, given the existence of social ties between employees. This thesis differs from their approach in an important aspect: The research question of this thesis is different. The research question of this thesis is: What are the consequences of not adequately accounting for social relations between employees? This thesis is purely descriptive of a situation where governance, i.e. the principal's handling of the firm, is suboptimal to start with. Governance is suboptimal only in that the principal does not take into account social relations. Given that it is suboptimal, the implications of positive, neutral and negative social relations are explored in a stylized framework.

Most existing literature in the field of organizational economics focuses on optimal

behavior and decisions. If people in real-world workplaces behaved optimally and foresaw and planned for any contingency, problems that regularly occur in reality would not occur. However, in a disconnect between practice and theory, such problems do occur. These problems then need to be studied, analyzed and explained before they can be solved. This thesis is descriptive and per se offers no solutions to problems that arise from social relations in firms. It is rather meant as an early attempt at a theoretical foundation for future empirical work and extensions of the model. This foundation and the extensions could then be utilized to find solutions for this complex problem. The thesis itself establishes a model in which ignoring social relations in workplaces, where they can influence decision making against the firm's interests, can lead to losses. These losses can be financial, losses of efficiency or opportunity costs. This simple observation from the model, and the testable hypotheses derived therefrom, are made in absence of useable data.[4]

The main result obtained in this thesis is as follows: Not taking social relations between workers in production planning into account leads to outcomes that are sub-optimal from the perspective of the firm. When social relations are pronounced enough in either direction, they can indeed lead to a selection of employees that is not in the principal's best interests for the second period of production. A supervisor who is friends with a low-type agent can lie in his evaluation to retain this agent and a supervisor who dislikes a high-type agent can lie to get that agent dismissed. Both outcomes stem from the monitor's expected utility maximization for himself with regard to the agent and not the firm. In the basic framework, the principal furthermore never finds out that he has been lied to. This is the case because both agent types can make mistakes and two periods are not enough to display a clear trend that would cause him to become suspicious. He inadvertently finances his employees' friendships and enmities at his own cost. Bad relations with bad workers and good relations with good workers have the same effect on the monitor's evaluation decision as none or neutral social relations; the monitor evaluates truthfully. Good relations with bad workers or bad relations with good workers can distort the monitor's incentives and lead to untruthful evaluations. The effect of social relations on the principal is then either non-existent or negative, but never positive.

This thesis also contains three short extensions of the basic model. First, it is

[4]It would have been desirable to support the conclusions of this thesis with empirical results. Ideally, survey evidence about effects of social relations in real workplaces could have been collected as well as experimental evidence. Several large firms (17 in total in four countries) were contacted and responded shortly. When asked about participation in a survey or allowing participation of their employees in an experiment, all either declined or demanded their conditions be met. Following these conditions would have been unethical or distorted data to the point of uselessness. This lead to the discarding of a more empirical focus before more firms were contacted. Even without these ethical concerns, such identifiers would have put into question the truthfulness of respondents. In lieu of time and financial means, other potential experiments were discarded too. Possible experimental designs are listed in Appendix A.3, while the difficulties of obtaining data are elaborated further in section 6.

allowed that the principal can find out that the supervisor lied in the evaluation of the agent. The monitor knows about this chance, while the agent does not Second, the basic model is extended to a third agent type. This type can steal from the firm and can engage in bargaining with the supervisor, who is meant to curb employee theft. These two extensions are then combined to give a third, wherein the agent can steal while the principal can detect it. The results of the extensions of the basic model are as follows: First, when the principal can randomly learn the agent's type, the monitor is less willing to lie for the agent. Second, a third type of worker is introduced. When this agent type is able to steal from the principal, he can offer a transfer to the monitor that can induce the monitor to lie. This transfer is lowest for positive, higher for neutral and highest for negative social relations. Third, when the principal can randomly learn type, and the monitor but not the agent is aware of this, and the agent can steal, the agent offers a transfer to the monitor that does not take into account the added risk to the monitor. This transfer may then not be sufficient to convince the monitor to collude with the agent. Furthermore, even when it is sufficient for positive relations, it is likely insufficient to induce a monitor who dislikes the agent to lie. In such a situation, contrary to the results of the basic model, negative social relations do have a positive effect on the principal's outcome.

The remainder of this thesis is arranged as follows. Section 2 gives an overview of the related literature. Section 3 describes the model set-up. Section 4 contains the analysis, while section 5 extends the model into an auditing and whistleblowing framework which allows the principal to be more involved. Section 6 describes and treats limitations of the research presented here. It also gives an overview of avenues for future research. Section 7 concludes the thesis. Some mathematical analysis and two possible experimental tests of the implications of this research are presented in the Appendix A.

2 Related Literature

This section gives an overview over literature relevant to social relations in the workplace.[5] As humans are social animals that live in social groups, social interactions between individuals clearly determine or influence a large variety of behaviors (cf. Cooper & Kagel, 2013). From an economics point of view, social relations in the workplace have been relatively under-studied,[6] especially given their role as a driver of individual

[5]Further literature is provided in section 5.2.1 in the context of existing (limited liability) companies.

[6]Some research implicitly does treat social relations. Relationship-specific investments can be understood as endogenous social relations: continued, repeated trading with some partners after a specific investment has been made may be beneficial for all involved under certain circumstances, and an investment of time or effort into the forging of a social relationship with an individual is an investment of high specificity. See Shleifer & Vishny (1989), Nunn (2007) or, more recently, Rasul &

behavior.

A reason for this may be that in the past, social relations were seen as unnecessary in explaining the behavior of economic agents. Much, if not most, theoretical work in previous decades has made little to no explicit assumptions about social dynamics. It still managed to provide, at least in parts, valuable information to academics, practitioners and policy-makers. However, as (Fehr & Fischbacher, 2002, p. C1) put it, "economists fail to understand core questions in economics if they insist on the self-interest hypothesis and rule out [...] social preferences."[7] With the introduction of psychological concepts into economics (cf. Rabin, 1998), behavior that had previously been considered as irrational, i.e. not (fully) explainable by the *homo oeconomicus* model, became a focus of the (new) field of behavioral economics.

The importance of the study of workplace social relations between colleagues stems from a dichotomy: on the one hand, friendships (strong animosities) between workers could lead to higher (lower) productivity, while on the other hand they could lead to (prevent) collusion against the principal. In recent years, economics research that *has* studied social relations in the workplace finds support for both a positive as well as negative impact of social relations on firms. Most of this literature is focused more on the design of optimal contracts and less so on the optimality of social relationships.

A comparatively early theoretical work on economic implications of social relationships in the workplace is Rotemberg (1994). Therein it is assumed that altruism as a behavior is driven by selfish considerations. It is therefore proposed that reciprocal altruism arises between colleagues if it results in a symbiotic relationship. In a hierarchy, horizontal relations (those between co-workers) are more strongly altruistic than vertical relations (those between superiors and subordinates). While Rotemberg (1994) investigates the roots and causes of social relations in the workplace, the goal of this thesis is to investigate their consequences. In what amounts to a model of a supervisor who is (positively or negatively) altruistic towards an agent, Prendergast & Topel (1996) show that favoritism leads to biased performance evaluations. This imposes additional costs on the organization, which may manifest themselves in suboptimal agent effort provision or the misidentification of skilled agents and consequently, in suboptimal promotion decisions. However, Dur & Tichem (2014) predict that altruism between management and workers can raise productivity because it increases worker motivation and makes promises of bonus payments credible.

Similarly, as in Sol (2010), when employees evaluate each other in team production, pronounced social relations reduce the effectiveness of such peer evaluation schemes. Team bonus effectiveness and good social relations are however positively related, as

Sonderegger (2010) for examples of the literature on specific investments.

[7] For an overview of the literature on social preferences (broadly defined as including, among others, altruism, spite, inequity aversion and intrinsic motivation) see Bowles & Polania-Reyes (2012) who provide a survey of fifty studies.

positive relations impose a cost on free-riding. Furthermore, bad social relations reduce team output. As (Sol, 2010, p. 14) puts it, "bad co-worker relations are never good for profits, while good co-worker relations can be bad." The model presented in this thesis does not allow for synergies in production that may arise from social relations. Therefore, bad social relations may not automatically be bad for profits. Nevertheless, Sol's statement is in line with the main result of this thesis, where good and bad relations can have negative outcomes when they lead to a misalignment of goals between workers and the firm. When goals are not misaligned through social relations, no problem arises for the firm.

When subjectivity plays a role in determining an individual's workplace outcomes, evaluators tend to take into account the feelings of the evaluated: Centrality and leniency bias are commonly observable and can curb worker reactions stemming from inequity aversion (Grund & Przemeck, 2012). Giebe & Gürtler (2012) report leniency bias as a consequence of supervisor altruism. Equivalently, Tichem (2013) finds that long-term relations between an agent and an altruistic supervisor can enforce truth-telling in the supervisor's evaluation of the agent. The supervisor values his job precisely due to the collaboration with the agent. Thiele (2013) shows that subjective performance evaluations —which leave ample room for evaluator discretion and may therefore be sensitive to ex ante existing social relations— can be optimal when employees have little to gain from collusion with the evaluator.

These theoretical results suggest that firms would sometimes be well-advised to avoid constellations where friends will be working with each other. However, there is empirical evidence that points to potential benefits of friendships in the workplace. This research also shows that social relations affect corporate decisions. While Bandiera et al. (2009) find a negative impact of favoritism on firm profits, Bandiera et al. (2010) suggest that careful considerations of friendship ties and individual skill levels could be leveraged to increase firm performance. Workers change their effort levels according to their friends' skill levels. When working with friends whose skill levels are higher (lower) than their own, workers have higher (lower) levels of productivity than in isolation. In total, workplace social relations between employees can be preferred or disliked by principals in a principal-supervisor-agent environment, depending on the circumstances. This is summed up by Lee & Persson (2011, p. 13): "On one hand, friendship undermines the supervisor's incentives to monitor the agent [..]. On the other hand, friendship makes the agent reluctant to place the supervisor in a bad light [if worker shirking is detected by the principal], and thereby more likely to voluntarily act in the principal's interest." Morrison & Nolan (2007) investigate the worker's point of view and find these difficulties to be common, with the addition of having to spend time on friendships.

This two-edgedness seems to apply to the effect of social relations in general and

is found in empirical research too, implying that sometimes bad relationships could be supported (avoided) by a rent-seeking principal. On the one hand, Agell (2004) finds that a majority of managers in small businesses rely relatively heavily on good relations with their workers as a motivational tool, and Berman et al. (2002) find a clear managerial bias towards good workplace relations. Yonker (2013) finds that local managers are biased in that they are less likely to lay off workers than their non-local counterparts, suggesting that social factors in the community influence corporate decisions. However, it should be kept in mind that not all managerial biases stem from social relations with employees. To name but one example, high search costs on the evaluator's part can also lead to biased evaluations (Bol, 2011). Also, few people would argue that animosity between agents cannot provide additional motivation to competitors when, for example, tournament incentives are used. Neither is conflict in the workplace necessarily always a negative factor: Disagreements in the workplace can preclude herd behavior (Bénabou, 2013) and may facilitate a process of creative destruction (Schumpeter, 1942), helping separate good ideas from bad ones. However, monitoring in itself can have negative effects on worker motivation that could offset the consequences of social relations (cf. Dickinson & Villeval, 2008; Falk & Kosfeld, 2006). Enzle & Anderson (1993) show that a monitor's intentions also influence worker motivation. Especially when the monitor holds a clear antagonistic position towards the surveilled, motivation plummets.

Depending on how broadly one defines "social relations," one may find a multitude of dimensions and causes for social relationships (cf. Gintis et al., 2003). Undoubtedly, the domain of social relations encompasses more than the (common) interpretation of the extent of positive or negative altruism. While not the underlying causes for a social relation but the relation's consequences are of interest here, potential determinants of other-regarding preferences still bear mention. Empirically, Jackson & Schneider (2011) show that existing social ties in the form of a shared country of origin curb moral hazard in taxi-leasing arrangements. Such reduction of moral hazard is achieved by social sanctions or the threat thereof. Similarly, Gómez et al. (2000) find that in-group membership leads to more positive peer evaluations.

Social relations can also serve as informal contract enforcement devices (cf. Fehr et al., 1997), where social relations implicitly complete an incomplete contract, leading to outcomes closer to first-best than the same contract in the absence of social relations. From an individual's career perspective, however, it is unsurprisingly detrimental to have an antagonistic superior (Moerbeek & Need, 2003). Similarly, Morrison (2008) finds empirically that the existence of negative social relationships at work has negative effects on job satisfaction, loyalty to the firm and group work.

The results of both the theoretical and empirical literature mirror the dichotomy of inter-worker relationships: When production positively depends on social relations,

good social relations are preferred by all parties. When there are no spillovers from social relations to productivity or there exists room for discretion —i.e. hidden action is possible or evaluations and promotion decisions are based on subjective criteria— good social relations can foster behavior that, if unchecked, harms the principal. Undoubtedly though, a lot of research on workplace relations still remains to be done.

The principal in the model presented here is unaware of the importance of social relations between his employees on his own (expected) utility. In recent years, there has been a growing literature on unawareness in principal-agent relationships. This literature takes as its starting point that at least one individual is unaware of the full scale of the action set of another (Zhao, 2008; Von Thadden & Zhao, 2012; Galanis, 2013). It then focuses on whether it is beneficial to somehow become aware or how to make such an individual aware of another's action set. This approach differs from the one in this thesis. The principal here can be aware of everything in his employees' action sets, but is unaware that the existence of social relations gives an incentive to undertake actions they would not otherwise undertake. The focus here is then on the consequences of the principal's misperceptions. The following section takes this specific version of unawareness as its starting point and develops the model.

3 The Model

This section develops the basic model. Before describing the monitor's decision problem in a work environment with a "socially blinkered" principal, such a principal's characteristics need be defined.

Definition 1. ***Socially Blinkered Principal:*** *A socially blinkered principal is unaware of the real incentive distortion a social relation between the agents can entail. He therefore does not take it into account in his expectations and choices.*

This definition differs slightly from Zhao (2008) and Von Thadden & Zhao (2012), where an "unaware" agent or principal does not take into consideration all of his own or another's action set because he is unaware of its full scale. Here, a socially blinkered principal is unaware of one of his employees' *characteristics* and the consequences this characteristic has for their choices. This differs from a problem of hidden characteristics, because the characteristic in question is not hidden, but has similar implications. The —inadequately accounted for— social characteristics makes it more probable that agents utilize certain parts of their action sets. The principal can be aware of problems that may arise from social relations, but believes wrongly that they do not apply to his situation. He is not unaware of his agents' action sets, but is unaware that agents may have an incentive to engage in some actions that are not in his own interest.[8] This can

[8]This is similar to being aware that agents can shirk, but wrongly believing that they will not

be understood as a specific case of the more general theory of bounded rationality (cf. Simon, 1972). The investigation of the exact reason why the principal is socially blinkered or how to create awareness is not of interest here. Yet there are many plausible explanations, including: The principal may be naive or cognitively limited, previous employees may have not had anything but neutral social relations with each other, the principal may have studied only classical economics and may therefore be unaware of other-regarding preference models, or, the principal may be aware of the possibility of social relations but wrongly expects them to not arise in this situation. The principal is not non-strategic, he is strategic given his (wrong) beliefs.

Consider now the following one shot, two-period production process. There is no discounting.[9]

A socially blinkered principal P owns an existing firm and employs a monitor M. In order to produce output $Q > 0$ in a given period $t = 1, 2$, P needs to hire an agent A in addition to M. Without M and A working together, production in t is zero and firm profits are negative given a wage w_M for M. The agents can be of low or high type, i.e θ_i where the subscript $i = \{H, L\}$ denotes high or low type, respectively.[10] Monitors do not have a type. Types are private information ex ante. A knows her type. P believes the likelihood that A is of either type is

$$\Pr(\theta_i = \theta_L) = \frac{1}{2}, \tag{1}$$

which is the same as the base rate of the underlying distribution of the eligible agent population.[11] (1) is common knowledge. Either agent type can make a monetarily costly mistake c in t, with probability $p_i > 0$. Let it be that $p_L > p_H$, to the effect that a mistake by a θ_H agent can be considered no more likely than some base rate, while θ_L agents have a comparatively high disposition to make mistakes. Assume that the value of c and the error propensity of either agent type are common knowledge. It can therefore be interpreted as a known weak spot in the production process that cannot cost-effectively be eradicated. Low-type agents are more likely to encounter (or less likely to avoid) this weak spot than high-type agents. It is not possible for P or M to know the value of θ_i ex ante to hiring A. The first period therefore is a trial period for A in order to allow employee selection to maximize the principal's expected utility

(because they are believed not to have an incentive to do so). The model presented here does not allow for shirking but focuses only on the consequences of social relations on the principal.

[9] Ignoring intertemporal discounting in the static context here allows the focus to be on the agents' social relations. Adding some discount factor would merely complicate the analysis without adding to the results in a meaningful way. For analyses of effects of social relations in dynamic contexts, see Tichem (2013) and Dur & Tichem (2014).

[10] A third type of agent, a criminal agent who can steal from the firm, is introduced in the extension in section 5.2.

[11] This distribution is assumed to be the distribution ex post to some imperfect screening process. In the total population, this distribution may be different. The 50:50 split between types applies only to qualified candidates for the job.

over the two periods of production. At the end of the first period the principal decides whether to retain or lay off A. An agent who works in the second period finishes the period without chance of being fired.

In her function as the monitor, M can perfectly observe the value of θ_i. She can also observe whether A makes a mistake that is costly to P while working. P also can observe that a mistake happens, but the monitor has an information advantage over the principal about θ_i: P can, for reasons of physical distance or lack of information about the tasks performed, not determine why a mistake occurred if it occurred. Therefore, P cannot distinguish between mere bad luck (for θ_H agents) or a predisposition towards mistakes (for θ_L agents). She therefore relies on M to communicate the value of θ_i. Assume for now that P trusts the message she receives about the agent's type to be truthful. Type cannot be hidden from M by the agent. Having observed the value of θ_i, M has some room for discretion due to her information advantage over P. During the evaluation at the end of the trial period, P's decisions rely on M's message. What message M does send depends on her incentives. She can lie or tell the truth.[12] Based on the message she receives from M, P updates her beliefs about A's type. This updated belief is denoted by

$$\Pr(\theta_i = \theta_H|\text{Message}) = \begin{cases} 0 & \text{if } M \text{ transmits that } \theta_i = \theta_L \\ 1 & \text{if } M \text{ transmits that } \theta_i = \theta_M \end{cases}. \tag{2}$$

Since P does not observe θ_i herself, she is utterly dependent on M's evaluation of A. As the likelihood of making a mistake p_i in period two depends on the agent type, and $p_L c > p_H c$, the principal expects to be better off replacing an agent A, which she believes to be of low type, by another agent. This holds regardless of whether a mistake has been made in the first period. So if $\Pr(\theta_i = \theta_H|\text{Message}) = 0$, P fires A and replaces her with another agent B, who is drawn from the same agent population as A, at the end of period one. B then works for the remainder of the production process. As production ends after period two, there is no benefit from monitoring and M works just to enable output Q. If fired, A receives her outside options for both periods. Outside options of any agent are normalized to zero. Assume that P cannot fire employees without reason because P, for example, may be legally barred from doing so. She needs the monitor's message in order to decide on firing. This prevents the principal from firing all employees at the end of a period in order to avoid paying them, minimizing wage costs but still reaping Q. This also means that P cannot fire M.[13]

P maximizes her per-period expected utility function

[12] The following tie-breaking assumption is made: If M is indifferent between lying and telling the truth, she tells the truth.

[13] This assumption is relaxed for the extensions in sections 5.1 and 5.3.

$$E(U_{Pt}) = Q - (w_k + w_M + E[b_M]) - \gamma_P E(c) \text{ for } i = A, B. \tag{3}$$

Her expected profit in t equals output minus wages and a bonus b_M paid to M, where w_k represents the agent's wage, either A or B. $\gamma_P > 1$ captures the principal's non-monetary costs of the mistake, i.e. administration time and opportunity costs.[14] The last term captures the expected monetary cost of an error, which is

$$E(c) = \begin{cases} \frac{(p_H + p_L)}{2} c & \text{at the beginning of } t = 1, \\ p_i c & \text{after } P \text{ has updated her beliefs at the end of } t = 1 \text{ and} \\ \frac{(p_H + p_L)}{2} c & \text{again at the beginning of } t = 2 \text{ if } B \text{ has been hired.} \end{cases} \tag{4}$$

By firing an agent A who is believed to be of low type, the principal cannot guarantee herself that B will be of high type. B comes from the same population as A, as in (1). However, by hiring B she has a 50% chance that the agent is of the higher type than the type she believes A to be.

The employees maximize their per-period expected utility function

$$E(U_{kt}) = E(u_{kt}) + \alpha E(u_{lt}), \text{where } k \neq l. \tag{5}$$

The employees' total expected utility $E(U_{kt})$ depends on their own expected work utility $E(u_{kt})$ plus a socially weighted work utility of the other agent. Note that the subscripts k, l identify the agent, while subscript k denotes the type of the agent in question. Employees A and M have exogenously given social relations with each other,[15] parameterized by $\alpha \in (-1, 1)$, where a value above zero represents altruism (friendship), a value below zero represents spite (enmity) and a zero value perfectly neutral relations between M and A.[16] This also means that despite the nature of the social relation, no agent cares as much about the utility of the other as she cares about

[14] If one wishes not to assume that a socially blinkered principal has administration costs, γ_P can instead be interpreted as a loss-aversion parameter.

[15] A similar dichotomy of the potentially positive or negative effects of social relations applies to relations between principals and agents. Better social relations between them may increase loyalty and lead to more effort on the agent's part, but they could just as well lead to less effort exerted because any threat of dismissal of the agent becomes less believable. The social relations between principals and their agents are however not the focus of this thesis. In the model presented here, principal and agent do not and cannot interact with each other, which is both cause and consequence of the existence of a supervisor. For an analysis of social relations between a principal and an agent without a supervisor, see Dur & Tichem (2014). The social relationship between principal and monitor is also not the focus. It too is taken as given. Because the monitor works for the principal, the principal trusts the monitor.

[16] This approach differs from the approach in, for example, Tichem (2013) and Dur & Tichem (2014), where the part of agent i's utility that depends on the others' is αU_j (as in Bergstrom, 1999). The utility functions here avoid second order effects, which would complicate analysis without added insight.

her own. The agents are unaware of the value of α before they work together.

For simplicity, the social relation is assumed to be symmetrical and time-invariant. It does not change through the course of the production process. It comes unexpected to the agents, i.e. they do not account for it when making their participation decisions. The social relation only realizes in period one. P does not have a social relationship with either of her employees. B and M also do not have a social relation with each other, i.e. for them $\alpha \equiv 0$. This too is assumed for simplicity but does not affect the decision problem. In period two, a social relation, whether A or B is working, affects only expected utility levels. It does not affect choices, because there are no more choices to be made. These expected utility levels do determine prior actions, but period two does not offer either the monitor or the agent the opportunity to choose actions based on the social relation. The loss of a previously existing relation can however influence strategic choices.

A and B's expected work utility in each period is given by

$$E(u_{kt}) = w_k - p_i \gamma_k c \tag{6}$$

where w_k is the agent's wage and $\gamma_k > 0$ represents a psychological feeling of guilt for having made a mistake of monetary value c (cf. Ellingsen et al., 2010).[17] M's expected work utility in each period is given by

$$u_{Mt} = w_M + b_{Mt} \tag{7}$$

where w_M is her wage, which is independent of her report and b_{Mt} her bonus in t. The bonus P pays M aligns M's financial interests with P's. P prefers paying c to M to paying for and dealing with a costly error. Assume therefore that P pays c to M when no error occurs. M's expected bonus then is

$$E(b_{Mt}) = (1 - p_i)c. \tag{8}$$

Note that M can, in this context, not be fired or punished for lying. This is a direct consequence of the principal's implied trust in her. The principal believes M tells the truth and has, in her blinkered view, no reason to believe otherwise. The bonus' size is assumed and may not be optimal. Still, the importance of the bonus is to align interests between P and M. Through the bonus, M has an incentive to care about the error propensity of A.

All employees are protected by limited liability. This ensures that P cannot "sell the firm" to M to (further) align their interests. The limited liability constraints are

[17] The assumption that A suffers from making a mistake is relaxed in section 5.2.

$$w_A \geq 0, \quad w_B \geq 0, \quad w_M \geq 0, \quad b_M \geq 0. \tag{9}$$

The socially blinkered principal and the incompleteness of contracts evident above are both a consequence of the model setup as well as a necessity for social relations to have consequences at all. In a complete contract environment or in the presence of an all-knowing principal, agent behavior for any state of nature would be predetermined and enforceable, rendering social relations either moot or incorporating them perfectly in the model ex ante (cf. Rotemberg, 2006). As the principal is blinkered, she does not do so.

The timing of the one-shot production process is as follows: First, in $t = 0$, A is hired. In the first period, M observes type. Then M sends the message to P, who makes the decision to retain or fire A. Simultaneously, A makes or does not make a mistake. Then, first-period payoffs are realized. If A is fired, it takes place at the end of period one, so production in the period is still Q. In period two, if A has been fired, she gets replaced by B, who then makes or does not make a costly mistake. After this, the second-period payoffs are realized and production ends. If A is retained, she again either makes or does not make a costly mistake, after which payoffs are realized and production ends. Figure 1 is a timeline of this process. The long vertical lines in $t = 1$ indicate the intermediate stages in period one, while the short bars that cross the timeline indicate the beginning and end of a period. In the next section, the monitor's decision problem is analyzed.

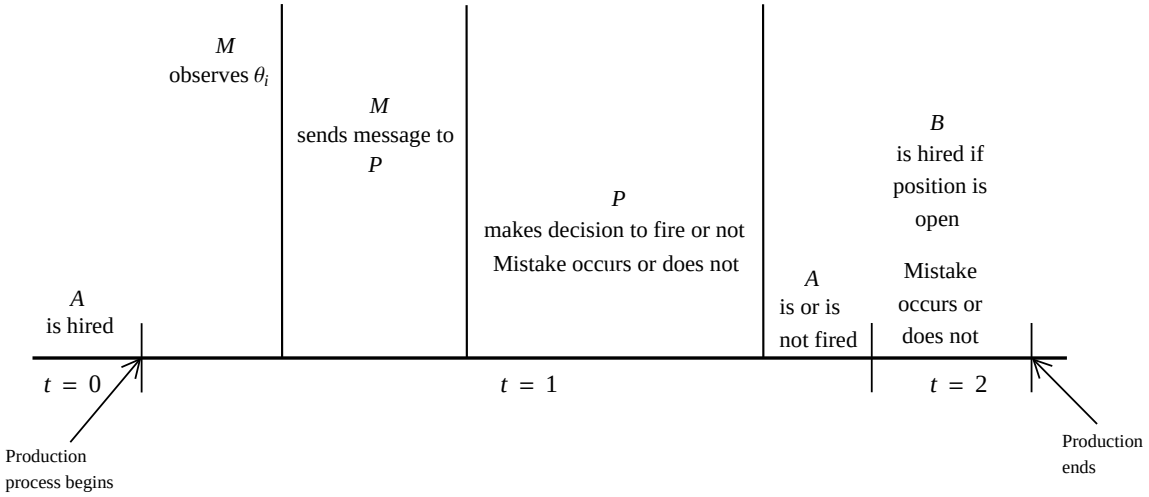

Figure 1: Timing of the Production Process

4 Analysis

In this section, the basic model is analyzed. This is done for discrete cases. First, the case of M working with a low-type friend is analyzed, followed by the case of M working with a high type foe. Third, comparative statics are reported. The implications of the results conclude this section. If the principal was not socially blinkered, she would expect the implications of a social relation, form beliefs about its consequences and then modify the job descriptions and incentives of her agents. This would lead to optimal contracts under the given constraints and would be done while factoring in all available information. The situation for a socially blinkered principal is much more constrained: For example, information that people can like or dislike each other and that this affects their decisions is hardly unavailable to the principal. However, while the information may technically be available, the principal does not take it into account (cf. Galanis, 2013, p. 43ff.). She forms her beliefs not on all available information, but on the information she *belives to be relevant*. Her beliefs are (preventably) wrong, but she is not aware of this. Because she is not aware of her folly but is otherwise rational, she acts optimally according to her beliefs, which she believes to be correct. In the absence of social relations, her beliefs *are* correct. It has previously been assumed for convenience that P trusts the message M sends. However, this assumption can be dropped here and replaced by the following proposition:

Proposition 1. *In the absence of social relations, or in the presence of a neutral social relation, it is optimal for P to trust M, for M always tells the truth.*

The proof for Proposition 1 is in Appendix A.1. However, the principal's beliefs about the non-existence of social relations are incorrect. By holding these incorrect beliefs the principal inadvertently allows M to take advantage of her. Therein is the underlying problem to be analyzed here. Clearly, in the absence of any other available information (that P is aware of), it is optimal for P to believe M's message.[18] As the monitor observes type, she knows more than the principal, and, as M has a (financial) interest in picking the right agent, P has no reason to distrust her —because she is unaware of the incentive distortion through α.

The environment described above may at first glance seem like a three player game. However, it should be noted that only M actually makes a decision. M's decision problem is shown in figure 2. She decides whether to lie or tell P the truth about A's type. All other situations are determined by nature: Nature picks A's type, whether a mistake occurs in period one, and regardless of what agent works in period two, picks

[18]The occurrence of a mistake could theoretically lead the principal to privately assume that the agent is of low type. While this may be true in other circumstances than those modeled here, it is nonsensical here. Given that she trusts the supervisor and that even a high type agent may experience bad luck, the occurrence of an error or its non-occurrence becomes irrelevant.

whether a mistake occurs then too. But ultimately, M can decide what to do in only two discrete cases, when type is either high or low, as drawn by nature. This decision problem holds no matter what the underlying distribution of agent types is.

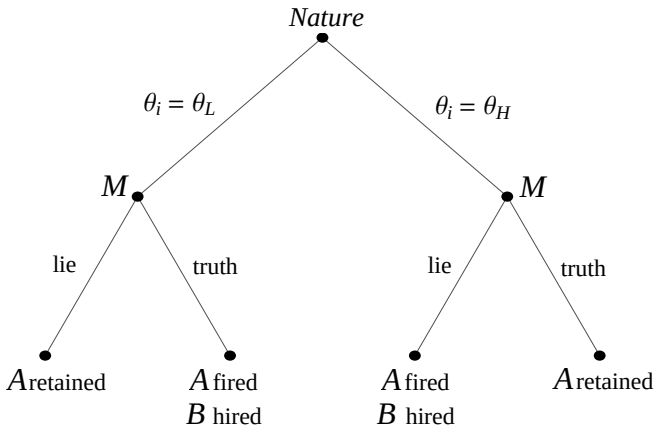

Figure 2: The Monitor's Decision Problem

It is assumed here that each individual's participation constraint holds. It is possible that the expected error cost would lead to P preferring not to engage in production at all. This special case is ignored in the analysis: It is assumed that Q is sufficiently high to induce P to engage in the production process. If P decides not to produce, the moral hazard problem that is investigated in this thesis does not arise. No co-workers or co-worker relations exist then. Because the workplace does not exist, this situation is not of interest here. Also, such high error costs may make it suddenly economically feasible to eliminate the production weak spot that causes the error. When an agent's type then makes no difference to the firm, the model does not apply because the moral hazard problem it describes does not exist.

It is therefore assumed that the respective expected utilities in (3) for P and in (5) for the agents are greater or at least equal to zero. If they are equal to zero, it is assumed that the agents still participate.

As can be seen in figure 2, there are two aspects to this moral hazard decision problem. P believes wrongly and also prefers that M gives a truthful evaluation of his agent. But M can use her information advantage in two different ways that are harmful to P's expected utility, depending on A's type. First, she could lie to prevent a bad agent from being fired. This can happen when the social benefits outweigh the cost of keeping A around, i.e. they have a good relationship. Second, she could lie to get a good agent fired. This would happen when the social benefits do not outweigh the benefit of keeping a good agent around, so social relations are negative. In either case, the social relation she has with A determines whether M lies or not. As M is aware of A's type before sending the message, these two cases are analyzed separately

below, starting with the low-type problem.

It is obvious that M's decisions are dependent on the value of α. It is the only factor that could change her decision from telling the truth to lying. The outcome of cases where M can keep a good agent whom she likes or can lead to the firing of a bad agent whom she dislikes are obvious. They shall therefore not be elaborated here. In these cases, α acts in the same direction as θ_i, positively influencing M's expected utility for θ_H and negatively for θ_L. Since the effect is in the same direction, they lead to M's truthtelling constraint —see (60) in Appendix A.1— to hold even more strictly. The effect on the principal is the same as if the agents had none or neutral social relations, for the monitor always tells the truth. The cases of interest arise only when expected profits (bonuses) act in opposition to social relations.

4.1 Monitoring a Low-Type Friend

Recall that M does not know B's type ex ante and has no social relations with her. In this case, the social relation between M and A is positive. So if $\theta_i = \theta_L$ and $\alpha > 0$ M's expected utility from telling the truth, leading to A's firing, is ex ante (before A has made a mistake)

$$E(U_{Mtotal}|\alpha > 0, \theta_i = \theta_L)_{truth} = 2w_M + (1-p_L)c + \alpha(-p_L\gamma_A c) + \frac{1}{2}(1-p_L)c + \frac{1}{2}(1-p_H)c \tag{10}$$

for periods one and two combined. The $\theta_i = \theta_L$ on the right of the vertical bar refers to agent A's *true* type, not the type P believes her to have. This notational convention is followed through the remainder of the thesis.[19] The beliefs of P can be inferred by the subscript *truth* or *lie*, which respectively mean that P's belief is correct or wrong. If M lies, her expected utility for both periods is

$$E(U_{Mtotal}|\alpha > 0, \theta_i = \theta_L)_{lie} = 2w_M + 2(1 - p_L)c + 2\alpha(w_A - p_L\gamma c). \tag{11}$$

Subtracting (11) from (10) gives

$$\alpha\left(c\gamma_A p_L - 2w_A\right) + \frac{1}{2}c\left(p_L - p_H\right) \tag{12}$$

which has to be larger than zero in order for M to tell the truth, or, equivalently,

$$\alpha\left(c\gamma_A p_L - 2w_A\right) \geq -\frac{1}{2}c\left(p_L - p_H\right). \tag{13}$$

[19] This always refers to A and never to B.

As $p_L > p_H$, $\frac{1}{2}c\,(p_L - p_H)$ is positive. Since no assumptions about w_A's size in relation to $c\gamma_A p_L$ has been made explicitly, the term $(c\gamma_A p_L - 2w_A)$ that α is multiplied with could theoretically be positive, zero or negative. If it is positive, the resulting constraint is a contradiction. This would imply that A's participation constraint in the absence of social relations does not hold. If it is zero, the variable that is of interest, α, disappears. Then, the (positive) social relation has the same effect as none at all. Since the truthtelling constraint always holds if social relations have no consequences or do not exist, this situation is already described by Proposition 1. It is therefore assumed that A's participation constraint for a given period holds with a strict inequality (i.e. that $w_A - c\gamma_A p_i > 0$).[20] Therefore, it has to be that

$$\alpha \leq \frac{c(p_H - p_L)}{2c\gamma_A p_L - 4w_A} \tag{14}$$

which is M's truthtelling constraint when dealing with a low-type agent she is friends with.

4.2 Monitoring a High-Type Foe

For $\alpha < 0$ and $\theta_i = \theta_H$, it again has to hold that the expected utility from being truthful

$$E(U_{Mtotal}|\alpha < 0, \theta_i = \theta_H)_{truth} = 2w_M + 2(1 - p_H)c + 2\alpha(w_A - p_H\gamma_A c) \tag{15}$$

is larger than the expected utility from lying, which is

$$E(U_{Mtotal}|\alpha < 0, \theta_i = \theta_H)_{lie} = 2w_M + (1 - p_H)c + \alpha(-p_H\gamma_A c) + \frac{1}{2}(1 - p_L)c + \frac{1}{2}(1 - p_H)c. \tag{16}$$

Subtracting (16) from (15) then results in the constraint

$$\alpha(2w_A - cp_H\gamma_A) + \frac{1}{2}c\,(p_L - p_H) \geq 0 \tag{17}$$

which, as before, assuming $(2w_A - cp_H\gamma_A) > 0$ and rewritten for α gives

$$\alpha \geq \frac{c(p_L - p_H)}{2c\gamma_A p_H - 4w_A} \tag{18}$$

which is M's truthtelling constraint when dealing with a high-type agent she dislikes. The next subsection elaborates on the truthtelling constraint of M.

[20] Recall that none of the participants take α into account when forming their expectations.

4.3 The Truthtelling Constraints of the Monitor

This subsection summarizes the results of the analysis above. From Proposition 1 it is clear that M's truthtelling constraint always holds for $\alpha = 0$. For the two cases described above, the expected bonus and α act in differing directions.

Proposition 2. *If $\alpha \neq 0$, M tells the truth only if α falls within a range. For monitoring a low-type friend, this range is*

$$0 < \alpha \leq \frac{1}{2}\frac{c(p_H - p_L)}{c\gamma_A p_L - 2w_A} \tag{19}$$

and when monitoring a high-type foe it is

$$0 > \alpha \geq \frac{1}{2}\frac{c(p_L - p_H)}{c\gamma_A p_H - 2w_A}. \tag{20}$$

Proof of Proposition 2. Follows directly from (14) and (18). □

The denominators of both (14) and (18) are negative by the assumption that regardless of type, A's participation constraint (while not taking α into account) holds with a strict inequality, i.e. $w_A - c\gamma_A p_i > 0$. Since in the cases described in Section 4.1 $1 > \alpha > 0$ and in Section 4.2 $-1 < \alpha < 0$, the truthtelling constraints for M depend on (the sign of) the numerators of (14) and (18). Since (14)'s denominator is negative but α is not, it has to hold that the numerator is positive. Similarly, since (18)'s denominator is negative and α is too, it has to hold that its numerator is negative as well. As can be easily seen, both cases always hold. Both numerators are independent of the type that A actually has but dependent on the difference in error likelihood between types. Although P does not know this, this is good news for her: The two constraints do not contradict each other outright. It *can* be that both constraints can hold at the same time, for both are a opposites of each other, in sign and in agent type they concern. It however does not have to be the case. Ultimately, it depends on the precise values of c, γ_A, w_A and p_i.

When monitoring a low-type friend, α has to be smaller or equal one half times the total expected losses M makes by lying divided by the difference between A's psychological costs of making a mistake and A's wage. When monitoring a high-type foe, α has to be larger or equal to one half the losses she makes from lying divided by the difference between A's psychological cost of making an error and her wage. How either constraint behaves according to the values of c, p_L and p_H depends on the value of γ_A. Per assumption, $\gamma_A > 0$, but if it is larger than one, (20) becomes harder and (19) becomes easier to satisfy in c, p_H and c and p_L, respectively. As is intuitive, both constraints become easier to satisfy in the difference in likelihood of making a mistake, if it becomes larger, the constraints become easier to satisfy. That A receives a wage which makes her participate is of greater importance than the actual wage she

receives, as soon as her participation constraint holds with inequality, it no longer has any influence on M's decision (but on her expected utility level) how large A's wage actually is.

The following subsection contains comparative statics for the case of M monitoring a low-type friend and the case of M monitoring a high-type foe.

4.4 Comparative Statics

This subsection gives comparative statics based on the analysis above. It is obvious that everyone is better off if no mistake occurs in both periods. The following comparative statics utilities are written in expected utility terms in order to allow differentiation of cases where no errors occur. The comparative statics for M reported below are from the perspective in time of M's decision to tell the truth or lie, i.e. from before a mistake has been made but after α has become known to the agents. First, the case of a low type friend being monitored is elaborated, followed by the case of monitoring a high type foe.

4.4.1 The Low-Type Friend Case

Comparative statics for A and B are omitted here for brevity. They can be found in Appendix A.2.2. If $\theta_i = \theta_L$ and α is in the range (19) prescribes, M again does not lie. P fires A according to the (truthful) message and hires B for period two. B and M do not have a social relation. For M then,

$$E(U_{M1}|\theta_i = \theta_L)_{truth} = w_M + (1 - p_L)c + \alpha(w_A - p_L\gamma_A c) \tag{21}$$

and, because B's type is distributed as A's and also unknown ex ante,

$$E(U_{M2}|\theta_i = \theta_L)_{truth} = w_M + \frac{(1 - p_L)c}{2} + \frac{(1 - p_H)c}{2} \tag{22}$$

which gives

$$E(U_{Mtotal}|\theta_i = \theta_L)_{truth} = 2w_M + \frac{3(1 - p_L)c}{2} + \frac{(1 - p_H)c}{2} + \alpha(w_A - p_L\gamma_A c). \tag{23}$$

P, due to M's message fires A, resulting in an expected utility for P in period one of

$$E(U_{P1}|\theta_i = \theta_L)_{truth} = Q - w_M - (1 - p_L)c - \gamma_P p_L c \tag{24}$$

and for period two of

$$E(U_{P2}|\theta_i = \theta_L)_{truth} = Q - w_M - w_B - \frac{(1-p_L)c}{2} - \frac{(1-p_H)c}{2} - \frac{\gamma_P p_L c}{2} - \frac{\gamma_P p_H c}{2} \quad (25)$$

giving a total expected utility of

$$E(U_{Ptotal}|\theta_i = \theta_L)_{truth} = 2Q - 2w_M - w_B - \frac{3(1-p_L)c}{2} - \frac{(1-p_H)c}{2} - \frac{3\gamma_P p_L c}{2} - \frac{\gamma_P p_H c}{2}. \quad (26)$$

P's expected losses through having hired a low-type agent are set off in part through the wage that is not payed to A in period one. Still, P is clearly better off having hired a good agent to start with. Since per assumption here M's truthtelling constraints hold, in both cases M is better off telling the truth. Obviously, M also prefers dealing with a good agent from period one over a bad agent in period one with a 50:50 chance of encountering a good agent in period two.

When M's truthtelling constraints do not hold in either of the discrete cases of working with a high-type foe or working with a low-type friend, P's beliefs are wrong. If M's message is that $\theta_i = \theta_H$, A does not get fired. B receives zero. M's first-period expected utility is

$$E(U_{M1}|\theta_i = \theta_L)_{lie} = w_M + (1-p_L)c + \alpha(w_A - p_L\gamma_A c) \quad (27)$$

and her second period expected utility is

$$E(U_{M2}|\theta_i = \theta_L)_{lie} = w_M + (1-p_L)c + \alpha(w_A - p_L\gamma_A c) \quad (28)$$

giving a combined expected utility of

$$E(U_{Mtotal}|\theta_i = \theta_L)_{lie} = 2w_M + 2(1-p_L)c + 2\alpha(w_A - p_L\gamma_A c). \quad (29)$$

P here believes that $\theta_i = \theta_H$, which is not true. She believes her expected utility to be higher than it actually becomes, but P does not know this and does never find out. The four final outcomes (no mistake in either period, a mistake in either period one or two, or a mistake in both periods) are not equally likely for both types, but P does not know which agent type actually leads to the outcome in question.

P believes her total expected utility to be

$$E(U_{Ptotal}|\text{wrong belief}, \theta_i = \theta_L)_{lie} = 2Q - 2w_M - 2w_A - 2(1-p_H)c - 2\gamma_P p_H c \quad (30)$$

while her actual expected utility is

$$E(U_{Ptotal}|\theta_i = \theta_L)_{lie} = 2Q - 2w_M - 2w_A - 2(1 - p_L)c - 2\gamma_P p_L c, \tag{31}$$

which, given that $p_L > p_H$, is lower than what she believes. P is obviously worse off when she is being lied to. B suffers from this situation too, foregoing second period wages involuntarily. M and A are better off in this situation than if M had told the truth.

4.4.2 The High-Type Foe Case

Comparative statics for A and B are again omitted for brevity. They can be found in Appendix A.2.1.

If $\theta_i = \theta_H$, and (20) holds, M does not lie, A does not get fired and a mistake occurs only with p_H. Regardless of the sign of the social relation, M then receives the following first-period expected utility

$$E(U_{M1}|\theta_i = \theta_H)_{truth} = w_A + (1 - p_H)c + \alpha(w_A - p_H\gamma_A c) \tag{32}$$

which is identical to her expected utility for period two, leaving her with an expected total utility of

$$E(U_{Mtotal}|\theta_i = \theta_H)_{truth} = 2\left(w_M + (1 - p_H)c + \alpha(w_A - p_H\gamma_A c)\right). \tag{33}$$

P's expected first period expected utility is

$$E(U_{P1}|\theta_i = \theta_H)_{truth} = Q - w_A - w_M - (1 - p_H)c - \gamma_P p_H c \tag{34}$$

which is identical to her expected utility for period two. This gives a total expected utility for the principal of

$$E(U_{Ptotal}|\theta_i = \theta_H)_{truth} = 2(Q - w_A - w_M - (1 - p_H)c - \gamma_P p_H c). \tag{35}$$

As is obvious, all involved here have a higher utility when no errors occur than when one error occurs than when two errors occur. In the absence of any errors, ceteris paribus, A and M are better off with positive rather than negative social relations. They are also better off with neutral sentiments than negative ones. If $\gamma_A c > w_A$ and a mistake occurs (regardless of in which period), M is better off with negative relations than with positive relations, while A is worse off regardless of whether a mistake occurs. P is unaffected by the value of α. P's expected utility is, in this case, exactly what she believes her expected utility to be. When M's truthtelling constraints do not hold, the utility she believes to be her expected utility is not her expected utility. Assume now that it does not hold. When (20) does not hold, M sends the message that $\theta_i = \theta_L$ and

A gets fired. B benefits, receiving $w_B - p_L\gamma_B c$ in period two. M's expected utility for period one is

$$E(U_{M1}|\theta_i = \theta_H)_{lie} = w_M + (1 - p_H)c + \alpha(w_A - p_H\gamma_A c) \tag{36}$$

and is

$$E(U_{M2}|\theta_i = \theta_H)_{lie} = w_M + \frac{(1 - p_L)c}{2} + \frac{(1 - p_H)c}{2} \tag{37}$$

for period two, giving her a total expected utility for the entire production process of

$$E(U_{Mtotal}|\theta_i = \theta_H)_{lie} = 2w_M + \frac{(1 - p_L)c}{2} + \frac{3(1 - p_H)c}{2} + \alpha(w_A - p_H\gamma_A c). \tag{38}$$

Given that this situation occurs, M is better off from lying than from telling the truth.

P's wrong beliefs are

$$E(U_{Ptotal}|\text{wrong belief}, \theta_i = \theta_H)_{lie} = 2Q - 2w_M - w_B - \frac{3(1 - p_L)c}{2} - \frac{(1 - p_H)c}{2} - \frac{3\gamma_P p_L c}{2} + \frac{\gamma_P p_H c}{2} \tag{39}$$

while truly

$$E(U_{Ptotal}|\theta_i = \theta_H)_{lie} = 2Q - 2w_M - w_B - \frac{(1 - p_L)c}{2} - \frac{3(1 - p_H)c}{2} - \frac{\gamma_P p_L c}{2} + \frac{3\gamma_P p_H c}{2}. \tag{40}$$

P believes herself to be better off than she actually is. P is the only one involved that is always better off when M tells the truth, although she wrongly believes herself to be in a better position than she is actually in.Obviously, the best possible outcome for M, A and P —though not for B— is that of a strong positive relation without mistakes occurring. Similarly, the worst outcome for them is that of A and M having negative relations, where A then gets fired after making a mistake and B also making a mistake. Note that these outcomes do not depend on types. They can occur for both types, and are indistinguishable for P. The writing of the equations above in expected utility terms is therefore necessary because a mistake's occurrence only determines (expected) utility levels and does not influence M's decision to lie or tell the truth.

4.4.3 Summary

P's main problem here is that she does not know that she may have an incentive problem among her workers. She expects M to behave always truthfully. Given that a mistake can occur for both types of agents, the principal does not know and cannot know when she is being taken advantage of. Table 1 shows which individuals are better off depending on type and M's decision. θ_i refers to A's type. A plus sign (+) marks the situation in which the individual is expected to be better off, while a minus sign (–) denotes the cases where the individual can be expected to be worse off. M, given the decision she makes, can expect to be either better off by lying or telling the truth. This depends on the range α is in. Whatever choice she makes indicates whether she expects to be better off in which situation.[21]

	$\theta_i = \theta_L$			$\theta_i = \theta_H$		
	A	P	B	A	P	B
M tells truth	–	+	+	+	+	–
M lies	+	–	–	–	–	+

Table 1: Expected Utility Comparison for M's Choices

Allowing for social relations between A and B does not change that M has an incentive to get an A she does not like fired. This would happen in the hope that the social relation with B is better. It also gives an incentive to keep an A who is friendly for fear that relations with B would be worse. The precise effect of such a social relation on M's truthtelling constraints ultimately depend on the probability distribution function one assumes for α. For the analysis conducted here it is sufficient to know that ceteris paribus, losing a friend hurts M's utility while losing an adversarial co-worker raises it. This also holds for the case of working with a high type A. It should be noted that M suffers (gains) in utility when the agent she is friends (enemies) with makes a mistake.

4.5 Implications

The implications of the results of the analysis are clear: When social relations between the monitor and the agent are ceteris paribus pronounced enough in either direction and lead to a misalignment of interests, a selection of agents that is not preferred by the principal happens during the first period. This has negative effects on the principal's expected utility in both periods.

The basic model without extensions then gives the following two hypotheses that can be tested experimentally:

[21] More detailed information about A and B's expected payoffs is contained in Appendix A.2.

Hypothesis 1. *A good worker that is disliked by his supervisor is more likely to receive a bad evaluation (and therefore get fired) than a good worker who is liked or on neutral terms with the supervisor. This leads to an efficiency loss from the firm's perspective.*

Hypothesis 2. *A bad worker that is liked by his supervisor is more likely to receive a good evaluation (and therefore will remain employed) than a bad worker who is disliked or on neutral terms with the supervisor. This leads to an efficiency loss from the firm's perspective.*

Another important aspect of the ranges for α in Proposition 2 is that it is ex ante possible for both constraints to hold at the same time. This may intuitively be seen as good news for P. However, given that P does not take into account the real incentive distortions of the social relation, this does not change anything. But even if the distortions would be taken into account, P's options in influencing the truthtelling constraints are limited. She cannot influence the denominator of either constraint. Her only choice variable there is w_A, which still has to be large enough to make A's participation constraint hold. While c in the denominator means the cost of the error, c in the numerator is refers to the bonus paid to M. The bonus is a choice variable for P, which is assumed to be c if no error occurs and zero otherwise. If one relaxes this assumption and allows P to pick a different bonus, P may need to pay a higher bonus to make the truthtelling constraints hold. In an extreme case, this bonus may become higher than the total cost of $\gamma_P c$ to the principal and lead to a situation where P may be willing to take the risk of employing a low type agent rather than not producing at all.

Bad relations with low-type workers and good relations with high-type workers have the same effect on the monitor's evaluation decision as none or neutral social relations. The monitor sends a truthful evaluation. Good relations with low-type workers or bad relations with high-type workers can distort the monitor's incentives and lead to untruthful evaluations. The effect of social relations on the principal is then either non-existent or negative, but never positive.

If empirical evidence supports the hypotheses stated above, they give valuable insights into both existing businesses as well as for budding entrepreneurs wanting to go into business. P, in the situation described above, is helpless, because she cannot influence M's decision. She never finds out that she is being taken advantage of and cannot fix a problem she does not know exists. The extensions in the next section relax this assumption.

5 Extensions

This section extends the basic model described and analyzed above. The basic model can be modified in different ways to apply to different circumstances. It is possible to make any number of small modifications. Two of these possible extensions are elaborated in the following subsections. First, it is allowed for the principal to be more involved than previously, by randomly learning the agent's true type. Then, a third type of agent is introduced and it is allowed for this type of agent to steal from the firm in a context without an involved principal. Then, the two extensions are combined. Since the impact of these modifications on the monitor's decision problem is the focus of these extensions, it is refrained from reporting the slightly modified comparative statics results. Instead, only the factors that influence the monitor's decision problem are reported.

5.1 A More Involved Principal

Previously, it was assumed that P is either physically removed from the production process or does not understand its particulars. It is assumed from here that P has a probability $p_P > 0$ to randomly learn the agent's type in period one after receiving the message but before making the firing decision. p_P can, for obvious reasons, not be a choice variable of P, but some exogenously determined probability. Otherwise, if P would have to choose to invest some positive amount of time and effort to find out the value of θ_i, M's monitoring task would be redundant. Also, given that P holds the belief that M never lies, P would never choose to undertake such an action.[22] Then, allowing P to check θ_i with some accuracy would not change M's incentives, because she knows that P will never check.

Assume that p_P is known to M, but A does not know that P has a chance of learning type at all. The non-strategic learning of θ_i needs to also imply some sanctions for M in order to impact his decision. In order to produce Q in a given period, two workers are still necessary. Assume that if M is caught lying, she is fired, receiving her outside option of zero for both periods and no bonus. When P learns type, a low-type A is still fired and a high-type A is retained, regardless of the message from M to P. P then can hire agent N to replace M who receives, who receives $w_N \geq 0$ in period two. While N replaces the monitor, she does not have monitoring tasks in period two, for monitoring does not effect the outcome. There are no strategic decisions to be made in period two, so N never receives a bonus. The decision problem is otherwise unchanged.

Assume that p_P is common knowledge and that M takes it into account when making her decision. P does not take it into account, for she never expects M to lie

[22] Assume the tie-breaker that P prefers not to exert costless effort if she believes that it has no added benefit.

given her social blinkeredness. Assume also that p_P does not affect M's contract.[23] Assume that if P observes agent type and finds out M lied, she can legally fire M. The analysis of the decision problem then is straightforward and follows the same basic pattern as in Section 4. The two cases of interest —when monitoring a low-type friend or monitoring a high type foe— also remain the same, and, clearly, M still tells the truth always if $\alpha = 0$. For M, her expected utility from telling the truth remains the same as in (10) for a monitoring low type friend and (15) for a high-type foe. P's random learning of type does not have any consequences for M if she is truthful.

Assume again that participation constraints hold and A's holds with a strict inequality. When monitoring a low-type friend, M's expected utility from lying if she is not caught lying —with probability $(1-p_P)$— is the same as in (11). With probability p_H however, her expected utility is $\alpha(-p_L\gamma_A c)$. This gives the following truthtelling constraint with regard to α

$$\alpha \leq \frac{p_L\left(c-4cp_P\right)-cp_H+2\left(2p_P\left(c+w_M\right)\right)}{2\left(p_P-1\right)\left(2w_A-c\gamma_A p_L\right)} \tag{41}$$

while it simultaneously has to hold that $\alpha > 0$. Again, both numerator and denominator are negative. As is unsurprising, the constraint becomes easier to satisfy the larger p_P, as the RHS of (19) is larger than the RHS of (41).[24]

When monitoring a high-type foe, M's expected utility from telling the truth is still the same as in (15). When lying, with $(1-p_P)$, her expected utility is the same as in (16) and when caught lying with p_P, it is $\alpha(w_A - p_H\gamma_A c)$. This gives the following truthtelling constraint with regard to α

$$\alpha \geq \frac{p_H\left(3cp_P+c\right)+cp_L\left(p_P-1\right)-2\left(2p_P\left(c+w_M\right)\right)}{2\left(c\gamma_A p_H+w_A\left(p_P-2\right)\right)}. \tag{43}$$

It has to hold that M and A dislike each other, i.e. $\alpha < 0$. Again, allowing for the possibility that P randomly learns the agent's type makes the constraint easier to satisfy.[25]

[23] This may not be optimal from the principal's perspective, for there now is a chance that the bonus paid to the monitor is redundant because the principal learns type anyway. Not changing M's contract can be interpreted as another consequence of the principal's blinkeredness, while the principals theoretically optimal choice may be to reduce the bonus or to pay out the bonus only if the principal does not learn type.

[24] p_P is not a choice variable of P. However, if she could chose the probability of randomly learning type in a situation where M and low type A are friends, she would have to choose

$$p_P \geq \frac{p_L\left(2\alpha c\gamma_A+c\right)-4\alpha w_A-cp_H}{2cp_L\left(\alpha\gamma_A+2\right)-4\left(\alpha w_A+c+w_M\right)} \tag{42}$$

in order to induce M to tell the truth.

[25] If P was to choose p_P in the situation where M works with a high type foe, she would have to choose such that

$$p_P \geq \frac{4\alpha w_A+cp_L-p_H\left(2\alpha c\gamma_A+c\right)}{2\alpha w_A+3cp_H+cp_L-4\left(c+w_M\right)} \tag{44}$$

When M is not caught lying or tells the truth, expected utilities for P, M, A and B are as before. For N, in these cases, expected utility is zero and $w_N \geq 0$ otherwise.

		$\theta_i = \theta_L$				$\theta_i = \theta_H$			
		A	P	B	N	A	P	B	N
p_P	M told truth	–	+	+	–	+	+	–	–
	M lied	–	+	+	+	+	+	–	+
$(1-p_P)$	M told truth	–	+	+	–	+	+	–	–
	M lied	+	–	–	–	–	–	+	–

Table 2: Outcomes for M's Choices given P's learning

Table 2 shows the outcomes for each participant given that P either learns type or does not, depending on whether M told the truth or not. θ_i refers to A's type. The passive tense is due to P's learning occurring after having received the message from M. For M, if she gets caught lying, her utility is zero. If she does not caught lying, her expected utilities are as in section 4.4.

For anyone who is not P, in table 2 a minus sign (–) signifies an expected utility of zero. Either they do not get hired (B and N) or get fired (A and M). P still receives $2Q$ in total minus the costs of errors and wages, but her expected utility is lower when she does not find out she has been lied to. When she does find out that she has been lied to[26] she may even be better off since she does not pay the monitor's wage then and no bonus to N in the second period. The ideal situation then would be to already have a θ_H worker who would have been fired otherwise. But even so, P cannot intentionally cause this situation.

Where there is a plus sign (+) in table 2, expected utilities for A, B and N,respectively, are equal to the other cases where there is also a plus sign for that individual. For P, as is obvious without further analysis, the cases with a plus sign are preferred to cases where there is not. Still, P cannot influence which case actually occurs. M reveals her highest expected utility by the choice she makes. If she decides to lie, her expected utility from lying is higher from lying, even though her realized utility may not be if she gets caught.

This allows the forming of another hypothesis, which can be tested experimentally.

Hypothesis 3. *If the principal has a chance of randomly learning whether the monitor lies, it becomes less likely that the monitor lies.*

Having now established that M's truthtelling constraint is easier to satisfy if there is a chance of being caught lying, the next section introduces a third agent type before

holds to induce M to tell the truth. However, P cannot choose this value intentionally. It is, without further assumptions also not clear whether (44) or (42) is larger, i.e. which constraint has to be binding for both to hold.

[26]Finding out that M told the truth changes nothing, her beliefs then turn out to be correct and her expected utilities are as before in section 4.4.

combining both that third agent and the principal's learning in section 5.3.

5.2 Allowing the Agent to Steal

One of the possible extensions is interesting from a high-level corporate and auditing perspective. It shall therefore be elaborated here: Allowing an agent to intentionally harm the firm in his own financial interest. So far it has been assumed that the agents' mistakes are unintentional. It can easily be imagined that an agent's tasks give her the opportunity to defraud or steal from the firm. When the monitor observes such criminal behavior, the job description would dictate reporting the misbehavior. If the agent is aware of this, he can try convincing the monitor not to report. They then essentially engage in bargaining over whether to collude or not.[27] This subsection is arranged as follows: First, it is explained how the basic model can be interpreted to apply to the context of limited liability companies. Then, the basic model is modified to allow for an agent to steal from the firm before results are reported. The extension in this subsection and subsection 5.1 are then combined in the following subsection.

5.2.1 Context

The bargaining[28] towards collusion is the focus of this extension. It will be shown that an agent can cause a monitor's truthtelling constraint not to hold.

Hierarchical social relations arise at virtually all levels of hierarchies. Room for discretion can be very large at high levels. The basic principal-supervisor-agent framework established in section 3 can relatively easily be applied to boards and CEOs of limited liability companies. This can be done if one sees shareholders as the principal (uninvolved in daily affairs), CEO or director as the monitor (involved in the daily affairs and employed to ensure maximum chances of success)[29] and the agent as some

[27] The economic literature on collusion between agents in organisations, including side contracting, is extensive (cf. Laffont & Rochet, 1997; Laffont & Martimort, 1998). Yet relatively little theoretical economic research focuses on the influence of social relations on collusion and side contracts. Notable exceptions are Itoh (1991), Spagnolo (1999) and the previously discussed Prendergast & Topel (1996). Collusion in Itoh (1991) is merely implicit. Itoh's model allows workers to engage in non-productive socializing at the cost of work effort, a weak form of collusion. It is then found that firms under these circumstances can never reach first-best effort levels. Spagnolo (1999) focuses on the effects of social relations on cooperation, not collusion, while acknowledging that social relations also allow negative cooperation, i.e. collusion, without elaborating much further. Spagnolo's model however finds a positive effect of social relations on productivity. This differs from the model in this thesis in that the model here does not allow for synergies in production that arise from social relations.

[28] Social influences on bargaining outcomes have long been accepted in the economics literature (Harsanyi, 1962). Empirical evidence that social ties of many sorts reduce selfishness (cf. Hoffman et al., 1996; Fehr & Fischbacher, 2002; Fehr & Schmidt, 2006). It could also be argued that in any bargaining that allows for communication, some social relations arises that then influences the outcome.

[29] This distinction is blurry as board members and CEOs as well as high-level employees commonly also hold stock or get partly paid in stock options.

employee of the limited liability company, for example an internal auditor. The agent may also be an external auditor, in which case he would not be an employee directly, but someone who still gets paid by the firm to execute a task. Performance gets evaluated after a predetermined amount of time has passed (i.e. at the end of a period or quarter) and stockholders can receive a fixed dividend. Furthermore, if one neglects the impact random chance has on CEO performance (cf. Mlodinow, 2008, p. 99ff.) it stands to reason that any CEO has an interest in hiring only those employees that maximize the chance of high returns. It can easily be imagined that a monitor (CEO) has a social relationship with, say, an auditor or someone whose errors can negatively impact the company's performance. It is also intuitive that losses to shareholders in case of fraud or accounting errors can be larger than the bonus paid to a CEO (i.e. $\gamma_P > 1$ holds). The assumption that the costs of potential fraud or accounting errors, c, are known ex ante is merely simplifying and does not hold in practice.[30]

Empirically, Hwang & Kim (2009) finds that a substantial proportion of "independent" directors are in truth socially connected with the CEO or the firm. Fracassi & Tate (2012) find that the more powerful a director, the more likely that he uses discretion to appoint friends to a board. This is not captured within the model here and its extensions, but would likely apply if the monitor knew the value of α ex ante and had influence on the hiring decision in $t = 0$. Chidambaran et al. (2011) find that social connections between CEO and a company board increase the likelihood of a firm committing financial fraud. Previous professional relations between board and CEO decrease this likelihood. Dey & Liu (2010) find that social ties overall have a negative effect on executive performance. The academic, practitioner's and legal debate about high-level employee and CEO independence is ongoing, not to mention attempts at ensuring that auditing best practices are applied (Bruynseels & Cardinaels, 2013).[31] The contribution of this extension and the extension in section 5.3 to the debate is that it models from an economic perspective the willingness of a CEO or director to accept inefficiencies within the company at the expense of shareholders. This willingness depends to an extent on social relations (this extension) and on the probability of detection (in the extension in the next subsection). In the following, the basic model

[30] Assuming c to be a random variable with some mean, standard deviation and distribution does not significantly change conclusions but complicates the analysis. It is therefore refrained from making such an assumption, realistic as it may be. It is enough for shareholders to know that costly errors can occur.

[31] Several solutions have been proposed to deter corruption. In a public sector context, Cooter & Garupa (2000) show that a government can deter bribes —i.e. collusion— by rewarding whistleblowers, essentially actively seeding distrust —a form of negative social relations— between agents. Kofman & Lawarree (1996) use a similar logic to develop a prisoner's dilemma designed to prevent collusion. In the context of the model here, these methods would imply paying a bonus to the monitor for reporting the theft. This is the same as trying to induce truthtelling by the monitor, which is what the bonus paid to the monitor already does. Any additional function that the bonus could perform is not anticipated by a principal who expects the monitor to always truthfully communicate the agent's type.

will be modified, before the outcome of the modification is discussed.

5.2.2 Modification of the Basic Model

Before the basic model established in section 3 is modified to allow for a bargaining period when a thieving agent is present, such an agent's characteristics need be defined. Since an agent who can and may want to steal from the firm has some characteristics of a gangster, let this agent's type be known as θ_G. Let this agent be identical to a low-type agent in that the probability of making a mistake is for her p_L. She can either work in a period and make a mistake with that probability or intentionally decide to steal c from the firm and then not make a mistake. This loss is for P indistinguishable from an unintentional error. M observes the nature of the error perfectly. This is why bargaining is necessary for A to get away with stealing. For simplicity, assume that this bargaining can only occurs in period one, i.e. there is no option to steal again in period two. The gangster agent suffers a positive cost from not stealing. When she does decide to steal however, this psychological cost becomes negative, i.e.

$$\gamma_A = \begin{cases} \gamma_A & \text{in } t = 1 \text{ if } A \text{ does not steal} \\ \gamma_G & \text{in } t = 1 \text{ if } A \text{ steals} \\ 0 & \text{in } t = 2 \end{cases} \tag{45}$$

for a θ_G agent, where $\gamma_G < 0$. This can be easily interpreted. When the agent has the opportunity to steal, but does not and then makes a mistake, she suffers from a missed opportunity: Instead of the deadweight loss of a mistake, she could have taken the money herself and would have been better off, with P being none the wiser. She then regrets not having stolen money. In order to make A decide decide to steal, she has to believe she will be better off from stealing, so she has to gain from it.[32] Then the parameter does not imply guilt aversion as before, but signifies a measure of regret (when not stealing) and a measure of utility gain from stealing (still potentially moderated by guilt, as γ_G is not defined being smaller or equal to -1 in this case). In $t = 2$, the gangster agent does not care anymore about the performance of the firm, first, because she cannot benefit more than through her wage, and second, given that she is in period two, she no longer can get fired. This holds even if B is of gangster type and gets hired in $t = 2$, she does not care about the firm but only herself. A gangster-type B cannot steal in $t = 2$, except for the absence of feelings of guilt, she then is identical to a low-type agent.

Since M observes type, she may need to be incentivized to lie for A not to get fired. This requires some transfer from A to M. Let this fraudulent transfer be denoted as

[32] The stealing process may also be not be perfectly efficient, i.e. she may not be able to use all of c for consumption (bargaining not withstanding yet).

f.

A's work utility becomes in $t = 1$

$$u_{A1} = \begin{cases} w_A - \gamma_A c & \text{if she does not steal and} \\ w_A - \gamma_G c & \text{if she steals.} \end{cases} \tag{46}$$

If she does not steal, she suffers the psychological cost of not stealing with certainty, regardless of whether she makes a mistake or not. f is not part of the agent's work utility, because it does not arise from work but extralegal activities. In $t = 2$, if she is not fired, her work utility becomes

$$u_{A2} = w_A. \tag{47}$$

Now assume that there are no low-type agents, only high and gangster type agents in the eligible agent population. High-type agents do not steal and behave as in section 3. Assume further that P does not know about θ_G and believes that only low and high types exist. This is a simple extension of the principal's blinkeredness.[33] P does not account for criminal agents because she believes they do not exist or because she cannot distinguish between them or low-type agents, while in truth, it is equiprobable that an agent is of gangster or high type, i.e.

$$\Pr(\theta_i = \theta_G) = \frac{1}{2}, \tag{48}$$

The θ_G agent knows her type and knows the value of α, and therefore knows that M will truthfully message that she is a bad worker and should be fired. Assume then that she can then decide whether to steal and if she decides to, engage in costless one-shot bargaining with M. In terms of timing, this bargaining process takes place after M observes type, but before a message is sent to P. A makes an offer to split the stolen money c with M, who either rejects or accepts. There is no second round of negotiation. While the timing of the bargaining matters, it does not matter when in period one A steals, as long as costs (caused by stealing or error) do not realize before P's decision. This is similar to, but not identical with, ultimatum game bargaining. If M takes the offer, A is safe and does not get fired. Then, the bargaining leads to collusion. If M turns down the offer and informs P, A gets fired.[34] When A steals and gets reported, assume M gets a bonus of c for certain and A receives zero in total.[35]

[33] Alternatively, assume that P does know that θ_G agents exist, but (wrongly) believes that they will never engage in stealing behavior.

[34] It can easily be argued that in an attempt of theft that is reported, P can pursue legal action against A. This potentially changes A's outside option. Assume here though that A's outside option is still zero.

[35] If A is an external auditor, reporting a gangster A —regardless of whether A attempts theft or not— would generally imply the limited liability company's unwillingness to work with that external auditor in the future.

Assume that if A gets fired in this case, Q is still produced in period one.

In the following, the bargaining outcome is analyzed.

5.2.3 Analysis

The following description and analysis of the bargaining outcome focuses on the case A is of gangster type and where M's truthtelling constraint holds ex ante to bargaining. Assume that M cannot dishonestly report a high-type agent for an attempt of stealing money. The θ_H case then is uninteresting because it does not change from the earlier case (cf. section 4.2). Also, if M is already willing to lie for A, then it seems clear that collusion can easily occur.[36] This case has been already studied extensively (see the literature in section 5.2.1). The aim of this extension is to show that while social relations have distorting incentive effects, they can also help curb collusion. Three cases are then of interest in which it is assumed that M's truthtelling constraint holds ex ante to bargaining: the case of neutral, weakly negative or weakly positive social relations. These weakly pronounced relations fit with the selfish orientation of a gangster agent, for it would be expected that such agents care less for others than agents of other types. In all three cases, M holds all bargaining power, even though she does not initiate bargaining and, in the case of positive (negative) social relations, she can be thought of as having less (more) power, because the amount she requires to become a liar is dependent on social relations. This however does not imply that she necessarily requires the entire stolen amount to be kept quiet.

Consider first the case of weakly positive social relations. A's utility from not stealing in this case is $-\gamma_A c$, for she will be fired A knows that she needs to offer f such that M's expected utility from lying is larger than her expected utility from telling the truth. A clearly has no incentive to steal and not bargain, for she knows that M's truthtelling constraint holds. Since she has to bargain when she steals, she will offer M just enough of a transfer to induce M to lie. As is obvious, it has to hold that $f \leq 2w_A - \gamma_G c$, as that is the maximum possible amount A has to offer over both periods.

A knows that when she steals, M's expected utility from telling the truth is $w_M + c + \alpha \times 0$ for period one, since M receives c for sure if she reports A. M then receives $w_M + \frac{1}{2}(1 - p_L)c + \frac{1}{2}(1 - p_H)c$ in period two, for a total expected utility of

$$E(U_{Mtotal}|\alpha > 0, \theta_i = \theta_G)_{truth} = 2w_M + c\left(\frac{1}{2}(1 - p_H) + \frac{1}{2}(1 - p_L) + 1\right) \tag{49}$$

When M lies, her expected utility in $t = 1$ is $w_M + f + \alpha(w_A - \gamma_G c)$ and $w_M + (1 -$

[36] If social relations in this case are pronounced enough, it is also possible that no or only a minimal transfer between A and M is necessary.

$p_L)c + \alpha(w_A)$ in $t = 2$, giving a total expected utility of

$$E(U_{Mtotal}|\alpha > 0, \theta_i = \theta_G)_{lie} = 2w_M + f + (1 - p_L)c + \alpha(2w_A - \gamma_G c). \tag{50}$$

Subtracting (49) from (50) gives the following collusion constraint in order to induce M to lie:

$$f_{(\alpha>0)} > \frac{c(p_L - p_H)}{2} + c - 2\alpha w_A + \alpha c\gamma_G, \tag{51}$$

recall the assumption that if M is indifferent between lying and telling the truth, she tells the truth. In order to convince M to lie, A has to offer her more than one half the difference between what M can expect to gain from telling the truth and what she gains through continued interaction with A. Whether $f_{(\alpha>0)} \leq 2w_A - \gamma_G c$ holds depends on parameter values. α has a negative effect on $f_{(\alpha>0)}$, making the constraint easier to satisfy.

Consider now the case of neutral social relations. M's utility from telling the truth remains unchanged for a combined expected utility as in (49). When M lies, her utility in $t = 1$ is $w_M + f$ and her expected utility is $w_M + (1 - p_L)c$ in $t = 2$, giving a total expected utility of $2w_M + (1 - p_L)c + f$. Subtracting her expected utility from telling the truth from her expected utility from lying gives a collusion constraint for neutral social relations as

$$f_{(\alpha=0)} > c + \frac{c(p_L - p_H)}{2} \tag{52}$$

Comparing (51) and (52), it is clear that without the moderating effect of positive social relations, the transfer required by M is larger. Whether $f \leq 2w_A - \gamma_G c$ holds still depends on the values of the parameters, but it becomes less likely than before.

Consider now the case of weakly negative social relations between M and A. Again, M's expected utility from reporting A is unchanged as in (49), as A's utility then becomes zero. M's expected utility from lying is as in (50), with the important difference that $\alpha < 0$, giving a collusion constraint of

$$f_{(\alpha<0)} > \frac{c(p_L - p_H)}{2} + c - 2\alpha w_A + \alpha c\gamma_G. \tag{53}$$

The interpretation of this constraint is the same as the one for weakly positive relations. An important difference between (53) and (51) is that $\alpha < 0$ changes the sign of the last two terms on the right hand side of the inequality. This implies that ceteris paribus, for two cases in which one social relation is positive and one negative, with the same absolute value for α, $f_{(\alpha<0)} > f_{(\alpha>0)}$.

It follows then that $f_{(\alpha<0)} > f_{(\alpha=0)} > f_{(\alpha>0)}$. While $f_{(\alpha>0)}$ is decreasing in α and $f_{(\alpha=0)}$ remains unaffected by it, $f_{(\alpha<0)}$ is increasing in more pronounced negative values

of the social relation. A positive social relation and f act in the same direction, while f optimally (from A's perspective) cancels out expected losses from working with the gangster agent int $t = 2$ for neutral or negative social relations.

Since in general, $-1 < \alpha < 1$, i.e. no one person cares as much about the other as one does about oneself, it is possible to require more than the entire stolen amount to lie in all or any of the three cases. In such a situation, A would anticipate this and then not steal.

While the one-shot take-it-or-leave-it offer has semblances with ultimatum game bargaining, the results here are not precisely in line with experimental findings on the effect of social influences on ultimatum giving (cf. Güth & Tietz, 1990; Handgraaf et al., 2003; Fehr & Schmidt, 2006), but fits with the assumed relative selfishness of the gangster agent. Empirically, it is usually found that social ties increase the amount given, i.e. generosity increases. Here, positive social relations decrease the amount required to be convinced to lie, while negative social relations increase it. This is a consequence of both the continuation utility from working with the thieving agent in $t = 2$ and the assumption that M must be made just slightly better off when lying than from telling the truth in order to lie. However, the reduced (increased) amount M requires in the case of positive (negative) social relations can be seen as an increase (decrease) in generosity by M. A takes advantage of this to choose the minimum transfer possible. If one modified the utility functions to account for some measure of inequity aversion, the results here would likely be closer to empirical results on ultimatum game bargaining. Then it would become less likely that A engages in theft, for it may not become possible to offer a split that M would accept.

Nevertheless, the findings of this extension allow for the forming of the following testable hypothesis:

Hypothesis 4. *When an agent can steal from the firm and the monitor can observe the theft, the bribe the monitor requires in order not to report the agent is lower when monitor and agent are friends than when monitor and agent dislike each other. In the absence of social relations, it lies in between the bribes required in the presence of positive or negative social relations.*

This also implies that if employee collusion to steal is possible, P may actually prefer social relations to be negative between her employees. This is in contrast with the results attained in section 4.4, where social relations either have no effect on P or a negative one. P may P can not intentionally prevent theft because she does not think it will occur. However, it is more likely that her employees do not collude against her when when social relations are negative. In the next extension, it will be shown briefly that a more involved principal can inadvertently prevent employee collusion.

5.3 A Thieving Agent with a More Involved Principal

Bringing the results of the above extensions together, this extensions briefly highlights that it is possible for *negative* social relations between employees to have a positive effect on the principal.

Assume as in section 5.1 that M can be fired and her outside option is zero. Assume further that p_P is the principal's chance of learning agent type and that the stealing process is finalized only at the end of $t = 1$, i.e. that if stealing is reported or collusion is detected, A has a total utility of zero. Assume also that $\frac{c(p_L - p_H)}{2} + c - 2\alpha w_A + \alpha c \gamma_G < f \leq 2w_A - \gamma_G c$, irrespective of the value of α, holds in order to make the criminal agent want to steal. This, holding all other variables constant, *can* happen for sufficiently small absolute values of α if one allows for $\gamma_G < -1$, i.e. the agent benefits more from stealing than what the stealing costs the firm.[37]

As the collusion constraint, regardless of the value of α, is $f > \frac{c(p_L - p_H)}{2} + c - 2\alpha w_A + \alpha c \gamma_G$, A will offer M exactly the lowest amount she can to induce collusion. Recall that in case of indifference between lying and telling the truth, M tells the truth. Let $f^* = \frac{c(p_L - p_H)}{2} + c - 2\alpha w_A + \alpha c \gamma_G + \epsilon$ be the amount A offers to M, where ϵ is some small positive amount that breaks the utility tie. Recall also that it is known to M that P has a chance of learning type, but that this is not known to A. This essentially makes A "detection blinkered." It is reasonable to assume that in the setup described, M has more knowledge about P than A does.

Assume now that A offers f^* to M. Regardless of what social relations they have, M's total expected utility from telling the truth is as in section 5.2 equal to $2w_M + c\left(\frac{1}{2}(1 - p_H) + \frac{1}{2}(1 - p_L) + 1\right)$. However, not knowing about p_P, A thinks M's expected utility in case of lying is $2w_M + (1 - p_L)c + f$, while it is zero with p_P and $2w_M + (1 - p_L)c + f^*$ only with $(1 - p_P)$. But, as the collusion constraint has changed without A's knowledge, M now requires

$$f_{p_P} > \frac{\begin{array}{c} p_P(\alpha(2w_A - c\gamma_G) + c(1 - p_L) + 2w_M) - \alpha(2w_A - c\gamma_G) \\ + c\left(\frac{1}{2}(1 - p_H) - \frac{1}{2}(1 - p_L) + 1\right) \end{array}}{1 - p_P}. \tag{54}$$

Unless ϵ is very large, f_{p_P} is larger than f^*. If ϵ, as it is assumed to be, is a small positve number, this means that the information asymmetry about p_P causes M to report theft, leaving A worse off, while P inadvertently becomes better off. If there is a non-zero chance that P randomly learns A's type, then it will become harder and

[37] This can happen if A's criminal desires are very pronounced of if A can sell what she steals for more than what it cost the firm. This could imply that M may only agree to the trade if she does not know this (and it should therefore not turn up in her utility function) or if she receives an additional cut of the profits. This may apply in cases of industrial espionage, where the cost of a prototype may be less than the black market value of that prototype. The investigation of this problem goes beyond the scope of this thesis. The assumption that A's participation constraint in the bargaining process holds is made merely to allow investigation of the effect of a more involved principal in this case.

likely impossible to achieve employee collusion. This leads to a testable hypothesis:

Hypothesis 5. *When the principal can randomly learn that his employees colluded to steal from the firm and the monitor knows this but the agent does not, this makes the monitor less likely to collude with the agent, leading to truthful monitoring.*

If one now assumes that the collusion constraint even with the given probability of detection holds for positive and neutral social relations,[38] i.e. holds in the two cases it is easiest to satisfy in, it is easy to see that the only factor that could influence M's decision towards telling the truth is the presence of a negative social relation. This implies that when the agent has the opportunity to steal, it may be desirable from P's perspective that employees have negative social relations with each other. This holds whether P can randomly learn type or not. This implication can be translated into another testable hypothesis:

Hypothesis 6. *Negative social relations are from the principal's financial perspective preferable to good and neutral relations when negative relations act as a safeguard against employee collusion. This is especially true when other safeguards are insufficient to deter collusion by themselves.*

A similar hypotheses can be derived for neutral social relations if one assumes the collusion constraint to only hold for positive social relations, but not neutral relations. While P is socially blinkered, it can be the case that P is actually better off with negative social relations. This gives M's message for each case, as displayed in table 3, where θ_i refers to A's type. She can work with either of three agent types. The θ_L case refers to M's decisions in the basic, unextended model.

	θ_L	θ_H	θ_G
$\alpha > 0$	Possibly Lie	Truth	Possibly Lie
$\alpha = 0$	Truth	Truth	Possibly Lie, but less likely
$\alpha > 0$	Truth	Possibly Lie	Possibly Lie, but least likely

Table 3: The Monitor's Message given Agent Type and the Social Relation

In table 3, "Truth" indicates a case where M will always tell the truth. In the other cases, she may lie, depending on the value of α, and in the θ_G case, f. P does not expect M to lie, but is obviously worse off when M does.

Having given possible practical applications of the model and some extensions, the next section discusses limitations of the basic model as well as the extensions.

[38] This could happen if ϵ is very large.

6 Limitations

This thesis and the model contained therein are subject to various limitations. These are elaborated in this section. The limitations fall into two categories: General limitations stemming from simplifying assumptions (which apply mostly to economic modeling in general) and more specific limitations that arise from the scope of the thesis.

There are a number of simplifying assumptions in the model that do not necessarily hold in real world environments. Some of the fundamental assumptions that this thesis makes are: The principal does not take into account the effects of social relations between his employees, the agent and principal do not interact with each other, and the principal completely relies on the monitor, there are only two periods without discounting. It can be argued that there exist many real world workplaces that meet these criteria. However, assumptions that are made in the model may not adequately reflect reflect reality. The 50:50 split between types in the eligible agent population is likely not to occur in practice, given the existence of sophisticated screening processes. Also, to give another example, there is no practical reason to assume that social relations have to be perfectly symmetrical. Even if they are not perfectly (a-)symmetrical, they likely differ in scale between individuals. Strongly asymmetrical relations are though likely to be unstable, especially in the long run. The model also does not account for the chances of being friends or enemies with the next colleague hired, while a social relationship between the principal and the supervisor is only implied through the principal's (potentially ill-fated) one-sided trust in the supervisor.

Loosening these assumptions to allow, for example, for employees with asymmetric social relations towards each other with some discount factor will undoubtedly change both the nature of the monitor's decision problem and its outcomes. It is, however, this author's belief that doing so does not change the underlying decision problem as it is described here. It would allow to investigate more specific aspects of the underlying problem and is therefore a potential avenue for future research. Their inclusion here would complicate the model without much added insight into the consequences of not adequately accounting for employee social relations. For example, one could assume that the monitor has lying costs (Kartik, 2009) that may or may not carry over into the second period. This would change the truthtelling constraints, but does not significantly alter the decision problem, for the bonus and lying cost would act in the same direction. Or, one could question why the size of the bonus here is taken as exogenously given. A bonus can be paid in many different ways and its optimal size would ideally be determined by a non-blinkered principal. However, the bonus exists merely to align the monitor's financial interests with the principal's (given that there are no social relations). This simplification is excusable because the bonus is not the focus of this thesis. The same applies if one tries to account for social relations

changing over time, i.e. revenge motifs that may influence the monitor's decision to lie or tell the truth. Similarly to how firms in some markets are price-takers, the principal in this model can be seen as a "relation-taker." She can, at least in the context here and even if she was aware of its existence, not influence the social relationships of her employees. For the employees, the value of α may however change depending on whether the agent actually makes a mistake, regardless of type. It may also change with the monitor's decisions themselves. Also, in the basic model, the monitor never runs the risk of being fired, which is, while logical in the basic model setup, hardly realistic. Furthermore, the principal is rather uninvolved, i.e. does not make strategic choices during the production process. Similarly, the agent could be more involved, i.e. have more control within the production process. This would turn it into a game and not a decision problem. Investigations of these facets present interesting research avenues in themselves that are not pursued in this thesis.

Regarding limitations based on the scope of this thesis: The main one, it can easily be argued, is that the model is rather specific in the circumstances it models. The model only applies to a situation where a supervisor and agent are far enough out of reach of the principal so that they enjoy relatively large freedoms or room for discretionary actions. It also only applies when the principal is somewhat myopic, i.e. unaware of the consequences of some of her employees' defining characteristics. This may seem at first glance to be an unlikely scenario. In practice however, such situations are likely to commonly arise all over the globe. This logically follows from the frequency the basic setup is employed in corporations: Firms *do* have owners, CEOs or executives (principals) who are, for a variety of reasons, somewhat unaware of the daily affairs of their company. They may simply have too many employees to adequately control all of them. Even when department heads are aware of some possible problems arising from their employee's misbehavior, standard contracts dictated by headquarters may prevent the flexibility that is necessary to offer each potential agent a fitting contract.

The model also does not take into account that an important predictor of retention is employee-organization fit. If an employee has conflicts with a supervisor, this fit may be low. Laying-off an employee with poor employee-organization fit may preserve the prevalent organizational culture. An intact culture retains the organizational attractiveness for current employees as well as for fitting future hires.

This thesis makes no attempt at presenting solutions for described decision problem stemming from vertical social relations. To name just one, it may be possible for a non-socially blinkered principal to give bad agents an incentive to quit voluntarily, as in Dur & Schmittdiel (2013). While the thesis presented here answers its own research question about the effects of not correctly accounting for social relationships in a stylized framework, the domain of problems that may arise from social relationships is much larger. While it clearly shows that unchecked vertical social relations present

a problem in a stylized workplace, there are numerous facets of social relations it does not explore. Horizontal relations are also of great importance. Social relations can take numerous forms, including nepotism and romantic entanglements, which are not exactly the same as altruism or spite but may influence behavior in other ways too. This is, however, a different research avenue. Nevertheless, especially given the rise of social networking, worker communication outside the workplace may prove to be a driving factor for both productivity and social relations in the workplace. These research avenues remain as of yet still to be explored.

The extensions presented in this thesis are also somewhat limited in that they also assume an imperfect principal, and in addition, a detection blinkered agent. Bargaining between agent and supervisor may not necessarily need to be ultimatum bargaining, and a thieving agent can likely try to hide his theft from his supervisor. The limitations of the basic model also apply here. However, the interesting result that negative social relations could be used —if the principal was not socially blinkered —to prevent employee theft presents a starting point for further research into the matter. Auditor rotation, for example, is set up to keep relations between audited and auditor relatively neutral to prevent fraud and collusion. It may then be that negative relations may be more successful at achieving this goal, notwithstanding the inordinate amount of other challenges that may then arise.

However, given the level of internal validity of the model, the main limitation of this thesis is that it is not possible without empirical work to find whether its implications are of any relevance in real workplaces. This relevance is assumed and supported by logic, yet empirics need to ultimately support it as well. The absence of natural and experimental data is explained through the nature of the subject matter: Social relations can lead to decisions that are directly in contrast with the interests of the firm. Firms obviously have an interest in participating in such research, for it may inform future personnel decisions. Especially when these decisions are to commit crimes or firable offenses, it is understandable that few employees are willing to communicate their attitudes or past behavior truthfully if they have committed such offenses. This holds especially if there is a real or suspected chance that truthful answers could be used against them, i.e. lead to their firing. Such a declaration-action-gap would have skewed any data that could have been collected in cooperation with firms. Cooperation with firms would have been crucial given the time and financial constraints of this thesis. Outside of existing firms, it would not have been possible to find enough reliable survey respondents that were working in relevant positions without using firms as subject pool suppliers, i.e. finding respondents independently of where they work. Although 17 companies with more than 500 employees in four countries (one French, four Dutch, six German and six U.S.-based firms) responded in a timely manner when contacted about their willingness to let employees in relevant positions participate in a survey,

the author decided against using survey data. Ten firms declined outright. Those that did not decline agreed to participate only under conditions. The author decided that these conditions (using identifiers in the survey that could have led the employers in question to the identity of the employee, using some form of supervision during survey completion or requiring full access to raw data) would not have met ethical standards if undisclosed to participants. If disclosed, employees may have had an incentive to under-report their own behavior. If asked about other's behavior that they have observed, there could not have been a control for distorted answers either. Answers could have been distorted in order not to implicate colleagues or not to admit to not having reported misbehavior. While a modification of a bayesian truth serum (cf. Prelec, 2004; Weaver & Prelec, 2013) may alleviate some of these issues, finding valid data remains a challenge to research into social relations in the workplace.

This leaves, for the near future, experimental data. Two possible experimental designs that could be employed to test the propositions brought forward in this thesis are included in appendix A.3. Financial, time and organizational constraints precluded such experiments from being carried out. With these designs, it should however prove to be relatively easy to do the experiments when such constraints do not apply.

Another possible limitation of the analysis conducted in this thesis is its simplicity. However, the model benefits from its simplicity and requires no sophisticated mathematics to solve. This thesis reaches its results by stripping down a moral hazard problem in a simplified work environment to an even simpler decision problem. Using more complex mathematical analysis would be inefficient, for the same results would be attained in a more difficult way as long as the underlying assumptions are kept the same.

7 Discussion & Conclusions

The goal of this thesis was to answer its own research question on the consequences of not adequately taking into account the real incentive distortions that may happen through social relations. Using a model of a stylized workplace, the monitor's decision problem was described and analyzed. It was found that social relations can indeed distort a supervisor's incentives and lead him to cause the firing of good agents and the retaining of bad agents. This has negative expected utility effects for a socially blinkered principal. While the effect of neutral relations —or the absence of social relations at all— has the same effect on the monitor's decisions as having positive relations with a high-type worker or negative relations with a low-type worker, negative relations with a high-type worker and positive relations with a low-type worker can harm the principal. In the extensions, it was found that if the principal has a random chance of finding out whether the monitor truthfully evaluated the agent, he can inadvertently

make the truthtelling constraints hold. When combined with a thieving agent, this random chance, when information about it is asymmetric, it can curb collusion between monitor and agent. If it is not enough to curb collusion between friends, it may be preferable for the principal that his employees dislike each other. This implies that when collusion is possible, negative social relations can have positive effects on the principal. Though in both the basic model and the extensions, the socially blinkered principal does not know that he may be able to take advantage of social relations or that he is being taken advantage of. The six hypotheses derived from the results of the basic model and its extensions can be empirically tested in future research.

The approach of this thesis can surely be considered to be unconventional. In a more classical approach, a non-blinkered, fully rational principal would analyze the situation at hand, see its potential pitfalls and redesign the incentive structure to prevent them. This approach, however, is blinkered in itself because it does not take into account cognitive limitations. Human beings have cognitive limits. They make mistakes. They do not consider every possible outcome of every possible action, consistently make errors of judgment or simply forget important information. Real humans in real environments are too often better described by the boundedly rational, satisficing individuals of Simon (1957) than by homo oeconomicus. Their utility often corresponds to prospect theory (Kahneman & Tversky, 1979) more than to expected utility theory.

Theoretical work under the assumption of clear, ideal circumstances certainly is of immense value to academia and practice. Yet, often the results of published papers focus on how to avoid undesirable situations. It is conceded that such situations exist and that not every solution is first-best. These situations are assumed not to occur or be preventable due to the intervention of some economic agents. This is not always helpful or informative. Given that the undesirable has happened, it is necessary to know how to best deal with the situation one *is* in. Anything else is sunk. Knowing what situation one *should* be in does not inform problem solving. It may just inspire feelings of regret, which are suboptimal in themselves, for they distract from solving the problem at hand. This thesis is limited in that it does not inform problem solving directly. It, as a first step, recognizes that the problems that arise if the monitor is not truthful can indeed occur. The solving of this fundamental problem, for example through targeted "anti-teambuilding" activities (which may backfire in public relation terms) or paying bad agents to quit voluntarily shall be left to others. It should be clear though that rotation (of auditors or supervisors) cannot fully curb the moral hazard problem that arises from social relations between employees. It is difficult to predict the precise nature and extent of a social relation that will arise between two individuals. It may then happen that the real incentive distortions —that rotation is set up to prevent —occur before the next rotation phase.

The principal is unaware of the potential negative consequences of his employees'

social relations. These social relations can influence decisions in a direction that is not desirable from the principal's perspective. He is worse off, and since he trusts the monitor, never realizes that he is being taken advantage of. In the spirit of Akerlof (1970), the presence of bad A(pples) in the agent population in combination with monitor preferences towards some agents but not others leads to a selection problem where those agents that the principal prefers may not be employed in the firm. This occurs when social relations are pronounced enough to make the truthtelling constraint of the monitor not hold. Those agents who are good workers may be fired while those agents who are bad workers may be retained.

As the principal in this production process is inactive to the extent that his involvement is irrelevant to the outcome, a logical question to ask would be why the agent is monitored at all. Lee & Persson (2011) ask the same question in the full-loyalty version of their principal-supervisor-agent model. Their answer is that a monitor that is fully loyal to the principal can be seen as the same person as the principal. Then, everything that applies to the principal applies to the monitor as well and their incentives are aligned. This logic could be applied here as well. But if one unites principal and monitor into the same person, a principal who monitors, then the moral hazard problem goes away. The principal-monitor person may still have incentives to keep on a bad worker he is friends with or fire a good worker he dislikes. But then his decision is led by his own utility, which may suffer on a monetary level and gain on a social level and vice versa. His decisions are then made at his own cost to his own benefit. With a separation between monitor and principal, the monitor's decisions are made to his own benefit at the principal's cost. This internally arising externality is the focus, so it makes little sense to define the principal away. The principal inadvertently and without intention (through his blinkeredness) makes the privately suboptimal decision to finance his employee's optimal decisions.

It should be clear though that these results do not imply that vertical social relations in the workplace should be avoided or are bad in general. It is possible —to some extent— to predict friendships based on some shared interests or characteristics. But it is very difficult and may even be impossible to successfully screen for social relations. Yet taking social relations in either direction into account can help productivity or prevent problems before they occur. When there are synergies to be gained from fruitful cooperation between supervisors and workers, social relations can be of great benefit to a firm. Negative social relations can have positive effects when mistakes are very costly and scrutiny inspired by spite helps catch these mistakes. The results imply merely that not adequately taking into account social relations can have negative consequences for the firm.

This may be especially relevant to new (small) businesses, whose founders are often somewhat naive, as the historically high failure rate of new businesses attests (Lane

& Schary, 1991). While most people are aware of the possibility and the effects of friendships, one can confidently assume that human beings in general, not just founders of new businesses, are in some ways cognitively limited. This could apply, for instance, in the "unawareness" sense of Zhao (2008), by being socially blinkered or by being boundedly rational in the sense of Simon (1972). In either case, new ventures — large or small — face an inordinate amount of internal and external challenges. They are also bound to undergo a process of trial and error before achieving some measure of success. During such a process, it may be easy to underestimate the internal challenge of aligning the social goals of employees with the goals of the venture. New ventures therefore may be among those who stand the most to gain from implementing the results attained here. This holds, in this author's belief, no matter whether they are aware of social relations and do not take them into account adequately or whether they are unaware of their existence.

Empirical support for the hypotheses derived here would have been desirable. It was not feasible however to do this with a high level of confidence in the validity of the data. This thesis therefore ends with a call for further empirical research. As the theoretical results clearly show, if only a few (realistic) assumptions hold, not adequately dealing with social relations between employees can have negative consequences for firms. Extrapolating the results of this thesis to relevant workplaces, which likely exist all over the world, it becomes clear that immense inefficiencies exist. These inefficiencies directly translate to large amounts of money lost unnecessarily. These losses arise from social situations in firms and could be better spent on solving (social) problems elsewhere.

A Appendix

A.1 Optimality of Trusting the Supervisor in the Absence of Social Relations

As before, M is the only player (except for nature) in the "game." M has complete information about the decision problem. It is assumed that participation constraints for A, B, M and P are met, i.e. wages are greater or equal to outside options and the expected error costs to P are not large enough to make non-production a better option than production. Proving that trusting the monitor is optimal for P if $\alpha = 0$ is trivial and follows the same steps as the analysis in section 4. Intermediary steps are omitted for brevity. Recall the notational convention of θ_i referring to the agents true type, while subscripts *truth* or *lie* refer to P's holding of correct or wrong beliefs, respectively.

Proof of Proposition 1. P wants M to tell the truth, no matter what type A has or whether a mistake actually occurs. For this to be the case, M's expected utility from telling the truth must be larger than her expected utility from lying. If $\theta_i = \theta_H$, A does not get replaced by B. Recall (8) and (7). M's expected utility over two periods from telling the truth in this situation is:

$$\begin{aligned} E(U_M|\theta_i = \theta_H)_{truth} = p_H\left((1-p_H)(c+w_M) + p_H w_M + w_M\right) \\ + (1-p_H)\left((1-p_H)(c+w_M) + c + p_H w_M + w_M\right) \\ = 2\left(c - cp_H + w_M\right). \end{aligned} \tag{55}$$

Recall (1) and that B is subject to it. B replaces A if A gets fired. M's expected utility from lying therefore is:

$$\begin{aligned} &E(U_M|\theta_i = \theta_H)_{lie} = \\ &p_H\left(\frac{1}{2}\left((1-p_H)(c+w_M) + p_H w_M\right) + \frac{1}{2}\left((1-p_L)(c+w_M) + p_L w_M\right) + w_M\right) \\ &+(1-p_H)\left(\frac{1}{2}\left((1-p_H)(c+w_M) + p_H w_M\right) + \frac{1}{2}\left((1-p_L)(c+w_M) + p_L w_M\right) + c + w_M\right) \\ &= \frac{1}{2}\left(4(c+w_M) - 3cp_H - cp_L\right). \end{aligned} \tag{56}$$

Subtracting (56) from (55) gives

$$E(U_M|\theta_i = \theta_H)_{truth} - E(U_M|\theta_i = \theta_H)_{lie} = \frac{1}{2}c\left(p_L - p_H\right). \tag{57}$$

Since per assumption $c > 0$ and $p_L > p_H$, it always holds that (57)> 0. The same

logic applies if $\theta_i = \theta_L$. Here, when telling the truth, M works with B in period two. M's expected utility from telling the truth is:

$$\begin{aligned} &E(U_M|\theta_i = \theta_L)_{truth} = \\ &(1-p_L)\left(\frac{1}{2}\left((1-p_H)(c+w_M)+p_H w_M\right)+\frac{1}{2}\left((1-p_L)(c+w_M)+p_L w_M\right)+c+w_M\right) \\ &\quad + p_L\left(\frac{1}{2}\left((1-p_H)(c+w_M)+p_H w_M\right)+\frac{1}{2}\left((1-p_L)(c+w_M)+p_L w_M\right)+w_M\right) \\ &\qquad = \frac{1}{2}\left(4(c+w_M)-cp_H-3cp_L\right). \end{aligned} \tag{58}$$

If M lies and A is retained, M's expected utility in this situation is:

$$\begin{aligned} &E(U_M|\theta_i = \theta_L)_{lie} = \\ &\quad (1-p_L)\left((1-p_L)(c+w_M)+c+p_L w_M+w_M\right) \\ &\quad + p_L\left((1-p_L)(c+w_M)+p_L w_M+w_M\right) \\ &\qquad = 2c+2w_M-2cp_L. \end{aligned} \tag{59}$$

Again, subtracting (59) from (58) gives

$$E(U_M|\theta_i = \theta_L)_{truth} - E(U_M|\theta_i = \theta_L)_{lie} = \frac{1}{2}c(p_L - p_H) \tag{60}$$

which is the same as (57). The truthtelling constraint for M is the same for both possible types of A. M never lies. It is therefore optimal for P to believe M's message. □

A.2 Comparative Statics for the Agents

A.2.1 The Case of a High-Type Foe

Assume here that M's truthtelling constraint holds. For A, expected utility in period one

$$E(U_{A1}|\theta_i = \theta_H)_{truth} = w_A - p_H\gamma_A c + \alpha(w_M + (1-p_H)c) \tag{61}$$

is equal to expected utility in $t = 2$ and therefore,

$$E(U_{Atotal}|\theta_i = \theta_H)_{truth} = 2(w_A - p_H\gamma_A c + \alpha(w_M + (1-p_H)c)). \tag{62}$$

B receives her outside option in both periods, i.e. zero, as she does not get hired.

Assume now that M's truthtelling constraint does not hold.

A receives

$$E(U_{A1}|\theta_i = \theta_H)_{lie} = -p_H\gamma_A c + \alpha(w_M + (1 - p_H)c) \tag{63}$$

in period one and zero in period two.

A.2.2 The Case of a Low-Type Friend

Assume here that M's truthtelling constraint holds.

Since A gets fired in period one, her second period utility is her outside option of zero. She also gets no wage for period one, but still suffers eventual guilt from having made a mistake. Her expected utility is also influenced by her social relation with M:

$$E(U_{A1}|\theta_i = \theta_L)_{truth} = -p_L\gamma_A c + \alpha(w_M + (1 - p_L)c) \tag{64}$$

which is equal to her total expected utility. Again, she is better off with positive social relations, regardless of whether she makes a mistake or not. If $\alpha = 0$, her total utility is zero. For B it is clear that she is better off when A gets fired, as she receives a wage for period two. B's utility is independent of M's decision to lie or tell the truth but depends only on the outcome of being hired.

$$E(U_{B2}) = \begin{cases} w_B - p_L\gamma_B c \text{ or } w_B - p_H\gamma_B c & \text{if } B \text{ is hired} \\ 0 & \text{if } B \text{ is not hired} \end{cases}. \tag{65}$$

It is obvious that B prefers being hired to not being hired.

Assume now that M's truthtelling constraint does not hold.

A then receives

$$E(U_{A1}|\theta_i = \theta_L)_{lie} = w_A - p_L\gamma_A c + \alpha(w_M + (1 - p_L)c) \tag{66}$$

in period one and

$$E(U_{A2}|\theta_i = \theta_L)_{lie} = w_A - p_L\gamma_A c + \alpha(w_M + (1 - p_L)c) \tag{67}$$

in period two, resulting in a total expected utility of

$$E(U_{Atotal}|\theta_i = \theta_L)_{lie} = 2(w_A - p_L\gamma_A c + \alpha(w_M + (1 - p_L)c)). \tag{68}$$

A.3 Possible Experiments

It was not possible within the scope of this thesis —keeping financial and time constraints of the author in mind— to perform experimental tests of the the first three hypotheses derived from the model and its first extension. Two possible experimental

setups are therefore described here. Since the goal of these experiments is to elicit dishonest behavior in subjects, the "no deception" precept that is commonly held in experimental economics (cf. Smith, 1976, 1989) is relaxed. Both experimental designs below are 2×3 between-subject designs and are meant to use university students as subjects. The designs differ only in the process that induces social relations and are otherwise identical. The descriptions below take knowledge about the model and its extensions as given to avoid repetition. Both designs can be used to test hypotheses 1, 2 and 3. Testing the other hypotheses would require asking subjects to engage in hypothetical illegal behavior and would therefore be unethical. Hypotheses 4, 5 and 6 need therefore be tested on naturally occurring data.

A.3.1 Possible Experiment 1

Consider the following setup for a lab experiment. In the first phase, subjects individually play a Nintendo Wii game with or against an actor, whom the experimenters introduce as another subject. The use of a Nintendo Wii as a tool to induce social relations has several advantages: First, the controls are intuitive can likely be learned quickly by all non-physically impaired subjects. Second, the number of competitive games as well as cooperative games on the Wii is extensive. Pretesting can show which games are most suitable to inducing which social relation. Third, the games can often easily be modified to be biased towards or against the subjects, by, for example, giving a clearly superior team to one player if, say, a soccer game is used. However, a game that depends on player skill rather than pure luck —with the exception of the possible bias— should be used to grant a level of control.

The actor's role is to induce a social relation with the subject. Deceptively using an actor to endear or antagonize subjects has several advantages: First, it reduces the number of subjects required and thereby reduces the costs of the experiment. Second, it introduces a measure of control into the social relation that arises between actor and subject. Also, the actor has an advantage over the subject, for he or she knows the game more in-depth than the subject does. The actor can taunt or insult, compliment or help the subject —in addition to losing or winning the game intentionally— in order to cause the subject to have positive or negative feelings towards him or her. A third benefit is that it allows focus on just the cases of interest, i.e. good relations with bad workers and bad relations with good workers, if that proves necessary after pretesting because other cases can then be excluded intentionally.

In the second phase, the actor and subject engage in a production process as described in this thesis. The actor takes the role of the agent, while the subject takes the role of the supervisor. The subject is separated from the actor and informed by an experimenter about her role as the monitor; her reporting duties to some unseen principal, that the actor has been assigned the role of either a good or a bad worker,

the external probabilities of making a mistake that destroys part of joint output for each type etc. Each worker type is identified by wearing a different color cap, say, white for high and black for low, and the subject is told about this distinction.[39] Actor and subject are then brought back together so that they can see each other but not communicate, with the actor wearing a colored cap.

For each period, subjects are given a very simple task that can be easily completed in the given time in order to control for any effects effort may have. This should ideally be a computer task such as clicking the mouse 100 times within three minutes. When production is completed, the subject is asked to communicate to the unseen principal what type the actor has. Then, at the end of the period, a random number generator determines whether the actor has made a mistake that costs the subject his or her bonus. This makes the evaluation decision independent of feelings of regret or anger. Depending on the condition the subject is in and on the evaluation he or she gives, the actor is replaced by another actor. The new actor's cap color is determined by a coin flip. The second stage task is then identical to the first period task without the evaluation stage. Depending on the outcomes of the twice-employed random number generator, the subject then receives her base reward plus eventual bonuses.

In the control group, subjects play the Wii game by themselves and engage in the process without being introduced to the actor. With three possible social relations and two possible worker types, this gives a total of six conditions for the experiment. These are displayed in table 4. The dependent variable of in each case is the proportion of subjects that report truthfully.

Worker Type/ Setting	$\alpha > 0$	$\alpha = 0$	$\alpha < 0$
$\theta_i = \theta_L$	Treatment 1	Control 1	Treatment 3
$\theta_i = \theta_H$	Treatment 2	Control 2	Treatment 4

Table 4: Expected Utility Comparison for M's Choices

The dependent variable in this setup is the proportion of liars in a treatment. The hypotheses can be tested using Mann-Whitney U tests by comparing the proportion of liars a treatment with the proportion of liars in the relevant control. This basic setup so far tests hypotheses 1 and 2. To test hypothesis 3, the production process can be easily modified to include some probability of lie detection. This would then make it a 4×3 design. Pretesting should determine whether the cases of negative relations with low type workers and positive relations with high type workers yield the expected results (no lying). If this is the case, the original design becomes a 2×2 design.

[39] Ideally, this color distinction should be randomized per subject to allow control for priming effects.

A priori power calculation using GPower with parameters $(1 - \beta) = 0.8$ (power, not bonus), $\alpha = 0.05$ (significance level, not social relation), $D = 0.5$ (effect size) for the experimental design described above for a Mann-Whitney U test yields a sample size of 51 per treatment, so for the entire experiment of 204 (controls plus positive relations low type worker condition and negative relations high type condition), 306 (all conditions with no probability of detection).

While phase two is relatively straightforward and replicable, phase one is crucially important and may be subject to uncontrollable influences. Generally, there is no solidly established, generally accepted way of inducing social relations in a laboratory. One recent approach is described below.

A.3.2 Possible Experiment 2

This experiment is identical to the above in phase two. The elicitation process for social relations in phase one however differs.

Harris et al. (2009) study the effect of group size on in-group favoritism, which is a form of (positive) social relations. In what they call a "triadic game," three players play an allocation game where two of the players are told that they belong in the same group while the other is a single player. This basic setup can be copied. One, two, or no actors may be used, with two or one or three subjects in each group then.

If only one actor is used, he or she shall be the single player and not be generous towards the two subjects. This can then inspire group identity, i.e. a positive social relation between the two subjects. The subjects can then either engage in the production process together to study positive social relations (then, replace the word "actor" in the above elaborated phase two, while the rest is constant), or engage in the production process with the actor, to study negative social relations. If implemented well, this will likely lead to the desired effects but leads to lop-sided treatment sizes.

If two actors are used, the subject can be the single player and then produce with one of the actors. This should lead to negative relations. If the actor is part of the in-group, the actor can engage with the subject in phase two as before.

If no actors are used, there is little control about the social relations that arise from this, but the social relations could be inferred from the allocation decisions subjects make (in a repeated game). Phase two can then proceed with any combination of the two, with one of them potentially taking the role of B. This would go further than testing the hypotheses derived in this thesis and on to studying the possible situation of having the opportunity to fire someone who is disliked to replace him or her with someone who is liked or vice versa.

There are many other ways that social relations could be induced. Whichever one chooses, as long as the process is reliably reproducable, phase two can be employed to study the effects of social relations in a laboratory production task.

References

Agell, J. (2004). Why are small firms different? Managers' views. *The Scandinavian Journal of Economics*, *106*(3), 437–452.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, *84*(3), 488–500.

Bandiera, O., Barankay, I., & Rasul, I. (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica*, *77*(4), 1047–1094.

Bandiera, O., Barankay, I., & Rasul, I. (2010). Social incentives in the workplace. *The Review of Economic Studies*, *77*(2), 417–458.

Becker, G. S. (1974). A theory of social interactions. *The Journal of Political Economy*, *82*(6), 1063–1093.

Bénabou, R. (2013). Groupthink: Collective delusions in organizations and markets. *The Review of Economic Studies*, *80*(2), 429–462.

Bergstrom, T. C. (1999). Systems of benevolent utility functions. *Journal of Public Economic Theory*, *1*(1), 71–100.

Berman, E. M., West, J. P., & Richter Jr, M. N. (2002). Workplace relations: Friendship patterns and consequences (according to managers). *Public Administration Review*, *62*(2), 217–230.

Bol, J. C. (2011). The determinants and performance effects of managers' performance evaluation biases. *The Accounting Review*, *86*(5), 1549–1575.

Bowles, S., & Polania-Reyes, S. (2012). Economic incentives and social preferences: substitutes or complements? *Journal of Economic Literature*, *50*(2), 368–425.

Bruynseels, L., & Cardinaels, E. (2013). The audit committee: Management watchdog or personal friend of the ceo? *The Accounting Review*, *89*(1), 113–145.

Chidambaran, N., Kedia, S., & Prabhala, N. (2011). Ceo director connections and corporate fraud. *Fordham University Schools of Business Research Paper 1787500*.

Cooper, D., & Kagel, J. H. (2013). Other regarding preferences: a selective survey of experimental results. In *Handbook of experimental economics* (Vol. 2). Princeton University Press.

Cooter, R. D., & Garupa, N. (2000). The virtuous circle of distrust: A mechanism to deter bribes and other cooperative crimes. *BLE Working Paper Series*.

Dey, A., & Liu, X. (2010). Social connections, stock-based compensation, and director oversight. *Available at SSRN 1581212*.

Dickinson, D., & Villeval, M.-C. (2008). Does monitoring decrease work effort?: The complementarity between agency and crowding-out theories. *Games and Economic Behavior*, *63*(1), 56–76.

Dur, R., & Schmittdiel, H. (2013). Paid to quit. *Tinbergen Institute Discussion Paper 2013-174/7*.

Dur, R., & Tichem, J. (2014). Altruism and relational incentives in the workplace. *Journal of Economics and Management Strategy*, forthcoming.

Ellingsen, T., Johannesson, M., Tjøtta, S., & Torsvik, G. (2010). Testing guilt aversion. *Games and Economic Behavior*, *68*(1), 95–107.

Enzle, M. E., & Anderson, S. C. (1993). Surveillant intentions and intrinsic motivation. *Journal of Personality and Social Psychology*, *64*(2), 257.

Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, *96*(5), 1611–1630.

Fehr, E., & Fischbacher, U. (2002). Why social preferences matter–the impact of non-selfish motives on competition, cooperation and incentives. *The Economic Journal*, *112*(478), C1–C33.

Fehr, E., Gächter, S., & Kirchsteiger, G. (1997). Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica*, *65*(4), 833–860.

Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism–experimental evidence and new theories. *Handbook of the Economics of Giving, Altruism and Reciprocity*, *1*, 615–691.

Fracassi, C., & Tate, G. (2012). External networking and internal firm governance. *The Journal of Finance*, *67*(1), 153–194.

Galanis, S. (2013). Unawareness of theorems. *Economic Theory*, *52*(1), 41–73.

Giebe, T., & Gürtler, O. (2012). Optimal contracts for lenient supervisors. *Journal of Economic Behavior & Organization*, *81*(2), 403–420.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, *24*(3), 153–172.

Gómez, C., Kirkman, B. L., & Shapiro, D. L. (2000). The impact of collectivism and in-group out-group membership on the evaluation generosity of team members. *Academy of Management Journal*, *43*(6), 1097–1106.

Grund, C., & Przemeck, J. (2012). Subjective performance appraisal and inequality aversion. *Applied Economics*, *44*(17), 2149–2155.

Güth, W., & Tietz, R. (1990). Ultimatum bargaining behavior: A survey and comparison of experimental results. *Journal of Economic Psychology*, *11*(3), 417–449.

Handgraaf, M. J., Van Dijk, E., & De Cremer, D. (2003). Social utility in ultimatum bargaining. *Social Justice Research*, *16*(3), 263–283.

Harris, D., Herrmann, B., & Kontoleon, A. (2009). Two's company, three's a group: The impact of group identity and group size on in-group favouritism. *CeDEx Discussion Paper Series No. 2009-13*.

Harsanyi, J. C. (1962). Measurement of social power, opportunity costs, and the theory of two-person bargaining games. *Behavioral Science*, *7*(1), 67–80.

Hoffman, E., McCabe, K., & Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review*, 653–660.

Hwang, B.-H., & Kim, S. (2009). It pays to have friends. *Journal of Financial Economics*, *93*(1), 138–158.

Itoh, H. (1991). Social relations and incentive contracts. *Kyoto University Economic Review*, *61*(1), 35–55.

Jackson, C. K., & Schneider, H. S. (2011). Do social connections reduce moral hazard? Evidence from the new york city taxi industry. *American Economic Journal: Applied Economics*, *3*(3), 244–67.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Kartik, N. (2009). Strategic communication with lying costs. *The Review of Economic Studies*, *76*(4), 1359–1395.

Kofman, F., & Lawarree, J. (1996). A prisoner's dilemma model of collusion deterrence. *Journal of Public Economics*, *59*(1), 117–136.

Laffont, J.-J., & Martimort, D. (1998). Collusion and delegation. *The Rand Journal of Economics*, 280–305.

Laffont, J.-J., & Rochet, J.-C. (1997). Collusion in organizations. *The Scandinavian Journal of Economics*, *99*(4), 485–495.

Lane, S. J., & Schary, M. (1991). Understanding the business failure rate. *Contemporary Economic Policy*, *9*(4), 93–105.

Lee, S., & Persson, P. (2011). Authority versus loyalty: Social incentives and modes of governance. *NYU Working Paper No. FIN-10-001*.

Mlodinow, L. (2008). *The drunkard's walk: How randomness rules our lives*. New York: Random House.

Moerbeek, H. H., & Need, A. (2003). Enemies at work: Can they hinder your career? *Social Networks*, *25*(1), 67–82.

Morrison, R. L. (2008). Negative relationships in the workplace: Associations with organisational commitment, cohesion, job satisfaction and intention to turnover. *Journal of Management and Organization*, *14*(4), 330.

Morrison, R. L., & Nolan, T. (2007). Too much of a good thing?: Difficulties with workplace friendships. *University of Auckland Business Review*, *9*(2), 32.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *The Quarterly Journal of Economics*, *122*(2), 569–600.

Polti, G. (1921). *The thirty-six dramatic situations*. Franklin, Ohio: JK Reeve.

Prelec, D. (2004). A bayesian truth serum for subjective data. *Science*, *306*(5695), 462–466.

Prendergast, C., & Topel, R. H. (1996). Favoritism in organizations. *The Journal of Political Economy*, *104*(5), 958–978.

Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature*, *36*(1), 11–46.

Rasul, I., & Sonderegger, S. (2010). The role of the agent's outside options in principal–agent relationships. *Games and Economic Behavior*, *68*(2), 781–788.

Rotemberg, J. J. (1994). Human relations in the workplace. *Journal of Political Economy*, *102*(4), 684-712.

Rotemberg, J. J. (2006). Altruism, reciprocity and cooperation in the workplace. *Handbook of the Economics of Giving, Altruism and Reciprocity*, *2*, 1371–1407.

Salas, E., Rozell, D., Mullen, B., & Driskell, J. E. (1999). The effect of team building on performance: An integration. *Small Group Research*, *30*(3), 309–329.

Schumpeter, J. (1942). *Capitalism, socialism and democracy.* London: Routledge.

Shleifer, A., & Vishny, R. W. (1989). Management entrenchment: The case of manager-specific investments. *Journal of Financial Economics*, *25*(1), 123–139.

Simon, H. A. (1957). *Models of man; social and rational.* New York: Wiley.

Simon, H. A. (1972). Theories of bounded rationality. *Decision and Organization*, *1*(1), 161–176.

Smith, V. L. (1976). Experimental economics: Induced value theory. *American Economic Review*, *66*(2), 274–79.

Smith, V. L. (1989). Theory, experiment and economics. *Journal of Economic Perspectives*, *3*(1), 151–69.

Sol, J. (2010). Peer evaluation: Incentives and co-worker relations. *Tinbergen Institute Discussion Paper 10-055/1*.

Spagnolo, G. (1999). Social relations and cooperation in organizations. *Journal of Economic Behavior & Organization*, *38*(1), 1–25.

Thiele, V. (2013). Subjective performance evaluations, collusion, and organizational design. *Journal of Law, Economics, and Organization*, *29*(1), 35–59.

Tichem, J. (2013). Leniency bias in long-term workplace relationships. *Tinbergen Institute Discussion Paper 13-196/7*.

Von Thadden, E.-L., & Zhao, X. (2012). Incentives for unaware agents. *The Review of Economic Studies*, 1-24.

Weaver, R., & Prelec, D. (2013). Creating truth-telling incentives with the bayesian truth serum. *Journal of Marketing Research*, *50*(3), 289–302.

Yonker, S. (2013). Do local managers give labor an edge? *US Census Bureau Center for Economic Studies Paper No. CES-WP-13-16*.

Zhao, X. (2008). Moral hazard with unawareness. *Rationality and Society*, *20*(4), 471–496.