Screening Ethics when Honest Agents Care about Fairness

Ingela Alger* and **Régis Renault****

First version: January 2001
This version: August 2004

Abstract

We analyze a principal's ability to discriminate between honest and dishonest agents, who have private information about the circumstances of the exchange. Honest agents reveal circumstances truthfully as long as the mechanism is sufficiently fair: the probability that an equilibrium allocation is chosen by an agent who is lying should not be too large. Without intolerance for lying the agent is given proper incentives if dishonest and zero rent if honest. With even a small intolerance for lying the optimal mechanism is discontinuously altered. It may still involve ethics screening whereby some allocation chosen by a dishonest agent is never chosen by an honest agent. It happens either when the dishonest is overstating circumstances, or when the principal is forced to allow for some suboptimal announcements due to an excessive intolerance for lying. With limited intolerance for lying, ethics screening allows for doing better than in the standard setup where the agent is dishonest with certainty, even if honesty is unlikely. However, if intolerance for lying is too strong, the principal cannot improve upon the standard setup.

Keywords: ethics, honesty, adverse selection, screening.

JEL classifications: D82

*Economics Department, Boston College, Chestnut Hill, MA 02467. E-mail: ingela.alger@bc.edu

**Université de Cergy-Pontoise, ThEMA. E-mail: regis.renault@eco.u-cergy.fr

Acknowledgement: We are especially thankful to one referee, whose remarks led to substantial improvements of the paper. We are also grateful to another referee, Antonio Merlo, Frédéric Koessler, and seminar participants at Université de Caen, Erasmus University Rotterdam, ESEM 99, Université de Lille 3, Université de Lyon 2, MIT, Université de Paris X, Université de Perpignan, University of St Andrews, Stockholm School of Economics, Université des Sciences Sociales de Toulouse, and University of Virginia for useful comments. The usual disclaimer applies.

1 Introduction

Adam Smith has taught us that we may trade with others without much regard for their ethics; "It is not from the benevolence of the Butcher (...) that we expect our dinner."[1] As long as the invisible hand is at work, ethics is irrelevant. However, in the extensive research devoted to the shortcomings of the invisible hand, it may no more be innocuous to postulate opportunistic economic agents, as is typically done. For instance, in the public goods provision problem, the emphasis has been on inefficiencies resulting from unrestrained opportunism. Yet there is some evidence of somewhat more scrupulous attitudes regarding public goods financing. The empirical studies on tax compliance surveyed by Andreoni, Erard and Feinstein (1998) find that a large number of taxpayers report their income truthfully, and that those who cheat do so by fairly small amounts. They further conclude that the IRS audit and penalty rates are too low to justify these findings if all taxpayers act strategically.[2] Similar conclusions have been reached in various experiments on voluntary public goods financing (see the survey by Dawes and Thaler, 1988, and the references therein). In one of the experiments, those who gave money "indicated that their motive was to 'do the right thing' irrespective of the financial payoffs." (p.194) In the context of work relations, the use of pre-employment integrity tests suggests that employers acknowledge a potential heterogeneity in ethics and find it useful to discriminate on this basis.[3] Our goal is to investigate the possibility of such screening among agents with different ethics, using standard tools of economic theory.

We consider a simple definition of an honest behavior within a simple and well-known framework. We present an extension of a one-period, adverse selection model with one principal and one agent, where the agent has private information about the circumstances

[1]Wealth of Nations, book I, Chapter II.

[2]See also Roth, Scholtz and Witte (1989) for survey evidence.

[3]See Ryan *et al.* (1997) for some references in psychology on the subject.

of the exchange, which may be more or less favorable.[4] The terms of the exchange are set within a mechanism, designed by the principal. We capture the idea of honest behavior by adding a second piece of private information, namely, the agent's ethics: for simplicity, the agent is either honest or opportunistic. Whereas an opportunistic agent is always willing to lie about circumstances if it increases his surplus in the standard exchange problem, an honest agent is not necessarily prepared to do this.[5] We take the individual's ethics as given, adopting a reduced form of a more complex model where the individual's preferences induce honest behavior. Formally we assume that an honest agent may be restricted in his announcements in a mechanism involving messages. This reflects the idea that the principal may ask certain questions such that an honest agent feels compelled to reveal circumstances in an unambiguous manner.

There are obvious benefits for the principal to try and screen on the basis of ethics. Ideally she should want to leave no rent to honest agents who need no incentive to disclose true circumstances. Ethics screening should enable her to do this while still achieving proper screening of circumstances for dishonest agents. Standard arguments show that this latter objective can only be achieved by leaving some rent to a dishonest agent. Throughout the paper we refer to this ideal outcome from the principal's viewpoint as "full ethics screening". Recently, Deneckere and Severinov (2001, 2003) have proposed a setup where full ethics screening is achieved. They assume that not only does an honest agent feel compelled to reveal true circumstances, he also feels compelled to reveal that he feels compelled to reveal his true circumstances. The dishonest agent may therefore communicate that he is dishonest in a credible way, simply by disclosing his willingness to misrepresent circumstances. As a result, the principal may put allocations involving an informational rent out of reach of the

[4]Below this will be referred to as the standard second-best problem.

[5]A possible extension of the present paper would be to adapt the framework used in the literature on costly state falsification, where misrepresenting circumstances is all the more costly that the discrepancy with the truth is large (see Lacker and Weinberg, 1989, Maggi and Rodriguez-Clare, 1989, and Crocker and Morgan, 1998). In that literature agents are homogeneous regarding falsification costs.

honest.

An honest agent would obviously consider full ethics screening as unfair since it implies that a dishonest agent in the same circumstances is treated better solely because he has been identified as being dishonest. A common theme in psychology, as for instance in the seminal study by Hartshorne and May (1928), is that honest and dishonest behavior depends more on the situations involved than on the individual's particular set of norms. A second and related observation by psychologists and sociologists is that those who engage in an unethical conduct usually resort to neutralization techniques providing a justification for a deviance from the common norm (Ryan *et al.*, 1997). In particular, agents weigh honesty against other moral values. One standard excuse for lying is that an individual is confronted with an inequitable situation.[6] For instance, Hollinger (1991), in a study of neutralization in the workplace, found that a significant predictor of deviant behavior (such as theft and counterproductive behavior) is what he calls "denial of victim," which is related to the worker's assessment of the inequity of the formal reward system: "Workers may elect to engage in unauthorized actions to redress the perceived inequities." (p. 182) It therefore seems unlikely that when confronted with full ethics screening, an honest agent would still feel a strong obligation to reveal circumstances.

By contrast to the existing literature involving ethics heterogeneity, we assume that honest behavior is conditional on the perceived equity of the proposed contract.[7] In his evaluation of a particular mechanism, an honest agent tries to assess whether the principal would take advantage of his honest behavior by openly treating a dishonest agent better. For

[6]Mueller and Wynn (2000) found that in the U.S. and in Canada, justice is the third most important workplace value, after the perceived ability to do the job, and the respect of the boss; pay came in ninth place only.

[7]Apart from Deneckere and Severinov (2001, 2003), several other authors have also analyzed principal-agent models featuring ethics heterogeneity with unconditionally honest agents: see Erard and Feinstein (1994), Kofman and Lawarrée (1996), Picard (1996), and Tirole (1992). In these papers, ethics screening is exogenously ruled out. Jaffee and Russell (1976) study the impact of honest borrowers on equilibria in credit markets. Ottaviani and Squintani (2002) have introduced unconditional honesty in a cheap talk environment.

instance, whenever full ethics screening is implemented in Deneckere and Severinov (2001, 2003), the principal openly endorses an unfair treatment of the honest agent, because any allocation that is meant for a dishonest receiving a rent, is chosen by no honest agent. In such a situation, an honest agent with fairness concerns might be willing to also misrepresent circumstances in order to rectify the perceived inequity. More generally, if the mechanism offered involves messages such that if they are announced in equilibrium it is likely that the agent is dishonest and is misrepresenting circumstances, then an honest agent may choose to give up honesty altogether. We model this conditional honesty by introducing a parameter measuring an honest agent's tolerance for misrepresenting circumstances, or lying. Loosely, it is a threshold probability of lying by others beyond which an honest agent would become opportunistic.

Here we say that there is ethics screening whenever some equilibrium allocation is not chosen by an honest agent in any circumstances. This definition allows for treating full ethics screening as an extreme case where none of the equilibrium allocations meant for dishonest agents is chosen by an honest agent under some circumstances.[8] If there is no ethics screening then it is as if the principal specifies only one allocation for each set of circumstances, as is assumed in much of the literature on ethics heterogeneity. No ethics screening does not guarantee a fully equitable outcome since it is possible that a dishonest agent earns a rent by misrepresenting circumstances; however, when he does so, there remains some uncertainty as to whether or not he is lying, so that a conditionally honest agent may still behave honestly if he has a high enough tolerance for lying.

As a benchmark we consider the limit case of unconditional honesty, which arises when there is no intolerance for lying. Given our definition of ethics screening, one would expect that a relevant distinction would be between messages that may be announced only by dishonest agents and messages that may be announced by an honest agent under some

[8]The optimal full ethics screening contract would specify the same allocation for the honest and the dishonest only in the worst circumstances, because there is no benefit in handing over a rent to the dishonest in these circumstances.

circumstances. Messages of the former type could for instance be messages that clearly identify the agent as being dishonest while an honest agent wishes not to misrepresent ethics, as in Deneckere and Severinov (2001, 2003). Intuition suggests that full ethics screening could only be achieved if there are enough such messages. However we show that as long as sufficiently many messages of the latter type may be used, it is possible to implement full ethics screening even if there are no messages that can be announced by dishonest agents alone. An implication of this is that full ethics screening may be implemented even if an honest agent does not feel compelled to reveal his ethics.

If lies are fully tolerated, then inducing lies entails no cost. By contrast, the introduction of some intolerance for lying, no matter how small, implies that all the allocations associated with messages available to an honest agent in some circumstances must yield the same surplus to the honest: if some allocation yields a strictly lower surplus, it is necessarily chosen only by some dishonest who is lying with certainty. These "acceptability constraints" would always be violated by the full ethics screening allocations: this means that the unconditional approach is not at all robust to the introduction of a slight intolerance for lying. Furthermore, these constraints imply that if an agent claims circumstances worse than his actual ones, he will be compensated as if circumstances were those he is announcing. Thus for such "downward" lies, the principal does not benefit from ethics screening and will treat all agents announcing the same circumstances in the same manner. We show that there will be some ethics screening in the optimal contract only if the dishonest is lying upwards by claiming that his circumstances are better than they are, or the principal is forced to allow for some suboptimal downward lies because of an excessive intolerance for lying.

Our model may endogenously generate no ethics screening as an optimal solution for the principal. In particular this is the case when there are only two sets of circumstances. However, our analysis also shows that it would be somewhat misleading to merely assume no ethics screening as a means of accounting for an honest agent's concern about fairness. First, if intolerance for lying is sufficiently strong, then the principal may be constrained to use a contract that would be suboptimal is she only had to satisfy a no ethics screening requirement. In particular, if the honest agent is sufficiently intolerant towards lying, then the optimal contract necessarily implements the standard second-best allocations even if honesty is sufficiently likely, a situation where the principal would find it optimal to let the dishonest lie, if she was only constrained to not using ethics screening. Thus with a strong enough intolerance for lying, the standard second-best approach is more robust than with no ethics screening exogenously imposed. Second, with more than two circumstances, ethics screening may be part of the optimal mix. The possibility to combine ethics screening with "upward" lies by a dishonest may be a very valuable option for the principal if intolerance for lying is not too strong. We find that the principal may be able to leave no rent to an honest agent even if his circumstances are not the worst and still screen circumstances as efficiently as in the standard second-best approach. This combination may dominate the standard second-best mechanism even if the probability of honesty is arbitrarily small. Without ethics screening, if honesty is too unlikely, the principal could not benefit from potentially dealing with an honest agent and would have to use the same mechanism as if the agent were opportunistic with certainty. Thus we find that with limited intolerance for lying, the standard second-best approach is less robust under conditional honesty than it would be if no ethics screening was imposed exogenously.

An alternative formulation of conditional honesty is explored in Alger and Renault (2004) where we use a two periods version of the current model, and in Alger and Ma (2003), who study optimal health insurance contracts when fraudulent insurance claims may be filed only if the physician is not honest: an agent truthfully reveals circumstances only if he feels committed to doing so (for instance, because he has signed a contract prior to learning circumstances). In the two models, there are two dates: whereas ethics is known to the agent (or to the physician) from the start, circumstances are only revealed in the second period. An honest agent reveals circumstances truthfully in the second period if a contract specifying allocations as a function of circumstances only is signed in the first period. Both papers find that no ethics screening may be optimal, but it may also be dominated, in particular when honesty is sufficiently likely, despite there being only two sets of circumstances.

The next section introduces the formal model. Section 3 is devoted to unconditional honesty, and we analyze conditional honesty in Section 4. Section 5 concludes.

2 The Model

Consider the following standard principal-agent setting. Preferences of both the principal (she) and the agent (he) depend on $y = (x, t)$, where $x \in \mathbb{R}^+$ is some decision variable, and $t \in \mathbb{R}$ is a monetary transfer from the principal to the agent. The agent's utility also depends on a parameter $\theta \in \Theta = \{\theta_i\}_{i \in I} \subset \mathbb{R}$, $I = \{1, 2, \ldots, n\}$; let α_i denote the probability that the agent's parameter is θ_i. The value of θ is a measure of how much benefit there is in contracting between the two parties. Let $\Pi(x, t) = \pi(x) - t$ be the surplus of the principal, with π strictly concave in x, and $V(x, t, \theta) = t - v(x, \theta)$ be the surplus of the agent, v being convex in x. Depending on the application, π and v are either both strictly increasing or both strictly decreasing in x. For instance, if the principal is an employer and the agent an employee, they are both increasing (and t is positive). They are on the contrary both decreasing (with t negative) if the agent is the principal's customer, x being the quantity supplied. Further assumptions ensuring existence and uniqueness of interior solutions are as follows (the prime indicates a partial derivative with respect to x): for any θ, $\pi(0) = v(0, \theta) = 0$, $\lim_{x \to 0} \pi'(x) = +\infty$ if $\pi' \geq 0$ (0 if $\pi' \leq 0$) and $\lim_{x \to +\infty} \pi'(x) - v'(x, \theta) < 0$. Finally, we assume that, for any x, $\dfrac{\partial v'(x, \theta)}{\partial \theta} < 0$, so that total surplus is increasing in θ. We will therefore say that circumstances are better, the larger is θ. We adopt the convention that $\theta_i < \theta_{i+1}$. With these assumptions, the first-best decision x_i^* under circumstances θ_i is uniquely defined by $\pi'(x_i^*) = v'(x_i^*, \theta_i)$.

The agent's ethics is denoted by k: he is dishonest ($k = d$) with probability γ, and honest ($k = h$) with probability $(1 - \gamma)$. An agent's type ω is therefore two-dimensional: $\omega = (\theta, k)$; let $\Omega = \Theta \times \{h, d\}$. The agent's ethics does not affect either party's preferences.

Only the agent knows his type. The principal acts as a Stackelberg leader and sets the terms of the transactions in a contract, or mechanism:

Definition 1 *[Mechanisms] A mechanism $M = (\mathcal{M}, g)$ defines a space $\mathcal{M}$ of messages μ, and a mapping $g : \mathcal{M} \rightarrow Y$, where $Y = \mathbb{R}^+ \times \mathbb{R}$ is the set of allocations. A mechanism is direct if $\mathcal{M} = \Omega$.*

The agent is free to accept or reject the offer. If there is no transaction, the agent's surplus is zero. If the principal could observe θ, she would therefore offer to implement the first-best allocation $y_i^* = (x_i^*, t_i^*)$ under circumstances θ_i, where $t_i^* = v(x_i^*, \theta_i)$.

Throughout the paper, a dishonest agent is assumed to have the standard opportunistic behavior, always selecting the message giving him the largest surplus. If there were only dishonest agents, the revelation principle would apply: without loss of generality, the principal could restrict her attention to direct revelation mechanisms. Standard analysis would show that the optimal mechanism implements second-best decisions for all circumstances but the best one (no distortion at the top), and leaves a rent to the agent for all circumstances but the worst one. For further use below we refer to this mechanism as the standard second-best one, and the allocation implemented under circumstances θ_i is denoted $y_i^s = (x_i^s, t_i^s)$.[9]

By contrast, an honest agent feels guilty if he misrepresents true circumstances. Here we do not make any *a priori* assumption as to the nature of messages that may be used in a mechanism. However, we assume that the principal may ask certain questions that would be somehow related to true circumstances, and to which an honest agent would feel obligated to provide only answers that are coherent with his private information. Formally, given a mechanism M honesty is defined by imposing restrictions on the set of messages available to the agent. These restrictions may be affected by changes in the set of messages proposed by the principal. For instance an honest agent may only feel restricted in his announcements

[9] For a detailed description of the standard second-best framework, see, e.g., Laffont and Martimort (2003).

if the mechanism is direct, or if he is asked to announce circumstances along with some other piece of information. Letting $\mathcal{R}_i(\mathcal{M}) \neq \emptyset$ denote the set of messages available to an honest agent under circumstances θ_i in mechanism $M = (\mathcal{M}, g)$, for any message space $\mathcal{M}$ either $\mathcal{R}_i(\mathcal{M}) = \mathcal{M}$ for all i, or $\mathcal{R}_i(\mathcal{M}) \cap \mathcal{R}_j(\mathcal{M}) = \emptyset$ for all i, j, $i \neq j$. We denote $\mathcal{M}_0$ any message space for which the honest agent is restricted in his announcements.[10] If the agent were honest with certainty the principal would achieve full revelation of θ_i at no cost, i.e., she could implement the first-best allocations by proposing a mechanism specifying some message space $\mathcal{M}_0$. In the following analysis the principal faces uncertainty concerning the agent's ethics: $\gamma \in (0, 1)$.

Our goal is to investigate to what extent the principal will be able or willing to screen on the basis of ethics. In order to achieve such a screening, she needs to induce agents to use a broad enough variety of messages in equilibrium. From our definition of honesty, with a message space $\mathcal{M}_0$ that induces restrictions for an honest agent, for each set of circumstances θ_i there is one message in $\mathcal{R}_i(\mathcal{M}_0)$ associated with these circumstances, that is announced in equilibrium by an honest in circumstances θ_i. If only these messages are used in equilibrium, a dishonest agent mimics an honest agent with circumstances either identical or different from his own. Such an outcome could be achieved by designing a mechanism where screening pertains to circumstances alone, as is the case in the early literature on ethics heterogeneity.[11] Ethics screening thus requires that the principal induces dishonest agents to select other messages, so that we define ethics screening as follows.

Definition 2 *There is ethics screening whenever equilibrium announcements involve more than one message in $\mathcal{R}_i(\mathcal{M}_0)$ for some i, or some messages that belong to no $\mathcal{R}_i(\mathcal{M}_0)$ for any i.*

[10] If there are some messages in $\mathcal{M}_0$ that may be announced by no honest agent, this could be because they are unrelated to circumstances or because, as in the next section, even though they contain truthful information about circumstances, they contain misleading information in some other dimension.

[11] In Erard and Feinstein (1998), Kofman and Lawarrée (1996), Picard (1996), and Tirole (1992), the message space is the set of circumstances so that ethics screening is exogenously ruled out.

Because honesty imposes restrictions on an agent's announcements, it is necessarily conditioned on the message space specified in the proposed mechanism. As we argue below, in order to properly account for equity motives on the part of honest agents, it is appropriate to also condition an honest behavior on the allocation rule g specified in the proposed mechanism. We therefore distinguish between unconditional honesty where an honest behavior does not depend on the allocation rule, from conditional honesty where honest agents may feel justified in behaving opportunistically whenever the allocation rule leads to an inequitable equilibrium outcome. We first consider unconditional honesty as a benchmark.

3 Unconditional honesty

In our formal definition an honest agent feels compelled to reveal true circumstances in so far as he is restricted to using different messages for different prevailing circumstances. Until now we have assumed nothing about an honest agent's attitude towards ethics revelation. Yet, it should be expected to have a major impact on the principal's ability to screen ethics and this ability should be the strongest when an honest agent feels compelled to reveal ethics as well as circumstances. We refer to this kind of honesty as being of the second order since, not only does an honest agent feel compelled to reveal his true circumstances, but he also feels compelled to reveal that he feels compelled to reveal his true circumstances. Intuition suggests that if honesty is of the second order, an honest agent will not be treated as well as a dishonest agent so that an honest agent might find it legitimate to misrepresent ethics in order to remedy such an unfair outcome. If this is the case then we say that honesty is of the first order. To fix ideas, consider a direct mechanism, where messages are of the form (θ_i, k). Then second-order honesty would imply $\mathcal{R}_i(\Omega) = \{(\theta_i, h)\}$ for all i, whereas first-order honesty would imply $\mathcal{R}_i(\Omega) = \{(\theta_i, h), (\theta_i, d)\}$ for all i.

Within the general framework that allows for non-direct mechanisms, we say that honesty is of the second order whenever there exists a message space $\mathcal{M}_0$ that contains at least

n messages that are not elements of $\mathcal{R}_i(\mathcal{M}_0)$ for any i. Each of these n messages could, for instance, be related to each set of circumstances θ_i, but would not be announced by a second-order honest agent in circumstances θ_i because he would consider it as misleading regarding ethics. The most obvious example is that of direct mechanisms where $\mathcal{R}_i(\Omega) = \{(\theta_i, h)\}$ for all i. The following example inspired by the "password" mechanisms introduced by Deneckere and Severinov (2001) also fits this definition. Consider some message space $\mathcal{M}_0$ such that the honest in circumstances θ_i would be restricted to messages in some subset $\mathcal{R}_i(\mathcal{M}_0)$. Now consider the message space $\widetilde{\mathcal{M}}_0 = \mathcal{M}_0 \times \{\mathcal{R}_1(\mathcal{M}_0), \ldots, \mathcal{R}_n(\mathcal{M}_0), \mathcal{M}_0\}$. Then honesty of the second order means that an honest agent in circumstances θ_i may only announce $(\mu_i, \mathcal{R}_i(\mathcal{M}_0))$, where μ_i denotes any message in $\mathcal{R}_i(\mathcal{M}_0)$.[12]

For both examples above, the optimal mechanism is the same. All that matters is that the principal may specify messages through which a dishonest agent may identify himself as such because these messages are out of reach of an honest, and there should be enough such messages so that the principal could specify as many allocations for a dishonest agent as there are circumstances: in short there needs to be at least n messages that may be announced by dishonest agents alone. To see this, consider some mechanism M where the message space $\mathcal{M}_0$ has this property, and let y_{ik} denote the equilibrium allocation of an agent with type (θ_i, k), and Y_M the set of equilibrium allocations. Then, independent of how the allocations in Y_M are associated with the messages in $\mathcal{M}_0$, the principal has to impose incentive constraints ensuring that a dishonest in circumstances θ_i prefers y_{id} to any other allocation in Y_M since the dishonest can choose any message in $\mathcal{M}_0$. By contrast, the set of allocations that an honest agent in any given circumstances may effectively choose from does depend on how the allocations in Y_M are associated with the messages in $\mathcal{M}_0$. Clearly, the best the principal may achieve consists in associating the n equilibrium allocations meant for the dishonest with messages that an honest agent may not announce: she then only

[12]Such a specification of restrictions in the message space $\widetilde{\mathcal{M}}_0$ would follow logically from restrictions in the original message space if restricted announcements were motivated by certifiable information. This point is explored in Forges and Koessler (2004).

needs to satisfy the honest agent's participation constraints.[13]

The following lemma summarizes some of the properties of the optimal contract.[14] Henceforth we will refer to the optimal allocations under second-order honesty as the full ethics screening allocations.

Lemma 1 *Under full ethics screening, a dishonest with circumstances better than θ_1 receives a strictly positive rent; a dishonest with circumstances θ_1 and an honest agent receive no rent. The allocation implemented under the worst circumstances is independent of ethics:* $y_{1h} = y_{1d}$.

If the honest agent truthfully reveals both circumstances and ethics, or more generally, if the principal may put the rent meant for the dishonest agent out of reach of the honest agent simply by using messages that an honest would not use, the principal leaves no informational rent to the honest agent. This does not mean, however, that the principal implements the first-best allocations for the honest agent. The traditional rent-efficiency trade-off still exists: since the dishonest may always claim to be honest, the principal may have to distort the decision associated with the honest agent in order to reduce the rent of the dishonest agent. For instance, if there are only two circumstances, the decision for the worst circumstances would be distorted downward compared to the first best; however, the distortion would be smaller than if the agent were dishonest with certainty, since the principal would now have to leave a rent only to the dishonest in good circumstances.

Under second-order honesty, the honest agent is subject to blatant discrimination: the dishonest receives a rent merely by claiming that he is dishonest. We believe that this is

[13] If the principal were restricted to using direct mechanisms, then results in Green and Laffont (1986) show that the revelation principle would apply for second-order honesty: the constraints that would apply would then be as described here. For non-direct mechanisms, if an honest agent is restricted in a "password" mechanism as is assumed in Deneckere and Severinov (2001) and (2003), then the principal's program would once again have the same structure.

[14] For a full characterization of the optimal contract with a continuum of circumstances, we refer to Deneckere and Severinov (2003).

inconsistent with the idea that honest agents evaluate the fairness of the situation to which they are confronted when deciding on whether to behave honestly. Honesty of the first order seems more appropriate since it at least allows an honest agent to misrepresent ethics whenever it guarantees him a larger surplus without having to misrepresent circumstances.

In a direct mechanism a first-order honest agent in circumstances θ_i would be willing to announce either (θ_i, h) or (θ_i, d). If the message space is $\widetilde{\mathcal{M}_0} = \mathcal{M}_0 \times \{\mathcal{R}_1(\mathcal{M}_0), \ldots, \mathcal{R}_n(\mathcal{M}_0), \mathcal{M}_0\}$, a first-order honest agent in circumstances θ_i would be willing to announce any (μ_i, δ), where μ_i is a message in $\mathcal{R}_i(\mathcal{M}_0)$, and $\delta \in \{\mathcal{R}_1(\mathcal{M}_0), \ldots, \mathcal{R}_n(\mathcal{M}_0), \mathcal{M}_0\}$. The key difference with second-order honesty is not that an honest may announce more messages, but rather that there are no messages that could be announced by a dishonest agent alone. It is thus no more possible for the principal to specify messages through which a dishonest could be identified as such. Formally, we define first-order honesty as follows: $\mathcal{M} = \bigcup_{i \in I} \mathcal{R}_i(\mathcal{M})$ for all $\mathcal{M}$. This trivially holds if the honest is unrestricted for message space $\mathcal{M}$; if he is restricted, this simply implies that there is no message in $\mathcal{M}$ that no honest agent may announce.

One would expect that the honest agent should be able to garner a rent if he is willing to claim to be dishonest. Surprisingly, this turns out not to be true in general. We now show that as long as the number of messages available to the honest agent for given circumstances is sufficiently large, the full ethics screening allocations may be implemented under first-order honesty.

Proposition 1 *Suppose that honesty is of the first order, and that for some message space $\mathcal{M}_0$, $\#\mathcal{R}_1(\mathcal{M}_0) \geq n$. Then the principal may implement the full ethics screening allocations by associating the allocation meant for the dishonest in circumstances θ_i with some message in $\mathcal{R}_1(\mathcal{M}_0)$.*

A simple argument proves this result. From Lemma 1 we know that only one message in $\mathcal{R}_1(\mathcal{M}_0)$ is needed for the allocation meant for an agent with circumstances θ_1, be he honest or dishonest. The remaining messages in $\mathcal{R}_1(\mathcal{M}_0)$ may then be associated with the full ethics screening allocations meant for the dishonest in circumstances other than θ_1; there are $n-1$ such circumstances. The honest agent under circumstances better than θ_1 may not select any of these messages, and the honest agent under circumstances θ_1 is not attracted to any allocation other than the one meant for him by virtue of the incentive constraints for the dishonest agent in circumstances θ_1.

Motivated by the question of whether an honest would be willing to misrepresent his ethics, we have come to analyze two rather extreme cases where the number of messages that may be announced by dishonest agents alone is either at least n or zero. Clearly all that really matters for the analysis above is whether this number is at least n or strictly less. Viewed in this way, Proposition 1 has a broader interpretation beyond considerations about an honest agent's willingness to lie about ethics. It says that the number of messages that can be announced only by dishonest agents is irrelevant, as long as it is possible to expand the number of different messages that an honest agent can announce under the worst circumstances.

We may now wonder whether and how restrictions on the number of messages available to an honest agent would prevent the principal from fully reaping the benefits of dealing with an honest agent. As suggested earlier, an honest agent may feel compelled to reveal circumstances truthfully only for certain message spaces. Perhaps he would be suspicious and refuse to volunteer his private information unless the mechanism is direct, so that there would be only two messages in every $\mathcal{R}_i(\Omega)$. Then from Proposition 1 we know that full ethics screening would be achieved if there are only two circumstances. With more than two circumstances, however, the principal may have to leave a rent to the honest agent if she wants to screen circumstances for the dishonest. Or worse still, maybe honesty would be ensured only if the message space is the set of circumstances Θ, so that there is only one message in every $\mathcal{R}_i(\Theta)$ as in the literature where ethics screening is exogenously ruled out. Then it is obvious that full screening of circumstances for the dishonest would guarantee a rent to the honest agent.

There is no clear theoretical foundation for restricting in one way or another the number of messages available to an honest agent for some given circumstances. Rather, we believe that it is more appropriate to build a theory that relies on the perception that an honest agent may have regarding the equilibrium outcome in the proposed mechanism. To illustrate, in any mechanism where the full ethics screening allocations are implemented under first-order honesty, an agent with circumstances θ_1 has a strict preference for one of the allocations associated with messages in $\mathcal{R}_1(\mathcal{M}_0)$. If another of these allocations is selected by some agent, he must have better circumstances than θ_1, so that only a dishonest agent could select such an allocation. Then the situation in equilibrium is not different from one where some messages are out of reach for an honest, as in second-order honesty. In both cases, it involves blatant discrimination between honest and dishonest agents; the principal may infer with certainty that when some allocations are selected, a dishonest agent is capturing a rent that is out of reach for an honest under similar circumstances.[15] Since the principal is openly endorsing the dishonest agent's behavior, the honest agent might feel vindicated in giving up honest behavior altogether. In the next section we allow for an honest behavior to be conditioned on the perceived fairness of the equilibrium outcome.

4 Conditional Honesty

Here we assume that an honest agent's behavior depends not only on the message space in the proposed mechanism, but also on his expectations about the behavior of the dishonest agent in the said mechanism. In his evaluation of a particular mechanism, an honest agent tries to assess whether the principal would take advantage of his honest behavior by openly treating a dishonest agent better. Under first-order unconditional honesty, such a discrimination was achieved by letting the dishonest get away with lying, where lying is

[15]This remark would still hold if the full ethics screening allocations were implemented by a mechanism where the principal asks the agent to announce circumstances, and where a unique allocation is offered for any message better than θ_1, whereas an agent announcing θ_1 is offered to choose among a menu of allocations (this mechanism was proposed by Deneckere and Severinov, 2001, as an alternative to the password mechanism which relies exclusively on messages).

defined as follows:[16]

Definition 3 *[Lying] Consider a mechanism $M = (\mathcal{M}_0, g)$. If for some i, (θ_i, d) announces $\mu \notin \mathcal{R}_i(\mathcal{M}_0)$, then he is lying.*

We introduce a parameter that measures the honest agent's tolerance for lying by others; it is a threshold probability such that if for some message the probability that the agent choosing this message is lying exceeds that threshold, then the honest agent would give up honest behavior altogether. Formally we assume that when confronted with a mechanism $M = (\mathcal{M}, g)$ an honest agent computes the Bayesian equilibrium that specifies announcements for all agent's types $\omega \in \Omega$, assuming that an honest agent in circumstances θ_i is restricted to messages in $\mathcal{R}_i(\mathcal{M})$. For any message μ announced in this equilibrium, he can then compute the probability $\lambda(\mu)$ that an agent who chooses this message is lying. A conditionally honest agent's tolerance for lying may then be described by specifying a value $\hat{\lambda} \in [0, 1]$ such that an honest agent in circumstances θ_i behaves honestly by choosing a message in $\mathcal{R}_i(\mathcal{M})$ if and only if $\lambda(\mu) \leq \hat{\lambda}$ for all μ. Otherwise, he behaves like a dishonest agent and chooses some message in $\mathcal{M}$.

For $\hat{\lambda} = 1$ the honest agent fully tolerates lies and we obtain as a limit case unconditional honesty that may be of the first or the second order. As shown in the following lemma, the introduction of some intolerance for lying, no matter how small, puts strong restrictions on equilibrium announcements in a given set $\mathcal{M}_0$, if the principal wishes to induce an honest behavior by the honest agent. Clearly, considering only mechanisms such that the honest agent behaves honestly involves no loss of generality: if a mechanism induces the honest to be opportunistic, the optimal allocations are the standard second-best ones; but if the agent is opportunistic with certainty, the revelation principle applies, so that these allocations may be implemented with a mechanism involving no lies.

[16]Although we use the term "lying," this definition is meant to capture a broader category of deceit. One can say that $\mathcal{R}_i(\mathcal{M}_0)$ represents the moral standards by which an honest agent feels that everybody should abide.

Lemma 2 *If $\hat{\lambda} < 1$, any message announced in equilibrium belongs to some $\mathcal{R}_i(\mathcal{M}_0)$. Moreover, for any $\mathcal{R}_i(\mathcal{M}_0)$, at most two messages may be announced in equilibrium.*

Proof: First, if there is a message μ that does not belong to $\mathcal{R}_i(\mathcal{M}_0)$ for any i, then if it is chosen in equilibrium we have $\lambda(\mu) = 1 > \hat{\lambda}$. Second, if for some $\mathcal{R}_i(\mathcal{M}_0)$ more than two messages were announced in equilibrium, there would be at least one that is announced by neither (θ_i, h) nor (θ_i, d) so that the probability of lying associated with this message would be one. *Q.E.D.*

It is therefore not possible to induce announcements of messages that could not be announced by an honest agent or to induce a wide variety of announcements in $\mathcal{R}_1(\mathcal{M}_0)$[17]. The following lemma further shows that in order to leave no rent to the honest agent, the principal would have to rely solely on "downward" lies by the dishonest.

Lemma 3 *A necessary condition for the principal to leave no rent to an honest agent in circumstances $\theta_i > \theta_1$ is that there exists no dishonest with circumstances $\theta_j \leq \theta_i$ selecting a message in $\mathcal{R}_i(\mathcal{M}_0)$.*

An important implication of Lemma 3 is that if the principal wishes to induce a dishonest agent to tell no lie she must leave a rent to an honest agent whenever his circumstances are not the worst. That rent must actually be equal to that of a dishonest under the same circumstances so that we may establish the following result.

Lemma 4 *An optimal mechanism where there is no lying implements the standard second-best allocations.*

[17]This would no more be the case if we allowed for mixed strategies on the part of the agent under circumstances θ_1 in which case it would be possible to have more than two messages in $\mathcal{R}_1(\mathcal{M}_0)$ announced in equilibrium. As will be seen shortly, conditional honesty puts restrictions on the mechanism used, other than those regarding announcements, that would render full ethics screening infeasible even with mixed strategies.

Proof: In a mechanism $M = (\mathcal{M}_0, g)$ where there is no lying, for all i there exists $\mu_{id} \in \mathcal{R}_i(\mathcal{M}_0)$ such that

$$(1) \qquad V(g(\mu_{id}), \theta_i) \geq V(g(\mu), \theta_i) \quad \forall \mu \in \mathcal{M}_0.$$

Now let μ_{ih} be a message that is chosen by (θ_i, h). Then we must have

$$(2) \qquad V(g(\mu_{ih}), \theta_i) \geq V(g(\mu_{id}), \theta_i).$$

Combining these constraints we get

$$(3) \qquad V(g(\mu_{ih}), \theta_i) \geq V(g(\mu), \theta_i) \quad \forall \mu \in \mathcal{M}_0.$$

Thus it is as if the principal should prevent the honest from lying. *Q.E.D.*

If the principal wishes to induce a dishonest not to lie, it is as if the agent were dishonest with certainty, and the optimal contract specifies the standard second-best allocations.

Lemma 3 also indicates that, in order to leave no rent to an honest agent, the principal must induce the dishonest to systematically claim that his circumstances are worse than they actually are. However, there are costs associated with such downward lies. In particular, if a dishonest in some circumstances $\theta_m > \theta_1$ lies downwards, only one message in $\mathcal{R}_m(\mathcal{M}_0)$ may be used, since if two messages were selected in equilibrium, then one would be chosen only by an agent who is lying. As a result, if an honest agent is to have zero rent irrespective of circumstances, there must be only one message used in $\mathcal{R}_i(\mathcal{M}_0)$ for all $i > 1$. This means that at most $n + 1$ different allocations may be implemented while leaving no rent to an honest agent. It is therefore not possible in general to implement full ethics screening, which involves up to $2n - 1$ allocations.[18]

Restriction on the number of different messages that may be announced in equilibrium would be somewhat relaxed by allowing mixed strategies on the part of the agent. There are

[18] As will become clear below, in the case of $n = 2$, full ethics screening could not be implemented, though for very different reasons.

however more fundamental restrictions on what can be implemented because of stringent constraints imposed by conditional honesty on the type of allocations that may be implemented. These "acceptability constraints" apply even if intolerance for lying is arbitrarily small. They are presented in the following lemma.

Lemma 5 *Suppose that $\hat{\lambda} < 1$. If for some θ_i there are two messages $\mu, \mu' \in \mathcal{R}_i(\mathcal{M}_0)$ that are announced in equilibrium, then $V(g(\mu), \theta_i) = V(g(\mu'), \theta_i)$.*

Proof: Consider a mechanism $M = (\mathcal{M}_0, g)$. Suppose that for some i there exists two messages $\mu, \mu' \in \mathcal{R}_i(\mathcal{M}_0)$, that are both announced in equilibrium. If $V(g(\mu'), \theta_i) > V(g(\mu), \theta_i)$, then $\lambda(\mu) = 1$. *Q.E.D.*

As a result of these constraints, even if more than two messages in any given set $\mathcal{R}_i(\mathcal{M}_0)$ could be used, the full ethics screening allocations could not be implemented. In fact, since the acceptability constraints would always be violated with the full ethics screening allocations, irrespective of how they might be distributed among the sets $\mathcal{R}_i(\mathcal{M}_0)$, $i = 1, \ldots, n$, there is a discontinuity at $\hat{\lambda} = 1$: unconditional honesty is not robust to the introduction of intolerance for lying, no matter how small.

The acceptability constraints further imply that if a dishonest in circumstances θ_m claims that his circumstances are worse than they really are by selecting a message in some $\mathcal{R}_i(\mathcal{M}_0)$, $i < m$, he would have to be compensated as though his circumstances actually were θ_i. Intuition suggests that the principal then would not design two different allocations for an agent in circumstances θ_i who is not lying and the dishonest in circumstances θ_m. The following result formally confirms this intuition.

Proposition 2 *Suppose that $\hat{\lambda} < 1$ and $\hat{\lambda}$ close to 1. If any agent selecting a message in $\mathcal{R}_i(\mathcal{M}_0)$ has circumstances at least as good as θ_i, then only one allocation is implemented using messages in $\mathcal{R}_i(\mathcal{M}_0)$.*

Acceptability implies that the principal would not want to treat a dishonest who lies downward differently from an agent who claims the same circumstances in a truthful manner, unless she had to. Proposition 2 indicates that with limited intolerance for lying she would only resort to ethics screening if she found it optimal to have a dishonest agent overstate circumstances and we will see below that it may be a quite beneficial option. In some situations, however, she may have to screen ethics although she does not induce upward lies, because of the honest agent's excessive intolerance towards lying. The following discussion provides some intuition for why this may be so.

Suppose that there are three circumstances, $\theta_1 < \theta_2 < \theta_3$, that the principal would ideally associate the first-best allocation y_i^* to each set $\mathcal{R}_i$,[19] and that among these allocations, the dishonest in the best circumstances would choose y_2^*. If the intermediate circumstances are relatively unlikely, this may result in an excessive probability of lying. In this case, the principal is led to choose a contract which gives her a lower surplus than the one specifying the first-best allocations; we now argue that this contract associates two different allocations with messages in $\mathcal{R}_1$. The graph in Figure 1 provides an illustration of the arguments. It shows four allocations, where the decision x_{ik} is meant for an agent with type (θ_i, k), and the corresponding indifference curves are labeled V_{ik} (the indifference curve of (θ_2, d) is not drawn). Here both π and v are strictly increasing in x. Since the principal would pick allocations inducing the dishonest in the best circumstances to announce in $\mathcal{R}_2$ if she were not constrained by $\hat{\lambda}$, the constraint ensuring that he prefers to announce in $\mathcal{R}_1$ rather than in $\mathcal{R}_2$ is binding: in the graph, the dishonest in the best circumstances is indifferent between y_{3d} and y_{2h}. Implementing the first-best allocation x_2^* for the honest in circumstances θ_2 would imply that $x_{3d} > x_1^*$, in which case it is clearly optimal to also associate the first-best allocation y_1^* with some message in $\mathcal{R}_1$ (in the graph it is assumed that (θ_1, h) would obtain this allocation). Although it will in general prove to be too costly to implement x_2^* for the honest in circumstances θ_2, the optimal contract will specify a

[19] From now on in our discussions of examples, we drop the reference to the overall message space $\mathcal{M}_0$ when designating restricted sets of messages.

decision x_{3d} which is larger than x_1^*, to optimally trade off the cost of decreasing x_{2d} below x_2^* against the cost of increasing x_{3d} above x_1^*.

[FIGURE 1 ABOUT HERE]

We now develop a numerical example reflecting exactly this situation. Suppose that $\pi(x) = x$ and $v(x, \theta) = \frac{x^2}{\theta}$, and that there are three sets of circumstances, $\theta_1 = 100/30$, $\theta_2 = 100/29$, and $\theta_3 = 100/14$, with $\alpha_1 = 0.65$, $\alpha_3 = 0.1$ and $\alpha_3 = 0.25$. Table 1 shows, for $\gamma = 0.2$, the message structure that is optimal depending on the degree of intolerance for lying. The third column provides a ranking of the relevant message structures, and the last column specifies the smallest value of $\hat{\lambda}$ for which each message structure is feasible. As in the example developed above, here the principal would ideally offer a contract with the first-best allocations, in which case the dishonest in circumstances θ_3 would choose y_2^*. However, this would require a $\hat{\lambda}$ of at least 0.385. The contracts yielding the second and third largest surpluses correspond to the situation illustrated in Figure 1: any dishonest picks a message in $\mathcal{R}_1$, and the principal associates two different allocations with messages in this set.

[TABLE 1 ABOUT HERE]

Intuition suggests that the principal would usually favor mechanisms involving some ethics screening. The above example however shows that she may end up screening ethics because her preferred solution with no ethics screening involves too much lying in the eyes of an honest agent. We will see below that there may be large benefits to resorting to ethics screening when the principal may induce upward lies on the part of dishonest agents who are not in the worst circumstances. We first consider a situation where ethics screening is never a preferred solution, namely, that with only two sets of circumstances.

Thus assume that there are two circumstances, θ_1 and $\theta_2 > \theta_1$. As a benchmark, and to illustrate the discontinuity at $\hat{\lambda} = 1$, we consider the limit case of unconditional honesty, i.e., $\hat{\lambda} = 1$. The principal would then be able to screen circumstances for the dishonest while leaving no rent to the honest as long as there are at least two messages in $\mathcal{R}_1$. It is straightforward to verify that the principal would then implement three allocations: the decision would be the first-best one x_2^* for an agent in circumstances θ_2, with only the dishonest agent receiving a rent, and the decision for an agent in circumstances θ_1 would be distorted downwards compared to the first-best decision x_1^*.

We now show that as soon as $\hat{\lambda} < 1$, there is no ethics screening, so that exactly two allocations are implemented, each being chosen by the honest in some circumstances. For now let us ignore potential constraints associated with intolerance for lying being too strong. Then it is straightforward to show that the two message structures involving an upward lie are dominated by the standard second-best mechanism.[20] From Proposition 2, as soon as $\hat{\lambda} < 1$, there is no ethics screening, so that exactly two allocations are implemented, each being chosen by the honest in some circumstances. Furthermore, the only two relevant message structures for a dishonest agent are no lying, or irrespective of his circumstances, he selects a message in $\mathcal{R}_1$. First, if there is no lying, Lemma 4 implies that the principal would offer the standard second-best mechanism, denoted M^s. Second, if a dishonest always chooses the allocation of an honest under bad circumstances, then it is as if the dishonest

[20] First, assume that, irrespective of his circumstances, the dishonest selects a message in $\mathcal{R}_2$, and further assume that two different allocations would be associated with messages in $\mathcal{R}_2$. Then the dishonest in circumstances θ_1 would be sharing an allocation with either the honest or the dishonest in circumstances θ_2, and constraints preventing (θ_2, h) and (θ_2, d) from selecting the same allocation as (θ_1, h) would have to be satisfied. But this would clearly be dominated by the standard second-best contract. Second, assume that the dishonest in the good circumstances θ_2 announces a message in $\mathcal{R}_1$, and *vice versa*: then there can be only one allocation associated with messages in either set $\mathcal{R}_i$, since if there were more than two one of them would be chosen by an agent who is lying. Let y_i denote the allocation associated with $\mathcal{R}_i$. Then both allocations y_1 and y_2 should satisfy the individual rationality constraint of an agent with circumstances θ_1. Absent constraints ensuring proper announcements by a dishonest, this would yield as a solution $y_1 = y_2 = y_1^*$ which is consistent with the specified message structure. This however could be achieved through a contract involving no lies, and is therefore dominated by the standard second-best mechanism.

were always in circumstances θ_1, so that the principal would propose a mechanism M^* specifying the first-best allocations.[21]

While M^s optimally screens circumstances for the dishonest at the cost of leaving a rent to the honest, M^* leaves no rent to the honest at the cost of forgoing any screening of circumstances for the dishonest. Clearly then, M^s is preferred to M^* for $\gamma = 1$, where γ is the probability that the agent is dishonest, and *vice versa* for $\gamma = 0$. Moreover, the principal's expected surplus is independent of γ with M^s whereas it is decreasing in γ with M^*. Hence there exists $\hat{\gamma} \in (0, 1)$ such that the principal would prefer the mechanism specifying first-best allocations to the standard second-best mechanism if and only if $\gamma \leq \hat{\gamma}$. However, the mechanism specifying the first-best allocations involves a lie: an agent selecting allocation y_1^* is lying with probability $\dfrac{\gamma\alpha_2}{\gamma\alpha_2 + \alpha_1}$. If the intolerance for lying is too strong so that $\hat{\lambda}$ is smaller than this, the honest would choose to behave dishonestly if the first-best mechanism were offered. In that case, the standard second-best mechanism is optimal since, as was shown above, it dominates any mechanism involving upward lies. The solution may thus be depicted as in Figure 2.

[FIGURE 2 ABOUT HERE]

With only two circumstances, conditional honesty makes it suboptimal to screen on the basis of ethics. Whereas this was imposed as an exogenous restriction in the early literature involving ethics heterogeneity, here it emerges as an endogenous property of the optimal contract. However, with a sufficiently strong intolerance for lying the standard second-best approach is used for much lower probabilities that the agent is dishonest than would be the case if ethics screening were exogenously ruled out. In fact, this is a general property of the optimal contract with any number of circumstances, as the following proposition shows.

[21] Here ethics screening with a downward lie cannot occur, since among the first-best allocations the dishonest always prefers y_1^*: as a result, the situation described in the previous example where an "upward" incentive compatibility constraint would be binding cannot arise.

Proposition 3 *For any* $\gamma > 0$*, there exists* $\hat{\lambda}_s(\gamma) > 0$ *such that for any* $\hat{\lambda} < \hat{\lambda}_s(\gamma)$ *the optimal mechanism implements the standard second-best allocations. Furthermore, the threshold value* $\hat{\lambda}_s(\gamma)$ *is strictly increasing and strictly concave in* γ*, and it tends to zero as* γ *tends to zero.*

Proof: Let α_{max} and α_{min} be the largest and the smallest value of the probabilities α_i, respectively. Then $\hat{\lambda}_s(\gamma) = \dfrac{\gamma\alpha_{min}}{\gamma\alpha_{min} + \alpha_{max}}$ represents the smallest possible probability of lying for some message given the probability distribution over circumstances. It is obtained by assuming that the dishonest with the least likely circumstances mimics the honest with the most likely circumstances, and the dishonest with the latter circumstances makes the same announcement. If $\hat{\lambda} < \hat{\lambda}_s(\gamma)$, any mechanism inducing the dishonest to lie in some circumstances leads the honest to behave in a dishonest manner, so that the optimal mechanism implements the standard second-best allocations. It is straightforward to verify that $\hat{\lambda}_s(\gamma)$ is strictly increasing and strictly concave in γ, tending to zero as γ tends to zero. *Q.E.D.*

The threshold value $\hat{\lambda}_s(\gamma)$ in the proposition is the smallest possible probability of lying given the distribution of types. If $\hat{\lambda}$ is smaller than that, any contract generating a lie would trigger an opportunistic behavior on the part of the honest agent. The optimal contract then involves no lies, and by Lemma 4 it is as if the agent were dishonest with certainty. The standard second-best approach is therefore robust to the introduction of honest agents as long as intolerance for lying is sufficiently strong. This is true even if honesty is very likely whereas if the principal only needed to satisfy an exogenous no ethics screening constraint, she would find it optimal to let the dishonest agent lie if dishonesty is unlikely (for instance by offering first-best allocations associated with the various possible circumstances). Thus for a strong enough intolerance for lying, the standard second-best approach is more robust in our conditional honesty framework than would suggest a model where no ethics screening is exogenously imposed.

Note that Proposition 3 only states a sufficient condition for the optimality of the standard second-best contract. It would seem intuitive that in our conditional honesty framework, it would also be optimal for large probabilities that the agent is dishonest. Surprisingly, although this is the case with only two sets of circumstances, we now show that it is not true generally. We also show that the honest agent's concerns for fairness does not necessarily rule out ethics screening, since as soon as there are more than two circumstances, it appears as a very attractive tool for the principal when combined with "upward lies". The discussion, however, further highlights how intolerance towards lying may prevent the principal from using this tool precisely in those situations when she would have benefited from it the most.

To illustrate these ideas, suppose that there are three circumstances $\theta_1 < \theta_2 < \theta_3$. Consider now a mechanism that specifies two messages in $\mathcal{R}_3$ along with one message in $\mathcal{R}_2$ and in $\mathcal{R}_1$. Suppose that the allocation associated with the unique message in $\mathcal{R}_1$ is the standard second-best allocation y_1^s and the allocation associated with the unique message in $\mathcal{R}_2$ specifies the standard second-best quantity with no rent under circumstances θ_2, $(x_2^s, v(x_2^s, \theta_2))$. Suppose further that the allocations associated with messages in $\mathcal{R}_3$ are the standard second-best allocations for circumstances θ_2 and θ_3, y_2^s and y_3^s. Incentive compatibility in the standard second-best implies that (θ_2, d) picks y_2^s, thus lying upwards, and (θ_1, d) picks y_1^s.

The key argument is that the acceptability constraint for allocations associated to messages in $\mathcal{R}_3$ is equivalent to the incentive constraint preventing an agent in circumstances θ_3 to claim that circumstances are θ_2 in the standard second-best mechanism. This is illustrated by the graph in Figure 3. As in Figure 1, the indifference curve for an agent with type (θ_i, k) is labelled V_{ik}. As in the standard second-best analysis the dishonest in the best circumstances is indifferent between his allocation and the allocation of the dishonest in circumstances θ_2, who in turn is indifferent between his allocation and that meant for an agent in the worst circumstances. The first of these incentive constraints being binding ensures that acceptability holds for allocations associated with messages in $\mathcal{R}_3$. Thus upward lies induce no additional costs relative to the standard second-best which would not be the case if (θ_2, d) lied downwards.

[FIGURE 3 ABOUT HERE]

Now agents in circumstances θ_3 being indifferent between y_2^s and y_3^s we may assume that the dishonest selects y_3^s while the honest selects y_2^s. The implemented allocations are thus those of the standard second-best except for the honest in circumstances θ_3 who picks y_2^s and the honest in circumstances θ_2 who gets no rent while being awarded the decision x_2^s. By switching from the standard second-best to this mechanism the principal may leave no rent to an honest in circumstances θ_2 at the cost of implementing a suboptimal decision for an honest in circumstances θ_3. Now this cost may be made arbitrarily small by shifting some probability weight from circumstances θ_3 to circumstances θ_2 so that α_3 goes to zero. The standard second-best mechanism could therefore be dominated even if γ is arbitrarily close to 1 (note that here the relative cost and benefit from switching to the mechanism with upward lies is independent of γ).

Leaving zero rent to the honest in circumstances θ_2 is all the more beneficial that these circumstances are likely; however, this in turn means that intolerance for lying cannot be too high in order for the principal to be able to induce such a behavior by a dishonest in circumstances θ_2 without upsetting the behavior of an honest agent. This suggests that intolerance for lying may prevent the principal from using upward lies precisely in those cases where she would have benefited from it the most.

We now illustrate these intuitions with a numerical example. Suppose again that $\pi(x) = x$ and $v(x, \theta) = \frac{x^2}{\theta}$. Further assume that $\theta_1 = 10/3$, $\theta_2 = 10/2.1$, and $\theta_3 = 5$, and that $\alpha_1 = \alpha_3 = .1$ and $\alpha_2 = .8$. The benefit from not having to give away the rent to an honest agent in circumstances θ_2 in the standard second-best is then 0.0146 whereas the cost of having an honest agent in circumstances θ_3 pick y_2^s is 0.0004. The standard second-best mechanism is therefore always dominated by a mechanism with ethics screening where the honest in circumstances θ_3 picks the same allocation as the dishonest in circumstances θ_2. Table 2 shows, for four different values of γ, the ranking of the message structures from the principal's perspective (the third column), as well as the values of $\hat{\lambda}$ that would be required to implement these message structures (the last column).[22] For the top three values of γ, down to $\gamma = 0.5$, the mechanism with ethics screening is optimal. This requires however that the honest agent is tolerant enough towards lying: even for $\gamma = 0.5$, $\hat{\lambda}$ should be at least 0.889. This example also shows that, even if the principal was exogenously restricted to no ethics screening, she could still improve upon the standard second-best by inducing upward lies (as shown on the third line for $\gamma = 0.9$ and $\gamma = 0.5$). As should be expected, this is not the case if γ is sufficiently close to 1, as for instance $\gamma = 0.99$ where the principal needs ethics screening to improve upon the standard second-best.

[TABLE 2 ABOUT HERE]

5 Concluding Remarks

We have introduced honesty in a principal-agent model in the simplest manner, by assuming that an honest agent is not willing to misrepresent his private information to increase his surplus. If such an honest behavior is unconditional, the principal is able to fully exploit it by leaving no informational rent to an honest agent, while still being able to trade off rent and efficiency by fully screening circumstances for the dishonest. This full ethics screening solution however means that the principal openly endorses discrimination against the honest agent. Psychological findings suggest that it may be unrealistic to expect honest behavior from an individual facing such an unfair treatment. We incorporate the idea that an honest behavior may be conditional on the contract involving no blatant discrimination against the honest. We introduce a parameter which can be interpreted as a measure of the honest

[22] The contract with ethics screening is the optimal one rather than the one based on standard second-best decisions that was used in the general discussion.

agent's intolerance for discrimination. Allowing for intolerance for discrimination, even if it is arbitrarily small, drastically affects the set of implementable allocations, and full ethics screening may no more be achieved regardless of the set of messages that could be used by the principal.

An important insight from our analysis is that ethics screening is not necessarily inconsistent with an honest agent's intolerance for discrimination as long as it is not too severe. This is true whether that intolerance is somewhat pronounced or limited, although ethics screening arises in either situation for very different reasons. If an honest agent is moderately tolerant for discrimination, ethics screening is used to ensure that the outcome does not look too discriminatory, whereas if he is more tolerant, ethics screening allows for achieving discrimination under circumstances that are relatively likely thus inducing massive lying on the part of the dishonest agent.

These favorable conclusions for ethics screening should be mitigated by the two following remarks. First, there is a difficulty with resorting to ethics screening because it requires that under some set of circumstances, agents in such circumstances are indifferent between several different allocations. Then the optimality of ethics screening may critically depend on how these agents pick and choose between these allocations. Here we have looked at the most favorable case for the principal but if she expects a different behavior by the agent, she may renounce using ethics screening. Second, the optimality of ethics screening when intolerance for discrimination is limited hinges on the possibility that a dishonest agent lies upwards whereas an honest agent would not. Yet, if we consider the underlying motivations for an honest behavior, upward lies may appear somewhat more palatable for an honest agent than downward ones. Think for instance of a workplace situation where circumstances are the agent's productivity: then an agent who feels guilty about pretending that his productivity is low in order to engage in shirking, may have no pang of conscience when picking some allocation designed for more productive workers. This last point suggests that it would be useful to reconsider the issues analyzed in the present paper in a framework where the underlying motivations of an honest agent would be more explicitly modelled.

Appendix

Proof of Lemma 1.

Let $y_{ik} = (x_{ik}, t_{ik})$ denote the equilibrium allocation of an agent with type (θ_i, k). Then the principal's problem is to determine the allocations $y_{ik},\ i \in I,\ \ k = h, d$ so as to maximize

$$\sum_{i=1}^{n} \alpha_i \left[\gamma[\pi(x_{id}) - t_{id}] + (1-\gamma)[\pi(x_{ih}) - t_{ih}]\right] \tag{4}$$

subject to

$$t_{ik} - v(x_{ik}, \theta_i) \geq 0 \quad \forall i \in I,\ k = h, d \tag{5}$$

$$t_{id} - v(x_{id}, \theta_i) \geq t_{jk} - v(x_{jk}, \theta_i) \quad \forall i, j \in I,\ k = h, d. \tag{6}$$

By a slight abuse of language, we will refer to the constraints in (5) as the individual rationality constraints, and the constraints in (6) as the incentive compatibility constraints. Moreover, let (IC_{ijk}) denote the constraint ensuring that the dishonest in circumstances θ_i prefers y_{id} to y_{jk}.

First, thanks to the single-crossing condition, the participation constraint for the dishonest in the worst circumstances θ_1 implies that for any $i > 1$, the dishonest has a strictly positive rent. Second, the rent of the honest should be set to zero for all i: indeed, if for some i the rent of the honest were strictly positive, the transfer t_{ih} could be reduced without jeopardizing any constraint, thus increasing the principal's expected surplus.

Next we show that the rent of the dishonest in circumstances θ_1 is zero. Assume that this rent were strictly positive. We need to consider two cases.

(i) First, assume that $y_{1d} \neq y_{id}$ for all $i > 1$ (no bunching for the dishonest). Then since the honest receives no rent for any i, any allocation y_{ih} for $i > 1$ would give the dishonest in circumstances θ_1 a strictly negative surplus; moreover, single-crossing implies that he strictly prefers y_{1d} to any allocation y_{id}, $i > 1$. As a result, the principal may decrease t_{1d} without affecting any incentive compatibility constraint adversely, thereby increasing her surplus.

(*ii*) Second, assume that there exists some $i > 1$ such that $y_{id} = y_{1d}$. Consider the largest such i and denote it ℓ; single-crossing implies that $y_{id} = y_{1d}$ for all $i \leq \ell$. If for all i and j such that $j < i \leq \ell$, (IC_{ijh}) is slack, then t_{1d} can be decreased without upsetting any incentive compatibility constraint. Next, suppose that for some i and j such that $j < i \leq \ell$, (IC_{ijh}) is binding. Then any rent that the principal leaves to a dishonest that depends on x_{jh} is increasing in x_{jh}, so that $x_{jh} \leq x_j^* < x_i^*$. Furthermore, since (θ_i, d) is indifferent between y_{1d} and y_{jh}, it must be that $x_{1d} < x_{jh}$ for (θ_1, d) to pick y_{1d}. As a result total surplus would increase if the dishonest in circumstances θ_i was reassigned to allocation y_{jh}; the principal could therefore increase her surplus by doing that, since the rent of the dishonest in circumstances θ_i would be unchanged. This contradicts the assumption that it would be optimal to let (θ_i, d) choose y_{1d}.

Finally, we show that $x(\theta_1, d) = x(\theta_1, h)$. Suppose that these decisions were different, and let $\bar{x}_1$ denote the largest one and $\underline{x}_1$ the smallest one. Since both the honest and the dishonest in circumstances θ_1 receive the same rent, the single-crossing condition implies that a dishonest in circumstances $\theta_i > \theta_1$ would prefer the allocation with $\bar{x}_1$. Furthermore, if a dishonest in circumstances $\theta_i > \theta_1$ receives a rent which depends on $\bar{x}_1$, it is increasing in $\bar{x}_1$. This implies that $\bar{x}_1 \leq x_1^*$. But then the principal may increase her expected surplus by increasing $\underline{x}_1$ while also increasing the corresponding transfer so as to leave the rent of the agent in circumstances θ_1 unchanged: since $\underline{x}_1 < \bar{x}_1 \leq x_1^*$ this increases total surplus, and jeopardizes no incentive compatibility constraint. *Q.E.D.*

Proof of Lemma 3.

Given a certain pattern of announcements, the equilibrium allocations $y_{ik} = (x_{ik}, t_{ik})$ must satisfy the individual rationality constraints (5) and the incentive compatibility constraints for the dishonest (6) specified in the proof of Lemma 1. In addition they must satisfy any relevant incentive compatibility constraints for the honest:

$$t_{ih} - v(x_{ih}, \theta_i) \geq t_{jd} - v(x_{jd}, \theta_i) \tag{7}$$

$$\forall i \in I, \forall j \text{ such that } (\theta_j, d) \text{ announces a message in } \mathcal{R}_i(\mathcal{M}_0).$$

These constraints ensure that among the allocations associated with messages in $\mathcal{R}_i(\mathcal{M}_0)$, the honest in circumstances θ_i prefers allocation y_{ih}; the constraints reflect the fact that the only relevant messages are those that are announced in equilibrium.

Consider some circumstances $\theta_i > \theta_1$, and assume that no dishonest selects a message in $\mathcal{R}_i(\mathcal{M}_0)$. Then, if the honest in circumstances receives a rent, the principal may increase her surplus by decreasing t_{ih} without jeopardizing any constraint. Thus, a necessary condition for an honest agent in circumstances $\theta_i > \theta_1$ to earn a strictly positive rent is that there exists θ_j such that (θ_j, d) announces a message in $\mathcal{R}_i(\mathcal{M}_0)$. A sufficient condition is that $\theta_j \leq \theta_i$: this is a direct implication of v being strictly increasing in θ. *Q.E.D.*

Proof of Proposition 2.

This proof uses the constraints (5)-(7) specified in the proofs of Lemmas 1 and 3. First note that if for some i any agent selecting a message in $\mathcal{R}_i(\mathcal{M}_0)$ has circumstances θ_i, then only one allocation is implemented using messages in $\mathcal{R}_i(\mathcal{M}_0)$. Next, consider some $i < n$ for which there exists some $m > i$ such that (θ_m, d) selects a message in $\mathcal{R}_i(\mathcal{M}_0)$ and there is no $j < i$ such that (θ_j, d) selects a message in $\mathcal{R}_i(\mathcal{M}_0)$.

Suppose there are two different allocations associated with messages in $\mathcal{R}_i(\mathcal{M}_0)$. Let y_m denote the allocation obtained by (θ_m, d) and $y_i \neq y_m$ the one obtained either by (θ_i, d) or (θ_i, h) (a necessary condition for two different allocations to be implemented using messages in $\mathcal{R}_i(\mathcal{M}_0)$ is that (θ_i, d) does not lie). Acceptability implies that an agent in circumstances θ_i must be indifferent between y_i and y_m. For (θ_m, d) to prefer y_m to y_i it must therefore be that $x_m > x_i$. We would be done if we could prove that $x_m \leq x_i^*$. Indeed, $x_i < x_m \leq x_i^*$ means that total surplus would increase if x_i were increased, so that the principal could increase her surplus by increasing x_i while keeping the surplus of the agent in circumstances θ_i who picks y_i unchanged. Such a manipulation would not jeopardize any constraints: first, as long as $x_i \leq x_m$, any dishonest with circumstances better than i who announces in $\mathcal{R}_i(\mathcal{M}_0)$ still prefers y_m to y_i; second, if for some $j < i$, (θ_j, d) is indifferent between his equilibrium allocation and y_i, then increasing x_i along the indifference curve of the agent in circumstances θ_i simply means that (θ_j, d) now strictly prefers his equilibrium allocation to y_i.

We now proceed to showing that if $\hat{\lambda}$ is sufficiently large for any lie to be feasible, it must be that $x_m \leq x_i^*$; it will become clear at the end of the proof why this would not necessarily be true for values of $\hat{\lambda}$ such that some lie is infeasible. The key to proving $x_m \leq x_i^*$ is to show that any incentive compatibility constraint containing x_m that might be binding ensures that an agent with circumstances better than θ_m prefers his equilibrium allocation to y_m. This in turn implies that any rent that depends on x_m is increasing in x_m; since x_m must be on the indifference curve of an agent in circumstances θ_i, at the optimum it must be that $x_m \leq x_i^*$.

Clearly, any incentive compatibility constraint ensuring that a dishonest agent with circumstances worse than θ_i strictly prefers his equilibrium allocation to y_m is implied by the fact that (θ_i, d) is indifferent between y_i and y_m and $x_i < x_m$. Thus any binding incentive compatibility constraint ensuring that some dishonest is not attracted to y_m would concern an agent with circumstances better than θ_i. Still, we cannot yet conclude that this would be the only binding incentive compatibility constraint containing y_m: there may exist some allocation $y' = (x', t')$ such that (θ_m, d) is indifferent between y_m and y', and this may in turn affect x_m in non-trivial ways. We thus need to prove that there exists no such allocation. We only need to consider allocations y' such that $x' > x_m$: if x' were smaller than x_m, since $\theta_i < \theta_m$ and an agent with circumstances θ_i is indifferent between y_i and y_m, an agent in circumstances θ_i would strictly prefer y' to y_i.

Thus, assume that there exists an allocation $y' = (x', t')$ with $x' > x_m$ such that (θ_m, d) is indifferent between y_m and y'. We now show that this cannot be part of an optimal mechanism. The argument uses two remarks that also apply to the standard incentive problem. First, incentive constraints imply that decisions in allocations chosen by dishonest agents are non decreasing in the dishonest agent's circumstances. Second, if some dishonest agent is indifferent between his equilibrium allocation and some other allocation specifying a higher decision, then any dishonest agent with better circumstances strictly prefers that other allocation to any allocation that involves a lower decision.

We now show that if (θ_m, d) is indifferent between y_m and y' with $x' > x_m$, then it would be possible to decrease t' while keeping all constraints satisfied. Constraints that may be violated are the individual rationality constraint for an honest choosing y' and incentive constraints preventing a dishonest who is choosing y' from choosing another allocation involving a higher decision. It is straightforward to keep all these constraints satisfied as long as whenever a dishonest is indifferent between his equilibrium allocation and some other allocation involving a higher decision, that other allocation yields a strictly positive surplus for an honest agent choosing it. Indeed, it is then possible to decrease t' without violating any individual rationality constraint, and if some dishonest is indifferent between y' and some other allocation, then the transfer in that other allocation may be decreased as well by the same amount. Thus to complete the proof we only need to show the following Lemma.

Lemma 6 *Let $\hat{y} = (\hat{x}, \hat{t})$ be an allocation associated with some message in $\mathcal{R}_r(\mathcal{M}_0)$, that yields zero surplus for an agent in circumstances θ_r. Then if a dishonest is indifferent between his equilibrium allocation $\tilde{y} = (\tilde{x}, \tilde{t})$ and $\hat{y}$, then $\hat{x} < \tilde{x}$.*

Proof: The proof proceeds by first showing that the result holds for the allocation $\hat{y}$ satisfying the assumptions of the lemma and for which the decision is the largest among all these allocations and then showing by induction that the result is true for all allocations $\hat{y}$.

Letting y_{rh} denote the equilibrium allocation of (θ_r, h), and Y the set of all equilibrium allocations, we define $\hat{Y}$ as the set of all allocations that satisfy the assumptions of Lemma 6: $\hat{Y} = \{y_{rh} \in Y : V(y_{rh}, \theta_r) = 0\}$.[23] Let us denote elements of $\hat{Y}$ as $\hat{y}^q$, where a larger q indicates a smaller decision.

[23] It is not necessary to consider allocations that are out of equilibrium, since they could simply be removed. Furthermore, there cannot be allocations that would yield zero surplus for some honest but that would only be chosen by a dishonest: since the dishonest in the same circumstances as the honest would choose some other allocation giving him a larger surplus, such a situation would involve a probability of lying equal to one.

Now consider $\hat{y}^1$. Then all equilibrium allocations involving a larger decision are associated with a message in $\mathcal{R}_s(\mathcal{M}_0)$, for some s such that an agent with circumstances θ_s earns a strictly positive surplus for these allocations. We proceed by contradiction to show that the result holds for $\hat{y}^1$.

Assume that for some θ_ℓ, the dishonest in circumstances θ_ℓ is indifferent between his equilibrium allocation $y_{\ell k}$ and $\hat{y}^1$ and that $\hat{x}^1 > x_{\ell k}$. Clearly, $\theta_\ell > \theta_r$, where θ_r denotes the circumstances of the honest receiving zero rent at $\hat{y}^1$. Note that any dishonest with circumstances worse than θ_ℓ (in particular (θ_r, d)) strictly prefers $y_{\ell k}$ to $\hat{y}^1$, whereas the opposite is true for any dishonest with circumstances better than θ_ℓ.

If $x_{\ell k} < \hat{x}^1 \leq x_r^*$, since $x_r^* < x_\ell^*$, the principal would be better off by letting (θ_ℓ, d) pick $\hat{y}^1$ instead of $y_{\ell k}$: this would increase total surplus, while the surplus of (θ_ℓ, d) would be unchanged. Note that it is precisely this argument which may fail if $\hat{\lambda}$ is small, as letting (θ_ℓ, d) pick $\hat{y}^1$ instead of $y_{\ell d}$ may imply a too large probability of lying.

We now prove by contradiction that $\hat{x}^1 \leq x_r^*$. If $\hat{x}^1$ were greater than x_r^*, total surplus would increase by reducing $\hat{x}^1$, so that the principal could increase her surplus by reducing $\hat{x}^1$ while also reducing $\hat{t}$ so as to keep the agent's surplus in circumstances θ_r unchanged; this surplus remaining unchanged ensures that no individual rationality constraint is violated. Furthermore, no incentive constraints would be jeopardized. First, the surplus that a dishonest in circumstances at least as large as θ_ℓ would derive from $\hat{y}^1$ is reduced. Second, a dishonest in circumstances worse than θ_ℓ would still prefer $y_{\ell k}$ to $\hat{y}^1$. Third, if $\hat{y}^1$ is chosen by some dishonest, he must have better circumstances than $\theta_{\ell k}$ so that none of the constraints preventing him from choosing an allocation with a smaller decision than $\hat{x}^1$ are binding. Now, if some incentive constraint involving an allocation with a larger decision than $\hat{x}^1$ were binding, since by assumption, all honest agents who would be choosing these allocations have a strictly positive surplus, it would be possible to reduce transfers while keeping their participation constraints satisfied. Then standard arguments may be used to show that transfers associated with allocations involving decisions above $\hat{x}^1$ may be reduced appropriately so as to keep all incentive constraints satisfied.

Therefore, any dishonest choosing an allocation whose decision is smaller than $\hat{x}^1$ strictly prefers his equilibrium allocation to $\hat{y}^1$, and to any allocation with a decision which is larger than $\hat{x}^1$. With this we can show that the same arguments as those used above may be applied to the allocation $\hat{y}^2$. Thus assume that for some θ_ℓ, (θ_ℓ, d) is indifferent between his equilibrium allocation $y_{\ell d}$ and $\hat{y}^2$ and that $\hat{x}^2 > x_{\ell d}$, and again let θ_r be the circumstances of the honest receiving zero rent at $\hat{y}^2$. Recall that any agent with circumstances larger than θ_ℓ would strictly prefer $\hat{y}^2$ to $y_{\ell d}$, and to any allocation with a decision smaller than $x_{\ell d}$. Then, if $\hat{x}^2$ were greater than x_r^*, the principal could increase her surplus by reducing $\hat{x}^2$ and $\hat{t}^2$. Now consider reducing the transfers associated with all the allocations with decisions between $\hat{x}^2$ and $\hat{x}^1$ by the same amount as $\hat{t}^2$ was reduced. Then these allocations become less attractive, so that any dishonest who picked some allocation with a decision which is either larger than $\hat{x}^1$ or smaller than $\hat{x}^2$ still does so. Finally, if the transfer decrease is sufficiently small, any dishonest agent who chose an allocation with some decision x such that $\hat{x}^2 \leq x < \hat{x}^1$ before the transfer decrease would still pick the same allocation: this is true since before the transfer decrease, any such agent strictly preferred the allocation meant for him to any allocation involving a decision which is either smaller than $x_{\ell d}$, or larger or equal to $\hat{x}^1$. But since $\hat{x}^2 \leq x_r^*$ and $\theta_\ell > \theta_r$, it would be better to let (θ_ℓ, d) pick $\hat{y}^2$ instead of $y_{\ell d}$. *Q.E.D.*

References

Alger, I. and A. C.-t. Ma (2003), "Moral Hazard, Insurance, and Some Collusion," *Journal of Economic Behavior and Organization*, 50:225-247.

Alger, I. and R. Renault (2004), "Screening Ethics when Honest Agents Keep Their Word," mimeo, Boston College and Université de Cergy-Pontoise.

Andreoni, J., B. Erard and J. Feinstein (1998), "Tax Compliance," *Journal of Economic Literature*, 36:818-860.

Bok, S. (1978), *Lying: Moral Choice in Public and Private Life*, New York: Pantheon Press.

Crocker, K.J. and J. Morgan (1998), "Is Honesty the Best Policy? Curtailing Insurance Fraud through Optimal Incentive Contracts," *Journal of Political Economy*, 106:355-375.

Dawes, R.M. and R.H. Thaler (1988), "Anomalies: Cooperation," *Journal of Economic Perspectives*, 2:187-197.

Deneckere, R. and S. Severinov (2001), "Mechanism Design with Communication Costs," mimeo, University of Wisconsin and Duke University.

Deneckere, R. and S. Severinov (2003), "Does the Monopoly Need to Exclude?," mimeo, University of Wisconsin and Duke University.

Erard, B. and J.S. Feinstein (1994), "Honesty and Evasion in the Tax Compliance Game," *Rand Journal of Economics*, 25:1-19.

Forges, F. and F. Koessler (2003), "Communication Equilibria with Partially Verifiable Types," mimeo, Université de Cergy-Pontoise.

Green, J.R. and J.-J. Laffont (1986), "Partially Verifiable Information and Mechanism Deisgn," *Review of Economic Studies*, 53:447-456.

Hartshorne, H. and M.A. May (1928), *Studies in deceit*, New York: McMillan.

Hollinger, R.C. (1991), "Neutralizing in the Workplace: An Empirical Analysis of Property Theft and Production Deviance," *Deviant Behavior: An Interdisciplinary Journal*, 12:169-202.

Jaffee, D.M. and T. Russell (1976), "Imperfect Information, Uncertainty, and Credit Rationing," *Quarterly Journal of Economics*, 90:651-666.

Kofman, F. and J. Lawarrée (1996), "On the Optimality of Allowing Collusion," *Journal of Public Economics*, 61:383-407.

Lacker, J.M. and J.A. Weinberg (1989), "Optimal Contracts under Costly State Falsification," *Journal of Political Economy*, 97:1345-1363.

Laffont, J.-J. and D. Martimort (2002), *The Theory of Incentives*, Princeton: Princeton University Press.

MAGGI G. and A. RODRIGUEZ-CLARE (1989), "Costly Distortion of Information in Agency Problems," *Rand Journal of Economics*, 26:675-689.

MUELLER, C.W. and T. WYNN (2000), "The Degree to Which Justice is Valued in the Workplace," *Social Justice Research*, 13:1-24.

PARILLA, P.F., R.C. HOLLINGER and J.P. CLARK (1988), "Organizational Control of Deviant Behavior: The Case of Employee Theft," *Social Science Quarterly*, 69:261-280.

PICARD, P. (1996), "Auditing Claims in the Insurance Market with Fraud: The Credibility Issue," *Journal of Public Economics*, 63:27-56.

ROTH, J.A., SCHOLZ, J.T., AND WITTE A.D., EDS. (1989), *Taxpayer compliance: An Agenda for Research*, Philadelphia: University of Pennsylvania Press..

RYAN, A.M., M.J. SCHMIDT, D.L. DAUM, S.BRUTUS, S.A. MCCORMICK and M.H. BRODKE (1997), "Workplace Integrity: Differences in Perceptions of Behaviors and Situational Factors," *Journal of Business and Psychology*, 12:67-83.

SPICER, M.W. and L.A. BECKER (1980), "Fiscal Inequity and Tax Evasion: An Experimental Approach," *National Taxation Journal*, 33:171-175.

TIROLE, J. (1992), "Collusion and the Theory of Organizations," in Laffont, J.-J.,ed., *Advances in Economic Theory: Proceedings of the Sixth World Congress of the Econometric Society*, Cambridge: Cambridge University Press.

Table 1: Ethics screening using a downward lie.

$\gamma = 0.2$	# of alloc.	Optimal message structure	$\hat{\lambda}$
	3	(θ_2, d) announces in $\mathcal{R}_1$, and (θ_3, d) in $\mathcal{R}_2$ (FB allocations)	0.385
	4	for any i, (θ_i, d) announces a message in $\mathcal{R}_1$ (θ_1, d) announces the same message as (θ_3, d)	0.350
	4	for any i, (θ_i, d) announces a message in $\mathcal{R}_1$. (θ_1, h) announces the same message as (θ_3, d)	0.119
	3	No lies (SSB allocations)	0.000

Table 2: Example with three sets of circumstances.

γ	# of alloc.	Optimal message structure	$\hat{\lambda}$
$\gamma = \mathbf{0.99}$	4	(θ_2, d) announces the same message in $\mathcal{R}_3$ as (θ_3, h)	0.999
	3	No lies (SSB allocations)	0.00
$\gamma = \mathbf{0.9}$	4	(θ_2, d) announces the same message in $\mathcal{R}_3$ as (θ_3, h)	0.986
	4	(θ_2, d) announces the same message in $\mathcal{R}_3$ as (θ_3, d)	0.889
	3	(θ_2, d) announces in $\mathcal{R}_3$	0.878
	3	(θ_3, d) announces in $\mathcal{R}_2$	0.101
	3	No lies (SSB allocations)	0.000
$\gamma = \mathbf{0.5}$	4	(θ_2, d) announces the same message in $\mathcal{R}_3$ as (θ_3, h)	0.889
	4	(θ_2, d) announces the same message in $\mathcal{R}_3$ as (θ_3, d)	0.889
	3	(θ_2, d) announces in $\mathcal{R}_3$	0.800
	3	(θ_3, d) announces in $\mathcal{R}_2$	0.059
	3	No lies (SSB allocations)	0.000
$\gamma = \mathbf{0.05}$	3	Any dishonest announces in $\mathcal{R}_1$ (FB allocations)	0.310
	4	(θ_2, d) announces the same message in $\mathcal{R}_3$ as (θ_3, h)	0.296
	3	(θ_2, d) announces in $\mathcal{R}_3$	0.286
	3	(θ_3, d) announces in $\mathcal{R}_2$	0.006
	3	No lies (SSB allocations)	0.000

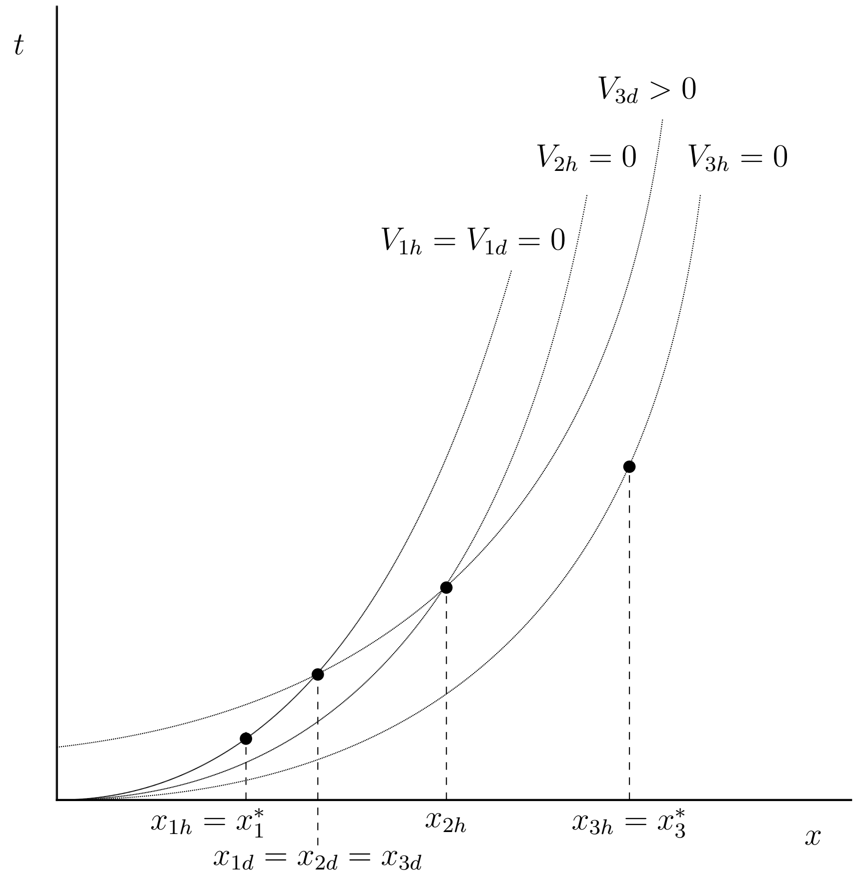

Figure 1. Ethics screening with a downward lie.

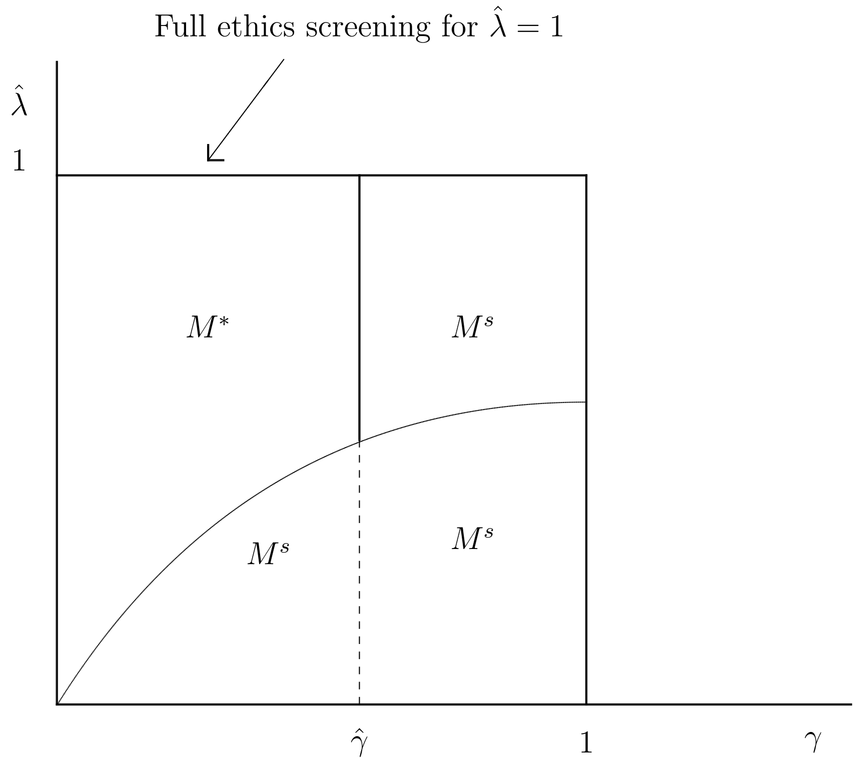

Figure 2. The solution with two circumstances.

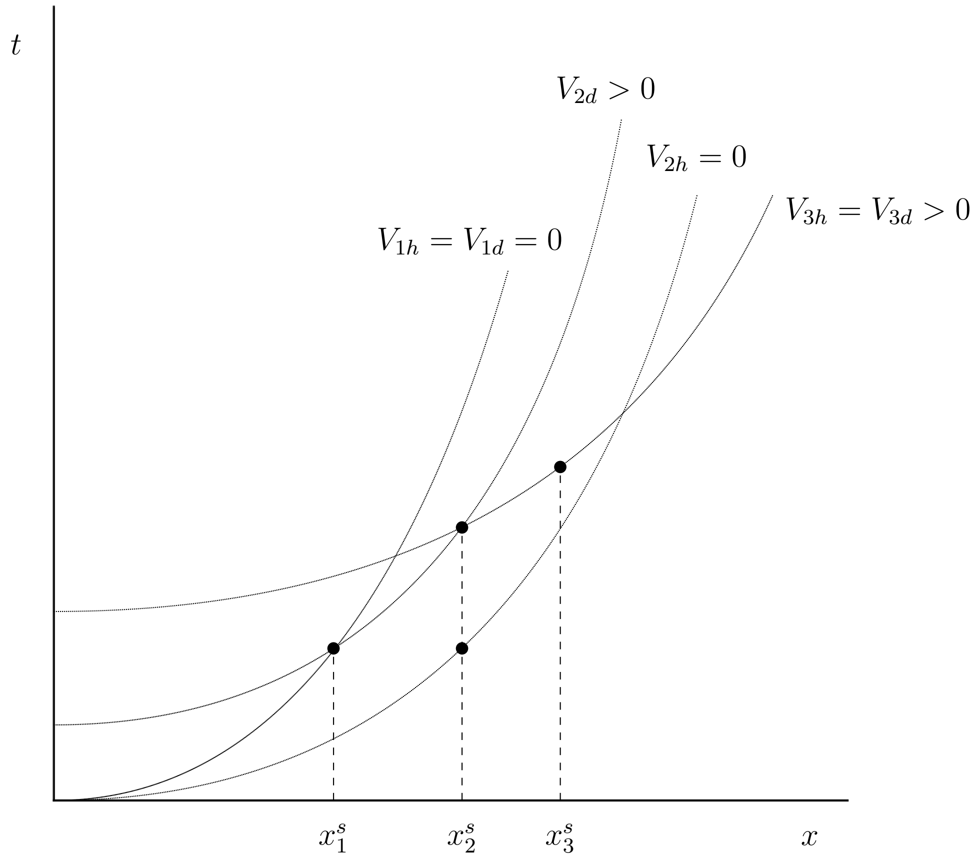

Figure 3. Ethics screening with an upward lie.