# Optimal Overspecified Contracts*

Francesco Squintani†

December 2019

**Abstract**

This paper identifies situations where desired outcomes may only be implemented with the stipulation of 'overspecified contracts', defined as contracts that include clauses that are not enforceable. The desired outcome is achieved in equilibrium by secretly breaching the overspecified contract. Such mechanisms are needed to circumvent legal restrictions or practical enforcement limitations that prevent implementation with enforceable contracts. A general model of contractual enforcement is formulated. Necessary and sufficient conditions for the optimality of overspecified contracts are derived. Overspecified contracts are redundant when Courts process information as standard Bayesian agents and the only legal restriction to contract enforcement is limited liability.

*I thank several seminar audiences including those at University of Pennsylvania, University of California at Los Angeles, Columbia University, Princeton University, European University Institute, University of Edinburgh, as well as Eddie Dekel, Faruk Gul, Philippe Jehiel, Steve Matthews, József Sákovics, and especially Meg Meyer for very valuable comments, as well as Sarthak Joshi for excellent research assistance. The usual caveats apply.

†Department of Economics, University of Warwick, f.squintani@warwick.ac.uk

# 1 Introduction

Regulations and contracts are common means to improve economic efficiency. While regulations and contracts may contain any sort of clauses in principle, some prescriptions may be too costly or impossible to enforce in Court. Legal restrictions on the enforcement of private contracts are also unavoidable: it would be too costly to check all clauses in all contracts individuals stipulate. Such legal or technical limitations impose a burden on the optimal use of private contracts and regulations ('social contracts').

This paper identifies situations where desired outcomes may only be implemented with the stipulation of 'overspecified contracts' that include clauses that are not enforceable. In equilibrium, the parties secretly breach the contract and implement the desired outcome. Overspecified contracts are needed to circumvent legal restrictions or practical enforcement limitations that prevent implementation of a desired outcome with simpler contracts.

As a concrete example, let me start with a casual observation. In many areas of the US, fast motorists drive on highways at about 10 mph above the speed limit, and such violations seem to be mostly tolerated by the Police. Many believe that this "equilibrium behaviour" is justifiable because speed limits are too restrictive in the first place. The police may choose to not enforce strict speed limits in case inflexible enforcement is unpopular, or if high fixed administrative costs make it impractical to do so. In such a case, it may indeed be optimal to overspecify the social contract by setting the speed limit to 10 mph below the optimal maximum highway speed, in the expectation that violations below 10 mph will seldom be sanctioned. In general, overspecified contracts can be a sensible strategy to overcome the impossibility or impracticality of precise enforcement of contractual clauses or norms.

Overspecified contracts are not uncommon in commerce and in professional services. As reported for example by Keating (1997), while sellers often commit to provide highest grade commodities, buyers will be satisfied with lower specifications. Consulting contracts include clauses such as "provision of service according to best practices," that are difficult to verify or quantify. Many clauses, especially small-print clauses, are usually ignored by the contractual parties.

A major reason for contract overspecification is the standard legal provision that rules out contractual transfers (liquidated damages clauses) that are punitive in nature. Even if it is verified that a party breached contract, she cannot be sanctioned in excess of verified harm or of the foregone profits suffered by her counterparts because of her breach.[1] Con-

[1] See the Uniform Commercial Code (sections 2-718 and 2A-504), the Restatement (Second) of Contracts

tract overspecification is a suitable approach to provide sufficiently strong incentives and deter breach. Law and economics scholars are well aware of the consequences of laws ruling out punitive transfers. (E.g., see the symposium on 'Economic Loss' published on the *International Review of Law and Economics,* 2007.) But the idea that overspecified contracts may be used to avert such issues has not been adequately fleshed out. An exception is a study by Edlin (1997) that I will discuss in details in Section 2.

A more subtle example of overspecified contracts may be found in multi-agent production problems. It is based on a broad legal principle, to which I refer as 'individual liability', here. Sanctioning an agent may be possible only when that specific agent is verified in breach of her contractual commitment. It is not sufficient that a claimant verifies in Court that the contracted outcome has not been implemented to enforce any individual sanctions.[2] Evidently, individual liability makes it difficult to motivate the agents to exert high effort in multi-agent production environments.

A possible solution to this problem is for the agents to form a company, or to hire an intermediary that acts as the sole liable legal entity vis-a-vis any contractual counterpart. But this entails non-trivial legal costs, and need not always be efficient. A different possible solution is for the agents to sign individual contracts by which they not only commit to exert high effort, but also to monitor each others' efforts. These overspecified contracts make each agent indirectly responsible for the other agents' efforts, thereby circumventing individual liability constraints. Furthermore, since each agent expects that all agents exert high effort in equilibrium, no one needs to waste time into monitoring her peers.

The usefulness of overspecified contracts may not come as a surprise. But their analysis poses a challenge to standard theoretical models.[3] Available contracts are usually defined

section 356, and 347 cmt. a: "Contract damages ... are intended to give [the plaintiff] the benefit of his bargain by awarding him a sum of money that will, to the extent possible, put him in as good a position as he would have been in had the contract been performed." In virtually all legal systems outside the United States, punitive damages are either excluded or play a very minor role. Even in the US, punitive damages are normally restricted to cases of reckless conduct, e.g. drunken driving.

[2]Contractual liability normally follows from individual contractual commitments, so that individual liability may be considered the default rule. An exception is outlined is the regulation of 'negotiable instruments' (i.e. unconditional promises or orders to pay a fixed amount of money). The Universal Commercial Code 3-116 states that "Except as otherwise provided in the instrument, two or more persons who have the same liability on an instrument [...] are jointly and severally liable [...]." While not predominant in contract law, joint and several liability may apply in tort law, e.g. the Comprehensive Environmental Response Compensation and Liability Act.

[3]It is useful to disclaim any relationship between my earlier leading examples and common mechanisms studied in contract theory. First, there is no contract renegotiation here. Overspecified contracts are stipulated and then breached in equilibrium. Second, my examples are one-shot interactions, with no repetition effects, nor relational contracts. Finally, what I describe is not a case of contract incompleteness. To the contrary, contracts are overspecified and include unenforceable clauses. The optimality of overspecified contracts is conceptually similar to a failure of the revelation principle, as I will make precise later.

as collections of transfers contingent on signals verifiable in Court. The contract space may be further restricted to capture technological or legal enforcement constraints. The implicit assumption is that it would be a waste of resources to include unverifiable or unenforceable provisions in contracts. This natural supposition is falsified in instances where the optimal contract is overspecified. First best is implemented by violating the stipulated contract, that hence necessarily contains unenforceable prescriptions.

I formulate a general model of contractual enforcement. Instead of focusing on contracts that only include enforceable clauses, I do not place any restrictions on the space of contracts available for stipulation. A contract $t$ is defined as an unconstrained collection of transfers contingent on the agents' actions and on all information available to them. I introduce an 'enforcement function' $F$, that maps stipulated contracts $t$ into enforced contracts $F \circ t$. I use this framework to ask: Under which conditions is the restriction to contracts whose clauses are all enforceable without loss?

Theorem 1 identifies a simple sufficient condition. Restricting attention to enforceable contracts is without loss when the enforcement function $F$ is idempotent, that is, $F \circ F \circ t = F \circ t$ for all contracts $t$. I show that this condition holds for functions $F$ that represent standard assumptions in contract theory (Proposition 1). Specifically, suppose that Courts process information just like any standard Bayesian agent, and that there are no legal constraints to contractual enforcement. Then the only enforcement constraint embedded in $F$ is that, regardless of the stipulated contract $t$, the transfers in the enforced contract $F \circ t$ depend only on signals verifiable in Court. Suppose now the contract $F \circ t$ is stipulated. By construction, all clauses included in $F \circ t$ are verifiable, and hence are enforceable in Court. The enforced contract $F \circ F \circ t$ equals $F \circ t$, and $F$ is idempotent. The same reasoning applies to the case in which the only legal constraint is limited liability: the enforced transfers cannot reduce any agent's utility below an exogenous lower bound.

A necessary condition for the optimality of enforceable contracts is characterized in Proposition 3. Its formulation is involved. I describe it here in a simplified counter positive form. Suppose that the first best strategy profile $s^*$ may only be implemented with a contract that includes unenforceable provisions. Then, there must exist an agent $i$ who cannot be deterred from playing a strategy $s_i$ different from $s_i^*$ with any contract prescribing that she plays $s_i^*$. Further, there must exist a strategy $s_i'$ such that the contractual prescription to play $s_i'$ deters agent $i$ from playing $s_i$ but not $s_i^*$. It is intuitive that there is a violation of transitivity: prescribing $s_i'$ does not deter $s_i^*$, and prescribing $s_i^*$ does not deter $s_i$, but prescribing $s_i'$ does deter $s_i$. Proposition 3 makes this intuition precise.

This necessity result is used to shed light on the situations identified above, in which the first best may only be implemented with overspecified contracts, i.e., contracts that include unenforceable prescriptions. In all these cases, the reason for the optimality of overspecified contracts is identified in a legal or technological contractual enforcement constraint. It is intuitive that slack enforcement makes deterrence intransitive. In the speeding example, setting the limit at $s^*$ mph does not deter from driving at a speed above $s^*$ and below $s^*+10$ mph. Any such a speed is deterred by setting limits at $s^* - 10$ mph, which does not deter driving at the optimal speed $s^*$ mph. A similar reasoning identifies the legal provision that rules out transfers in excess of verified damage as the source of the optimality of overspecified contracts in commerce and in professional services.

The characterization in Proposition 3 may be also related to costs associated with evidence verification in Court. As was earlier suggested by Geanakoplos (1989) and Shin (1993), the assumption that Courts process information in the same way as standard Bayesian agents is not always defensible. Courts produce decisions based on hard evidence, and not on personal opinions or equilibrium beliefs. As I will explain in Section 5 in details, evidence verification costs may lead to violations of an 'axiom of positive introspection,' and this leads to intransitive deterrence. Specifically, it may be that Courts cannot verify that an agent $i$ did not play $s_i'$ when she played $s_i^*$, nor that $i$ did not play $s_i^*$ when she played $s_i$, but it is verifiable that $i$ did not play $s_i'$ when she played $s_i$. Deterrence is intransitive, and the first best strategy profile $s^*$ may only be implemented with an overspecified contract prescribing that agent $i$ plays $s_i'$.

A similar logic explains how individual liability may lead to intransitive deterrence even when Courts reason as standard Bayesian agents. Consider the case of multi-agent production with peer monitoring described earlier. In case of low quality output, the agent $i$ who shirked cannot be identified in Court. But say that agents contractually committed to both exert productive effort and monitor each other. In case of low output, each agent $i$'s breach of contract is verified in Court. Even if agent $i$ did not shirk in the production, she is held accountable for not monitoring other agents' actions. Thus, although agents' liability is individual, overspecified contractual commitments make them jointly liable in case of low output. Moreover, overspecified commitments do not deter first best play. When agents exert high effort, the fact that they do not monitor each other is not verifiable.

The next section relates my findings to the extant contract-theoretic literature. Section 3 makes precise the earlier described instances where overspecified contracts are optimal. The general model of contract enforcement is presented in Section 4, and the analysis follows in Section 5.

## 2 Literature Review

While this paper studies when optimal contracts are overspecified, contract theorists since Williamson (1985), Grossman and Hart (1986), and Hart and Moore (1990) have devoted much attention to the opposite possibility of contract incompleteness. Conceptually closest to my model, Bernheim and Winston (1998) investigate when the optimal contracts exclude verifiable prescriptions. To the contrary, optimal contracts include unenforceable and unverifiable clauses in my analysis.

The idea that overspecified contracts may be useful to provide incentives in commerce may be related to insights in the literature on relational contracts that build on Bull (1987), and MacLeod and Malcomson (1989). Specifically, Schmidt and Schnitzer (1995) suggest the possibility that formal contracts may complement informal agreements in repeated interactions. The idea is that the overspecified clauses of a formal contract may be ignored in practice, and used only if the relationship between buyer and seller ends. My contribution here is entirely distinct however, because I consider non-repeated interactions.

My analysis of overspecified contracts may bear some resemblance with mechanisms based on contract renegotiation. Renegotiation of an incomplete contract after uncertainty has resolved is understood as a simple means to achieve desirable outcomes (Huberman and Kahn, 1988). Here however, there is no resolution of uncertainty, nor contract renegotiation. Overspecified contracts are signed and then breached on the equilibrium path.

A sizeable literature in law and economics describes various rules Courts use to enforce contractual transfers and the consequent incentive effects (see, for example, Kaplow, 2000). Legal restrictions such as the ones I consider here need not always be suboptimal. Aghion and Bolton (1987) demonstrate the anti-competitive effects of allowing for punitive damage provisions in contracts. Besides being detrimental to defendants, Kornhauser and Revesz (1994) show that the joint and several liability rule may also be detrimental to plaintiffs because it may stifle or complicate out-of-Court settlement.

The question of which liability rule is economically efficient in different environments has also received significant attention.[4] The idea that contracts may be overspecified in order to overcome restrictive transfer rules has not be sufficiently fleshed out. An exception is the work by Edlin (1997) who documents how overspecified contracts are used to solve hold up problems generated by legal rules that void punitive transfers clauses in contracts. A seller's unverifiable investment determines the quality of a good. After the purchase,

[4] Anderlini, Felli and Postlewaite (2007) take the different approach of studying optimal rules of incomplete contract enforcement, instead of taking the possible enforcement rules as exogenous.

a buyer may make an unverifiable investment, whose returns depend on the quality of the good. The good's quality is verifiable ex-post. But because punitive transfers are ruled out, stipulating optimal quality good provision in the sale contract leads to quality underprovision.

A solution of this hold up problem features an overspecified contract in which the seller commits to deliver a good of excessively high quality. In equilibrium, the seller breaches contract, and gives the buyer a discount for the lower quality provision. In order to compensate the buyer for the expected lower sale price, the seller pays the buyer an upfront fee. While the motivation and the logic of our works is similar, there are some key technical differences. Most importantly, the overspecified contracts described by Edlin (1997) are enforceable, according to my paper's definition, whereas the overspecified contracts that I consider are not. I will explain these differences in details in Section 3.

Of course, mine is not the only paper suggesting peer monitoring as a possible means to alleviate moral hazard in teams (for example, see Che and Yoo, 2001). While not explicitly discussed, individual liability is often represented by the assumption that the principal stipulates a single bilateral contract with each agent (e.g., Prat and Rustichini, 2003, Segal and Whinston, 2003). But in all these papers, peers monitor each other in equilibrium with positive probability, so that some effort is sub-optimally diverted from output production. Here instead, peer monitoring is used only as a latent threat to make agents exert productive effort.

# 3 Overspecified contracts

This section describes examples where the optimal contract is overspecified. The first example is loosely inspired by the casual observations about speed limits presented in the introduction. Enforcement of contractual clauses is constrained by fixed administrative costs and leniency considerations. As a result, the optimal social contract sets limits that are tighter than optimal, in the expectation that individuals will violate these limits and first best will be achieved.

**Example 1 (Slack enforcement).** Consider a simple free-rider problem. Each player $i = 1, ..., n$ chooses a contribution $c_i \geq 0$, and payoffs are symmetric across players: for any $i, j$, $u_i(c_i, \mathbf{c}_{-i}) = u_j(c'_j, \mathbf{c}'_{-j})$ if $c_i = c'_j$ and $\mathbf{c}_{-i} = \mathbf{c}'_{-j}$. The society's welfare $W(\mathbf{c})$ is symmetric across contributions $c_i$, and maximal when $c_i = c^*$ for all $i$. However, each player $i$'s payoff $u_i(c_i, \mathbf{c}_{-i})$ is decreasing in $c_i$ for any opponents' contributions $\mathbf{c}_{-i}$.

Consider social contracts that include a prescription that each player must make at least a contribution $\hat{c}$. Let $b$ denote the (possibly negative) net social benefit of enforcing a sanction $p > 0$ on a player $i$ who shirks on her social duties and plays $c_i < \hat{c}$. The net social benefit $b$ is a continuous function of the contract violation $d_i = c_i - \hat{c}$, of the sanction level $p$, and possibly also of $\mathbf{c}_{-i}$ and $\hat{c}$. It weighs social benefits of enforcement such as deterrence and revenue collection, against administrative costs, leniency, and other social costs.

A sanction $p$ is not enforced unless the net social benefit $b$ is positive. For any contract violation $d_i = \hat{c} - c_i$ and opponents' contributions $\mathbf{c}_{-i}$, consider the sanction levels $p$ that would deter $d_i$, i.e., $p \geq u_i(c_i, \mathbf{c}_{-i}) - u_i(\hat{c}, \mathbf{c}_{-i})$. To represent slack enforcement, I assume that there is an "enforcement margin" $m > 0$, possibly function of $\hat{c}$ and $\mathbf{c}_{-i}$, such that $b(d_i, p, \hat{c}, \mathbf{c}_{-i}) > 0$ for some $p$ that would deter $c_i$ if and only if $d_i > m(\hat{c}, \mathbf{c}_{-i})$.[5] Although the society supports sanctions on large violations of the social contract, leniency considerations and fixed administrative costs dominate the social benefits of sanctioning smaller violations.

As a result, the first best outcome $\mathbf{c}^*$ is not implemented by setting the minimal contribution $\hat{c} = c^*$ in the social contract, but by stipulating the minimal contribution $\hat{c} = m(\hat{c}, \mathbf{c}^*_{-i}) + c^*$. (If $m$ is a function of $\hat{c}$, such a minimal contribution $\hat{c}$ exists as long as there is a uniform bound $\delta < 1$ such that $\partial m(\hat{c}, \mathbf{c}^*_{-i})/\partial \hat{c} \leq \delta$ for all $\hat{c} \geq c^*$.) In equilibrium, all players $i$ violate the excessive stipulation $\hat{c}$, and contribute exactly $c_i = c^*$. Thus, in this case, first best can only be implemented by means of an overspecified social contract that is not enforceable, and that is indeed violated in equilibrium. $\diamond$

The second example is more involved. It is motivated by laws that rule out 'punitive' contractual transfers. Even if a contract breach can be verified, the breaching party usually cannot be sanctioned in excess of the verified harm, or the foregone profits suffered by her counterparts because of her breach.

Such laws may be the reason why overspecified contracts are not uncommon in consulting and other similar professions. These contracts usually include clauses such as "provision of service according to best practices," or other overly demanding stipulations that are difficult to verify. Typically, it is not easy to assess how effectively a consultant conducts research about his client's company and market, or how useful his meetings with the company's customers and suppliers are. It may be even more difficult to quantify precisely the

[5]This assumption holds under mild conditions, e.g. when the function $b$ strictly increases in $d_i$, there exist $\varepsilon > 0$ for which $b(d_i, p, \hat{c}, \mathbf{c}_{-i}) < 0$ for all $d_i < \varepsilon$, $p$, $\hat{c}$, $\mathbf{c}_{-i}$, and $b$ does not decrease too fast in $p$ at $p = \max\{p' : b(d_i, p', \hat{c}, \mathbf{c}_i) \geq 0\}$, specifically $-\frac{\partial b(d_i,p,\hat{c},\mathbf{c}_{-i})/\partial d_i}{\partial b(d_i,p,\hat{c},\mathbf{c}_{-i}/\partial p} \geq -\frac{\partial u_i(c_i,\mathbf{c}_{-i})}{\partial c_i}$ at $c_i = \hat{c} - d_i$. (Note that by limited liability, $b(d_i, p, \hat{c}, \mathbf{c}_{-i}) < 0$ for all $p > u_i(c_i, \mathbf{c}_{-i}) - \underline{u}$, for some given $\underline{u}$). Details are in Appendix.

effect of professional advice on a company's metrics such as profit, market shares and so on.

Ostensibly, the reason for including unrealistically demanding clauses in professional contracts is to put the client in the strongest possible position if she is not satisfied with the consultant's advice. Overspecified contracts are a form of insurance for the client that is needed in these situations because penalties are bounded by law. Else, it would just be possible to include punitive transfers in the consultant's contract.

The next example provides a formal analysis of overspecified contracts in consulting and similar professions. The analysis is deliberately kept simple, so as to fit in this paper as an example.

**Example 2 (Damage compensation)** An agent chooses effort $e_i \in [0, \bar{e}]$, in two separate tasks $i = A, B$ to produce value for a client. The cost of efforts $\mathbf{e} = (e_A, e_B)$ is denoted by $c(\mathbf{e})$, symmetric in $e_A$ and $e_B$, twice differentiable, increasing and strictly convex in both arguments, with boundary conditions $c(\mathbf{0}) = 0$, $\partial c(\mathbf{0})/\partial e_i = 0$, for both $i = A, B$. Towards the formulation of valuable advice, for instance, consultants need to conduct several separate tasks, such as meeting with the client's employees, external suppliers and clients, and researching facts about the company, industry and market.

The value $v_\omega(\mathbf{e})$ produced by the agent is uncertain ex-ante. In addition to the efforts $e_A, e_B$, it depends also on an unknown state of the world $\omega \in \{0, 1\}$ such that $\Pr(\omega = 1) = q$. For both $\omega = 0, 1$, the value function $v_\omega(\mathbf{e}) > 0$ is symmetric in $e_A$ and $e_B$, twice differentiable, increasing and strictly concave in $e_A$ and $e_B$, with $0 < \partial v_\omega(\bar{\mathbf{e}})/\partial e_i < \partial c(\bar{\mathbf{e}})/\partial e_i$, for $i = A, B$. A substantive assumption is that efforts in the two tasks are complement: $\partial^2 v_\omega / \partial e_A \partial e_B > 0$. As is intuitive in consulting for example, well-conducted meetings would not lead to valuable advice in the absence of adequate research about the client's company, market and industry. To simplify notation, set $v(\mathbf{e}) \equiv (1-q)v_0(\mathbf{e}) + q v_1(\mathbf{e})$.

The first-best levels of effort $\mathbf{e}^* = (e^*, e^*)$ is symmetric and solves $\max_{\mathbf{e} \geqq 0} v(\mathbf{e}) - c(\mathbf{e})$. It is pinned down by the first-order condition, for either $i = A, B$,

$$\partial v(\mathbf{e}^*)/\partial e_i = \partial c(\mathbf{e}^*)/\partial e_i. \tag{1}$$

I assume that $v(\mathbf{e}^*) > c(\mathbf{e}^*)$: the optimal effort $\mathbf{e}^*$ yields a value to the client that is sufficient to remunerate the agent, assuming quasi-linear utilities as customary.

While the value produced by the agent is non-verifiable, the Court may verify the agent's effort in performing tasks $A$ and $B$. In the case of consulting, for example, it is usually very hard to quantify the value of professional advice to the client. But it may be possible in some cases to document and assess a consultant's preparatory work more accurately. Specifically, I assume that effort levels are never verified when $\omega = 0$, whereas each effort $e_i$ is verified with probability $p \in (0,1)$ when $\omega = 1$, independently across $i = A, B$.[6]

Punitive contractual transfers cannot be enforced. The agent cannot be sanctioned in excess of the damage suffered by the client when the agent breaches contract. Suppose that a simple contract is stipulated, in which the agent commits to exert some levels of effort $\hat{\mathbf{e}} = (\hat{e}_A, \hat{e}_B)$ in exchange for a fee $f$. The analysis in Section 5 considers general contracts defined as a schedule of remunerations for exerted effort. Such general contracts include the possibility of bonuses for producing value above $v_\omega(\hat{\mathbf{e}})$.

If the agent commits to efforts $\hat{\mathbf{e}}$ and then exerts efforts $\mathbf{e} = (e_A, e_B)$, the expected damage compensation paid to the client is:

$$d(\mathbf{e}; \hat{\mathbf{e}}) = qp^2 \max\{0, v_1(\hat{\mathbf{e}}) - v_1(\mathbf{e})\} + qp(1-p) \sum_{i=A,B; j \neq i} \max\{0, v_1(\hat{\mathbf{e}}) - v_1(e_i, \hat{e}_j)\},$$

so that the expected agent's net remuneration, ex-ante, is $f - d(\mathbf{e}; \hat{\mathbf{e}})$.[7] The function $d(\mathbf{e}; \hat{\mathbf{e}})$ embeds the presumption that if the Court cannot verify the agent's effort level $e_i$ for a task $i$, it will presume that the agent fulfilled her contract obligation and exerted effort $e_i = \hat{e}_i$.

Because punitive transfers are ruled out, I will now show that first best cannot be implemented with a contract in which the agent commits to exert the optimal efforts $\mathbf{e}^*$. Nevertheless, the agent can be induced to exert the efforts $\mathbf{e}^*$ by committing to some "overly demanding" levels of effort $\hat{\mathbf{e}} > \mathbf{e}^*$, that are calibrated so that the expected damage compensation $d(\mathbf{e}; \hat{\mathbf{e}})$ makes the agent exert the first best efforts $\mathbf{e}^*$.

In fact, given any contractually committed levels of effort $\hat{\mathbf{e}}$, the agent chooses $\mathbf{e} = (e_A, e_B)$ so as to solve $\max_{\mathbf{e} \geqq 0}[f - d(\mathbf{e}; \hat{\mathbf{e}}) - c(\mathbf{e})]$. The agent's choice $\mathbf{e}$ is pinned down by the first order condition

$$-\partial d(\mathbf{e}; \hat{\mathbf{e}})/\partial e_i = \partial c(\mathbf{e})/\partial e_i. \tag{2}$$

If the agent contractually committed to the first best efforts $\mathbf{e}^*$, it is easy to verify that her chosen effort levels $\mathbf{e}$ would be smaller than $\mathbf{e}^*$. For $\hat{\mathbf{e}} = \mathbf{e}^*$ and any $\mathbf{e} < \mathbf{e}^*$, the marginal remuneration of effort $e_i$ is: $-\partial d(\mathbf{e}; \mathbf{e}^*)/\partial e_i = qp^2 \partial v_1(\mathbf{e}^*)/\partial e_i + q(1-p)p \partial v_1(e_i, e_j^*)/\partial e_i$. In

[6]This example can be elaborated in several ways, without changing the main results. For instance, the value of service can be made partially verifiable, or one can consider more nuanced forms of partial verifiability of effort, as well as asymmetric value and cost functions.

[7]While it is not uncommon that agents such as consultants are sued by their clients, the net payments $v_1(\hat{\mathbf{e}}) - v_1(\mathbf{e})$ and $v_1(\hat{\mathbf{e}}) - v_1(e_i, \hat{e}_j)$, for $i = A, B$, $j \neq i$, need not necessarily represent Court sanctions in a trial or out of Court settlements. They may also be interpreted as "fee discounts" paid by the agent to a client who can verify that the agent's shirked his contractual levels of effort $\hat{\mathbf{e}}$.

the limit for any $\mathbf{e}$ approaching $\mathbf{e}$ from below, the marginal renumeration of effort $e_i$ is:

$$-\lim_{\mathbf{e}\uparrow\mathbf{e}^*} \partial d(\mathbf{e};\mathbf{e}^*)/\partial e_i = qp\partial v_1(\mathbf{e}^*)/\partial e_i < \partial c(\mathbf{e}^*)/\partial e_i,$$

where the inequality is a consequence of the facts that $\mathbf{e}^*$ solves (1), $\partial v_0(\mathbf{e}^*)/\partial e_i > 0$, $0 < q < 1$ and $0 < p < 1$. As the agent's marginal remuneration of effort $e_i$ is smaller than the marginal cost for all $\mathbf{e} < \mathbf{e}^*$ close to $\mathbf{e}^*$, the agent's choice efforts $\mathbf{e}$ must be such that $e_i < e_i^*$ for both $i = A, B$.

However, first best efforts $\mathbf{e}^*$ are implemented by having the agent contractually commit to the symmetric effort levels $\hat{\mathbf{e}} = (\hat{e}, \hat{e}) \neq \mathbf{e}^*$ that satisfy the above first-order condition (2) when $\mathbf{e} = \mathbf{e}^*$:

$$-\partial d(\mathbf{e}^*;\hat{\mathbf{e}})/\partial e_i = qp[p\partial v_1(\mathbf{e}^*)/\partial e_i + (1-p)\partial v_1(e_i^*, \hat{e}_j)/\partial e_i] = \partial c(\mathbf{e}^*)/\partial e_i, \tag{3}$$

in exchange of a fee $f$ such that $c(\mathbf{e}^*) < f - d(\mathbf{e}^*, \hat{\mathbf{e}}) < v(\mathbf{e}^*)$. Because $\partial v_0(\mathbf{e}^*)/\partial e_i > 0$ for both $i = A, B$, $0 < q < 1$, $0 < p < 1$, and the efforts are complement across tasks, the effort levels $\hat{\mathbf{e}}$ are in excess of the first best effort levels: $\hat{\mathbf{e}} > \mathbf{e}^*$.

As in Example 1, first best can only be achieved by means of an overspecified contract. The agent commits to effort levels $\hat{\mathbf{e}} > \mathbf{e}^*$, and then implements the desired outcome (choosing efforts $\mathbf{e}^*$) by breaching the contract.[8]

I conclude Example 2 with a remark on a key assumption of the above analysis. When effort $e_i$ in task $i$ is not verified, the Court presumes that the agent fulfilled her contractual obligation and exerted effort $e_i = \hat{e}_i$, independently of what is verified about effort $e_j$ in task $j$. This is not the presumption most favourable to the agent. Suppose that the Court presumes that the agent acted in the client's best interest, unless the contrary is proved. Then upon verifying an insufficient effort $e_j < \hat{e}_j$ in task $j$, the Court would presume that the agent compensated the insufficient effort $e_j$ by overproviding effort $e_i > \hat{e}_i$ in task $i$, so as to minimize the client's damage $v_1(\hat{\mathbf{e}}) - v_1(\mathbf{e})$.[9]

As in the analysis above, first best cannot be implemented in this case with the contract by which the agent commits to exert optimal efforts $\mathbf{e}^*$. And again, first best can be implemented with an overspecified contract. I prove in the appendix that efforts $\mathbf{e}^*$ are

[8] The existence and uniqueness of effort levels $\hat{\mathbf{e}}$ that satisfy (3), and the fact that the agent chooses effort levels $\mathbf{e}^*$ when contractually committing to $\hat{\mathbf{e}}$, follow from concavity of $v_1$, convexity $c$, together with the boundary condition: $qp[p\partial v_1(\mathbf{e}^*)/\partial e_i + (1-p)\partial v_1(e_i^*, \bar{e}_j)/\partial e_i] \geq \partial c(\mathbf{e}^*)/\partial e_i$. If such a condition fails, it is optimal that the agent commits to the maximum effort levels $\bar{\mathbf{e}}$. While first best is not achieved, the optimal contract is still overspecified.

[9] Indeed, it may be that the agent underprovided effort $e_j$ in task $j$ only because of unforeseen contingencies beyond the agent's control.

achieved when the agent commits to the "overly demanding" symmetric efforts $\hat{\mathbf{e}} = (\hat{e}, \hat{e}) > \mathbf{e}^*$ that solve $v_1(\hat{\mathbf{e}}) = v_1(e^*, \bar{e})$. $\diamond$

As already pointed out in Section 2, Example 2 may be related to the work by Edlin (1997) on how overspecified contracts may be used to solve hold up problems in commercial transactions. But there are some important modelling differences. The optimal contracts in Edlin (1997) are enforceable. The seller (agent) commits to provide quality above the optimal one, and the value of its provided good is verifiable. On the equilibrium path, the seller breaches contract, and needs to compensate the buyer (principal) for the "verified damages." The contract "overspecification" only manifests in the feature that the quality contractual commitment is excessive and sub optimal. Instead, the contracts that parties optimally choose to adopt in my Example 2 are not enforceable. In exchange for a fee, the agent commits to exert excessive levels of effort. This contractual commitment is unenforceable because effort is not always verifiable.

The final example of this section is based on a broad legal principle, to which I refer to as 'individual liability,' here. A contractual party subject to individual liability may not be held responsible for the actions of others. She may be sanctioned in Court only if it is verified that she violated her own individual contractual obligations. Consider a situation in which multiple agents' efforts are needed for the provision of output to a client, but it is difficult to distinguish individual agents' inputs. For example, the joint efforts of a carpenter, a plumber, and an electrician may be needed for a house renovation project. Individual liability constraints make it hard to motivate individual agents to exert high efforts. A possible solution is for the agents to form a company, or to hire an intermediary, that acts as the sole liable legal entity vis-a-vis any contractual counterpart. But this entails non-trivial legal costs, and need not always be efficient.[10]

Example 3 below describes a different possible solution. Agents sign contracts by which they not only commit to exert high effort, but also to monitor each others' efforts. These overspecified contracts make each agent indirectly responsible for the other agents' efforts, thereby circumventing individual liability constraints. Furthermore, each agent expects that all agents exert high effort in equilibrium, so that none of them needs to waste time into monitoring her peers. As is the case in Example 1 and 2, first best is implemented only with a contract that the agents breach in equilibrium. After presenting the example, I will briefly discuss economic environments to which the suggestions uncovered here may apply.

[10]Incentivizing multiple agents is easier in environments in which contractual enforcement is based on joint and several liability. Their principal may fully recover her damages from the agents upon demonstrating breach of contract, without the need to identify the agent who shirked her contractual obligations.

**Example 3 (Individual liability)** I begin by laying down a simple multi-agency problem. A client may hire two agents for the joint provision of output. The agents' individually unverifiable efforts jointly determine output quality. Quality is verifiable in Court, and it is high if and only if both agents exert high effort (action $H$). In the first best, the client hires the two agents, and they both play $H$. However, each agent has a private incentive to shirk and exert low effort (action $L$).

If the agents are jointly liable vis-a-vis the client, first best is simply implemented with a contract that sanctions both agents if output quality is low, regardless of which agent shirks. The case in which joint commitments are not legally or practically enforceable is more interesting. Each agent cannot be sanctioned unless she is verified in breach of her own individual contractual obligations, and hence effort underinvestment cannot be effectively deterred. Suppose that agent $i$ contractually commits to play $H$, but then breaches the agreement and plays $L$. If taken to Court, agent $i$ can defend herself by blaming agent $j \neq i$ for the low output. Since individual effort cannot be identified, there is no way to show that agent $i$ breached her contract.[11]

Let us now elaborate this example, and suppose that each agent $i$ can choose to monitor the other agent $j$ at some cost (action $M$), or choose not to monitor her ($N$). The choice of monitoring, like productive effort, is a private action that cannot be directly verified. Nevertheless, if agent $i$ monitors agent $j$, then she is able to document that agent $j$ exerted low effort, whenever this is the case. Specifically, there is a binary verifiable signal $s_i$, such that $s_i = 1$ if and only if agent $j$ plays $L$ and agent $i$ plays $M$, and otherwise $s_i = 0$. For instance, this may represent agent $i$ gathering evidence that $j$ exerted low effort in a report, without including any evidence that $j$ worked hard. If $i$'s report fails to document that $j$ exerted low effort, it can be either because $j$ worked hard or because $i$ did not spend enough effort in preparing the report. I will demonstrate that, when available, this monitoring technology makes it possible to implement first best, by means of an overspecified contract.[12]

The verifiable information associated with the agents' choices is described in Figure 1. For future reference, note first that upon observing high quality output, all that a Court can verify is that the agents' action profile belong to the upper-left box $B$. Both agents played $H$ and it cannot be verified whether either played $M$ or $N$. Second, suppose a Court

[11]Discretionary bonuses for joint high effort also fail to achieve first best. This is not a repeated game, and hence the client has no reason to pay bonuses for which she has no contractual obligation. And if payments of joint "bonuses" were included in a contract, their withholding would be a violation of individual liability. The detailed analysis of bonuses and general contracts is in Section 5.

[12]Instead, first best cannot ever be implemented if agents' monitoring efforts are directly verifiable, as I will show later.

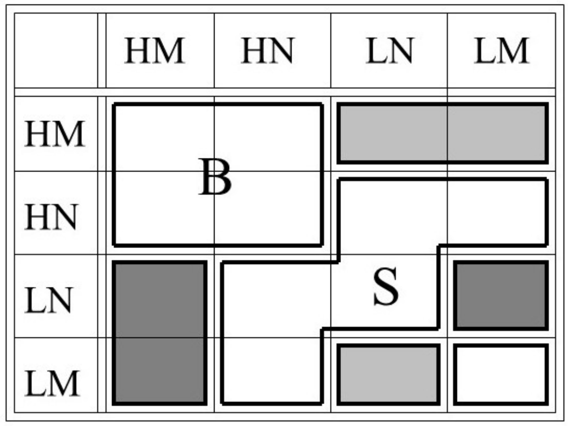


Figure 1: verifiable information in Example 3

observes low output quality and signals $s_1 = s_2 = 0$. This is the case when neither agent $i$'s report can document that $j$ played $L$. Then, the Court can only verify that the agents' actions belong to the step-like shaped set $S$. Importantly, it can be verified that neither agent $i$ played $HM$.

In the first-best outcome, the agents play $(HN, HN)$: they exert high efforts without wasting resources monitoring each other. However, as was the case in the simpler multi-agency problem above, a contracts by which each agent commits to high effort without monitoring the other agent fails to deter underinvestment. Again, if an agent breaches her contract and shirks, there is no way to verify that she exerted low effort.

The first-best outcome $(HN, HN)$ is implemented if and only if each agent contractually commits to action $HM$: in addition to exerting high effort, each agent commits to monitoring the other agent. In equilibrium, each agent secretly breaches her contract and plays $HN$, so that high quality output is produced without wasting time on monitoring. The agents' contract violations cannot be verified in Court: since output is high, no agent can be blamed for exerting low effort, and the absence of monitoring cannot be verified either.

Moreover, the contracts with which each agent commits to action $HM$ are effective in deterring effort underinvestment. Suppose that agent $i$ deviates from the profile $(HN, HN)$ to play $LN$. Then, the client is able to verify in Court that agent $i$ did not play $HM$ and hence breached her individual contract. Output quality is low and signal $s_i = 0$, implying that agent $i$'s report is ineffective in showing that the other agent $j$ has shirked in production. Thus, it must be case that either agent $i$ did not monitor agent $j$, or that agent $i$ is the one who exerted low effort. In either case, it is verified that agent $i$ breached her own contract, as she committed to both exert high effort and to monitor agent $j$.[13]

As in Example 1 and 2, first best may be implemented only with an overspecified contract which is breached in equilibrium. As anticipated earlier, I conclude the example by showing that first-best cannot be implemented with a monitoring technology with which the Court verifies directly whether an agent has monitored the other agent or not. In this case, if an agent $i$ does not monitor agent $j$, then the client can verify that $i$ breached her contract by not monitoring, and can potentially collect a fine. Thus, only a second-best outcome is implemented in case the Court can verify agents' monitoring: both agents commit to action $HM$, exerting high effort and monitoring each other, and then fulfill their contractual obligations to avoid being found in breach.

The same reasoning shows that the first-best cannot be implemented with any contract by which agents commit to a wasteful verifiable action unrelated with output. The inclusion of monitoring in the contract does not just serve the role of "an excuse" to sanction agents when they shirk unverifiable efforts. It is important that the contract permits sanctioning the agents if and only if output is low. $\diamond$

While presenting detailed economic applications of Example 3 is out of scope of this paper, it is easy to think of cases and anecdotes in which agents' cooperation and effort are incentivized through the instruction that they monitor each other. One such anecdote can be taken from university life. One of the aims of the students' code of honor is to prevent cheating by copying from others. However, it is usually difficult to directly observe whether a student cheats, and cheating is most often indirectly verified when two or more students turn in identical exam or homework papers. But in this event, it is impossible to establish who copied from whom. The aim of preventing cheating is implemented by overspecifying the code of honor (i.e. the students' "contract"). Students are not only asked not to copy from others, but also to prevent others from copying their work. Students are penalized only if there is evidence that cheating has occurred, and in practice, little effort is spent in preventing others from copying one's work.

From the corporate world, consider the "Collaborative Anytime Feedback" policy of peer monitoring that has been employed by Amazon already for a few years. As reported in a series of articles on the New York Times in August 2015, peer monitoring is a institu-

[13]In fact, both signals $s_1$ and $s_2$ are equal to zero, so that the Court verifies that neither agent played $HM$. Although each agent is only individually liable, both agents are penalized off the equilibrium path when one of them exerts low effort (as is the case when they are jointly liable). This occurs because the client can independently verify that both agents breached their contractual obligations. Under joint liability, it would be possible to sanction both agents if the client could show that at least one of them breached contract.

tionalized practice at Amazon:

> "The Anytime Feedback Tool is a widget in the company directory that allows employees to send praise or criticism about colleagues to management. (While bosses know who sends the comments, their identities are not typically shared with the subjects of the remarks.)" The New York Times, 15 August 2015

Amazon workers' package preparation tasks are a form of joint production, because Amazon relies on online reputation for keeping standards of fast and precise goods' delivery, and loses more from a negative review than it gains from a positive review. Workers operate in open spaces and can monitor each other. There is evidence that a minority of workers actively engage in monitoring each other, or even pervert peer monitoring to sabotage other workers. But by and large, it appears that peer monitoring serves mostly as a latent threat for shirking, as I detailed in Example 3.[14]

Professional ethic codes are another manifestations of overspecified "contracts." Medical ethical codes often set unrealistically high standards, to which doctors do not abide literally in practice. Instances arise in which a patient suffers from a condition that requires separate therapies performed by different doctors. Unless all therapies are performed effectively, the patient's condition will not improve. In these instances, none of the doctors involved is legally liable for the other doctors' actions. But her ethical code includes a duty that she monitors the effectiveness of the other doctors' therapy. In practice, the main effort is placed in administering therapy, rather than on monitoring the other doctors' decisions.

This section has presented examples in which contractual parties implement first best by stipulating overspecified contracts that they breach in equilibrium. The next section presents a general model of contractual enforcement to investigate the conditions under which overspecified contracts are needed to implement first best.

# 4 A general model of contract enforcement

Consider an economic interaction represented as a Bayesian game $G = (I, \Omega, \rho, A, \Theta, u)$: $I$ is the set of players, $\Omega$ is the set of states, $\rho$ is the move of nature over $\Omega$, and for each player $i$, $A_i$ is the set of actions, $\Theta_i$ (a partition of $\Omega$) is a set of types, $u_i$ is the payoff function over pairs $(a, \omega)$, where $a$ is an action profile and $\omega$ is a state of nature, so that $A = \times_{i \in I} A_i$ is the action space, $\Theta = \times_{i \in I} \Theta_i$ is the type space, and $u = (u_i)_{i \in I}$ is the payoff

[14] Amazon can also directly monitor workers, for example, by videotaping and inspecting their activities. But these forms of direct monitoring are costly, and their feasibility that does not diminish the role of peer monitoring in increasing workers' efficiency.

function.[15] I assume that $I$ is finite, and that $\Omega$ and $A_i$ are compact subsets of Euclidean spaces, for all $i \in I$. For example, they may be finite, or closed intervals. For each player $i$, the strategy set is $S_i$ is the set of strategies $s_i : \Omega \to A_i$ measurable with respect to $\Theta_i$, i.e., such that for each type $\theta_i \in \Theta_i$, the strategy function $s_i$ is constant on $\theta_i$. The interim payoff of player $i$ of type $\theta_i$ who plays $s_i$ is $u_{\theta_i}(s_i, s_{-i}) \equiv \sum_{\omega \in \theta_i} u_i(s_i(\theta_i), s_{-i}(\omega); \omega)\rho(\omega|\theta_i)$, where the profile of opponents' strategies is given by $s_{-i}$.

Contract theory investigates how to implement desirable strategy profiles by means of social or private contracts. The players commit to make transfers to each other, contingent on the realized outcomes $(a, \omega)$, thereby changing the strategic incentives in $G$.[16] It is standard to formulate the contract space so that it includes only enforceable contracts. Typically, including unenforceable clauses into such contracts is considered wasteful. However, the leading examples presented in Section 3 show that this presumption is not always correct. In these examples, first best is implemented when players stipulate contracts that are not enforceable, and breach them in equilibrium.

My general model of contract enforcement does not assume that players only sign enforceable contracts, nor does it place any restrictions on the contract space. Instead, I explicitly describe how stipulated contractual transfers translate into enforced transfers, because of legal or practical enforcement constraints. This allows me to single out enforceable contracts as a subset of all possible contracts, and to test whether or not assuming that players only stipulate enforceable contracts restricts the set of implementable strategy profiles, and whether this entails a welfare loss. I will later show how my model subsumes more standard models of contracts. For ease of exposition, I only consider implementation of pure strategy profiles $s$. It is a simple exercise to generalize my construction to mixed strategies profiles.

In general, a contract may include sanctions against deviations from the strategy profile $s$ that the contract aims to implement, as well as transfers across players to incentivate the play of $s$. Given any game $G = (I, \Omega, \rho, A, \Theta, u)$, I generally define contracts as collections of net transfers among the players, contingent on the realized play $a$ and state $\omega$. For

[15] Because the underlying game $G$ may be expanded so as to include the possibility of sending messages, my model subsumes the message game approach adopted, for instance, by Green and Laffont (1977).

[16] This investigation is formulated either by presuming that the contract is crafted by a player (a 'social planner' or a 'principal') whose type is common knowledge, or that contracts are stipulated ex-ante, before players learn their types. A separate, and largely unexplored question is how players may signal their types by bargaining over contracts that are chosen at the interim stage, i.e., after they learn their types. Some progress has been made for the case of a privately informed principal (e.g., Maskin and Tirole, 1990, 1992, and Mylovanov and Tröger, 2014). The question of signalling through contracts is orthogonal to the scope of this paper and I stay within the simplest formulation where contracts are formulated ex-ante. The distinction between ex-ante and interim contracts is immaterial in the leading examples of Section 3 in any case, because each player's type is common knowledge.

every pair of players $i$ and $j$, a contract $t$ prescribes a net transfer $t_{ij}(a,\omega) \in \mathbb{R}$ from $i$ to $j$ if the realized contingency is $(a,\omega)$. By construction, $t_{ij}(a,\omega) = -t_{ji}(a,\omega)$ and $t_{ii}(a,\omega) = 0$ for all contracts $t$. Given contract $t$, the net aggregate transfers are denoted by $\tau|t : A \times \Omega \to \mathbb{R}^I$ such that $\tau_i(a,\omega)|t \equiv \sum_{j\in I} t_{ji}(a,\omega)$ for every player $i$ and contingency $(a,\omega)$.[17,18] The expected net transfers received by any player $i$ of type $\theta_i$ who plays $s_i$ is $\tau_{\theta_i}(s_i, s_{-i})|t \equiv \sum_{\omega\in\theta_i}[\tau_i(s_i(\theta_i), s_{-i}(\omega);\omega)|t]\rho(\omega|\theta_i)$ if the opponents play $s_{-i}$. For future reference, let $T \equiv \{t \in \mathbb{R}^{I\times I\times A\times\Omega} : t_{ij}(a,\omega) = -t_{ji}(a,\omega)$ and $t_{ii}(a,\omega) = 0$ for all $(a,\omega) \in A \times \Omega\}$ be the set of contracts $t$ associated with game $G$.

Were a contract $t$ to be enforced, the players' payoffs when choosing their strategies would be $u+\tau|t$, and I introduce the notation $G|t = (I, \Omega, \rho, A, \Theta, u+\tau|t)$. In my framework, the stipulated transfers and enforced transfers may differ, as described by an *enforcement function* $F$.

**Definition 1** *Given a game $G = (I, \Omega, \rho, A, \Theta, u)$, an enforcement function $F$ maps every stipulated contract $t \in T$ into an enforced contract $F\circ t \in T$.*

Hence, when stipulating contract $t$ in game $G$ with enforcement function $F$, the players' payoffs are $u+\tau|F\circ t$. I assume that $F\circ t^0 = t^0$, where $t^0$ is the null-contract, for which $t^0_{ij} = 0$ for all $i, j$. I do not place any further restrictions on $F$. I also note that the appropriate enforcement function $F$ may depend on the details of the game $G$ played. For example, it may depend on the payoff functions.[19] Given a game $G$ and enforcement function $F$, a strategy profile $s$ is implemented with a contract $t$ if it is a Bayesian Nash equilibrium of the game $G|F\circ t = (I, \Omega, \rho, A, \Theta, u + \tau|F\circ t)$, i.e., if for all $i$, all $\theta_i \in \Theta_i$ and all $s'_i \in S_i$, $u_{\theta_i}(s) + \tau_{\theta_i}(s)|F\circ t \geq u_{\theta_i}(s'_i, s_{-i}) + \tau_{\theta_i}(s'_i, s_{-i})|F\circ t$.

By considering any possible contract $t$, including contracts that cannot be enforced, and then representing how stipulated contracts are mapped into enforced contracts through $F$, I can single out enforceable contracts as a subset of the set of all possible contracts. Plainly, a contract $t$ is enforceable if all its stipulated transfers are enforced by a Court, despite the constraints imposed by $F$. More weakly, a contract $t$ is enforceable at a specific outcome $(a,\omega)$ if the stipulated transfers $t(a,\omega)$ contingent on $(a,\omega)$ are enforced in Court.

[17] My definitions detail the direction of the transfer between $i$ and $j$, instead of just considering net aggregate transfers, because I want to cover cases like individual liability, in which the identity of the player committing to making a transfer is crucial to assess whether a contractual clause is legal or not.

[18] By construction, each contract $t$ is budget-balanced: $\sum_{i\in I} \tau_i(a,\omega)|t = 0$, for all $t$ and $(a,\omega)$. This is without loss of generality, because any game can always be expanded to include an idle player whose only role in the game is to collect transfers and break the budget constraint.

[19] For instance, contractual transfers that are not justified as compensation of damages are usually not enforced by Courts. Whether a transfer is enforced or not depends on the payoff functions in the game.

**Definition 2** *Given a game $G$ and a enforcement function function $F$, a contract $t$ is enforceable at $(a, \omega)$ if $t(a, \omega) = F \circ t(a, \omega)$, whereas a contract $t$ is enforceable if $t(a, \omega) = F \circ t(a, \omega)$ for all $(a, \omega)$.*

This paper's research question is under which conditions overspecified contracts are needed to implement first best. I call *enforceability principle* the intuitive supposition that they are not needed, so that first best may be implemented with enforceable contracts. I note that the analysis of Section 3 identifies instances where such enforceability principle fails.

**Enforceability Principle.** *Given game $G$, and enforcement function $F$, any implementable strategy profile $s$ can also be implemented by an enforceable contract $t$.*

A weaker version of the enforceability principle only requires that any implementable strategy profile $s$ can also be implemented by a contract $t$ enforceable at any outcome supported by $s$, i.e., such that $F \circ t(s(\omega), \omega) = t(s(\omega), \omega)$ for every state $\omega$. In this form, the enforceability principle is not violated when the only contracts that implement a profile $s$ fails to be enforceable only because $F \circ t(a, \omega) \neq t(a, \omega)$ for some outcomes $(a, \omega)$ off the play of strategies $s$. To make my result stronger, when presenting conditions under which the enforceable principle holds, I will refer to its stronger version, and I will consider the weaker version when identifying instances where the enforceability principle fails (including the leading examples of Section 3).

The next section identifies general conditions under which the enforceable principle holds or fails. I will also relate such conditions to the findings of Section 3, and to more standard contract theoretical environments. As anticipated earlier, I devote the last part of this section to demonstrating that my general model of contract enforcement subsumes more standard contract theory models that assume contracts are enforceable.

An indispensable requirement for a contract to be enforceable is that its prescriptions do not depend on events that the Court cannot verify. Such verifiability constraints are typically modelled by introducing verifiable signals $x$, whose realizations $\hat{x} \in X$ depend possibly stochastically on the outcomes $(a, \omega)$. (Enforceable) contracts are then defined as either pairwise transfers $t$, or directly as aggregate net transfers $\tau$, that are functions of only the verifiable evidence $X$, instead of all the possible outcomes $A \times \Omega$.

Let me demonstrate how my framework subsumes this model. I begin by noting that with simple game theoretical operations, the definitions of states $\omega$ and move of nature $\rho$ of game $G$ can be expanded to include also the verifiable signals $x$ and their stochastic dependence on the outcomes $(\omega, a)$ of the "unexpanded" game $G$.[20] Once the set of all verifiable evidence $X$ is included in the set of states $\Omega$, it is easy to see that modelling contracts as functions of only verifiable evidence $X$ corresponds to a restriction on the set of general contracts $T$ defined earlier.

In fact, the information available to the Court during a trial can be described as a partition $P$ of the outcome space $A \times \Omega$ such that each element $p_{\hat{x}}$ of $P$ corresponds to a signal realization $\hat{x} \in X$, and viceversa. The Court knows that the realized outcome $(a, \omega)$ belongs to $p_{\hat{x}}$, but cannot determine which outcome $(a', \omega') \in p_{\hat{x}}$ has realized. The assumption that contract $t$ be contingent only on verifiable evidence $X$ is then just a measurability condition with respect to $P$. For each element $p_{\hat{x}}$ of $P$, this condition requires that $t(a, \omega)$ be constant $p_{\hat{x}}$. Notably, the analysis in the next section shows that none of the optimal contracts discussed in the leading examples presented in Section 3 would satisfy this restriction.

As I allow for general contracts, that are functions of action profiles $a$ and states $\omega$, verifiability constraints are modelled within my framework by stipulating that a transfer $t_{ij}(a, \omega)$ from a player $i$ to another player $j$ is enforced, $F \circ t_{ij}(a, \omega) = t_{ij}(a, \omega)$, only if $t_{ij}(a, \omega) = t_{ij}(a', \omega')$ for all $(a', \omega') \in P(a, \omega)$. As in more standard models, when the Court cannot verify that the realized contingency for which transfer $t_{ij}(a, \omega)$ is stipulated, it does not enforce $t_{ij}(a, \omega)$. Unlike more standard models, I allow for Courts' information to be represented by non-partitional correspondences $P : A \times \Omega \to 2^{A \times \Omega}$. While non-partitional information may seem esoteric, I will explain in the next section why it may appear quite naturally in the context of Court decisions. I assume throughout the paper that the verification correspondence $P$ is 'truthful:' $(a, \omega) \in P(a, \omega)$ for all $(a, \omega)$. In words, the Court cannot mistakenly conclude that $(a, \omega)$ did not realize when in fact it did.

The enforcement function $F$ does not only represent under what conditions a Court does not enforce stipulated contractual transfers $t_{ij}(a, \omega)$. It also describes what transfers $F \circ t_{ij}(a, \omega)$ are enforced in lieu of the stipulated ones. Depending on the context, the Court may base the determination of such transfers $F \circ t_{ij}(a, \omega)$ on different legal rules. One simple possibility is that the Court voids all transfers that it does not enforce: $F \circ t_{ij}(a, \omega) = 0$

[20] Complemented with the verifiable signals $x$, the Bayesian game $G$ can be represented in extended form with a tree in which, first, nature chooses $\omega$ according to $\rho$ and informs each player $i$ of $\theta_i(\omega)$. Then, each player $i$ chooses $a_i$ without knowing the opponents' choices $a_{-i}$. Finally, nature selects the verifiable realizations $\hat{x} \in X$ according to a distribution $\lambda(\cdot|a, \omega)$. This representation is equivalent to one in which nature moves first selecting a state $\omega$ and a profile of evidence realizations $\hat{\mathbf{x}} = (\hat{x}_a)_{a \in A}$ according to the joint probability distribution $\rho'(\omega, \hat{\mathbf{x}}) = \prod_{a \in A} \lambda(\hat{x}_a|a, \omega)\rho(\omega)$ without informing any player of $\hat{\mathbf{x}}$, then the players are informed of their types $\theta$, choose $a$, and finally they are informed of the verifiable signals $\hat{x}_a$. This extensive form game is equivalent to an "expanded" Bayesian game $G'$ where the state space is $\Omega' = \Omega \times X^A$ and the move of nature is the distribution $\rho'$ defined above.

when $F \circ t_{ij}(a,\omega) \neq t_{ij}(a,\omega)$. I denote this transfer determination rule as the 'voidance rule.' I will later consider more sophisticated 'transfer determination rules,' leading to more complex enforcement functions $F$. For the moment I present an assumption that all such rules need satisfy.

**Assumption 1** *For every game $G$, verification correspondence $P$, and contract $t$, the enforcement function $F$ is such that if $P(a,\omega) = P(a',\omega')$, then $F \circ t(a,\omega) = F \circ t(a',\omega')$.*

The motivation for this assumption is evident. When a contingency $(a,\omega)$ has realized, the Court does not have any information about the realized contingency other than it belongs to $P(a,\omega)$. Two outcomes $(a,\omega)$ and $(a',\omega')$ such that $P(a,\omega) = P(a',\omega')$ are indistiguishable for the Court, and hence the Court must enforce the same transfers.

Verifiability is a necessary condition for enforcing contractual transfers. In the presence of legal constraints, it may be that some verifiable transfers $t_{ij}(a,\omega)$ are not enforced by the Court. Thus it may be that $F \circ t_{ij}(a,\omega) \neq t_{ij}(a,\omega)$ despite $t_{ij}(a,\omega) = t_{ij}(a',\omega')$ for all $(a',\omega') \in P(a,\omega)$. However, the opposite cannot ever happen. While not very realistic, the case in which verifiability is the only requirement for contract enforcement (i.e. all contracts are legal) is an important benchmark for the analysis. I therefore pay special attention to this case in the discussion that follows.

**The case in which all contracts are legal.** In this case, for any game $G$, given the Court's information $P$, stipulated contract $t$, and outcome $x \in A \times \Omega$, a transfer $t_{ij}(x)$ from a player $i$ to another player $j$ is enforced—i.e., $F \circ t_{ij}(x) = t_{ij}(x)$—if and only if $t_{ij}(x) = t_{ij}(y)$ for all $y \in P(x)$.

When $t_{ij}(x) \neq t_{ij}(y)$ for some $y \in P(x)$, the enforced transfer $F \circ t_{ij}(x)$ is based on an underlying transfer determination rule. Depending on the context, different rules may be appropriate. A voidance rule my be characterized simply by: $F \circ t_{ij}(x) = 0$ when $t_{ij}(x) \neq t_{ij}(y)$ for some $y \in P(x)$. A more sophisticated possibility is to invoke a 'conservative transfer determination rule,' under which the Court voids transfers between $i$ and $j$ when it cannot verify who should pay whom, and enforces the smallest possible transfer when it can. Formally, $F \circ t_{ij}(x) = 0$ if $t_{ij}(y) < 0 < t_{ij}(z)$ for some outcomes $y, z$ both in $P(x)$, and $F \circ t_{ij}(x) = \min_{y \in P(x)} t_{ij}(y)$ if $t_{ij}(y) > 0$ for all $y \in P(x)$. In words, the Court cannot verify the precise transfer $t_{ij}(x)$ to $j$ stipulated by $i$ with contract $t$, but it is certain that $i$ has committed to transfer at least $\min_{y \in P(x)} t_{ij}(y)$ to $j$, and the Court enforces that transfer.

The transfers enforced when contingencies are not verified may often depend on which player holds the burden of proof to verify contingencies in Court. For example, suppose that player $j$ is an agent hired by a principal $i$. It is then reasonable to presume that unless the principal verifies that the agent shirked his assignments, the Court will enforce the agent's stipulated remuneration with no penalties. Thus, the principal holds the burden of proof in this case. This case is represented by what I denote as 'seniority transfer determination rules,' defined as follows. The players are ranked with a linear order, so that if $i$ succeds $j$, then $i$ is less favored than $j$ in case the net transfer $t_{ij}(x)$ cannot be verified. When considering a net transfer $F \circ t_{ij}(x)$ from $i$ to $j$, the Court enforces the net transfers most favorable to $j$: $F \circ t_{ij}(x) = \max_{y \in P(x)} t_{ij}(x)$. ◇

While it serves an an important benchmark, the case in which all contracts are legal is not realistic in practice. As I detailed in section 3, there are many ways in which the law restricts the enforcement of contractual transfers. I will formalize such cases within my general framework in the next section. A contractual restriction that usually appears in contract theory is 'limited liability.' This restriction places upper bounds on the (net) transfers that can be enforced on any player. When entering a contractual situation, individuals are not legally allowed to be expropriated of fundamental inalienable rights, and this is despite of what is stipulated in the contract they sign. The case in which the only legal constraint is limited liability is thus of special interest for the analysis. I formalize it below.

**The case of limited liability.** For each player $i$, define a bound $\bar{u}_i$ as the minimum utility enjoyed by $i$ when retaining (only) her inalienable rights. Suppose that $u_i(a, \omega) \geq \bar{u}_i$ for all action profiles $a$ and states $\omega$: player $i$ can always appeal to retain her rights even after the game $G$ is played. For any (admissible) game $G$, Court's information $P$, stipulated contract $t$, and outcome $x \in A \times \Omega$, limited liability prescribes that player $i$'s transfers to other players $j$ are enforced, $F \circ t_{ij}(x) = t_{ij}(x)$, if and only if $t_{ij}(y) = t_{ij}(x)$ for all $y \in P(x)$ and $\min_{y \in P(x)} u_i(y) + \tau_i(x)|t \geq \bar{u}_i$.

When the latter condition fails and the former holds, the Court verifies that a contingency has realized in which $i$ is committed to transfer $t_{ij}(x)$ to player $j$, but does not verify that the net aggregate transfer $\tau_i(x)|t$ satisfies player $i$'s limited liability constraint. The Court does not enforce transfers $t_{ij}(x)$ to any player $j$, and may determine transfers $F \circ t_{ij}(x)$ according to different rules in different contexts. The voidance rule simply prescribes $F \circ t_{ij}(x) = 0$. More sophisticated rules prescribe that $\tau_i(x)|F \circ t = \bar{u}_i - \min_{y \in P(x)} u_i(y)$ and reduce each net positive transfer $t_{ij}(x)$ to a different player $j$ either equally, or proportionally, or according to a seniority order. ◇

This section has presented a general model of contracts and contractual enforcement. I have shown how it subsumes the standard methodology that assumes contracts are enforceable. I stated an enforceability principle, according to which it is without loss to assume that players only sign enforceable contracts. The next section uses my formalization to investigate under which conditions this enforceable principle holds or fails.

# 5 Analysis of the enforceability principle

This section is divided in three parts. The first one investigates the mathematical structure of the enforcement function $F$ introduced earlier, so as to identify a general sufficient condition for the enforceability principle. This condition holds under "standard" contract theoretical assumptions, namely when the Court's information correspondence $P$ is partitional, and either all contracts are legal, or the only legal restriction is limited liability. In these cases, there is no loss when adopting the usual methodology that restricts attention to enforceable contracts.

In the second part of this section, I first demonstrate the relevance of non-partitional Courts' information correspondences $P$. Then I show that the enforceability principle still holds when all contracts are legal, if the correspondence $P$ satisfies a transitivity condition. However, I argue that transitivity may fail because Courts' information is based on evidence verification that may entail substantial costs. Moving beyond Courts' information, I introduce a general concept of 'transitivity of contractual deterrence,' and show that it is a necessary condition for the enforceability principle.

The third part of this section revisits the leading examples of Section 3. The enforceability principle is shown to fail because contractual deterrence is not transitive. The source of such intransitive deterrence lies in administrative costs and leniency considerations that render contract enforcement slack (Example 1), or in legal principles such as individual liability (Example 3) or the prohibition of punitive transfers in excess of verified damages (Example 2). When contractual deterrence is not transitive, a player who contractually commits to play first best is not deterred from shirking. However, she is deterred from shirking by an overspecified contract, which in turn does not deter her from playing first best.

## 5.1 Fundamental results and a sufficient condition

Let me begin by noting that any contract $t$ is enforceable whenever $F \circ t = t$, i.e., whenever $t$ it is a *fixed point* of $F$. The enforceability principle can then be determined as follows.

For every strategy profile $s$, let $T(s) = \{t : s \text{ is a Bayesian Nash equilibrium of } G|t\}$ be the set of *enforced* contracts that implement $s$.

**Lemma 1** *For any Bayesian game $G$ and enforcement function $F$, the enforceability principle holds if and only if, for any strategy profile $s \in S$ such that $T(s)$ is non-empty, the function $F$ has a fixed point on $T(s)$.*

**Proof.** The if part is obvious. For any $s$ such that $T(s) \neq \varnothing$, take any any fixed point $t$ of $F$ on $T(s)$. Because $t \in T(s)$, it implements $s$, and becase $t$ is a fixed point of $F$, it is enforceable. For the only if part, suppose there is $s$ such that $T(s) \neq \varnothing$ and the function $F$ has no fixed point on $T(s)$. Hence any enforced contract $t$ that implements $s$ is such that there exists a stipulated contract $\hat{t} \neq t$ for which $F \circ \hat{t} = t$. The contract $\hat{t}$ is not enforceable. Hence $s$ can only be implemented by means of unenforceable contracts. ■

Lemma 1 translates the validity of the enforceability principle into the existence of fixed points of the function $F : T \to T$ on the restricted domains $T(s)$ associated with implementable strategy profiles $s$. Whenever $A$ and $\Omega$ are finite spaces and the enforcement function $F$ is continuous, this suggests using Brouwer's fixed point theorem to establish the validity of the enforceability principle.[21] To verify the hypothesis of Brouwer's theorem, I note that every non-empty set $T(s)$ is a convex subset of the Euclidean space $\mathbb{R}^{I \times I \times A \times \Omega}$ and can be made compact by placing uniform bounds on the transfers $t_{ij}(a, \omega)$.[22]

Besides introducing the possibility to use Brouwer's theorem to verify the enforceability principle, the logic behind Lemma 1 yields the following simple sufficient condition that will provide the foundations for my subsequent results. Theorem 1 below shows that if $F \circ F \circ t = F \circ t$ for all $t$ (i.e., if $F$ is idempotent), then all implementable strategy profiles can be implemented with an enforceable contract.

**Theorem 1** *For any Bayesian game $G$, if the enforcement function $F$ is* idempotent, *i.e. $F \circ F = F$, then any implementable strategy profile $s$ can be implemented with an enforceable contract, so that the enforceability principle holds.*

[21] Brouwer's theorem states: Every continuous function from a convex compact subset $K$ of a Euclidean space to $K$ itself has a fixed point. (See, for example Aliprantis and Border, 1999). Suitable generalization of the Brouwer's theorem exist for the case that $A$ and $\Omega$ are not finite spaces.

[22] The contract space $T$ is a hyperplane of $\mathbb{R}^{I \times I \times A \times \Omega}$, hence convex; and each non-empty $T(s) \subseteq T$ can be characterized as a set of contracts $t$ such that appropriate linear combinations of the transfers $t_{ij}(a, \omega)$ are smaller than given constants: $T(s) = \{t \in T : \sum_{\omega \in \theta_i} \sum_{j \in I} [t_{ij}(a_i, s_{-i}(\omega), \omega) - t_{ij}(s(\omega), \omega)] \rho(\omega|\theta_i) \leq \sum_{\omega \in \theta_i} [u_i(a_i, s_{-i}(\omega), \omega) - u_i(s(\omega), \omega)] \rho(\omega|\theta_i)$, for all $i$, all $\theta_i \in \Theta_i$ and all $a_i \in A_i\}$.

**Proof.** For every contract $t$, it is the case that $F \circ F \circ t = F \circ t$ as $F$ is idempotent, and hence that contract $F \circ t$ is enforceable. Consider any strategy profile $s$ implemented by a contract $t$, i.e., such that $s$ is a Bayesian Nash equilibrium of $G|F \circ t$. Because $u + \tau|F \circ t = u + \tau|F \circ F \circ t$, the profile $s$ is also a Bayesian Nash equilibrium of game $G|F \circ F \circ t$, i.e., $s$ is also implemented by contract $F \circ t$. Hence the enforceability principle holds for all $s$. ■

It is not difficult to identify instances where $F$ is not idempotent. Such instances include the examples of Section 3. In Example 1, the failure of idempotence of $F$ arises because administrative costs and leniency considerations make enforcement of sanctions slack. Consider 'simple social contracts,' each identified by a minimal contribution threshold $\hat{c}$ and a sanction schedule $p$. If any player $i$ contributes $c_i < \hat{c}$, the simple contract $t[\hat{c}, p]$ prescribes that an external enforcer, player $e$, collects sanction $p(c_i, \mathbf{c}_{-i})$ from $i$. Formally, $t_{ie}[\hat{c}, p](c_i, \mathbf{c}_{-i}) = p(c_i, \mathbf{c}_{-i})$ if $c_i < \hat{c}$, and $t_{ie}[\hat{c}, p](c_i, \mathbf{c}_{-i}) = 0$ if $c_i \geq \hat{c}$.

Suppose enforcement is slack in the sense defined in Example 1. A sanction $p \geq \Delta u_i(c_i, \hat{c}, \mathbf{c}_{-i}) \equiv u_i(c_i, \mathbf{c}_{-i}) - u_i(\hat{c}, \mathbf{c}_{-i})$ is enforced on any player $i$ if and only if $i$'s contract violation $d_i = \hat{c} - c_i$ exceeds the margin $m$, which I assume is independent of $\hat{c}$ and $\mathbf{c}_{-i}$ to simplify notation. The corresponding enforcement function $F$ is such that $F \circ t_{ie}[\hat{c}, p](c_i, \mathbf{c}_{-i}) = 0$ if $c_i \geq \hat{c} + m$. Hence, for any stipulated simple contract $t[\hat{c}, p]$ such that $\hat{c} > m$, and $p(c_i, \mathbf{c}_{-i}) \geq \Delta u_i(c_i, \hat{c}, \mathbf{c}_{-i})$ for all $c_i < \hat{c}$ and $\mathbf{c}_{-i}$, the enforced contract $F \circ t[\hat{c}, p]$ coincides with the simple contract $t[\hat{c} - m, p]$. Specifically, $F \circ t_{ie}[\hat{c}, p](c_i, \mathbf{c}_{-i}) = 0$ if $c_i \geq \hat{c} - m$, and $F \circ t_{ie}[\hat{c}, p](c_i, \mathbf{c}_{-i}) = p(c_i, \mathbf{c}_{-i})$ for $c_i < \hat{c} - m$. Applying the enforcement function $F$ to contract $F \circ t[\hat{c}, p]$, I obtain the contract $F \circ F \circ t[\hat{c}, p]$ such that $F \circ F \circ t_{ie}[\hat{c}, p](c_i, \mathbf{c}_{-i}) = 0$ if $c_i \geq \hat{c} - 2m$. Because $F \circ t_{ie}[\hat{c}, p] = p(c_i, \mathbf{c}_{-i}) \geq \Delta u_i(c_i, \hat{c}, \mathbf{c}_{-i}) > 0$ for all $c_i \in [\hat{c} - 2m, \hat{c} - m)$, I conclude that $F \circ F \circ t[\hat{c}, p] \neq F \circ t[\hat{c}, p]$. The enforcement function $F$ is shown not to be idempotent. Similar failures of idempotence of $F$ can be demonstrated for the other examples of Section 3.

The property of idempotence is known in several branches of mathematics, such as linear algebra and group theory. In my setup, idempotence of an enforcement function $F$ can be interpreted as follows. When applied to a contract $t$, the function $F$ enforces some contingent transfers $t(a, \omega)$ and modifies some other transfers according to a given determination rule. The function $F$ fails to be idempotent if, after the Court has modified all contingent transfers $t(a, \omega)$ that should not be enforced, a contract $F \circ t$ is obtained in which some transfers $F \circ t(a, \omega)$ should not be enforced: $F \circ F \circ t \neq F \circ t$. Because of how $F$ modified contract $t$ at contingencies $(a', \omega')$ different from $(a, \omega)$, the form taken by enforced transfers at $(a, \omega)$ has changed. This can only be possible if the transfers $F \circ t(a, \omega)$ enforced at $(a, \omega)$ depend on the shape of the whole stipulated contract $t$, and not only on the "local" properties of $t$ at $(a, \omega)$.

I will now show that this observation leads to important implications in the "standard cases" presented at the end of Section 4, in which the Court's information is partitional, and either all contracts are legal, or the only legal constraint is limited liability.

**Proposition 1** *For any Bayesian game $G$, and partitional verification correspondence $P$, absent legal constraints, or when the only legal constraint is limited liability, the enforceability principle holds.*

The intuition for this result is easier to see in the case when all contracts are legal, so that the only requirement for the enforcement of transfers is that they are verifiable. In this case, the enforcement function $F$ is such that, for any contract $t$ and contingency $x \in A \times \Omega$, the transfers $t(x)$ are enforced if and only if $t$ is constant on $P(x)$. When not enforcing a transfer $t_{ij}(x)$ from a player $i$ to another player $j$, the Court determines the transfer $F \circ t_{ij}(x)$ according to a rule that must be constant on $P(x)$, in line with Assumption 1. Hence, while the arbitrary contract $t$ need not be measurable with respect to the partition $P$, the enforced contract $F \circ t$ must be measurable with respect to $P$. Suppose now that contract $F \circ t$ is stipulated. Because all contracts are legal and the contract $F \circ t$ is constant on $P(x)$ for all $x$, all transfers $F \circ t(x)$ are enforced. The enforcement function $F$ is thus idempotent, and the enforceability principle holds. Intuitively, this is because when all contracts are legal and the verification correspondence $P$ is partitional, the enforcement of any contractual transfer $t(x)$ depends only on the "local" shape of contract $t$ at $x$. Specifically, whether $t(x)$ is enforced or not depends only on whether or not $t$ is constant on the set $P(x)$, which does not overlap any other element of $P$ because $P$ is partitional.

The reasoning is analogous for the case when the only legal constraint is limited liability. Then, for any stipulated contract $t$, the enforcement function $F$ is such that transfers are enforced, $F \circ t(x) = t(x)$, if only if $t$ is constant on $P(x)$ and for all $i$, it is the case that $\min_{y \in P(x)} u_i(y) + \tau_i(x)|t \geq \bar{u}_i$ for some given bounds $\bar{u}_i$ with the property that $u_i(x) \geq \bar{u}_i$ for all $x \in A \times \Omega$. When not enforcing contractual transfers $t(x)$, the Court implements transfers $F \circ t(x)$ that must be constant on $P(x)$, and that must satisfy $\min_{y \in P(x)} u_i(y) + \tau_i(x)|F \circ t \geq \bar{u}_i$ for all $i$. Again, when contract $F \circ t$ is stipulated, all transfers $F \circ t(x)$ meet the requirements embedded in $F$ for their enforcement. Because the enforcement of transfers $t(x)$ depends only on the shape that the contract $t$ takes on $P(x)$ and the verification structure $P$ is partitional, the enforcement function $F$ is idempotent, and the enforceability principle holds.

## 5.2 The role of transitivity and a necessary condition

Let me start by arguing that, while partitional information correspondence are often considered standard, it is worth considering non-partitional Court's information correspondences, in the context of this paper.

Restricting attention to partitional correspondences $P$ is without loss if assuming that the Court is a "standard" Bayesian player and that evidence verification is costless. In this case, for any realized outcome is $(a, \omega)$, the Court can costlessly figure out what its knowledge $P(a', \omega')$ would be in case any other possible outcome $(a', \omega')$ had realized (that is to say, the Court costlessly knows $P$). If on top that, the Court is also allowed to render judgments based on standard Bayesian reasoning, then it can be shown that $P$ must be a partition.[23] Indeed, Geanakoplos (1989) proved that a Bayesian agent's truthful information correspondence $P$ is partitional if and only if, whenever any outcome $(a, \omega)$ realizes, the agent knows all what she would know and would not know had any other outcomes $(a', \omega')$ realized.[24]

However, evidence verification is a costly and often gruelling task. It is not plausibly costless for a Court to know its information correspondence $P$. This requires not only that it examines the evidence presented for the realized outcome $(a, \omega)$, but also that the Court figures out what evidence would be presented for any possible other outcome $(a', \omega')$. Also, it does not always appear plausible to model Court judgments as standard Bayesian decisions. When rendering a sentence, a Court is required to provide *"reasons for judgment"* that explain its motivations. As Shin (1993) pointed out, the requirement that one 'knows all that she does not know' may be too stringent when knowledge of an event equates with the capability of proving that the event has occurred.[25] In sum, assuming that Courts' information correspondences are partitional seems difficult to justify in general.

While earlier work on non-partitional information has focused on the axiom that one 'knows all that she does not know,' this paper uncovers a crucial role for the complementary axiom that 'one (the Court, here) knows all that she knows.' It is intuitive that this axiom

[23] Consider outcomes $x$ and $y$ in $A \times \Omega$ and a Bayesian Court who knows that $P(x) = \{x, y\}$ and $P(y) = \{y\}$. Suppose that $x$ realizes. Because $y \in P(x)$, the Court cannot exclude the possibility that $y$ realized. It also knows that, had $y$ realized, it would know that $x$ did not realize. The Court concludes that $y$ did not realize, and refines its information correspondence to $P'(x) = \{x\}$ and $P'(y) = \{y\}$.

[24] One knows an event $E \subseteq A \times \Omega$ at any outcome $x \in A \times \Omega$ if $P(x) \subseteq E$. Let $KE \equiv \{x : P(x) \subseteq E\}$ and $K^C E \equiv A \times \Omega \setminus KE$. $P$ is partitional if she knows all that she knows and that she does not know: $KE \subseteq KKE$ and $K^C E \subseteq KK^C E$, for all events $E$.

[25] Consider for example the event $E =$ "Eve can play the piano." Eve can prove $E$ by demonstration, but cannot disprove $E$ by not playing a piano. Letting $N = (\Omega \times A) \setminus E$, a Court's information structure is $P(E) = E$, $P(N) = E \cup N$. The axiom that the Court 'knows all that it does not know' fails: the event that it does not know $E$ is $N$, but if $N$ realizes, the Court does not know $N$, as $P(N) \not\subseteq N$.

may fail if knowledge is based on costly verification. When each round of verification and counterfactual reasoning entails a cost, it may be cost effective to know (verify) an event without "verifying to know the event." And this may result in instances of failure of the enforceability principle. This is demonstrated in Example 4 below, which is analogous to Example 1 of Section 3. Costly verification here plays the same role as costly enforcement and leniency considerations there, in making contract enforcement slack. And again as a result, first best may be achieved only by means of overspecified contracts.

Before presenting the example, let me note that, as shown by Geanakoplos (1989), the axiom that 'one knows all that she knows,' is equivalent to transitivity of her information structure $P$: for any outcomes $x, y, z$ in $A \times \Omega$, $x \in P(y)$ and $y \in P(z)$ imply $x \in P(z)$. I also note that if $P$ is partitional, then $P$ is transitive, but not necessarily viceversa.

**Example 4 (Costly verification and intransitivity).** Consider the same free-rider problem of Example 1. Each player $i = 1, ..., n$ chooses a contribution $c_i \geq 0$, the optimal outcome is $c_i = c^*$ for all $i$, but each player $i$'s payoff $u_i(c_i, \mathbf{c}_{-i})$ is decreasing in $c_i$, for all opponents' contributions $\mathbf{c}_{-i}$. Suppose that evidence verification is costly. When player $i$'s plays $c_i$, a Court can only rule that $i$'s choice $c_i'$ belongs to the interval $[\max\{0, c_i - m'\}, c_i + m']$, but is unable to verify $c_i'$ precisely. The quantity $m' > 0$ is a "margin of error" that could be made equal to zero only with a socially excessive expenditure on evidence verification by the legal system. Most importantly, it would be too costly for the Court to engage in counterfactual reasoning to refine its information structure $P$, as this would require figuring out what evidence would be presented for any possible choice $c_i'$ of agent $i$. As a result, the Court's information correspondence $P$ is neither partitional, nor transitive. For any player $i$, consider any triple of choices $a_i > b_i > c_i > m'$ such that $a_i - b_i < m'$ and $b_i - c_i < m'$, but $a_i - c_i > m'$. Then $(a_i, \mathbf{c}_{-i}) \in P(b_i, \mathbf{c}_{-i})$ and $(c_i, \mathbf{c}_{-i}) \in P(b_i, \mathbf{c}_{-i})$, but $(c_i, \mathbf{c}_{-i}) \notin P(a_i, \mathbf{c}_{-i})$, for any opponents' choices $\mathbf{c}_{-i}$.

Due to the lack of transitivity in the Court's information correspondence $P$, the enforceability principle fails. Suppose the social contract $t$ prescribes that each player $i$ contribute at least $\hat{c} = c^*$. Then, any contribution $c_i \in [\max\{0, c^* - m'\}, c^*)$ cannot be deterred. The Court cannot establish that $i$ did not contribute $c^*$, because $c^* \in [\max\{0, c_i - m'\}, c_i + m']$. The only contract that implements the optimal contributions $\mathbf{c}^*$ is the one that prescribes that each player $i$ contributes at least $\hat{c} = c^* + m'$. In equilibrium, each player $i$ contributes $c^*$ without being sanctioned. The enforceability principle fails. The margin of error $m'$ that makes the information structure $P$ intransitive plays the same role here as the enforcement margin $m$ in Example 1. $\diamond$

The above example shows how the enforceability principle may fail because of intransitivity of the Court's information. Conversely, Proposition 2 shows that if the Court's information correspondence $P$ is transitive, then when all contractual transfers are legal, the enforceability principle holds for all the transfer determination rules presented in Section 4. The enforceability principle generalizes beyond the standard case of partitional Court's information considered in Proposition 1. The decisive requirement turns out to be transitivity.[26]

**Proposition 2** *For any Bayesian game $G$, and transitive verifiability structure $P$, absent legal constraints, the enforceability principle holds for all the transfer determination rules presented in Section 4.*

This result is proved in the Appendix by invoking Theorem 1 and by showing that if all contractual transfers are legal and the Court's information correspondence $P$ is transitive, then the resulting function $F$ is idempotent. Here, I present a heuristic argument that highlights the role played by transitivity of $P$ in ensuring that the enforceability principle holds. I simplify the argument by focusing on the case with no state and type uncertainty, and simplify notation accordingly.

Proceeding by contradiction, suppose that the enforceability principle fails. There is an implementable strategy profile $a$ that cannot be implemented with any contract $t$ by which the parties commit to play $a$. Then, there must be a player $i$ who has a profitable deviation $a'_i$ that cannot be verified as a breach of contract $t$ by the Court when the opponents play $a_{-i}$, i.e., it must be that $a \in P(a'_i, a_{-i})$. Because the outcome $a$ is implementable, there must exist a contract $t'$ supposed to deter $a$ (i.e., a contract with which the players commit to play a profile $b$ different from $a$), but such that $a$ is nevertheless a Nash equilibrium when $t$ is stipulated. Then, it must be that $a$ is not verified as a breach of contract when the players contractually committed to $b$, i.e., that $b \in P(a)$. And it must also be that $i$'s deviation $a'_i$ is deterred when the opponents play $a_{-i}$ and the players committed to $b$, i.e., that $b \notin P(a'_i, a_{-i})$. Thus, the correspondence $P$ is shown to not be transitive: $b \in P(a)$ and $a \in P(a'_i, a_{-i})$, but $b \notin P(a'_i, a_{-i})$.

The above argument describes the role played by the intransitivity of the Court's information structures $P$ to generate failure of the enforceability principle, when all contractual

[26] This result may be mathematically related to the necessary conditions by Green and Laffont (1986) for the revelation principle to hold. They consider a message game with an informed agent and a committed principal. But unlike in the standard set up by Myerson (1979), the agent's message is not cheap talk. Each agent's type may only send a message out of a subset of the type space. Green and Laffont (1986) show that the revelation principle holds in this setup if and only if the correspondence describing these constraints is transitive.

transfers are legal. In this case, the only enforcement constraint is that stipulated transfers be verifiable. Intransitivity of $P$ is the only possible reason for failure of the enforceability principle. However, when it is not the case that all possibly stipulated transfers are legal, the enforceability principle may also fail when the Court's information structures P is transitive, or even partitional, as in Examples 1 - 3. In general, the source of the failure of the enforceability principle can be identified in the intransitivity of a suitably defined order.

Let us introduce the concept of 'transitivity of contractual deterrence,' which can be described informally as follows: "Contractually stipulating the play $b$ deters the play $(a_i', a_{-i})$ but not play $a$, whereas contractually stipulating $a$ does not deter $(a_i', a_{-i})$." Formally, given a Bayesian game $G = (I, \Omega, \rho, S, \Theta, u)$ and enforcement function $F$, define the associated 'contractual deterrence order' $L_i$, for any type $\theta_i$ of any player $i$, and any opponents' strategy profile $s_{-i}$ as follows: $(s_i, s_i') \in L_i[\theta_i, s_{-i}; G, F]$, or $s_i L_i[\theta_i, s_{-i}; G, F] s_i'$, whenever there exist contract $t$ such that:

$$u_{\theta_i}(s_i', s_{-i}) + \tau_{\theta_i}(s_i', s_{-i})|F \circ t \leq u_{\theta_i}(s_i, s_{-i}) + \tau_{\theta_i}(s_i, s_{-i})|t.$$

The notation $(s_i, s_i') \in L_i[\theta_i, s_{-i}; G, F]$ means that "given that the other players play $s_{-i}$ in game $G$, type $\theta_i$ of player $i$ does not deviate to $s_i'$ when the contract $t$ enforceable at $s$ stipulates that she plays $s_i$."

The importance of contractual deterrence order $L_i$ is that the failure of its negative transitivity generates a failure of the enforceability principle at some strategy profile $s$, for the case of interest in which $s$ is implementable in the first place. In other terms, negative transitivity of $L_i[\theta_i, s_{-i}; G, F]$ for all $i, \theta_i$ and $s_{-i}$ is a necessary condition for the enforceability principle to hold in game $G$ and for enforcement function $F$.

**Proposition 3** *For any Bayesian game $G = (I, \Omega, \rho, S, \Theta, u)$ and enforcement function $F$, if there exists a player $i$, and opponent's strategy profile $s_{-i}$ such that the contractual deterrence order $L_i[\theta_i, s_{-i}; G, F]$ fails to be negative-transitive, then for some $s_i \in S_i$, either the outcome $s = (s_i, s_{-i})$ is not implementable, or the enforceability principle fails at $s$.*

This result shows how the role played by intransitivity in generating failures of the enforceability principle, identified as intransitivity of the verifiability correspondence $P$ in absence of legal constraints to contract enforcement, can be generalized to environments where legal or practical constraints limit contractual enforcement. In such cases, the role of intransitivity is identified by failure of negative transitivity of the contractual deterrence order $L_i[\theta_i, s_{-i}; G, F]$. Interestingly, because $L_i[\theta_i, s_{-i}; G, F]$ is a strict and complete preference order, for any $i$, $\theta_i$, $s_{-i}$, $G$ and $F$, a failure of negative transitivity of $L_i[\theta_i, s_{-i}; G, F]$

is a equivalent to a failure of rationality for order $L_i[\theta_i, s_{-i}; G, F]$ in the sense of Savage (1954).

In the context of Example 4, failure of negative transitivity of $L_i$ for any player $i$ is a consequence of failure of transitivity of $P$. Consider any contribution $c_i \in [c^* - m', c^*)$, and let $\hat{c} = c^* + m'$, note that $\mathbf{c}^* \in P(c_i, \mathbf{c}^*_{-i})$ and $(\hat{c}, \mathbf{c}^*_{-i}) \in P(\mathbf{c}^*)$ but $(\hat{c}, \mathbf{c}^*_{-i}) \notin P(c_i, \mathbf{c}^*_{-i})$. Hence, no contract $t$ that is enforceable at $\mathbf{c}^*$ may deter player $i$ from contributing $c_i$, because $F \circ t(c_i, \mathbf{c}^*_{-i}) = F \circ t(\mathbf{c}^*) = t(\mathbf{c}^*)$. By the same reasoning, no contract enforceable at $(\hat{c}, \mathbf{c}^*_{-i})$ may deter $i$ from contributing $c^*$. But playing $c_i$ is deterred by the simple contract $t$ that prescribes $i$ pays sanction $t_{ie}(c'_i, \mathbf{c}_{-i}) \geq \Delta u_i(c'_i, \hat{c}, \mathbf{c}_{-i})$ to an external enforcer $e$ for any contribution $c'_i < \hat{c}$. Hence, omitting the here irrelevant dependence on $\theta_i$, the contractual deterrence order $L_i[\mathbf{c}^*_{-i}; G, F]$ fails to be negative transitive: $(c_i, c^*) \notin L_i[\mathbf{c}^*_{-i}; G, F]$ and $(c^*, \hat{c}) \notin L_i[\mathbf{c}^*_{-i}; G, F]$, but $(c_i, \hat{c}) \in L_i[\mathbf{c}^*_{-i}; G, F]$. And as a consequence, the enforceability principle fails at the first best outcome $\mathbf{c}^*$, which can only be implemented with contracts not enforceable at $\mathbf{c}^*$, such as the simple contract $t$ described above.

## 5.3 Revisitation of the leading examples

I now return to the leading examples of Section 3 to explain in depth why the enforceability principle fails, using the results of this section. Because the revisitation of Example 1 provides the least added insight, in the interest of brevity, I present it in the Appendix. I directly turn to Example 2, where the enforceability principle fails because enforceable transfers are limited to compensation of verified damages caused by breach of contract.

**Damage compensation and Example 2 revisited.** First, I construct an enforcement function $F$ to represent laws that rule out punitive contractual transfers. Enforceable transfers are limited to compensation of verified damages due to breach of contract. Unlike in Example 2, I do not place any restrictions on the form of the stipulated contracts. Consider any possible contract $t$, defined as a menu of contingent transfers. In the context of Example 2, a contract $t$ may correspond to any schedule of remunerations for provision of effort, thus allowing for bonuses and penalties.

Consider any pair of players $i, j$ such that $j$ holds seniority rights relative to $i$, and the net transfers $t_{ij}$ from $i$ to $j$ induced by an arbitrary contract $t$. Because punitive transfers are illegal, and $j$ is senior to $i$, the Court does not enforce a reduction of the net transfer $t_{ij}$ if it is in excess of verified loss borne by player $i$. In the example where $j$ is an agent of principal $i$, the net transfer $t_{ij}$ is the agent's remuneration stipulated in contract $t$, and the burden of proof to justify withholding or reducing $t_{ij}$ lies with the principal.

Because $t$ is a general menu contract, it associates different transfers $t_{ij}(\hat{a},\omega)$ to different outcomes $(\hat{a},\omega)$, and damage needs to be assessed relative to any possible action profile $\hat{a}$. In the example where $j$ is an agent of principal $i$, the contract $t$ may stipulate a different remuneration for different provisions of effort. The assessment of whether reducing remuneration is justified as damage compensation or not should then be based on the whole profile of remunerations.

Let me momentarily assume that the Court's assessment of player $i$'s verified damage consists of the minimum damage compatible with all the information available. When an outcome $(a,\omega)$ realizes, all the Court knows about the realized outcome $(a',\omega')$ is that it belongs to $P(a,\omega)$. Further, the Court is shown all clauses of the stipulated contract $t$, and hence it knows the stipulated net transfer $t_{ij}(\hat{a},\omega')$ from $i$ to $j$ contingent on any possible action profile $\hat{a}$ and state $\omega'$. For any possibly realized outcome $(a',\omega')\in P(a,\omega)$, player $i$'s loss relative to any action profile $\hat{a}$ and state $\omega'$ is $\max\{0,u_i(\hat{a},\omega')-u_i(a',\omega')\}$. The Court enforces the transfer $t_{ij}(a,\omega)$ contingent on $(a,\omega)$ if and only if it is not in excess of player $i$'s verified loss relative to any action profile $\hat{a}$, i.e., if $t_{ij}(\hat{a},\omega')-t_{ij}(a,\omega)\leq \max\{0,u_i(\hat{a},\omega')-u_i(a',\omega')\}$ for all $\hat{a}$ and $(a',\omega')\in P(a,\omega)$. These assumptions lead to the enforcement function $F$ such that, for every $(a,\omega)$,

$$F\circ t_{ij}(a,\omega)=\max\left\{t_{ij}(a,\omega),\max_{\hat{a},(a',\omega')\in P(a,\omega)}\{t_{ij}(\hat{a},\omega')-\max\{0,u_i(\hat{a},\omega')-u_i(a',\omega')\}\}\right\}. \quad (4)$$

This formulation embeds the requirement that $t_{ij}(\hat{a},\omega')-F\circ t_{ij}(a,\omega)$ be smaller than the minimum possible loss $\max\{0,u_i(\hat{a},\omega')-u_i(a',\omega')\}$ for any possible outcome $(a',\omega')\in P(a,\omega)$. In some cases, this formulation may be too favourable to agent $j$. In Example 2 for instance, it implies that when the Court verifies effort $e_i$ but not $e_j$, it must presume that the agent acted so as to minimize the principal's loss and exerted the maximum level of effort $\bar{e}_j$ regardless of his contractual commitment $\hat{e}_j$. As noted earlier, it may be more realistic that the Court presumes that the agent exerted effort $e_j=\hat{e}_j$, instead.

Such a presumption may be generalized beyond Example 2 so as to cover any game and any contract $t$. For any action profile $\hat{a}$, let me modify the requirement of expression (4) that $t_{ij}(\hat{a},\omega')-F\circ t_{ij}(a,\omega)$ be smaller than $\max\{0,u_i(\hat{a},\omega')-u_i(a',\omega')\}$ for any $(a',\omega')\in P(a,\omega)$ as follows. For any $\omega'$, I stipulate that $t_{ij}(\hat{a},\omega')-F\circ t_{ij}(a,\omega)$ be smaller than $\max\{0,u_i(\hat{a},\omega')-u_i(a'',\omega')\}$ for any action profile $a''\in C(\hat{a},\omega';a,\omega)\equiv \arg\min_{a':(a',\omega')\in P(a,\omega)}||a'-\hat{a}||$. Instead presuming that $j$ acted so as to minimize $i$'s loss relative to the outcome $(\hat{a},\omega')$, the Court presumes that $i$'s deviation $a''$ from $\hat{a}$ is as small as possible. With this modification, I obtain the following expression for the enforced net transfers:

$$F \circ t_{ij}(a, \omega) = \max \{ t_{ij}(a, \omega), \max_{\hat{a}, \omega'; a'' \in C(\hat{a}, \omega'; a, \omega)} \{ t_{ij}(\hat{a}, \omega') - \max\{0, u_i(\hat{a}, \omega') - u_i(a'', \omega'')\}\}\}. \quad (5)$$

We are now ready to revisit Example 2. Let player $i = 1$ be the client and $j = 2$ her agent, so that the transfer $t_{12}$ is the remuneration for the agent's effort stipulated with contract $t$. The agent has seniority rights over the transfer: the burden of proof to withhold or reduce the remuneration $t_{12}$ lies with the principal. Every outcome $(a, \omega)$ is composed of an agents' effort choice $a \equiv \mathbf{e} = (e_A, e_B) \in [0, \bar{e}]^2$, a pair of effort levels $e_A$ and $e_B$, and of the state of the world $\omega \in \{0, 1\} \times \{v, n\}^2$. The first component $\omega_1 = 0, 1$ determines the principal's payoff $u_1(a, \omega) \equiv v_{\omega_1}(\mathbf{e})$, whereas the other two components $\omega_A$ and $\omega_B$ determine whether $e_A$ and $e_B$ are verified by the Court or not. Specifically, $\omega_i = n$ means that effort $e_i$ is not verified, and $\omega_i = v$ that effort $e_i$ is verified, for $i = A, B$. The agent's payoff is simply $u_2(a, \omega) \equiv -c(\mathbf{e})$.

The move of nature $\rho$ is such that $\rho(\omega) = \rho_1(\omega_1)\rho_A(\omega_A)\rho_B(\omega_B)$, where $\rho_1(\omega_1) = q$ for $\omega_1 = 1$, and in this case, $\rho_i(\omega_i) = p$ for $\omega_i = v$ and both $i = A, B$. Instead, if $\omega_1 = 0$, $\rho_i(\omega_i) = 0$ for $\omega_i = v$ and both $i = A, B$. The Court's information structure is such that it always observes $\omega_1$, but in case $\omega_i = n$ it cannot verify effort $e_i$. Formally, for every effort levels $\mathbf{e}$, $P(\mathbf{e}; 1, v, v) = \{(\mathbf{e}; 1, v, v)\}$, $P(\mathbf{e}; 1, v, n) = \{(e_A, e'_B; 1, v, n) : e'_B \in [0, \bar{e}]\}$, $P(\mathbf{e}; 1, n, v) = \{(e'_A, e_B; 1, n, v) : e'_A \in [0, \bar{e}]\}$, $P(\mathbf{e}; \omega_1, n, n) = \{(\mathbf{e}'; \omega_1, n, n) : \mathbf{e}' \in [0, \bar{e}]^2\}$ for $\omega_1 = 0, 1$.

Armed with this formalization, I recover the results of Example 2, generalizing them beyond simple contracts to consider any general contract $t$. I prove in the appendix that, with the enforcement functions $F$ described by either (4) or (5), the first best effort levels $\mathbf{e}^*$ cannot be implemented with any enforceable contract, and can be implemented with the overspecified contracts identified in Example 2. In the Proposition below, for any levels of effort $\hat{\mathbf{e}}$, the simple contract $t[\hat{\mathbf{e}}, f]$ is defined as the contract $t$ such that $t_{12}(\mathbf{e}, \omega) = f$ if $\mathbf{e} \geq \hat{\mathbf{e}}$ and $t_{12}(\mathbf{e}, \omega) = 0$ if $\mathbf{e} \not\geq \hat{\mathbf{e}}$.

**Proposition 4** *The first best effort levels* $\mathbf{e}^*$ *cannot be implemented with any contract* $t$ *enforceable at* $\mathbf{e}^*$ *when the enforcement function* $F$ *is described by either (4) or (5). For the latter, first best efforts* $\mathbf{e}^*$ *may be implemented with an overspecified simple contracts* $t[\hat{\mathbf{e}}, f]$ *where* $c(\mathbf{e}^*) \leq f \leq v(\mathbf{e}^*)$ *and* $\hat{\mathbf{e}} > \mathbf{e}^*$ *solves (3); for the former,* $\mathbf{e}^*$ *may be implemented with contracts* $t[\hat{\mathbf{e}}, f]$ *where* $\hat{\mathbf{e}} > \mathbf{e}^*$ *solves* $v_1(\hat{\mathbf{e}}) = v_1(e^*, \bar{e})$.

This result is proved by first focusing on simple contracts and showing that the enforcement function $F$ described by either (4) or (5) reproduces the analysis of Example 2. As a result, first best efforts $\mathbf{e}^*$ can be achieved with the overspecified simple contracts described in Proposition 4, but cannot be achieved with any simple contract enforceable at $\mathbf{e}^*$ (i.e., with any simple contract $t$ such that $t(\mathbf{e}^*, \omega) = F \circ t(\mathbf{e}^*, \omega)$ for all $\omega$).

Now consider any general contract $t$ enforceable at $\mathbf{e}^*$. For all deviations $\mathbf{e} \leq \mathbf{e}^*$, when assessing if a transfer $t_{12}(\mathbf{e}, \omega)$ is punitive or justified by damage compensation, a Court compares $t_{12}(\mathbf{e}, \omega)$ with the whole schedule of payments $t_{12}(\hat{\mathbf{e}}, \omega)$ for all $\hat{\mathbf{e}} \in [0, \bar{e}]^2$, including $t_{12}(\mathbf{e}^*, \omega)$. Hence, the enforced transfers $t_{12}(\mathbf{e}, \omega)|F \circ t$ cannot be smaller than with the simple contracts $t[\mathbf{e}^*, f]$. Because $t_{12}(\mathbf{e}^*, \omega)|F \circ t = t_{12}(\mathbf{e}^*, \omega)$ the difference in expected payoffs when playing $\mathbf{e} \leq \mathbf{e}^*$ instead of $\mathbf{e}^*$ cannot be less favorable to playing $\mathbf{e}$ than with the simple contracts $t[\mathbf{e}^*, f]$. Because such contracts fail to achieve first best efforts $\mathbf{e}^*$, so do any general contracts $t$ enforceable at $\mathbf{e}^*$.

The violation of the enforceability principle in Example 2 is a consequence of a failure of negative transitivity of the agent's contractual deterrence order $L_2$. Consider the enforcement function $F$ described in (5)—analogous arguments deal with the function $F$ characterized in (4). Omitting the here irrelevant dependence on $\theta_2$ and $s_{-2}$, for every efforts $\mathbf{e}', \mathbf{e}$, it is the case that $(\mathbf{e}', \mathbf{e}) \in L_2[G, F]$ whenever there exist a contract $t$ such that

$$\sum_{\omega \in \{0,1\} \times \{v,n\}^2} F \circ t_{12}(\mathbf{e}; \omega)\rho(\omega) - c(\mathbf{e}) \leq \sum_{\omega \in \{0,1\} \times \{v,n\}^2} t_{12}(\mathbf{e}'; \omega)\rho(\omega) - c(\mathbf{e}').$$

The proof of Proposition 4 shows that negative transitivity of $L_2[G, F]$ is violated. Any contract $t$ enforceable at $\mathbf{e}^*$, i.e., such that $t_{12}(\mathbf{e}^*; \omega)|F \circ t = t_{12}(\mathbf{e}^*; \omega)$, fails to deter the agent from underproviding efforts, i.e., $(\mathbf{e}^*, \mathbf{e}) \notin L_2[G, F]$ for some $\mathbf{e} \leq \mathbf{e}^*$. All contracts $t$ enforceable at the effort levels $\hat{\mathbf{e}} > \mathbf{e}^*$ that solve (3) fail to deter the agent from exerting efforts $\mathbf{e}^*$, that is $(\hat{\mathbf{e}}, \mathbf{e}^*) \notin L_2[G, F]$. But the simple contract $t[\hat{\mathbf{e}}, f]$ such that $c(\mathbf{e}^*) \leq f \leq v(\mathbf{e}^*)$ successfully deters the agent from exerting any effort $\mathbf{e} \leq \mathbf{e}^*$, so that $(\hat{\mathbf{e}}, \mathbf{e}) \in L_2[G, F]$. Negative transitivity of the agent's contractual deterrence order $L_2$ is violated. $\diamond$

I conclude this section by reconsidering Example 3, where the enforceability principle fails because agents are individual liable, their individual input cannot be identified, and high quality output requires that both exert high effort.

**Individual liability and Example 3 revisited.** Consider any pair of players $i, \ell$ such that $i$ is individually liable and holds seniority rights relative to $\ell$. For example, $i$ may be one of the agents of principal $\ell$. For any stipulated contract $t$ and contingency $(a, \omega)$, individual liability implies that the contingent transfer $t_{\ell i}(a, \omega)|F \circ t$ from $\ell$ to $i$ enforced in Court can only depend on player $i$'s verified action $a_i$, and not on the actions $a_{-i}$ of the other players. Because $i$ is senior to $\ell$, the Court enforces the largest net transfer

$t_{\ell i}(a', \omega')$ compatible with her knowledge about $i$'s action $a_i$ and state $\omega$. For example, when $i$ is an agent of $\ell$, she is paid her contractual fees even in case of controversy, whereas exacting penalties on $i$ for shirking requires verification of $i$'s actions. And because agent $i$ is individually liable, her net remuneration $t_{\ell i}(a, \omega)|F \circ t$ cannot depend on the actions $a_{-i}$ of the other players.

The above considerations lead to the following formalization of contractual enforcement under individual liability. For any set of outcomes $B \subseteq A \times \Omega$ , let me define the 'projection' $\pi_i(B) \equiv \{(a_i, \omega) \in A_i \times \Omega : (a_i, a_{-i}; \omega) \in B$, for some $a_{-i} \in A_{-i}\}$ of $B$ onto $A_i \times \Omega$. Individual liability constraints can then be represented by the enforcement function $F$ such that, for any contract $t$ and contingency $(a, \omega)$, the enforced net transfer from any player $\ell$ to an individually liable senior player $i$ is:

$$F \circ t_{\ell i}(a, \omega) = \max_{a'_{-i}, (a'_i, \omega') \in \pi_i(P(a, \omega))} t_{\ell i}(a'_i, a'_{-i}, \omega). \tag{6}$$

In words, the Court enforces the largest net transfer $t_{\ell i}(a', \omega')$ compatible with her knowledge about player $i$'s action $a_i$ and state $\omega$, regardless of the actions $a_{-i}$ of the other players. I note that, as a consequence, the only contract $t$ where $\ell$'s net transfers $t_{\ell i}$ to $i$ may have any effect on player $i$'s strategies are such that the transfers $t_{\ell i}(a_i, a_{-i}; \omega)$ are constant on $a_{-i}$ for all $a_i$ and $\omega$.

Now, reconsider Example 3. In the first best, the client (player 3) hires the agents $i = 1, 2$ who play $(HN, HN)$, i.e., they exert high effort without monitoring each other. Consider any contract $t$ enforceable at the first best outcome $(HN, HN)$, i.e. such that $F \circ t_{3i}(HN, HN) = t_{3i}(HN, HN)$ for both agent $i$—simplifying notation because there no uncertainty about the state $\omega$. I first note that such enforceable contracts include all contracts by which each agent $i$ commits to exert high effort without monitoring the other agent $j$. Identifying the net transfer $t_{3i}(HN, HN)$ with agent $i$'s fee $f$ in fact, any such contract $t$ yields each agent $i$ the net remuneration $t_{3i}(a) = f$ if $a_i = HM, HN$, and $t_{3i}(a) < f$ if $a_i = LM, LN$. Because $t_{3i}(HN, HN) = \max_{a \in A} t_{3i}(a)$, the enforceability function $F$ described by (6) is such that $F \circ t_{3i}(HN, HN) = t_{3i}(HN, HN)$.

I now show that the first best outcome $(HN, HN)$ cannot be implemented with any enforceable contract. Say that agent $j$ works hard without monitoring $i$ and plays $HN$, but agent $i$ shirks on effort and plays $LN$. Then, the Court's information is the set $P(LN, HN) = \{(LM, HN), (LN, HN), (LN, LN), (HN, LN), (HN, LM)\}$, and the projection onto player $i$'s action space $\pi_i(P(LN, HN)) = \{HN, LN, LM\}$ includes the action $HN$. Hence, agent $i$'s net remuneration is $F \circ t_{3i}(LN, HN) = F \circ t_{3i}(HN, HN) = f$, and $i$'s deviation from $HN$ to $LN$ is not deterred.

Letting each agent's cost of high effort be $c$, consider now instead a contract $t$ with which each agent $i$ commits to pay a penalty $p_i$ such that $c < p_i < f$ to the client unless she exerts high effort and monitors the other agent $j$. That is, each agent $i$'s net remuneration is $t_{3i}(a) = f$ if $a_i = HM$, and $t_{3i}(a) = f - p_i$ if $a_i = HN, LM, LN$. Such overspecified contracts succeed in implementing the first best outcome $(HN, HN)$. If agent $j$ works hard without monitoring agent $i$, and agent $i$ shirks, the outcome $(LN, HN)$ again yields the projection $\pi_i(P(LN, HN)) = \{HN, LN, LM\}$ onto agent $i$'s action space. Because the projection $\pi_i(P(LN, HN))$ does not include the action $HM$, the penalty $p_i$ will now be enforced by the Court, and agent $i$'s net remuneration will be $F \circ t_{3i}(LN, HN) = f - p_i$. The agent $i$'s deviation from $HN$ to $LN$ is deterred. The enforceability principle fails.

The violation of the enforceability principle in Example 3 is a consequence of the failure of negative transitivity of deterrence. With the function $F$ defined as above for each player $i$, the contractual deterrence order $L_i$ takes the simple form:

$$(a_i', a_i) \in L_i[a_{-i}; G, F] \text{ if and only if } a_i' \notin \pi_i(P(a_i, a_{-i})).$$

In words, given that the opponents play $a_{-i}$, the Court can conclude that agent $i$ did not play $a_i'$, if $i$ played $a_i$. (Again, I omit the dependence of $L_i$ on $\theta_i$.) I note that, for any opponents' actions $a_{-i}$, the mapping $\pi_i(P(\cdot, a_{-i})) : A^i \to 2^{A_i}$ can be understood as the Court's information correspondence restricted to agent $i$'s action $a_i$.

In Example 3, while the information correspondence $P$ is partitional, and hence transitive, the information correspondence $\pi_i(P(\cdot, a_{-i}))$ is not transitive. Specifically, $HN \in \pi_i(P(HM, HN))$ and $LN \in \pi_i(P(HN, HN))$ but $LN \notin \pi_i(P(HM, HN))$ for both $i = 1, 2$. As a result, $HN \notin L_i[HN; G, F]$, $LN \notin L_i[HN; G, F]$, but $LN \in L_i[HN; G, F]$, the order $L_i[HN; G, F]$ fails to be negative transitive. Each agent $i$ can be deterred from deviating from $HN$ to $LN$ when contractually committing to play $HM$, as long as the other agent $j$ plays $HN$. However, $i$ cannot be deterred from playing $LN$ when committing to play $HN$, nor for playing $HN$ when committing to play $HM$. $\diamond$

# 6 Conclusion

This paper has identified situations where desired outcomes may only be implemented with the stipulation of 'overspecified contracts', defined as contracts that include clauses that are not enforceable. The desired outcome is achieved in equilibrium by secretly breaching the stipulated contract. Such mechanisms are needed to circumvent legal restrictions or practical enforcement limitations that prevent implementation with enforceable contracts.

I have developed a general model of contract enforcement that does not presume that parties only sign enforceable contracts. I have used this framework to ask: When is the assumption that parties sign only enforceable contracts without loss? Within the standard domain in which Courts process information as standard Bayesian agents and the only legal restriction to contract enforcement is limited liability, usual contract theoretical modelling (which assumes that parties sign only enforceable contracts) is without loss. My richer formalization identifies two legal principles because of which the enforceability principle may fail in economically relevant situations. The first one is the standard provision that limits contractual transfers to compensation of verified damages and foregone profits, thus ruling out clauses that are punitive in nature. The second principle prescribes that any individual agent may only be sanctioned if identified in breach of her own contractual commitment, and hence rules out collective punishments.

# References


Aghion, Philippe and Patrick Bolton (1987): "Contracts as a Barrier to Entry," *American Economic Review,* 77: 388-401

Aliprantis, Charlambos and Kim C. Border (1999): *Infinite Dimensional Analysis: A Hitchhiker's Guide (3rd ed.).* Ch. 17, Springer-Verlag: Berlin Heidelberg.

Anderlini, Luca, Leonardo Felli and Andrew Postlewaite (2007): "Courts of Law and Unforeseen Contingencies," *Journal of Law, Economics, & Organization,* 23: 662–684.

Bernheim, B. Douglas, and Michael D. Whinston (1998): "Incomplete Contracts and Strategic Ambiguity," *American Economic Review*, 88: 902-32.

Bull, Clive (1987): "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics,* 102(1): 147-59.

Geanakoplos, John G. (1989): "Game Theory without Partitions, and Applications to Speculations and Consensus," mimeo, *Cowles Foundation for Research at Yale University.*

Green, Jeremy R., and Jean Jacques Laffont (1977): "Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods," *Econometrica,* 45: 427-438.

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

Grossman, Sanford J., and Oliver Hart (1986): "The Costs and Benefits of Ownership: a Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94: 691-719

Hart, Oliver, and John Moore. 1990. "Property Rights and the Nature of the Firm." *Journal of Political Economy,* 98(6):1119-1158.

Hart, Oliver D. and Jean Tirole (1988): "Contract Renegotiation and Coasian Dynamics," *Review of Economic Studies,* 55(4), 509-540.

Huberman, Gur, and Charles Kahn (1988): "Limited Contract Enforcement and Strategic Renegotiation," *American Economic Review* 78(3): 471-484.

Kaplow, Louis (2000): "General Characteristics of Rules," in B. Bouckaert and G. DeGeest, eds., *Encyclopedia of Law and Economics. Vol. V.* Cheltenham, UK: Edward Elgar.

Keating, Daniel (1997): "Measuring Sales Law against Sales Practice: a Reality Check," *The Journal of Law and Commerce*, 17: 99

Kornhauser Lewis A., and Richard L. Revesz (1994): " Multi-Defendant Settlements, the Impact of Joint and Several Liability," *Journal of Legal Studies* 41-76.

Maskin Erik, and Jean Tirole (1990), "The Principal-Agent Relationship with an Informed Principal: The Case of Private Values," *Econometrica,* 58, 379–409.

Maskin Erik, and Jean Tirole (1992), "The Principal-Agent Relationship with an Informed Principal, II: Common Values," *Econometrica,* 60, 1–42.

MacLeod, Bentley and James Malcomson (1989), "Implicit contracts, incentive compatibility, and unvoluntary unemployment," *Econometrica*, 57, 447-480.

Myerson, Roger B. (1979): "Incentive Compatibility and the Bargaining Problem," *Econometrica*, **47(1)**, 61–73.

Mylovanov, Tymofiy and Thomas Tröger (2014): Mechanism Design by an Informed Principal: Private Values with Transferable Utility," *Review of Economic Studies,* 81(4): 1668–1707.

Prat, Andrea and Aldo Rustichini (2003), "Games Played through Agents," *Econometrica,* 71, 989-1026.

*Restatement (Second) of Contracts,* copyright 1981-2003 by the *American Law Institute*

Savage, Leonard J. (1954): *The Foundations of Statistics,* New York: John Wiley & Sons.

Segal, Ilya, and Michael D. Whinston (2003): "Robust predictions for bilateral contracting with externalities," *Econometrica,* 71(3): 757-791.

Shin, Hyun Song (1993): "Logical Structure of Common Knowledge," *Journal of Economic Theory* **60**: 1-13.

*Uniform Commercial Code,* copyright 1978-2001 by the *American Law Institute* and the *National Conference of Commissioners on Uniform State Laws*

Williamson, Oliver. 1985. *The Economic Institutions of Capitalism.* New York: Free Press

# Appendix for online publication

**Notes on Example 1.** I earlier assumed that when a player $i$ contributes less than the minimal contribution $\hat{c}$ stipulated in the social contract, she is sanctioned if and only if the magnitude of her contract violation $d_i = \hat{c} - c_i$ is above a given "enforcement margin" $m > 0$, possibly function of $\hat{c}$ and of the opponents' contributions $\mathbf{c}_{-i}$. Specifically, the net social benefit $b$ of enforcing some sanction $p \geq u_i(\hat{c} - d_i, \mathbf{c}_{-i}) - u_i(\hat{c}, \mathbf{c}_{-i})$ on player $i$ is such that $b(d_i, p, \hat{c}, \mathbf{c}_{-i}) > 0$ if and only if $d_i > m(\hat{c}, \mathbf{c}_{-i})$. Here, I present conditions on the net benefit and utility functions $b$ and $u_i$ such that this assumption holds.

Simple sufficient conditions are that the function $b$ is strictly increasing in $d_i$, there exist $\varepsilon > 0$ for which $b(d_i, p, \hat{c}, \mathbf{c}_{-i}) < 0$ for all $d_i < \varepsilon$, $p$, $\hat{c}$, $\mathbf{c}_{-i}$, and $-\frac{\partial b(d_i,p,\hat{c},\mathbf{c}_{-i})/\partial d_i}{\partial b(d_i,p,\hat{c},\mathbf{c}_{-i}/\partial p} \geq \frac{\partial u_i(c_i,\mathbf{c}_{-i})}{\partial c_i}$, for $p = \max\{p' : b(d_i, p', \hat{c}, \mathbf{c}_i) \geq 0\}$ and $c_i = \hat{c} - d_i$.

For any $d_i, \hat{c}, \mathbf{c}_{-i}$, let $P(d_i; \hat{c}, \mathbf{c}_{-i}) \equiv \{p : b(d_i, p; \hat{c}, \mathbf{c}_{-i}) \geq 0\}$ and $\overline{p}(d_i; \hat{c}, \mathbf{c}_{-i}) = \max P(d_i; \hat{c}, \mathbf{c}_{-i})$. By limited liability, $b(d_i, p; \hat{c}, \mathbf{c}_{-i}) < 0$ for all $p > u_i(\hat{c} - d_i, \mathbf{c}_{-i}) - \underline{u}$, for some given threshold $\underline{u}$. Hence, $\overline{p}(d_i; \hat{c}, \mathbf{c}_{-i})$ is well defined for all non-empty $P(d_i; \hat{c}, \mathbf{c}_{-i})$.

Because the function $b$ strictly increases in $d_i$, and $b(d_i, p, \hat{c}, \mathbf{c}_{-i}) < 0$ for all $d_i < \varepsilon, p, \hat{c}, \mathbf{c}_{-i}$, it follows that there exists a threshold $\hat{m}(\hat{c}, \mathbf{c}_{-i}) > 0$ such that $P(d_i; \hat{c}, \mathbf{c}_{-i})$ is non-empty if and only if $d_i \geq \hat{m}(\hat{c}, \mathbf{c}_{-i})$. In other terms, the domain $D(\hat{c}, \mathbf{c}_{-i}) \equiv \{d_i : P(d_i; \hat{c}, \mathbf{c}_{-i}) \neq \emptyset\}$ of function $\overline{p}(\cdot; \hat{c}, \mathbf{c}_{-i})$ is such that $D(\hat{c}, \mathbf{c}_{-i}) = [\hat{m}(\hat{c}, \mathbf{c}_{-i}), \hat{c}]$.

There exists $p \geq \Delta(d_i; \hat{c}, \mathbf{c}_{-i}) \equiv u_i(\hat{c} - d_i, \mathbf{c}_{-i}) - u_i(\hat{c}, \mathbf{c}_{-i})$ such that $b(d_i, p, \hat{c}, \mathbf{c}_{-i}) > 0$, if and only if $\overline{p}(d_i; \hat{c}, \mathbf{c}_{-i}) \geq \Delta(d_i; \hat{c}, \mathbf{c}_{-i})$. Suppose I show $\frac{\partial \overline{p}(d_i;\hat{c},\mathbf{c}_{-i})}{\partial d_i} > \frac{\partial \Delta(d_i;\hat{c},\mathbf{c}_{-i})}{\partial d_i} = -\frac{\partial u_i(\hat{c}-d_i,\mathbf{c}_{-i})}{\partial c_i}$ for all $d_i$ that belong to the domain $D(\hat{c}, \mathbf{c}_{-i})$ of $\overline{p}(\cdot; \hat{c}, \mathbf{c}_{-i})$. Then, this would imply that on the domain $D(\hat{c}, \mathbf{c}_{-i})$ there exists a unique threshold $m(\hat{c}, \mathbf{c}_{-i})$ such that $\overline{p}(d_i; \hat{c}, \mathbf{c}_{-i}) \geq \Delta(d_i; \hat{c}, \mathbf{c}_{-i})$ if and only if $d_i \geq m(\hat{c}, \mathbf{c}_{-i})$. Because $m(\hat{c}, \mathbf{c}_{-i}) \geq \hat{m}(\hat{c}, \mathbf{c}_{-i}) > 0$, this proves the result sought after: there exists $p \geq \Delta(d_i; \hat{c}, \mathbf{c}_{-i})$ such that $b(d_i, p, \hat{c}, \mathbf{c}_{-i}) > 0$, if and only if $d_i > m(\hat{c}, \mathbf{c}_{-i})$.

Indeed, by the implicit function theorem, $\frac{\partial \overline{p}(d_i;\hat{c},\mathbf{c}_{-i})}{\partial d_i} = -\frac{\partial b(d_i,p,\hat{c},\mathbf{c}_{-i})/\partial d_i}{\partial b(d_i,p,\hat{c},\mathbf{c}_{-i}/\partial p} > -\frac{\partial u_i(c_i,\mathbf{c}_{-i})}{\partial c_i}$ for $p = \overline{p}(d_i; \hat{c}, \mathbf{c}_{-i})$ and $c_i = \hat{c} - d_i$, where the latter inequality is the final condition considered.

**Omitted analysis of Example 2.** Suppose that the Court presumes that the agent acted in good faith to minimize the client's damage $v_1(\hat{\mathbf{e}}) - v_1(\mathbf{e})$. Then, upon not verifying effort $e_i$ in either task $i = A, B$, the Court presumes that the agent exerted the maximum level of effort possible $\bar{e}$ in task $i$. Hence, given contractual obligation efforts $\hat{\mathbf{e}}$ and actual efforts $\mathbf{e}$, the client's expected damage compensation is:

$$\tilde{d}(\mathbf{e}; \hat{\mathbf{e}}) = qp^2 \max\{v_1(\hat{\mathbf{e}}) - v_1(\mathbf{e}), 0\} + qp(1-p) \sum_{i=A,B; j \neq i} \max\{v_1(\hat{\mathbf{e}}) - v_1(e_i, \bar{e}_j), 0\}.$$

Suppose the agent contractually committed to the first best effort levels $\mathbf{e}^*$. The agent's choice $\mathbf{e}$ satisfies the first order condition:

$$-\partial \tilde{d}(\mathbf{e}; \mathbf{e}^*)/\partial e_i = \partial c(\mathbf{e})/\partial e_i, \text{ for } i = A, B, \ j \neq i.$$

For all effort levels $\mathbf{e}$ sufficiently close to $\mathbf{e}^*$, it is the case that $v_1(\mathbf{e}^*) < v_1(e_i, \bar{e}_j)$, and hence that

$$-\partial \tilde{d}(\mathbf{e}; \mathbf{e}^*)/\partial e_i = qp^2 \partial v_1(\mathbf{e})/\partial e_i < \partial c(\mathbf{e}^*)/\partial e_i,$$

where again the inequality follows because $\mathbf{e}^*$ solves (1), $\partial v_0(\mathbf{e}^*)/\partial e_i > 0$, $0 < q < 1$ and $0 < p < 1$. As the agent's marginal remuneration of effort $e_i$ is smaller than the marginal cost for all effort levels $\mathbf{e}$ close to $\mathbf{e}^*$, the agent's chosen efforts $\mathbf{e}$ must be such that $e_i < e_i^*$ for both $i = A, B$.

First best efforts $\mathbf{e}^*$ can be implemented with the "excessive" contractual obligations $\hat{\mathbf{e}} > \mathbf{e}^*$ that solve:

$$v_1(\hat{\mathbf{e}}) = v_1(e^*, \bar{e}).$$

For any $i = A, B$, and $j \neq i$, fix $e_j = e^*$ and consider the agent's choice $e_i$. For all $e_i > e^*$, it is the case that $v_1(\tilde{\mathbf{e}}) < v_1(e_i, \bar{e})$, the marginal remuneration of effort is below marginal cost:

$$-\partial \tilde{d}((e_i, e^*); \hat{\mathbf{e}})/\partial e_i = qp^2 \partial v_1(e_i, e^*)/\partial e_i < \partial c(e_i, e^*)/\partial e_i.$$

For any $e_i \leq e_i^*$, it is the case that $v_1(\hat{\mathbf{e}}) \leq v_1(e_i, \bar{e})$, hence the agent's marginal remuneration of effort $e_i$ is:

$$-\partial \tilde{d}((e_i, e^*); \hat{\mathbf{e}})/\partial e_i = qp[p\partial v_1(e_i, e^*)/\partial e_i + (1-p)\partial v_1(e_i, \bar{e})/\partial e_i].$$

Because $\partial^2 v/\partial e_i \partial e_j > 0$, and the contractual commitment $\hat{\mathbf{e}}$ optimal in the earlier case is smaller than $\bar{\mathbf{e}}$,

$$-\partial \tilde{d}((e_i, e^*); \hat{\mathbf{e}})/\partial e_i > -\partial d((e_i, e^*); \hat{\mathbf{e}})/\partial e_i = qp[p\partial v_1(e_i, e^*)/\partial e_i + (1-p)\partial v_1(e_i, \hat{e})/\partial e_i] = \partial c(e_i, e_j^*)/\partial e_i.$$

As a result, the overburdening contractual obligations $\hat{\mathbf{e}} > \mathbf{e}^*$ lead the agent to choose the optimal effort levels $\mathbf{e}^*$.

**Proof or Proposition 1.** Suppose that all transfers are legal. For any contract $t$, and outcome $x \in A \times \Omega$, the contract $F \circ t$ is such that $F \circ t(x) = t(x)$ whenever $t$ is constant on $P(x)$, and $F \circ t(x) \neq t(x)$ otherwise. For any enforcement function $F$ that satisfies Assumption 1, $F \circ t(x)$ is constant on $P(x)$. Then, evidently, $F \circ F \circ t(x) = F \circ t(x)$.

Suppose that the only legal constraint is bounded liability, with bounds $\bar{u}_i$ for all $i$. Here, for any outcome $x \in A \times \Omega$, it is the case that $F \circ t(x) = t(x)$ if and only if $t(x)$ is constant on $P(x)$ and $\min_{y \in P(x)} u_i(y) + \tau_i(x)|t \geq \bar{u}_i$. When this is not the case, again, $F \circ F \circ t(x) = F \circ t(x)$ for any enforcement function $F$ that satisfies Assumption 1. ■

**Proof of Proposition 2.** I first show that if all contracts are legal, and $P$ is transitive, then $F$ is idempotent.

By contradiction, suppose that there is a contract $t$ such that $F \circ F \circ t(x) \neq F \circ t(x)$, for some $x \in A \times \Omega$. I.e., there is $i, j$ such that $F \circ F \circ t_{ij}(x) \neq F \circ t_{ij}(x)$. Under the voidance rule, for this to be the case, it must be that $0 = F \circ F \circ t_{ij}(x) \neq F \circ t_{ij}(x) = t_{ij}(x)$. The second equality implies that, for all $(a, \omega) \in P(x)$, it is the case that $t_{ij}(a, \omega) = t_{ij}(x)$. The inequality implies that there exists $y \in P(x)$ such that $F \circ t_{ij}(y) \neq F \circ t_{ij}(x)$. Because $F \circ t_{ij}(x) = t_{ij}(x)$ and, as concluded above, $t_{ij}(y) = t_{ij}(x)$, this means that $0 = F \circ t_{ij}(y) \neq t_{ij}(y)$. Hence, there must be $z \in P(y)$ such that $t_{ij}(y) \neq t_{ij}(z)$. But I earlier found that $t_{ij}(a, \omega) = t_{ij}(x)$ for all $(a, \omega) \in P(x)$; hence it must be that $z \notin P(x)$, and I have derived a failure of transitivity.

Turning to the conservative rule, in which $F \circ t_{ij}(x) = \min_{(a,\omega) \in P(x)} t_{ij}(a, \omega)$, first I note that when $F \circ F \circ t_{ij}(x) \neq F \circ t_{ij}(x)$ and $F \circ t_{ij}(x) = t_{ij}(x)$ for some $x \in A \times \Omega$ and $i, j$, the proof is the same as the one above with the voidance rule. Continuing to show the contradiction, suppose that, for some $x \in A \times \Omega$ and $i, j$, it is the case that $F \circ F \circ t_{ij}(x) \neq F \circ t_{ij}(x)$ and that $F \circ t_{ij}(x) \neq t_{ij}(x)$. Hence, $F \circ t_{ij}(x) = \min_{(a,\omega) \in P(x)} t_{ij}(a, \omega) < t_{ij}(x)$ and there must be a $w \in P(x)$ such that $t_{ij}(w) = \min_{(a,\omega) \in P(x)} t_{ij}(a, \omega) = F \circ t_{ij}(x)$. Further, $F \circ F \circ t_{ij}(x) = \min_{(a,\omega) \in P(x)} F \circ t_{ij}(a, \omega) < F \circ t_{ij}(x)$ and there must be a $y \in P(x)$ such that $F \circ t_{ij}(y) = \min_{(a,\omega) \in P(x)} F \circ t_{ij}(a, \omega) = F \circ F \circ t_{ij}(x)$. By construction, $t_{ij}(w) \leq t_{ij}(y)$ and $F \circ t_{ij}(y) < t_{ij}(w)$. Jointly, these two results imply that $F \circ t_{ij}(y) < t_{ij}(y)$. Hence, there must exist $z \in P(y)$ such that $t_{ij}(z) = \min_{(a,\omega) \in P(y)} t_{ij}(a, \omega) = F \circ t_{ij}(y)$. If $P$ were transitive, $y \in P(x)$ and $z \in P(y)$ would imply $z \in P(x)$. This would imply $t_{ij}(z) \geq F \circ t_{ij}(x)$, by construction. But $t_{ij}(w) > F \circ t_{ij}(y) = t_{ij}(z)$ and $t_{ij}(w) = F \circ t_{ij}(x)$ imply $t_{ij}(z) < F \circ t_{ij}(x)$, reaching a contradiction.

It is easy to see that an analogous chain of arguments establishes the claimed result also for the other rules defined in Section 4. ■

**Proof of Proposition 3.** Simplifying notation, suppose that there is a triple $s_i, a_i, b_i$ such that $(a_i, s_i) \notin L_i[\theta_i, s_{-i}]$, $(s_i, b_i) \notin L_i[\theta_i, s_{-i}]$ and $(a_i, b_i) \in L_i[\theta_i, s_{-i}]$ for some $\theta_i$ and $s_{-i}$: negative transitivity for $L_i[\theta_i, s_{-i}]$ is violated.

Because $(s_i, b_i) \notin L_i[\theta_i, s_{-i}]$, it is the case that: $u_{\theta_i}(b_i, s_{-i}) + \tau_{\theta_i}(b_i, s_{-i})|F \circ t > u_{\theta_i}(s_i, s_{-i}) + \tau_{\theta_i}(s_i, s_{-i})|t$, for every contract $t$. For any contract $t$ enforceable at $s$, it is the case that $\tau_{\theta_i}(s_i, s_{-i})|F \circ t = \tau_{\theta_i}(s_i, s_{-i})|t$. Hence, $s$ cannot be implemented with any contract enforceable at $s$.

Because $(a_i, b_i) \in L_i[\theta_i, s_{-i}]$, there exist contracts $t$ such that $u_{\theta_i}(b_i, s_{-i}) + \tau_{\theta_i}(b_i, s_{-i})|F \circ t \leq u_{\theta_i}(a_i, s_{-i}) + \tau_{\theta_i}(a_i, s_{-i})|t$. Because $(a_i, s_i) \notin L_i[\theta_i, s_{-i}]$, $u_{\theta_i}(s_i, s_{-i}) + \tau_{\theta_i}(s_i, s_{-i})|F \circ t > u_{\theta_i}(a_i, s_{-i}) + \tau_{\theta_i}(a_i, s_{-i})|t$ for all contracts $t$. Subtracting the first inequality from the second, and simplifying I realize that there exist contracts $t$ such that $u_{\theta_i}(b_i, s_{-i}) + \tau_{\theta_i}(b_i, s_{-i})|F \circ t < u_{\theta_i}(s_i, s_{-i}) + \tau_{\theta_i}(s_i, s_{-i})|F \circ t$. Under all such contracts $t$, type $\theta_i$ of player $i$ prefers to play $s_i$ than $b_i$ if the opponents play $s_{-i}$.

This verifies the stated claim because either it is the case that none such contracts $t$ implements $s$, or that at least one contract $t$ does, in which case the enforceability principle is violated at $s$. ■

**Slack enforcement and Example 1 revisited.** I represent slack contract enforcement simply by assuming that for every game $G$, contract $t$, players $i, j$ and outcome $x \in A \times \Omega$, the transfer $t_{ij}(x) = p$ is enforced if and only if the social net benefits $\hat{b}_{ij}(p; x, t, G)$ of enforcement are positive. By consistency, the function $\hat{b}_{ij}$ is such that $\hat{b}_{ij}(p; x, t, G) = \hat{b}_{ji}(-p; x, t, G)$. For any transfer $t_{ij}(x) > 0$, the enforcement function $F$ is such that $F \circ t_{ij}(x) = t_{ij}(x)$ if $\hat{b}_{ij}(t_{ij}(x); x, t, G) > 0$, and else $F \circ t_{ij}(x)$ is equal to some transfer value $p$ such that $\hat{b}_{ij}(p; x, t, G) > 0$. For example, if the Court adopts a 'voidance transfer determination rule', then $F \circ t_{ij}(x) = 0$ whenever $\hat{b}_{ij}(t_{ij}(x); x, t, G) \leq 0$. Conservative and seniority transfer determination rules are defined analogously.

Returning to Example 1, each player $i = 1, ..., n$ chooses a contribution $c_i \geq 0$, the first best is $c_i = c^*$ for all $i$, but each player $i$'s payoff $u_i(c_i, \mathbf{c}_{-i})$ decreases in $c_i$. There is an idle player $e$ (the "enforcer") whose role is to collect the sanctions $t_{ie}(\mathbf{c})$ stipulated in the social contract $t$, from any player $i$. For simplicity, say that no other player $j$ is allowed to collect sanctions from $i$. The enforcer has no budget to subsidize virtuous contributions outside of what she collects. As the contribution game is symmetric, let me assume that the enforcer must treat all agents in the same way. Hence, for any symmetric profile $\mathbf{c}$, it must be that the transfer $t_{ie}(\mathbf{c})$ is weakly positive and equal across all agents $i$.

The Court is perfectly informed of every player $i$'s contribution $c_i$, but enforcement of sanctions is slack. Given a social contract $t$, suppose that a player $i$ contributes less than all the other players $j$, specifically $c_j = \hat{c}$ for all $j \neq i$ and $c_i < \hat{c}$. Consider the net social benefits $\hat{b}_{ij}(p, (c_i, \hat{\mathbf{c}}_{-i}), t, G)$ of enforcing a sanction $p$ that would deter $i$'s play $c_i$, i.e. a sanction $p \geq u_i(c_i, \hat{\mathbf{c}}_{-i}) - u_i(\hat{\mathbf{c}}) + t_{ie}(\hat{\mathbf{c}})$. To represent slack enforcement, let me say that $\hat{b}_{ie}(p, (c_i, \hat{\mathbf{c}}_{-i}), t, G) > 0$ if and only if the under-contribution $\hat{c} - c_i$ is above the threshold $m(\hat{c}, \hat{\mathbf{c}}_{-i})$ defined in Example 1.

As a result, the first best profile $\mathbf{c}^*$ cannot be implemented with any contract $t$ enforceable at $\mathbf{c}^*$, i.e. such that $F \circ t(\mathbf{c}^*) = t(\mathbf{c}^*)$. Plainly, any player $i$ cannot be deterred from making a contribution $c_i < c^*$ such that $c^* - c_i \leq m(c^*, \mathbf{c}^*_{-i})$, when the opponents play $\mathbf{c}^*_{-i}$. The social net benefits of enforcing any sanction $p \geq u_i(c_i, \hat{\mathbf{c}}_{-i}) - u_i(\hat{\mathbf{c}}) + t_{ie}(\hat{\mathbf{c}})$ are negative. Hence, $F \circ t_{ie}(c_i, \mathbf{c}^*_{-i}) < u_i(c_i, \mathbf{c}^*_{-i}) - u_i(\mathbf{c}^*) + t_{ie}(\mathbf{c}^*) = u_i(c_i, \mathbf{c}^*_{-i}) - u_i(\mathbf{c}^*) + F \circ t_{ie}(\mathbf{c}^*)$, as $t$ is enforceable at $\mathbf{c}^*$. Rearranging, I obtain: $u_i(\mathbf{c}^*) + F \circ t_{ie}(c_i, \mathbf{c}^*_{-i}) < u_i(c_i, \mathbf{c}^*_{-i}) + F \circ t_{ie}(\mathbf{c}^*)$: player $i$ cannot be deterred from playing $c_i$. But we know from Example 1 that there exist simple social contracts $t$ that implement $\mathbf{c}^*$. Such simple contracts prescribe a sanction $p(c_i, \mathbf{c}^*_{-i}) \geq u_i(c_i, \mathbf{c}^*_{-i}) - u_i(c^* + m(\hat{c} - c^*, \mathbf{c}^*_{-i}), \mathbf{c}^*_{-i})$ on any player $i$ for contributing less than $\hat{c} \equiv c^* + m(\hat{c} - c^*, \mathbf{c}^*_{-i})$, when the opponents contribute $\mathbf{c}^*_{-i}$. The enforceability principle fails.

As predicted by Proposition 3, this violation of the enforceability principle is a consequence of failure of negative transitivity of contractual deterrence. With the enforcement function $F$ defined above, for each player $i$, the contractual deterrence order $L_i$ takes the following simple form for $c'_i < c_i$ (I omit the dependence on $\theta_i$ as it is irrelevant here):

$$(c'_i, c_i) \in L_i[\mathbf{c}^*_{-i}; G, F] \text{ if and only if } c_i - c'_i > m(c_i, \mathbf{c}^*_{-i}).$$

It is easy to show that the order $L_i[\mathbf{c}^*_{-i}; G, F]$ is not negative transitive. Proceeding as with Example 4, for any $c_i \in [c^* - m(c^*, \mathbf{c}^*_{-i}), c^*)$, it is the case that $(c_i, c^*) \notin L_i[\mathbf{c}^*_{-i}; G, F]$ and $(c^*, \hat{c}) \notin L_i[\mathbf{c}_{-i}; G, F]$, but $(c_i, \hat{c}) \in L_i[\mathbf{c}_{-i}; G, F]$. A player $i$ cannot be deterred from playing $c_i$ by a contract $t$ that sanctions contributing less than $c^*$, nor to play $c^*$ by sanctions for contributing less than $\hat{c}$, but such sanctions succeed in deterring contribution $c_i$. The role of enforcement margins $m$ here in making $L_i$ violate negative transitivity is the same as the role played by the "margin of errors" $m'$ of the intransitive verification correspondence $P$ in Example 4. ⋄

**Proof of Proposition 4.** The analysis in the first part of this proof considers the enforcement function $F$ in (5). When the Court verifies that the agent broke its contractual commitment, it presumes that the agent made the smallest possible deviation from the commitment.

Pick a contract $t : A \times \Omega \to R_+^{I \times I}$, the notation $t_{12}(\mathbf{e}, \omega)$ stands for the agent's remuneration for effort levels $\mathbf{e} = (e_A, e_B)$. In line with the exposition of Example 2, I first calculate the enforced transfer $t_{12}(\mathbf{e}, \omega)|F \circ t$ according to the rule in (5). That is, when assessing whether a transfer $t_{12}(\mathbf{e}, \omega)$ is justified by damage compensation relative to stipulated efforts $\hat{\mathbf{e}}$, the Court presumes that the agent's efforts $\mathbf{e}'$ were as close as possible to $\hat{\mathbf{e}}$, compatibly with her knowledge $P(\mathbf{e}, \omega)$. Let me rewrite expression (5) for any contingency $(\mathbf{e}, \omega)$ as follows: $t_{12}(\mathbf{e}, \omega)|F \circ t = \max\{t_{12}(\mathbf{e}, \omega), \max_{(\hat{\mathbf{e}}, \omega') \in P(\mathbf{e}, \omega)} R(\hat{\mathbf{e}}, \omega'; \mathbf{e}, \omega)|t\}$. Here, the expression $R(\tilde{\mathbf{e}}, \omega'; \mathbf{e}, \omega)|t \equiv t_{12}(\hat{\mathbf{e}}, \omega') - \min_{\mathbf{e}'' \in \arg\min_{(\mathbf{e}', \omega') \in P(\mathbf{e}, \omega)} ||e' - e||} d(\hat{\mathbf{e}}; \mathbf{e}'', \omega')$ denotes the agent's remuneration $t_{12}(\hat{\mathbf{e}}, \omega')$ associated with contingency $(\hat{\mathbf{e}}, \omega')$, net of the minimum possible damage to the principal $d(\hat{\mathbf{e}}; \mathbf{e}'', \omega') \equiv \max\{0, u_1(\hat{\mathbf{e}}, \omega') - u_1(\mathbf{e}'', \omega')\}$ compatible with the Court's presumption that the agent's effort $\mathbf{e}''$ is as close as possible to $\hat{\mathbf{e}}$ within the knowledge that $(\mathbf{e}'', \omega') \in P(\mathbf{e}, \omega)$. Because the Court knows $\omega$, I can further simplify the above expressions as:

$$t_{12}(\mathbf{e}, \omega)|F \circ t = \max\{t_{12}(\mathbf{e}, \omega), \max_{(\hat{\mathbf{e}}, \omega) \in P(\mathbf{e}, \omega)} R(\hat{\mathbf{e}}; \mathbf{e}, \omega)|t\} \tag{7}$$

where $R(\hat{\mathbf{e}}; \mathbf{e}, \omega)|t \equiv t_{12}(\hat{\mathbf{e}}, \omega) - \min_{\mathbf{e}'' \in \arg\min_{(\mathbf{e}', \omega) \in P(\mathbf{e}, \omega)} ||e' - e||} d(\hat{\mathbf{e}}; \mathbf{e}'', \omega)$.

If the Court cannot verify $e_A$ nor $e_B$, $\omega_A = n = \omega_B$, then $R(\hat{\mathbf{e}}; \mathbf{e}, \omega) = t_{12}(\hat{\mathbf{e}}, \omega)$, and using (7), the enforced transfer is $t_{12}(\mathbf{e}, \omega)|F \circ t = \max\{t_{12}(\mathbf{e}, \omega),\ t_{12}(\hat{\mathbf{e}}; \omega)\}$. This happens either when $\omega_1 = 0$, an event of probability $1 - q$, or when $\omega_1 = 1$ an event of probability $q(1-p)^2$. With probability $qp^2$, $\omega_1 = 1$, $\omega_A = v = \omega_B$, and the Court verifies both $e_A$ and $e_B$. Then, $R(\hat{\mathbf{e}}; \mathbf{e}, \omega) = t_{12}(\hat{\mathbf{e}}, \omega) - \max\{0, v_1(\hat{\mathbf{e}}) - v_1(\mathbf{e})\}$ and the enforced transfer is $t_{12}(\mathbf{e}, \omega)|F \circ t = \max\{t_{12}(\mathbf{e}, \omega),$ $t_{12}(\hat{\mathbf{e}}, \omega) - \max\{0, v_1(\hat{\mathbf{e}}) - v_1(\mathbf{e})\}\}$. Finally, for both $i = A, B$, $j \neq i$, it is the case that $\omega_1 = 1$, $\omega_i = v$ and $\omega_j = n$ with probability $q(1-p)$. The Court verifies $e_i$ but not $e_j$, and then $R(\hat{\mathbf{e}}; \mathbf{e}, \omega) = t_{12}(\hat{\mathbf{e}}, \omega) - \max\{0, v_1(\hat{\mathbf{e}}) - v_1(e_i, \hat{e}_j)\}$ and the enforced transfer is $t_{12}(\mathbf{e}, \omega)|F \circ t = \max\{t_{12}(\mathbf{e}, \omega),$ $t_{12}(\hat{\mathbf{e}}, \omega) - \max\{0, v_1(\hat{\mathbf{e}}) - v_1(e_i, \hat{e}_j)\}\}$.

Let me momentarily focus on the simple contracts of Example 2, expressed as contracts $t$ such that $t_{12}(\mathbf{e}, \omega) = f$ if $\mathbf{e} \geq \hat{\mathbf{e}}$ and $t_{12}(\mathbf{e}, \omega) = 0$ if $\mathbf{e} \ngeq \hat{\mathbf{e}}$. I note that for any $\hat{\mathbf{e}} \ngeq \hat{\mathbf{e}}$, it is the case that $t_{12}(\hat{\mathbf{e}}, \omega) = 0$ and hence $R(\hat{\mathbf{e}}; \mathbf{e}, \omega) = 0$. If instead $\hat{\mathbf{e}} \geq \hat{\mathbf{e}}$ then $t_{12}(\hat{\mathbf{e}}, \omega) = f$, and then the largest possible remunerations $R(\hat{\mathbf{e}}; \mathbf{e}, \omega)$ are realized with $\hat{\mathbf{e}} = \hat{\mathbf{e}}$, for any $(\mathbf{e}, \omega)$, and I can restrict attention to the case that $\hat{\mathbf{e}} = \hat{\mathbf{e}}$ without loss of generality.

For any effort levels $\mathbf{e} \geq \hat{\mathbf{e}}$, it is the case that $t_{12}(\mathbf{e}, \omega)|F \circ t = t_{12}(\mathbf{e}, \omega) = f$. Consider any pair of effort levels $\mathbf{e} \ngeq \hat{\mathbf{e}}$. Because $t_{12}(\mathbf{e}, \omega) = 0$, it is the case that $t_{12}(\mathbf{e}, \omega)|F \circ t = f$ with probability $(1-q) + q(1-p)^2$, $t_{12}(\mathbf{e}, \omega)|F \circ t = f - \max\{0, v_1(\hat{e}) - v_1(\mathbf{e})\}$ with probability $qp^2$ and $t_{12}(\mathbf{e}, \omega)|F \circ t = f - \max\{0, v_1(\hat{e}) - v_1(e_i, \hat{e}_j)\}$ with probability $qp(1-p)$ for both $i = A, B$. Hence, for all $\mathbf{e} < \hat{\mathbf{e}}$, I have reconstructed the expected net remuneration $f - d(\mathbf{e}; \hat{\mathbf{e}})$ of the analysis of Example 2 in Section 3, where

$$d(\mathbf{e}; \hat{\mathbf{e}}) = qp^2[v_1(\hat{\mathbf{e}}) - v_1(\mathbf{e})] + q(1-p)p[v_1(\hat{\mathbf{e}}) - v_1(\hat{e}_A, e_B)] + q(1-p)p[v_1(\hat{\mathbf{e}}) - v_1(e_A, \hat{e}_B)].$$

It is then straightforward that the analysis of Section 3 applies here. The only simple contract by which the optimal effort levels $\mathbf{e}^*$ are implemented is such that the agent commits to exert the excessive effort levels $\hat{\mathbf{e}} > \mathbf{e}^*$ in exchange for a $f$ such that $c(\mathbf{e}^*) < f < v(\mathbf{e}^*)$.

Let me now consider any general contract $t : A \times \Omega \to R_+^{I \times I}$, including contracts that include bonuses and penalties. I prove in general that the optimal effort levels $\mathbf{e}^*$ cannot be implemented with any enforceable contract $t$. Consider any contract $t$ such that $t_{12}(\mathbf{e}^*, \omega)|F \circ t = t_{12}(\mathbf{e}^*, \omega)$, and let $f_{\omega_1} = t_{12}(\mathbf{e}^*, \omega)$ for both $\omega_1 = 0, 1$ regardless of $\omega_A$ and $\omega_B$.

Consider any effort levels $\mathbf{e}$ such that $v(\mathbf{e}) < v(\mathbf{e}^*)$. With probability $(1-q) + q(1-p)^2$, the Court cannot verify $e_A$ nor $e_B$, then using (7), the enforced transfer is $t_{12}(\mathbf{e}, \omega)|F \circ t = f_{\omega_1}$, because $\max_{(\hat{\mathbf{e}}, \omega) \in P(\mathbf{e}, \omega)} R(\hat{\mathbf{e}}; \mathbf{e}, \omega) \geq R(\mathbf{e}^*; \mathbf{e}, \omega) = t_{12}(\mathbf{e}^*, \omega) = f_{\omega_1}$. With probability $qp^2$, the Court verifies both $e_A$ and $e_B$, and then the enforced transfer is $t_{12}(\mathbf{e}, \omega)|F \circ t \geq \max\{t_{12}(\mathbf{e}, \omega),$ $f_1 - \max\{0, v_1(\mathbf{e}^*) - v_1(\mathbf{e})\}\}$. With probability $qp(1-p)$ the Court verifies $e_i$ but not $e_j$ and the enforced transfer is $t_{12}(\mathbf{e}, \omega)|F \circ t \geq \max\{t_{12}(\mathbf{e}, \omega),\ f_1 - \max\{0, v_1(\mathbf{e}^*) - v_1(e_i, e_j^*)\}\}$.

The agent's expected payoff $E[u_2(\mathbf{e}) + t_{12}(\mathbf{e}, \omega)|F \circ t]$ for playing $\mathbf{e}$ is thus:

$$\begin{aligned} E[u_2(\mathbf{e}) + t_{12}(\mathbf{e}, \omega)|F \circ t] \quad \geq \quad & (1-q)f_0 + q(1-p)^2 f_1 + qp^2[f_1 - \max\{0, v_1(\mathbf{e}^*) - v_1(\mathbf{e})\}] \\ & + qp(1-p) \sum_{i=A,B} [f_1 - \max\{0, v_1(\mathbf{e}^*) - v_1(e_i, e_j^*)\}] - c(\mathbf{e}). \end{aligned}$$

The agent's expected payoff for playing $\mathbf{e}^*$ is instead $E[u_2(\mathbf{e}^*) + t_{12}(\mathbf{e}^*, \omega)|F \circ t] = (1-q)f_0 + qf_1 - c(\mathbf{e}^*)$. The expected payoff difference $E[u_2(\mathbf{e}^*) + t_{12}(\mathbf{e}^*, \omega)|F \circ t] - E[u_2(\mathbf{e}) + t_{12}(\mathbf{e}, \omega)|F \circ t]$ is thus smaller than the payoff difference in the analysis of Example 2 in Section 3:

$$qp^2[v_1(\mathbf{e}^*) - v_1(\mathbf{e})] + qp(1-p) \sum_{i=A,B} [v_1(\mathbf{e}^*) - v_1(e_i, e_j^*)] + q(1-p)^2[v_1(\mathbf{e}^*) - v_1(e)] + c(\mathbf{e}) - c(\mathbf{e}^*).$$

A fortiori, the algebra of the analysis of Example 2 in Section 3 concludes that none of the considered contracts $t$ can make agent choose the optimal effort levels $\mathbf{e}^*$.

The analysis in the second part of this proof considers the enforcement function $F$ in (4). When the Court verifies that the agent broke its contractual commitment, the Court presumes that the agent acted to minimize the damage to his client. Let me rewrite (4) as follows, again using the fact that the Court observes $\omega$:

$$t_{12}(\mathbf{e}, \omega)|F \circ t = \max\{t_{12}(\mathbf{e}, \omega), \max_{\hat{\mathbf{e}}, (\mathbf{e}', \omega) \in P(\mathbf{e}, \omega)} \{t_{12}(\hat{\mathbf{e}}, \omega) - \max\{0, u_1(\hat{\mathbf{e}}, \omega) - u_1(\mathbf{e}', \omega)\}\}\}.$$

Again, the analysis starts by considering simple contracts. With the same arguments as in the first part of the proof, I show that it is without loss of generality to restrict attention to $\hat{\mathbf{e}} = \hat{\mathbf{e}}$, and I reconstruct the expected remuneration $f - \tilde{d}(\mathbf{e}; \hat{\mathbf{e}})$ of the analysis of example 2 in Section 3, where

$$d(\mathbf{e}; \hat{\mathbf{e}}) = qp^2[v_1(\hat{\mathbf{e}}) - v_1(\mathbf{e})] + qp(1-p)[v_1(\hat{\mathbf{e}}) - v_1(\bar{e}_A, e_B)] + qp(1-p)[v_1(\hat{\mathbf{e}}) - v_1(e_A, \bar{e}_B)].$$

Then again, the analysis of Section 3 straightforwardly applies here. The first best $\mathbf{e}^*$ can only be implemented with the overburdening contractual commitment $\hat{\mathbf{e}} > \mathbf{e}^*$ such that $v_1(\hat{\mathbf{e}}) = v_1(e^*, \bar{e})$.

I conclude by considering any general contract $t : A \times \Omega \to R_+^{I \times I}$, to prove again that the optimal outcome $\mathbf{e}^*$ cannot be implemented with a contract $t$ enforceable at $\mathbf{e}^*$, i.e. a contract $t$ such that $t_{12}(\mathbf{e}^*, \omega)|F \circ t = t_{12}(\mathbf{e}^*, \omega)$, for all $\omega$. Proceeding as in the first part of the proof, I conclude that the difference $E[u_2(\mathbf{e}^*) + t_{12}(\mathbf{e}^*, \omega)|F \circ t] - E[u_2(\mathbf{e}) + t_{12}(\mathbf{e}, \omega)|F \circ t]$ in the agent's expected payoff for playing $\mathbf{e}^*$ and playing any profile effort $\mathbf{e}$ such that $v(\mathbf{e}) < v(\mathbf{e}^*)$ is smaller than the payoff difference in the analysis of Example 2 in Section 3:

$$qp^2[v_1(\mathbf{e}^*) - v_1(\mathbf{e})] + qp(1-p) \sum_{i=A,B} [v_1(\mathbf{e}^*) - v_1(e_i, \bar{e}_j)] + q(1-p)^2[v_1(\mathbf{e}^*) - v_1(e)] + c(\mathbf{e}) - c(\mathbf{e}^*).$$

Again a fortiori, the algebra of the analysis of Example 2 in Section 3 concludes that none of the considered contracts $t$ can make agent choose the optimal effort levels $\mathbf{e}^*$. ■