Rochester Center for
Economic Research

Optimal Contracts under Generalized Verifiability Correspondences

Francesco Squintani

Working Paper No. 477
January 2001

UNIVERSITY OF
ROCHESTER

Optimal Contracts under Generalized Verifiability Correspondences*

Francesco Squintani
University of Rochester
Department of Economics†

November, 2000

Abstract

Before playing a game, the players may sign a contract that prescribes not to take some actions. Following the methodology introduced by Bernheim and Whinston (1998), henceforth BW, this paper models verifiability as a correspondence mapping actually played actions into sets of actions that cannot be ruled out by a court. BW characterize optimal contracts in various settings both static and sequential, and show instances where optimal contracts must necessarily leave some verifiable prescriptions unspecified. This paper focuses on static settings. We extend the line of research of BW by considering also non-partitional and non-product correspondences, and by introducing different liability regimes in the framework. A complete characterization of optimal contracts is derived. We identify instances where, because of liability constraints, the optimal contracts must explicitly include unverifiable prescriptions. In some cases, the optimal outcome may be achieved only by signing a contract that cannot be enforced. Our analysis may be of some relevance for the foundations of incomplete contracts, as it shows that it is not necessarily the case that the players should sign an incomplete contract when complete contracts are unenforceable.

JEL CLASSIFICATION: C72, K40

*I thank Larry Blume, Betsy Caucutt, Samuel Danthine, Eddie Dekel, Steven Matthews, Juuso Valimaki, John Woodruff, two referees and an Associate Editor. The usual caveat applies.

†Harkness Hall, University of Rochester, 14627 NY.

1 Introduction

Before playing a game, the players may sign a contract that prescribes not to take some actions. In order to punish a player violating such an agreement, a court must be able to verify the actions played.[1] Following the methodology introduced by Bernheim and Whinston (1998), henceforth BW, we model verifiability constraints as a correspondence mapping each action profile actually played into a set of action profiles that cannot be ruled out by the court. In particular, correspondences whose ranges include non-singleton sets represent cases of imperfect verifiability, and product correspondences represent cases where the action taken by each player can be verified independently of the opponents' choices.[2]

BW focus on product and partitional correspondences. They show that in static settings the players can restrict attention without loss to the most restrictive contracts that do not include any unverifiable prescription. In sequential settings, they show that the players may be better off by signing less restrictive contracts that leave some verifiable aspects of the interaction unspecified. We focus on static settings, extend the analysis to non-product and to non-partitional correspondences, and derive a complete characterization of the optimal contracts. The use of the methodology of verifiability correspondences, instead of the representation of evidence as stochastic signals, greatly simplifies the analysis and allows us to derive sharper results.

Our analysis may be of some interest for the foundations of contract incompleteness.[3] Imperfect verifiability has been advocated by many as a straightforward justification for

[1]Maskin and Tirole (1999a) identify sufficient conditions under which subgame-perfect implementation message mechanisms (see Moore and Repullo 1998) can alleviate verifiability constraints. The main difficulties that they highlight consist of the requirements of welfare neutrality and of renegotiation proofness. For more on the issue, see also Hart and Moore (1999) and Maskin and Tirole (1999b).

[2]Non-product verifiability structures are fairly common in contract theory. The hold-up models following Hart and Moore (1990), for example, consider a trading model in which the court can verify whether trade takes place or not, but cannot verify which party refused to trade. In team problems such as Holmstrom (1982), the court can only verify whether the members of the team cooperated or not. In the latter case, it cannot tell which member of the team did not cooperate.

[3]The literature on contract incompleteness that follows the seminal contributions of Grossman and Hart (1986) and Hart and Moore (1990) is vast. The reader can consult one of the many surveys, such as Hellwig (1996), or Tirole (1999).

incomplete contracts. Complete contracts may be unenforceable as they may include prescriptions that concern unverifiable aspects of the interaction. It is natural to think that, at least when writing contracts is not costless, it would be a waste of resources to stipulate unverifiable prescriptions.[4] While this intuition is correct if the verifiability correspondence is product and partitional, this paper shows that it is also possible to identify instances where, because of liability constraints, the players can achieve first best only by signing a contract that explicitly includes unverifiable prescriptions. In some cases, the players should sign a contract that cannot be enforced, and secretly agree to breach it in equilibrium.[5]

The key requirement for a contract to achieve first best is not that its prescriptions are verifiable, but that they are effective in deterring individual deviations from first best, without at the same time preventing the players from playing the optimal outcome. Thus, it is not always enough to stipulate that a deviation should not be taken, it may also be necessary to rule out actions that are not distinguishable in court from the deviation. This implies that some unverifiable prescriptions may be necessary in order to achieve first best. Moreover, the optimal contract does not need to prescribe that the players play the optimal outcome, it may be enough that the prescribed play is not distinguishable in court from first best. In some instances, in order to effectively deter deviations from first best, it may be necessary to sign a contract that prescribes that the players sacrifice themselves for the social good more than it would be optimal. First best can then be achieved if the players are willing to violate the agreement and play the optimal outcome, and if such violations are not punished by the court.

In its extreme simplicity, the following example should be appropriate to introduce the analysis for the case of non-partitional verifiability structures.

[4] Transaction and complexity costs have been introduced in the contract literature by Williamson (1975), and further studied by Anderlini and Felli (1999) and by Segal (1999), among others.

[5] Following BW, this paper defines contracts as mappings that associate outcomes enforceable by a court to the actions taken by the players. Alternatively one may define contracts as functions of the evidence verifiable in court. But in that case, the comparison of complete and incomplete contracts becomes meaningless. Unless verifiability is perfect, such a definition of contracts precludes complete contracts from the analysis, and is equivalent to *assuming* that the players may only sign contracts that do not include unverifiable prescriptions.

Example 1 Each of two players must decide a private contribution to a joint project.[6] At the social optimum both players contribute c^*, but each player has a private incentive to under-contribute, so that in absence of a contract they will settle on the Nash Equilibrium e. We assume that, when asked to verify a player's action, the court will be able to produce a statement that is correct only up to a (possibly wide) margin of error M, and that the action played cannot be pinned down. Specifically, we say that when a player provides the contribution c, the court will only be able to conclude that her contribution belongs to the set $[c - M, c + M]$, but will not be able to determine which amount in the set has been contributed. The verifiability correspondence is non-partitional because, if two distinct values c and c' are less than M apart, the sets $[c - M, c + M]$ and $[c' - M, c' + M]$ overlap but do not coincide.

The contract prescribing that both players play c^* will not achieve the social optimum. The contract fails to deter the players from making a contribution c smaller than c^*, and larger than $c^* - M$. Since $c + M$ is larger than c^*, it cannot be concluded in court that they did not comply with the agreement. Moreover, regardless of the size of M, it is easy to see that any enforceable contract will implement only the Pareto-dominated Nash Equilibrium outcome e. In fact, any contract prescribing a contribution c larger than e will fail to deter contributions smaller than c and at least as large as e. However, it is immediate to notice that the players will be able to achieve the social optimum, by contractually committing to contribute at least $c^* + M$. The smallest contribution that each player can make without verifiably breaching such an agreement is in fact c^*. The optimal contract prescribes actions different from the first-best ones. These prescription are not enforceable, and the contract will be violated in equilibrium to achieve first best. $\diamond$

Our treatment of product non-partitional verifiability structures first relates them to the contract theory literature and to the knowledge literature. We argue that, while they have not been formally analyzed with the model introduced by BW, these verifiability structures

[6] The literature on (static) voluntary contribution games includes, for instance, Palfrey and Rosenthal (1984), Bernheim (1986), Bergstrom, or Blume and Varian (1986).

have not been uncommon in contract theory. In order to identify the features of non-partitional verifiability that imply the suboptimality of enforceable contracts, we then look more in depth at the knowledge axioms that generate non-partitional information structures. Geanakoplos (1989) shows that truthful information correspondences are partitional if and only if they are transitive and euclidean. We show that the key axiom is the first one: the players can restrict attention without loss to contracts that do not include any unverifiable prescriptions if and only if the verifiability correspondence is transitive.[7]

This result naturally begs the question of how likely it is to encounter intransitive verifiability structures in contractual settings. We thus conclude our treatment by relating non-partitional verifiability to a simple model of a Bayesian court: while not uncontroversial, the assumption that the court makes use of statistical evidence and of the Bayes rule is often found in the law and economics literature.[8] We show that if the players are unlimitedly liable, our model of court decision generates transitive verifiability structures. When the players have limited liabilities, instead, one can reconstruct examples where the players should sign a contract that will be breached in equilibrium, because they play as if facing an intransitive verifiability structure.[9]

Non-product verifiability structures arise when the court cannot verify the action taken by a player independently of the other players' choices. When the verifiability structure is non-product it may then be the case that a player violates a contract, and the court verifies the violation but cannot identify the violator. In order to analyze these instances, we introduce a distinction based on the liability regime. We denote by *individual liability* regime the case where each player can only be held accountable for her own actions, and cannot be punished by the court for a contract violation that may have been committed by

[7]We also show that the suboptimality of enforceable contracts may arise in a very unrestricted class of games. Unlike BW, the distinction between games with strategic substitutes and games with strategic complements does not play a role in our analysis.

[8]See for example Milgrom and Roberts (1986) or Rubinfeld and Sappington (1987). A discussion of the issues that arise when statistical assessments are presented as evidence in court is in Fienberg (1989).

[9]The relevance of limited liability constraints is studied, for instance, by Legros and Matthews (1993) in the case of moral hazard in teams.

another player.[10] If a player may be punished by the court regardless of the identity of the contract violator, we will say that the regime is one of *joint liability.*

For the case of individual liability, this paper establishes that the court will not settle the trial on the basis of the verifiability structure, but on the basis of the projections of the verifiability structure onto the players' action spaces. Even if the verifiability structure is partitional, it may then be the case that its projections onto the players' action spaces are intransitive. Therefore we can identify instances where, in order to achieve first best, the players should sign a contract that cannot be enforced. The following example should be helpful in substantiating our point.

Example 2 Consider a double hold-up problem, in which each of two agents individually makes a private investment specific to the provision of a good or a service to a counterpart.[11] In the socially optimal outcome, the two agents cooperate to maximize their joint profits. However each of the two agents has a private incentive to underinvest. Say that the court can observe the joint outcome of the agents' choices, but cannot distinguish private investments. If each one of the two agents is only liable for her own actions, and cannot be punished for a contract violation that may have been committed by the other agent, it will not be possible to effectively deter underinvestment.

Suppose now that each agent may also, at a cost, gather evidence so that she will be able to prove in court that the other agent has underinvested, whenever this is the case. Specifically, say that there are two signals verifiable in court. The court observes the first (second) signal if and only if the first (second) agent has underinvested and the other agent has monitored her. The first-best outcome is such that the two agents invest at the co-operative levels, without wasting resources monitoring each other. But again, the contract

[10] To the best of our knowledge, individual liability constraints have not been explicitly considered in the contract literature. A simple real-life example in which they play a major role is the following. Suppose that one hires a carpenter and a plumber to renovate her house. While the quality of the renovation depends on both agents' efforts, if something goes wrong, the two agents will not typically be held jointly liable in court.

[11] There is a very large literature on the hold up problem in contract theory, in particular in relation with incomplete contracts, see for instance Grossman and Hart (1984), Hart and Moore (1988), Holmstrom and Tirole (1989), Noldeke and Schmidt (1995); a review is in Hart (1995).

prescribing cooperative investments without monitoring will fail to deter underinvestment.

The players can achieve first best only by signing a contract that prescribes cooperative investment levels, and that requires the agents to monitor each other. In equilibrium, the agents will violate the contract, and cooperate without wasting their efforts on monitoring. This violation cannot be verified in court: since neither agent has underinvested, the court does not observe any of the two signals, and thus the absence of monitoring efforts cannot be verified. Moreover in equilibrium this apparently contradictory contract is effective in deterring underinvestment. Suppose that the first agent reduces her investment level. If she is taken to court by the counterpart, she will not be able to discharge herself of the accusation of breaching the contract. Since she is unable to show that the other agent has underinvested, the court will conclude that either she has failed to monitor the other agent or that she is really the one who underinvested. In either case she is liable for breaching the contract.

As in Example 1, the first-best outcome can be achieved only by signing a contract that explicitly prescribes an action profile different from the first-best one. This prescription is not verifiable, and the optimal contract is not enforceable. In fact it will be breached in equilibrium so that first-best is achieved. $\diamond$

The key feature of Example 2 is that the players cannot be held accountable for their opponents' choice, and that, at the same time, their actions cannot be verified independently of their opponents' play. The role of the individual liability regime is crucial. In economic instances where the players are subject to a joint liability regime, the agreement prescribing cooperative investments without monitoring would achieve first-best. If either of the agents violates the agreement, the counterpart will take them jointly to trial, and the court will be allowed to punish them jointly. Following this lead, we will show that, when the players are subject to a joint liability regime, and the verifiability structure is non-product (but transitive), they may restrict attention without loss to contracts that only include verifiable prescriptions.

In conclusion, our analysis shows that what matters for a contract is not whether it describes verifiable aspects of the interaction, but whether it deters deviation from first-best, without at the same time preventing the optimal play. This paper shows that optimal contracts cannot be determined by looking only at the imperfections in the process of gathering information by the court. The effect of a contract does not only depend on the extent of the verifiability constraints, but on the entire relation between the profiles taken by the players and the punishment enforced by the court. This underlines the relevance of the constraints imposed by a limited or an individual liability regime.

While in principle the players could always stipulate an agreement that holds them jointly liable for unlimited sums of money, we believe that in some economic circumstances these contracts are not practically feasible. This paper does not attempt to explain the existence of liability constraints. Rather, by extending the methodology introduced by BW to the domain of non-partitional and non-product verifiability structures, we characterize optimal contracts both when contracting is constrained and when this is not the case. This allows us to identify some liability assumptions required for the optimality of contracts that only specify verifiable aspects of the economic interaction.

The paper is presented as follows. In the second section we review the BW model for normal form games. In the third section we extend the analysis to non-partitional product correspondences. In the fourth section, non-product correspondences (partitional or not) are considered. The fifth section concludes the paper.

2 Review of BW Normal-Form Model

Consider a normal-form game $G = (I, A, u)$, where I is the set of players, A is the (finite) action space, and u are the utility functions. In anticipation of playing this game, the players may choose to stipulate a contract that restricts their choices. BW introduce the concept of *simple contracts*: agreements that prescribe each player not to take some actions. Formally a simple contract is a profile of allowable actions $C = (C_i)_{i \in I}$, where for each player i, $C_i \subseteq A_i$.

The contract C prescribes each player i not to take any action $b_i \notin C_i$. The model of BW implicitly assumes that each player is unlimitedly liable for breaking the contract, so that if the court can verify that she violated the contract, it will punish her in an arbitrarily harsh manner.

BW represent the court's verifiability constraints with the information correspondence $P : A \rightarrow 2^A$; if the players play profile a, then the court cannot distinguish it from any other action profile contained in $P(a)$.[12] They assume the correspondence P to be partitional, product, and truthful. Formally, P is *partitional* if the range of P is a partition of A, P is *product* if each player is verified independently of the others (i.e, $\exists (P_1, \cdots, P_I)$ with $P_i : A_i \rightarrow 2^{A_i}$ such that $\forall a \in A$, $P(a) = \times_{i=1}^{I} P_i(a_i)$), and P is *truthful* if $\forall a \in A, a \in P(a)$.

The traditional definition of complete contracts is that a single action is prescribed to each player in each state of the world (cf. Laffont and Maskin 1982). Once imperfect verifiability is taken into account, however, it is natural to focus attention on contracts that restrict the players' actions to the maximal extent possible, without including any unverifiable prescriptions. BW reinterpret the issue of contract incompleteness and determine whether the players can improve upon such contracts by leaving some verifiable prescriptions unspecified. They say that a simple contract is *enforceable* only if the court can determine whether each player has played an action allowed by the contract.[13] They define a contract as *complete* if it is one of the most restrictive enforceable simple contracts. Their definition allows one to assign to each action profile a the complete contract $P(a) = (P_i(a_i))_{i \in I}$: the agreement that allows each player i only to take actions that the court cannot distinguish from a_i.

Definition 1 (Bernheim and Whinston 1998) *A simple contract C is* enforceable *if for*

[12]Following BW, in most of the paper we take the correspondence P as primitive, instead of deriving it as the outcome of an explicitly modeled trial game. Among the models of trials that appear in the law and economics literature: Milgrom and Roberts (1986), Rubinfeld and Sappington (1987), Andreoni (1991).

[13]The model of BW does not require the court to punish a violator only if it assesses that the violation has occurred with probability 1. In fact, the correspondence P may be derived also from p-belief operators (see Monderer and Samet 1989), where p is arbitrary. The court concludes that a violation has occurred whenever it assesses that probability larger than p.

any set B in the range of P, either $B \subseteq C$ or $B \cap C = \emptyset$. A simple contract C is complete *if it belongs to the range of P.*

Whenever a player signs an enforceable contract, she knows that she will be punished if she takes an action not allowed by the contract. Therefore, after signing the enforceable contract C the parties behave as if they were playing the game G restricted to the set of action profiles C. Specifically, BW assume that they will coordinate on a *pure-strategy* Nash Equilibrium of the game $G|_C := (I, C, u|_C)$. When the verifiability structure is partitional and product, BW establish that in normal form games the players never have a reason to sign an incomplete contract: if any outcome a may be achieved by signing a simple enforceable contract, then it may be achieved also by signing its associated complete contract $P(a)$. At the same time, the players have no reason to sign a contract that includes prescriptions that cannot be verified in court.

Proposition 1 (Berheim and Whinston 1998) *If there is a simple contract C such that the action profile a is a Nash Equilibrium of the game $G|_C$, then a is a Nash Equilibrium of $G|_{P(a)}$.*

3 Non-Partitional Product Verifiability

3.1 Related Literature

While they have not been formally represented as non-partitional verifiability correspondences, these verifiability structures have not been uncommon in the contract theory literature. Consider for instance a simple revisitation of Example 2 in Okuno-Fujiwara, Postlewaite and Suzumura (1990). Two oligopolists may or may not make investments that reduce marginal costs. It is very simple to verify in court that a firm's marginal cost is low, for instance by running its production line very fast. Running the line slowly however does not demonstrate that it cannot run faster, and thus that the costs of production are high.[14]

[14] As in BW, we take the correspondence P as primitive instead of deriving it as the outcome of an explicitly modeled trial game. In particular, we do not explicitly model the players' incentives to produce or conceal

Non-partitional verifiability structures can also be found in Tirole (1986, 1992). Independently of the player's moves, a verifiable signal is observed with positive probability if the cost of production is low, but nothing is observed if the cost is high. Therefore, while it can be verified with positive probability that the cost is low, it can never be verified that the cost is high. The structure is non-partitional because of the stochastic structure of the signal (a more detailed description is in the proof of Lemma 3).

Among other motivations for the study of non-partitional correspondences, we point out that not all evidence in a trial may be admissible in court, that the judicial system may be unable to gather all available evidence (possibly because of the costs involved), and that the court may be unable to fully exploit the informational content of evidence. Seidmann (1997), for instance points out that whereas a defendant's testimony discharging herself from an accusation is admissible evidence in court to support the defendant's innocence, the absence of a discharging testimony is not admissible as evidence that the defendant is guilty.

In the literature on knowledge (see Dekel and Gul 1997 for a review), a common criticism of non-partitional information correspondences is that it is unclear whether they are compatible with the epistemic knowledge of the players in a game. Specifically, a player's information structure may be non-partitional only if she does not precisely know her structure; otherwise, she will resolve it into a partition. Suppose in fact that there are states a, b such that $\{a, b\} \subset P(a)$, and $a \notin P(b)$, $b \in P(b)$. When observing $P(a)$, the agent knows that a may have occurred, and that if b had occurred, she would know that a has not occurred. This counterfactual argument allows her to conclude that b has not occurred, and to refine her information $P(a)$ to $P(a) \diagdown \{b\}$. If instead, when a occurs, the individual does not know $P(b)$, then she will not be able to conclude that b has not occurred.

When a player does not know her information structure, as underlined by Brandenburger,

evidence during the trial. In principle, we could expand the game G endowing the players with the additional action of producing or conceiling evidence. But in that case, we would deal with a game with sequential actions, rather than simultaneous moves, and that line of development would lead us beyond the scope of the present paper. It should be pointed out that the information correspondence defined on this expanded action set will be partitional whenever the only source of non-partitional information is the private control of conclusive evidence.

Dekel, and Geanakoplos (1992), the ex-ante expected utility calculation is problematic, because ex-ante the information processing pathology has not yet occurred. Moreover, they argue, an ex-ante equilibrium solution depending on the ex-post non-partitional information is meaningless. As its derivation requires common knowledge of each other's information structure, each player must also know her own information structure, but then she can solve it into a partition.

In this paper, the non-partitional structure is not of the players but of the court, an external institution constrained by the rules of legal procedure. These rules may limit the use of counterfactual arguments in trial decisions, thus preventing the court from resolving a non-partitional verifiability structure. Our players instead have epistemic knowledge, and they know the legal procedure that constrains the court. It is thus ex-ante common knowledge among the players that the court has a non-partitional information structure. It is therefore logically consistent to require the players to calculate ex-ante utility by compounding interim utilities from the non-partitional information structure, and it makes sense to find an ex-ante equilibrium constructed in the same way.

3.2 The Model

As BW point out, the definition of complete contracts contained in Definition 1 is not appropriate for dealing with non-partitional verifiability structures. In the example by Okuno-Fujiwara, Postlewaite and Suzumura (1990), if an oligopolist plays L, the court will be able to verify that she did so, but if an oligopolist plays H, the court will not be able to verify that H has been played, and it will conclude that either H or L has been played. The court's verifiability structure is $P(L) = \{L\}$ and $P(H) = \{H, L\}$: according to Definition 1, $\{L\}$ would be an enforceable complete contract. However, suppose that an oligopolist contractually commits to play L, but then plays H. The court cannot rule out that she played L instead, and so cannot punish her. The contract $\{L\}$ is thus unenforceable and the only enforceable contracts are $\{H\}$ and $\{H, L\}$.

While Definition 1 is inappropriate, the logic of Proposition 1 is still valid. Even though the contract $\{H, L\}$ contains the contract $\{H\}$, it is the only enforceable contract that allows the action L, because the contract $\{L\}$ is not enforceable. Thus the contract $\{H, L\}$ is to be considered the most restrictive enforceable contract (in fact the only enforceable contract) that contains H. In that sense, we can still conclude that the players cannot improve upon the most restrictive enforceable contracts. In fact, if the action L can be achieved at all, then it can also be obtained signing the contract $\{H, L\}$.

In order to extend the model of BW to account for general non-partitional, product verifiability structures, first note that, since each player's action can be verified independently of her opponents' move, we can still restrict attention without loss of generality to contracts that prescribe a set of allowable actions for each player. So we suppose that before playing the game G, the players sign a simple contract $C = (C_i)_{i \in I}$ (we shall henceforth drop the term simple). Consider an arbitrary player i. Any action $b_i \notin C_i$ is a violation of the contract C. However, if the profile b_i is played, and $P_i(b_i) \cap C_i \neq \emptyset$, the court cannot rule out that an action allowed by the contract was played. Thus we say that b_i is a *verifiable violation* of the contract C if $P_i(b_i) \cap C_i = \emptyset$. Player i knows that she will be punished if she verifiably violates the contract, and she takes account of this when choosing her action. Specifically, for all $a_{-i} \in A_{-i}$, we set $u_i(b_i, a_{-i}|C) = -\infty$ if b_i is a verifiable violation of C_i, and $u_i(b_i, a_{-i}|C) = u_i(b_i, a_{-i})$ otherwise. Armed with these definitions, we can formally define achievable outcomes.[15]

Definition 2 *The action profile a is* achievable *if there exist a contract C such that a is a pure strategy Nash equilibrium of the game $G|C = (I, A, u(\cdot|C))$, and that $u(a|C) = u(a)$.*[16]

Since this framework is supposed to capture a pre-game agreement by the players, it is also assumed that the parties select an outcome that is not Pareto dominated by any other

[15]Following BW, we restrict attention to pure-strategy outcomes. As is well known, any mixed-strategy or correlated equilibrium can be represented as a pure strategy equilibrium of a game with a mediator who privately observes the realization of a correlating device.

[16]If this last condition does not hold, then it would be the case that $u_i(a|C) = -\infty$, and that a is an equilibrium of $G|C$ only because player i is always punished when her opponents play a_{-i}. Obviously, player i would never sign the contract C. While in principle contract C could achieve profile a, this outcome would not be individually rational.

achievable outcome, and that makes each player better off with respect to the payoff she obtains by not signing any contract.

The logic of Definition 1 is that a contract is enforceable if all its prescriptions can be verified by a court. But, as our results will show, there are games where in equilibrium the players do not violate agreements that include unverifiable prescriptions. In some sense, such contracts can be considered "enforceable." To avoid any confusion, we will distinguish between enforceable and inviolable contracts. A contract is enforceable in a game if in equilibrium the players will not violate it. The prescriptions of an inviolable contract, instead, must be followed by the players regardless of their payoff in the game. Specifically a contract C is inviolable if, for any player i, any contract violation $b_i \notin C_i$ is verifiable by the court. Otherwise the court will not be able to punish a player who has taken an action ruled out by the contract. Obviously any inviolable contract is enforceable in all games.

Definition 3 *A contract C is* inviolable *if for any i, and any $b_i \notin C_i$, it is the case that $P_i(b_i) \cap C_i = \emptyset$. A contract C is* enforceable in game G *if the game $G|C$ has a pure-strategy Nash Equilibrium a, such that $a \in C$.*

Following BW, we want to assign to each action profile a the most restrictive inviolable contracts that allows the players to play a. The next Lemma shows that for any profile a, there exists only one such contract, which is denoted by $C(a)$. The proof demonstrates a simple algorithm for deriving the most restrictive inviolable contract associated with each action profile. Initially take the contract that consists only of the action profile. In the second step, for each player, include all actions that cannot be distinguished by the court, in the third step add all actions that cannot be distinguished from the actions included in the second stage, and proceed iteratively until no more actions need to be added.

Lemma 1 *For any verifiability structure P, for any action profile $a \in A$, there exists a unique minimal (in terms of set inclusion) inviolable contract $C(a)$ such that $a \in C(a)$.*

Proof. Let the contract $C(a)$ be the limit of the sequence $\{C_n\}_{n\geq 0}$ constructed as follows. For any i, let $C_i^0 = \{a_i\}$ and $C^n = \times_{i=1}^I f_i(C_i^{n-1})$ for any $n \geq 1$, where each relation $f_i : 2^{A_i} \rightarrow 2^{A_i}$ is such that:

$$f_i(C_i) = \{b_i \in A_i | P(b_i) \cap C_i \neq \emptyset\}, \text{ for every } C_i \subseteq A_i. \tag{1}$$

Note that the limit always exists because each relation f_i is well-defined and non-decreasing and each 2^{A_i} is finite.

Suppose that C is an arbitrary inviolable contract and that $a \in C$. For any player i, any action b_i such that $a_i \in P(b_i)$ must be included in C_i, by the definition of inviolable contracts. Once b_i is included in C_i, also any c_i such that $b_i \in P(c_i)$ must be included, and so on. Therefore, it must be the case that $C(a) \subseteq C$. ■

In order to determine the optimal contracts, we proceed as follows. For any game G, and for any action profile a, we construct the list $\mathcal{C}(a)$ that includes all the contracts that achieve a. For any specified social ranking over profiles, we obtain a ranking over lists of contracts, and thus a complete characterization of optimal contracts.

First of all, we let $D_i(a) = \{d_i | u_i(d_i, a_{-i}) > u_i(a)\}$ denote the list of individually-profitable deviations from a. In order to achieve a, a contract must deter each player i from playing an action in $D_i(a)$, and thus rule out all actions indistinguishable from any action in $D_i(a)$. So we associate with each profile a a contract $E(a)$ that we denote as the *least restrictive deterrent* contract. The contract $E(a)$ is the least restrictive agreement which is effective in punishing each player i if she takes an individually-profitable deviation from a_i, when her opponents play the profile a_{-i}.

Definition 4 *For each profile a, the* least restrictive deterrent *associated with a is the contract $E(a) = (E(a)_i)_{i\in I}$ where for each i,*

$$E(a)_i = A_i \backslash \left[\cup_{d_i \in D_i(a)} P_i(d_i)\right].$$

Our analysis shows that the players may achieve the profile a by signing the least restrictive deterrent contract $E(a)$, if this does not prevent the players from playing a. This is equivalent to requiring that $E(a)$ includes at least one action profile that cannot be distinguished from a by the court. We thus conclude that the profile a can be achieved with any contract C contained in $E(a)$, such that C includes at least one action profile that belongs to $P(a)$, and with no other contract.

Lemma 2 *For any game G, any action profile a can be achieved with any contract C that belongs to the collection*

$$\mathcal{C}(a) = \{C | C \subseteq E(a) \text{ and } P(a) \cap C \neq \emptyset\},$$

and with no other contracts.

Proof. For any contract C, the profile a is a Nash Equilibrium of $G|C$ with $u_i(a|C) = u_i(a)$ for each player i if and only if for any player i, it is the case that $P_i(a_i) \cap C_i \neq \emptyset$, and that $u_i(d_i, a_{-i}|C) = -\infty$ for any deviation $u_i(d_i, a_{-i}) > u_i(a)$. By definition of $D_i(a)$ and $u(\cdot|C)$, the second requirement translates into the requirement that $C_i \cap \left[\cup_{d_i \in D_i(a)} P_i(d_i)\right] = \emptyset$. This is equivalent to require that $C_i \subseteq A_i \backslash \left[\cup_{d_i \in D_i(a)} P_i(d_i)\right]$, for any i. Thus the profile a is a Nash Equilibrium of $G|C$ with $u_i(a|C) = u_i(a)$ for all i if and only if $P(a) \cap C \neq \emptyset$ and $C \subseteq E(a)$. ■

A straightforward consequence of Lemma 2 is that an action profile a is achievable if and only if the collection $\mathcal{C}(a)$ is non-empty. If a is achievable, then the least restrictive contract that achieves a is $E(a)$, and any singleton contract $\{c\}$ such that $c \in P(a) \cap E(a)$ is one of the most restrictive contracts that achieves a.

3.3 The Optimality of Enforceable and Inviolable Contracts

Example 2 in the introduction shows that with general product, non-partitional verifiability structures, the players may need to sign unenforceable contracts. One can also describe instances where the players may achieve first best only by signing a contract that will be enforced in equilibrium, but that necessarily includes unverifiable prescriptions. Consider a two-player symmetric game where each player's action set is $A = \{b, c, d\}$. The strategy c denotes cooperation, the strategy d denotes defection, and the strategy b identifies a strategy that fares worse than c against c. The only Nash Equilibrium is (d, d), and it is Pareto-dominated by the outcome (c, c). For example, one can focus attention on the following game.

G	b	c	d
b	$1, 1$	$2, 2$	$1, 0$
c	$2, 2$	$3, 3$	$0, 4$
d	$0, 1$	$4, 0$	$2, 2$

The players would like to implement the action profile (c, c), by signing a contract that deters the deviation d (note that the deviation b does not need to be contractually deterred). Suppose that $c \in P(b)$, $b \in P(d)$ and $c \notin P(d)$. In order to find the most restrictive inviolable contract associated with action c, we start by considering the contract that prescribes to play action c only. Since $c \in P(b)$, this contract contains an unverifiable prescription: the action b cannot be distinguished in court from c. After we remove the unverifiable prescription not to play b, we obtain the contract that prescribes to play either b or c, but not d. This contract also contains an unverifiable prescription: since $b \in P(d)$, the action d cannot be distinguished in court from b. The most restrictive contract associated with c does not restrict the players' choice at all, and thus it will fail to deter action d.

The players however may achieve the outcome (c, c), by agreeing to the contract $\{(c, c)\}$. While it cannot be determined in court whether any player breached the contract and played b, it is not in the best interest of either player to deviate from c and play b. The prescription not to play b is not verifiable, but it must be explicitly included in the optimal contract. Or else, since defection (action d) cannot be distinguished from b in court, each player will violate the agreement and play d.

In order to identify general conditions under which inviolable or even enforceable contracts may fail to be optimal, we need to look more in depth at the knowledge axioms that generate non-partitional verifiability. Geanakoplos (1989) shows that an individual has a non-partitional truthful information structure if and only if she does not necessarily know what she knows, or she does not necessarily know what she does not know. In the language of information structures, these two axioms are respectively called the transitivity and euclidean axiom.

Definition 5 *The information structure P is* transitive *if $a \in P(b)$ and $b \in P(c)$ implies $a \in P(c)$. The information structure P is* euclidean *if $a \in P(b)$ and $c \in P(b)$ implies $a \in P(c)$.*

Our analysis shows that the key axiom for our results is transitivity. Specifically Proposition 2 shows that when the verification structure is transitive and product, the most restrictive inviolable contract is optimal. Conversely, Proposition 3 shows that in strategic environments that cannot be reduced to a single-agent decision problem, if a product verification structure fails to be transitive, one can construct generic games in which there are no enforceable (hence no inviolable) optimal contracts. Finally, Corollary 4 shows that for any game where the action space is rich enough, one can construct product, intransitive verifiability structures for which all enforceable contracts are suboptimal. This result shows that enforceable contracts may be suboptimal in a very unrestricted class of games, and underlines that the source of this suboptimality does not reside in the specifics of the game played, but rather in the intransitivity of the verifiability structure.

Proposition 2 *For any game G, and any* product, transitive *verifiability structure P, any achievable action profile a can be achieved with the most restrictive inviolable contract $C(a)$.*

Proof. Recall from Lemma 1 that $C(a)$ is constructed as the fixed point $C^N = \times_{i=1}^I f_i(C_i^N)$ of the sequence defined by $C_0 = \{a\}$, and $C_i^n = f_i(C_i^{n-1}), \forall n \geq 1$ with $f_i(C_i) = \{b_i \in A_i | P_i(b_i) \cap C_i \neq \emptyset\}$. Take any player i, and consider $f_i(\{a_i\}) = \{b_i \in A_i | a_i \in P_i(b)\}$. Take any action $b_i \neq a_i$, by transitivity, if $a_i \in P_i(b_i)$ and $a_i \notin P_i(c_i)$, then $b_i \notin P_i(c_i)$. Thus for any $c_i \notin f_i(\{a_i\})$, we can say that $\forall b_i \in f(\{a\})$, $b_i \notin P_i(c_i)$, and that $a_i \notin P_i(c_i)$ by the definition of f_i. In conclusion, $\forall b_i \in f_i(\{a_i\})$, $\forall c_i \notin f_i(\{a_i\})$, $b_i \notin P_i(c_i)$. Thus $f_i(f_i(\{a_i\})) = f_i(\{a_i\})$, so that $C(a) = \times_{i=1}^{I} f_i(\{a_i\})$.

Say that there is a contract C such that a is a Nash Equilibrium of $G|C$ with $u_i(a|C) = u_i(a)$ for each player i. Thus $P_i(a_i) \cap C_i \neq \emptyset$, and for any player i, and any $d_i \in D_i(a)$, it is the case that $P(d_i) \cap C_i = \emptyset$.

Since $a_i \in P_i(a_i) \cap C_i(a_i)$, it is the case that $u_i(a|C(a)) = u_i(a)$. Thus, if $u_i(a) \geq u_i(d_i, a_{-i})$, player i will not take the deviation d_i after $C(a)$ is signed. So we just need to consider players i, and deviations d_i, such that $u_i(d_i, a_{-i}) > u_i(a)$ and $P_i(d_i) \cap C_i = \emptyset$. For any action $c_i \in C_i$, $c_i \notin P_i(d_i)$. But since $P_i(a_i) \cap C_i \neq \emptyset$, there is at least a $c_i \in C_i$ such that $c_i \in P_i(a_i)$. Transitivity then implies that $a_i \notin P(d_i)$. One further application of transitivity then implies that for any b_i such that $a_i \in P_i(b_i)$, it must the case that $b_i \notin P(d_i)$. Since $C_i(a_i) = f_i(\{a_i\}) = \{b_i | a_i \in P(b_i)\}$, we have obtained that $P(d_i) \cap C_i(a_i) = \emptyset$. Player i will be deterred from taking the deviation d_i also if she signs $C_i(a_i)$. ■

Proposition 3 *Fix any pair* (I, A) *such that there are at least* 2 *players whose action sets are non-singleton. For any* product, intransitive *verifiability structure* P, *there is a generic payoff function* $u : A \rightarrow \mathbb{R}$ *such that the game* $G = (I, A, u)$ *has a Pareto-undominated achievable profile that Pareto-dominates all Nash equilibria of* G *and that cannot be achieved with any enforceable contract.*

Proof. For $P = (P_i)_{i \in I}$ to be intransitive, it must be the case that there is at least a player i, and a triple $\{b_i, c_i, d_i\}$ such that $c_i \in P_i(d_i)$, $b_i \in P_i(c_i)$, $b_i \notin P_i(d_i)$.

Construct u as follows. Say that there is a profile c_{-i} such that the profile (c_i, c_{-i}) (hereafter denoted as c) satisfies $u_i(a_i, c_{-i}) < u_i(c) < u_i(d_i, c_{-i})$ for $a_i \notin \{c_i, d_i\}$ and that, for any $j \neq i$ and any $a_j \neq c_j$, $u_j(c) > u_j(a_j, c_{-j})$. To make c individually rational we construct payoffs such that G has a unique Nash Equilibrium e Pareto-dominated by c. Specifically, we select a profile e such that for at least one $j \neq i$, $e_j \neq c_j$, and require that, for any i,

$u_i(c) > u_i(e) > u_i(a_i, e_{-i})$ for all $a_i \neq e_i$, and that for any $a \neq e$, there is a player j and an action $a'_j \neq a_j$ such that $u_j(a'_j, a_{-j}) > u_j(a)$. In particular we are saying that there is a player $j \neq i$ and an action $a'_j \neq c_j$ such that $u_j(a'_j, d_i, c_{-ij}) > u_j(c_j, d_i, c_{-ij})$. To make c Pareto undominated, we further extend the requirements on u, by assuming that, for each $a \neq c$ there is a player j such that $u_j(a) < u_j(c)$. In particular, there must be a player $j \neq i$ such that $u_j(d_i, c_{-i}) < u_j(c)$.

Given our payoff construction, the players would like to implement the profile c. By definition any enforceable contract (a fortiori any inviolable contract) that achieves c must contain c. For any contract C such that $c_i \in C_i$, since $c_i \in P_i(d_i)$, however, $u_i(d_i, c_{-i}|C) = u_i(d_i, c_{-i}) > u_i(c) = u_i(c|C)$. Thus the profile c is not a Nash Equilibrium of the game $G|C$. We have shown that the profile c cannot be achieved with any enforceable contract.

Consider however the contract $C^* = (\{b_i\}, A_{-i})$. Since $b_i \notin P_i(d_i)$, and $C^*_i = \{b_i\}$, it follows that $u_i(d_i, c_{-i}|C^*) = -\infty$. Moreover, since $b_i \in P_i(c_i)$, it is the case that $u_i(c|C^*) = u_i(c)$. Thus we obtain that $u_i(c|C^*) > u_i(d_i, c_{-i}|C^*)$. By assumption, $u_i(c|C^*) = u_i(c) > u_i(a_i, c_{-i}) \geq u_i(a_i, c_{-i}|C^*)$, for any $a_i \notin \{c_i, d_i\}$. At the same time, for any player $j \neq i$, $u_j(a|C^*) = u_j(a)$, and by assumption, for any $a_j \neq c_j$, $u_j(c) > u_j(a_j, c_{-j})$. Therefore, c is a Nash Equilibrium of $G|C^*$ and $u_i(c|C^*) = u_i(c)$ for all i. We have shown that c is achievable.

■

Corollary 4 *Fix any game $G = (I, A, u)$ such that there is a Pareto-undominated profile c that Pareto-dominates all Nash equilibria of G, and such that at least one player i has actions b_i and d_i satisfying $u_i(b_i, c_{-i}) < u_i(c) < u_i(d_i, c_{-i})$. For some product, intransitive verifiability structure P, the outcome c is achievable only with unenforceable contracts.*

Proof. Construct $P = (P_j)_{j \in I}$, by assuming perfect verifiability for any $j \neq i$: let $P_j(a_j) = \{a_j\}$, for any $a_j \in c_j$. Let $P_i(c_i) = \{b_i, c_i, d_i\}$, and $P_i(b_i) = \{b_i, c_i\}$, and for any other $a_i \in A_i$, let $P_i(a_i) = \{a_i\}$. The remainder of the proof is identical to the last part of the proof of Proposition 3, and is thus omitted. ■

3.4 Bayesian Courts

This part of the section relates product non-partitional verifiability structures to a simple Bayesian model of the court's decision process. While this model is in many respects a very coarse representation of the intricacies of trial decisions, we believe that it may give some indications on how likely it is to encounter intransitive verifiability structures in contractual settings. We show that the model generates transitive verification structures when the players are unlimitedly liable, but that this need not be the case when they are only limitedly liable.

First we assume unlimited liability: any player found guilty can be punished very harshly by the court. We suppose that the evidence is stochastic, that all evidence realized is admissible in a trial (regardless of its contextual meaning), that it will be brought to trial, and that the court will be able to fully exploit its informational content. Without loss of generality, the evidence can be represented as an I-dimensional random vector that can take a value in the finite set of realizations X. When player i plays action a_i, the realization x_i is brought to court with probability $\Pr(x_i|a_i) \in [0, 1]$.

Say that player i plays an action b_i. Whenever a signal $x_i \in X_i$ such that $\Pr(x_i|b_i) > 0$ and $\Pr(x_i|a_i) = 0$ is brought to trial, the court can conclude that it is impossible that a_i was played.[17] Suppose that player i has contractually committed to play action a_i. When playing action b_i, she anticipates that, with positive probability, the signal x_i will be brought to court and that she will be punished. Since the punishment is very harsh, she will behave as if she were sure that when taking action b_i the court would be able to rule out action a_i. In other terms, she behaves as if she was facing a court with verifiability structure P such that $a_i \notin P_i(b_i)$.

The above discussion allows us to derive the verifiability structure P_i from the triple (A_i, X_i, $\{\Pr(x_i|a_i)|a_i \in A_i, x_i \in X_i\}$), which we henceforth define a *Bayesian model,* with the

[17] For the sake of simplicity, this model assumes that the court holds a full-support prior and that it concludes that an event has occurred only if it assesses that probability equal to 1. We could also require the court to state that an event has occurred when it assesses that probability larger than p, for some given p.

rule that for any pair of actions $a_i, b_i \in A_i$,

$$a_i \in P_i(b_i) \text{ if and only if } \forall x_i \in X_i, \ [\Pr(x_i|b_i) = 0 \text{ or } \Pr(x_i|a_i) > 0, \text{ or both}]. \qquad (2)$$

Any verifiability structure derived from a Bayesian model according to rule (2) is transitive, but not necessarily partitional.

Lemma 3 *If the correspondence P_i is derived from a Bayesian model, then P_i is transitive, but not necessarily partitional.*

Proof. For simplicity we drop subscripts. We will show that if $c \in P(a)$ and $c \notin P(b)$ then $a \notin P(b)$. Since $c \notin P(b)$, there must exist an $x \in X$ such that $\Pr(x|b) > 0$ and $\Pr(x|c) = 0$. As $c \in P(a)$, it follows that $\forall y \in X$, $\Pr(y|a) = 0$ or $\Pr(y|c) > 0$ or both: in particular, $\Pr(x|c) = 0$ implies that $\Pr(x|a) = 0$. So we have concluded that $\Pr(x|b) > 0$ and $\Pr(x|a) = 0$, i.e. $a \notin P(b)$.

We now present an example where P_i is not partitional. Set $A = \{c, d\}$, and $X = \{x, y\}$, with $\Pr(y|c) = 0$, $\Pr(x|d) = 2/3$. When x is presented in trial, the court will know that c has occurred with 60% probability and d with 40% probability, so neither action can be ruled out. However, when the signal y occurs, the court will know that c has occurred with probability 0. The contract $\{c\}$ is thus enforceable, because if the player takes d, she will be harshly punished when y occurs. The contract $\{d\}$ instead is not enforceable. The player behaves as if she were facing a court with the verifiability structure $P(c) = \{c, d\}$, $P(d) = \{d\}$. ■

We conclude this section by showing that the verifiability structures induced by Bayesian models are not necessarily transitive, when the players are subject to limited liability. In the following example, the stochastic structure of signals and the lower bound on liability allow us to formulate a simplified version of Example 1.

Example 1 Revisited. Each player $i = 1, 2$ contributes a sum $c_i \in A_i = \{0, 1, ..., 6\}$ to a joint project. The production function is $y = 2\sqrt{2(c_1 + c_2)}$, and the utility function is $u_i(c_i, c_{-i}) = y - c_i$. The social optimum is given by $c^* = 4$ and the Nash Equilibrium by $e = 1$.

For each player i, the signal space is $X_i = \{-1, 0, 1, ..., 7, N\}$, where $x_i = N$ denotes the absence of any evidence, and each signal $x_i \in \{0, 1, ..., 6\}$ is associated with one of the actions in A_i. The evidence is correct up to a margin of error, specifically, for any c_i, the probabilities are described by $\Pr\{x_i = c_i|c_i\} = 0.25$, $\Pr\{x_i = c_i - 1|c_i\} = \Pr\{x_i = c_i + 1|c_i\} = 0.05$, and $\Pr\{x_i = N|c_i\} = 0.7$.

To simplify the calculations, we assume that the court has a uniform prior. If a signal $x_i \in \{1, 2, ..., 5\}$ is taken to court, then the court will perform simple calculations and conclude that $\Pr\{c_i = x_i|x_i\} = 71.4\%$, and $\Pr\{c_i = x_i + 1|x_i\} = \Pr\{c_i = x_i - 1|x_i\} = 14.3\%$. While it is more likely that player i contributed amount x_i, the court cannot exclude the possibility that player i contributed $x_i - 1$ or $x_i + 1$ instead.

If the players are unlimitedly liable, then the contract prescribing each of them to contribute at least c^* will achieve the social optimum. Suppose in fact that a player i violates the contract and contributes $c_i < 4$. With positive probability, a signal $x_i < 3$ will realize, the court will conclude that player i has played $c_i < 4$ and assign player i a very harsh punishment. Since the punishment is very harsh and can be received with positive probability, player i's violation will be deterred.

If the players are only limitedly liable, instead, the probability that the violation is detected is important. Whenever the contract prescribes that they contribute at least c, the probability of being punished if playing $c - 2$ is 5 times larger than the probability of being punished when playing $c - 1$. So it may be the case that in order to deter a contribution smaller than c^* it is necessary to contractually commit to contribute at least $c^* + 1$. For instance, when the harshest punishment that can be given to each player is $\underline{u}$=2.5, simple calculations show that if the players commit to provide at least c^*, they will in fact play $c^* - 1$ in equilibrium. They may improve upon such an outcome by contracting to play at least $c^* + 1$. In equilibrium, they will violate the agreement and play c^* instead. $\diamond$

4 Non-Product Verifiability

4.1 Joint Liability

When verifiability is non-product, a player's action may be verifiable depending on the opponents' choices. Thus it is possible that one player violates a contract, and the court is able to verify that a contract has been breached but cannot identify who has violated it. Under a regime of joint liability, the court is allowed to punish any player regardless of whether the violation can be imputed to her. The breach of the agreement is then effectively deterred as long as there is an agent who is willing to go to court when the contract is violated. This is equivalent to requiring that there is an agent contractually appointed with the right of collecting fines when the contract is violated.

The collector of fines does not need to be a player that directly participates in the game. In Holmstrom (1982) for instance, the members of the team achieve the cooperative outcome by appointing an agent, external to a team, with the right of collecting fines from the members of the team if any of them has failed to cooperate. There are also instances where the collector of fines may be the court itself. Consider for example a group of farmers signing a partnership to jointly produce grocery goods. If any partner violates a health regulation, and the contaminated groceries are sold in the market, then all the partners will be prosecuted and held jointly liable for the violation. Thus each partner will comply with the agreement of producing in accordance with the health regulations.

An agreement that appoints an external collector of fines is always at least as effective in deterring individual deviations as a contract setting the allocation of fines only among the players who directly participate in the game. Therefore, any action profile that can be achieved without an external collector of fines can also be obtained with an external collector of fines, but not viceversa. We thus proceed by describing contracts that implicitly include an agent or an institution that does not directly participate in the game, with the function of taking the players to trial when the agreement is breached.

Suppose that, before playing the game G, the players sign the contract $C \subseteq A$. We say that b is a *verifiable violation* of C if $P(b) \cap C = \emptyset$. In case the players verifiably violate the contract they will all be jointly liable vis-a-vis the collector of fines. For each player i, we thus set $u_i(b|C) = -\infty$ if b is a verifiable violation of C, and $u_i(b|C) = u_i(b)$ otherwise. We say that the action profile a is *achievable* if there exists a contract C such that a is a pure strategy Nash equilibrium of the game $G|C = (I, A, u(\cdot|C))$, and that $u(a) = u(a|C)$. We say that a contract C is *enforceable* if there is a pure strategy Nash equilibrium a of the game $G|C = (I, A, u(\cdot|C))$ such that $a \in C$.

In general, a contract C should be considered inviolable if all its prescriptions can be enforced by the court. This implies however that a contract is not to be deemed inviolable if it admits unverifiable joint violations. Such a requirement may be perceived as over-restrictive because the model does not allow the players to coordinate their deviations, as they must play a Nash Equilibrium. Since this paper aims to identify in which instances the players can restrict attention without loss to inviolable contracts, an over-restrictive definition of inviolability may undermine the results. To avoid any ambiguity, we introduce a definition that requires only individual deviations to be verifiable.

Definition 6 *The contract C is inviolable under* joint liability *if for any $a \in C$, any $i \in I$, and any $(b_i, a_{-i}) \notin C$, it is the case that $P(b_i, a_{-i}) \cap C = \emptyset$.*[18]

Extending Lemma 1, we can show that for any action profile a there exist a unique minimal contract $C(a)$ that is inviolable under joint liability and for which $a \in C(a)$. The contract $C(a)$ is the limit of the sequence $\{C_n\}_{n \geq 0}$ such that $C^0 = \{a\}$ and $C^n = f(C^{n-1})$ for any $n \geq 1$, where the relation $f : 2^A \to 2^A$ is defined by the rule:

$$f(C) = \cup_{i=1}^{I} \{(b_i, a_{-i}) | \exists a_i : (a_i, a_{-i}) \in C \text{ and } P(b_i, a_{-i}) \cap C \neq \emptyset\}, \text{ for every } C \subseteq A. \quad (3)$$

We associate with any profile a the least restrictive deterrent contract, defined as:

$$E(a) = A \backslash \left[\cup_{i \in I} \left[\cup_{d_i \in D_i(a)} P(d_i, a_{-i}) \right] \right].$$

[18]When an inviolable contract is also immune to joint verifiable deviations, inspired by the concept of strong equilibrium, we will define it as strongly inviolable. Formally C is *strongly inviolable* if $\forall b \notin C$, $P(b) \cap C = \emptyset$. When proving our result, we will show that, in fact, the distinction between inviolability and strong inviolability is irrelevant in our analysis.

A simple extension of Lemma 2 allows us to show that for any game G, any action profile a can be achieved with all contracts C such that $C \subseteq E(a)$ and $P(a) \cap C \neq \emptyset$, and with no other contract.

If the verifiability correspondence is partitional, it can be shown that the most restrictive inviolable contracts are necessarily optimal. In order to strengthen the result and connect it with the analysis presented in the previous section, we show that such a result holds also for all transitive verifiability correspondences.

Proposition 5 *For any game G, and* transitive *verifiability structure P, if the action profile a is achievable under* joint liability, *then it can be achieved with the contract $C(a)$.*

Proof. Fix any C, note that a is a Nash Equilibrium of $G|C$ such that for all i, $u_i(a|C) = u_i(a)$ if and only if $P(a) \cap C \neq \emptyset$ and $\forall i, \forall d_i \in D_i(a),\ P(d_i, a_{-i}) \cap C = \emptyset$.

We want to show that a can be achieved with both the minimal inviolable contract $C(a)$, and with the minimal strongly-inviolable contract, which we denote by $J(A)$, and that is uniquely derived as the fixed point of $C^n = j(C^{n-1}), \forall n \geq 1$ with $C^0 = \{a\}$ and $j : 2^A \to 2^A$ such that $j(C) = \{b|P(b) \cap C \neq \emptyset\}$. It is immediate that for any C, $f(C) \subseteq j(C)$, and thus $C(a) \subseteq J(a)$.

Take any pair of profiles b and c different from a. By transitivity, if $a \in P(b)$ and $a \notin P(c)$, then $b \notin P(c)$. By definition, $j(\{a\}) = \{b|a \in P(b)\}$. Thus if $c \notin j(\{a\})$ then $c \notin j(j(\{a\}))$, and so $J(a) = j(\{a\})$.

Observe that $a \in P(a) \cap J(a)$, so $u_i(a|J(a)) = u_i(a)$. Thus, if $u_i(a) \geq u_i(d_i, a_{-i})$, player i will not play the deviation d_i from a, when $J(a)$ is signed. So we only need to consider pairs i and d_i, such that $u_i(d_i, a_{-i}) > u_i(a)$ and $P(d_i, a_{-i}) \cap C = \emptyset$. For any $c \in C$, $c \notin P(d_i, a_{-i})$. But since $P(a) \cap C \neq \emptyset$, there is a $c \in C$ such that $c \in P(a)$. Transitivity then implies $a \notin P(d_i, a_{-i})$. One further application of transitivity implies that for any b such that $a \in P(b)$, $b \notin P(d_i, a_{-i})$. Since $J(a) = \{b|a \in P(b)\}$, we have obtained that $P(d_i, a_{-i}) \cap J(a) = \emptyset$.

To show that a is also a Nash Equilibrium of the game $G|C(a)$, and that for all i, $u_i(a|C(a)) = u_i(a)$, since $a \in C(a)$ and $a \in J(a)$, we just need to show that, $\forall i, \forall d_i \in D_i(a)$, if $P(d_i, a_{-i}) \cap J(a) = \emptyset$ then $P(d_i, a_{-i}) \cap C(a) = \emptyset$. This fact follows from the relation $C(a) \subseteq J(a)$. ■

4.2 Individual Liability

Suppose that the players are subject to an individual liability regime, and sign a contract C before playing the game. If player i violates the contract, the court will be allowed to punish her when the following two conditions are met. First of all, it must be verified that a contract violation has occurred. Secondly, the court must rule out the possibility that the contract has been violated by a group of players that does not include i. These two requirements translate into the following definition of liability.

Definition 7 *Given any contract C, and any profile b, we say that player i is* liable *when*

1. *it is the case that $P(b) \cap C = \emptyset$,*
2. *for any profile $c \in C$, any group of players $J \subseteq I \setminus \{i\}$, and any violation $b'_J \in A_J \backslash C$, it is the case that $(b'_J, c_{-J}) \notin P(b)$.*

The following Lemma allows us to make players' liability more transparent, and will be key in studying the properties of optimal contracts. We will show that, in order to determine whether a player's innocence can be ruled out, the court will not explicitly consider the set of all action profiles that cannot be ruled out. Instead, the court will inspect the projection of this set onto the player's action space. Formally, for any set of profiles $B \subseteq A$, the *projection* of B on A_i consists of the set $\pi_i(B) = \{a_i \in A_i | (a_i, a_{-i}) \in B \text{ for some } a_{-i} \in A_{-i}\}$.

Lemma 4 *Given any contract C, and profile b, player i is liable if and only if $\pi_i(P(b)) \cap \pi_i(C) = \emptyset$.*

Proof. Fix any contract C, and profile b. By definition, player i is liable when $P(b) \cap C = \emptyset$, and for any $c \in C$, any $J \subseteq I \backslash \{i\}$, and any $b'_J \in A_J \backslash C$, it is the case that $(b'_J, c_{-J}) \notin P(b)$.

This is equivalent to requiring that for any $c \in C$, and any $a_{-i} \in A_{-i}$, it is the case that $(c_i, a_{-i}) \notin P(b)$. By using the definition of projection, we obtain that for any $c_i \in \pi_i(C)$, for any $a_{-i} \in A_{-i}$, it is the case $(c_i, a_{-i}) \notin P(b)$. This is equivalent to saying that for any $c_i \in \pi_i(C)$, it is the case that $c_i \notin \pi_i(P(b))$. ■

Because of the above result, we can again restrict attention without loss of generality to product contracts. Consider in fact any contract C and any profile b. A player i is liable for the profile b under the contract C if and only if she is liable for the profile b under the contract $C^* = (\pi_i(C))_{i \in I}$.

Whenever a player is found liable for a contract violation, she will be punished by the court. So, for any contract $C = (C_i)_{i \in I}$ and profile b, we assign the payoffs $u_i(b|C) = -\infty$ if $\pi_i(P(b)) \cap C_i = \emptyset$. Given this payoff assignment, the definitions of achievable outcomes and enforceable contracts are not different from the case of joint liability. For the contract C to be inviolable, we need to assure that each player is liable for all her individual violations of C.

Definition 8 *A contract $C = (C_i)_{i \in I}$ is inviolable under* individual liability *if for any player i, for profile $a \in C$, and any violation $b_i \notin C_i$, $\pi_i(P(b_i, a_{-i})) \cap C_i = \emptyset$.*

It can be proven that, for any action profile a, there exist a unique minimal contract $C(a)$, that is inviolable under individual liability, and for which $a \in C(a)$. The contract $C(a)$ is the limit of the sequence $\{C_n\}_{n \geq 0}$ such that $C^0 = \{a\}$ and $C^n = f(C^{n-1})$ for any $n \geq 1$, where the relation $f : 2^A \to 2^A$ is defined by the rule:

$$f(C) = \times_{i=1}^{I} \{b_i | \pi_i(P(b_i, a_{-i})) \cap C_i \neq \emptyset, \text{ for some } a_{-i} \in C_{-i}\}, \text{ for every } C \subseteq A. \quad (4)$$

For any profile a, we define the least restrictive deterrent contract as $E(a) = (E(a)_i)_{i \in I}$, where for each i,

$$E(a)_i = A_i \backslash \left[\cup_{d_i \in D_i(a)} \pi_i(P(d_i, a_{-i})) \right]. \quad (5)$$

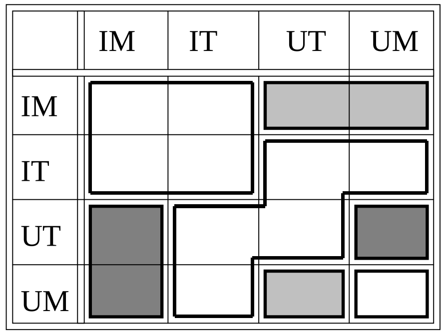

Figure 1: Agents' Actions Verifiability Structure in the Double Hold-Up Problem

One can extend the proof of Lemma 2 to show that the collection of contracts that achieve a is

$$\mathcal{C}(a) = \{C| \text{ for any } i,\ C_i \subseteq E(a)_i \text{ and } \pi_i(P(a)) \cap C_i \neq \emptyset\}.$$

We have shown in Lemma 4 that all that matters for determining a player's liability is the projection of the set of verifiable profiles onto the player's action space. When a verifiability structure is partitional but induces non-partitional projections on the players' action spaces, it may then be the case that inviolable and enforceable contracts are suboptimal. In order to demonstrate this, we will proceed by presenting two (counter)examples. First it will be useful to formalize our leading Example 2 and demonstrate a case where all enforceable contracts are suboptimal.

Example 2 Revisited. Each of two agents makes a private unverifiable investment that determines the quality of jointly produced output sold to a buyer, who may choose to buy (action B) or not (action N). The quality of the output is verifiable in court, and it is high if and only if both agents invest (action I). Each agent can costly choose to monitor the other agent (action M) or she may trust her (action T). There are two signals verifiable in court. The court will observe the first (second) signal if and only if the first (second) agent underinvests (action U) and the second (first) agent plays M. The verifiability structure associated with the agents' choices is described in Figure 1.

The payoffs are as follows. Each agent pays a cost c when playing I, and suffers a loss l if playing M. If the output is of high quality, it is worth y to the buyer, otherwise it is worth 0. In either case, it is worth 0 to the agents. If the transaction is concluded, the buyer receives the output and pays off p to each agent. To make the problem meaningful we assume that $p > c$, and that $y > 2p$.

The first-best outcome is (IT, IT, B). Any contract prescribing that each agent plays IT fails to achieve first best. Suppose that the first agent violates the agreement and plays UT. The projection of $P(UT, IT, B)$ onto the first agent's action set includes action IT, and so the first agent cannot be identified as the contract violator. Since it cannot be achieved with any contract prescribing that the agents play IT, the outcome (IT, IT, B) cannot be achieved with any enforceable contract, or with the most restrictive inviolable contract associated with (IT, IT, B).[19]

Suppose however that each agent contractually commits to play IM. In equilibrium the agents will secretly violate the contract and play IT, the buyer will play B, and first best will be achieved. Say in fact that the first agent deviates from equilibrium to play UT and is taken to court by the buyer. Since the projection of $P(UT, IT, B)$ onto the action set of the first player does not include the action IM, the court will conclude that the first agent has violated the agreement. If instead both agents play IT, since $(IM, IM, B) \in P(IT, IT, B)$, it will be impossible to show in court that they did not make the effort to monitor each other. $\diamond$

Now we show an instance where all inviolable contracts are suboptimal, but the players can achieve first best by signing a contract that will be enforced in equilibrium.

Example 3 Consider the hold-up problem presented in Figure 2. Player 1 may make an unverifiable costly investment that increases the quality of a durable good, to be supplied to

[19]By using the correspondence f introduced in Equation (4), it can be shown that the most restrictive inviolable contract associated with (IT, IT, B) is the contract A, that places no restrictions on the players' choices. It is immediate that also the most restrictive contract immune to all deviations (including joint ones) coincides with A.

	B	N
L	4, 0	0, 2
H	3, 3	1, 1

Figure 2: Hold Up Problem

player 2, who in turn may decide whether to buy the good or not. When the good is of high quality, the buyer prefers to buy it, but when the good is of low quality, the buyer would rather not. The court cannot verify whether the transaction has taken place, or ascertain the quality of the good before it is given to the buyer. However, if the good is of low quality, it will deteriorate shortly after being bought, and the buyer will be able to take the seller to court. In that case the court will conclude that the buyer has bought the good, and that the good was of low quality.[20]

It is straightforward to show that the efficient outcome (H, B) may be achieved by signing the contract $\{(H, B)\}$. In fact, when player 2 plays B, player 1 would be found guilty of violating the contract if she played L. At the same time, when player 1 plays H, playing B is in the best interest of player 2. The contract $\{H, B\}$ however is not inviolable, because in principle player 2 may play N, and this violation is not verifiable. Since the projection of $P(H, B)$ onto each player's action set covers the entire set, the most restrictive inviolable contract associated with (H, B) does not restrict the players' choice at all, and thus can only achieve the Nash equilibrium (L, N). $\diamond$

We conclude this section by pointing out that our previous analysis can be easily extended to show that the most-restrictive inviolable contracts are necessarily optimal if and only if the projections of the verifiability structure onto each player's action set are all transitive.

[20]We could argue that, while the court can verify that the profile (L, B) has occurred, she may not be able to verify that the profile (L, B) has not occurred. In that case the information structure P would satisfy $P(H, B) = P(H, N) = P(L, N) = A$. The analysis would not change with respect to the partitional information version presented here.

We omit the details for brevity.

5 Conclusion

This paper has extended the normal-form model of BW to account for non-product and non-partitional verifiability structures. We have also shown how to modify the model to account for joint and individual liability. In each instance, we have derived a complete characterization of the optimal contracts and of the most-restrictive contracts that do not include unverifiable prescriptions. Our analysis has shown that the natural intuition that the players always stipulate agreements which depend only on verifiable aspects of the interaction is not entirely correct. Moreover the issue may not be settled by looking only at the imperfections in the gathering of information by the courts. We have identified some critical liability assumptions that are required for the optimality of the contracts which specify only verifiable prescriptions.

References

[1] Anderlini L. and L. Felli [1999]: "Incomplete Contracts and Complexity Costs" *Theory and Decision* **46**: 23-50

[2] Andreoni J. [1991]: "Reasonable Doubt and the Optimal Magnitude of Fines: Should the Penalty Fit the Crime?" *Rand Journal of Economics* **22**: 385-395

[3] Bergstrom T, L. Blume and H. Varian [1986]: "On the private provision of public goods" *Journal of Public Economics* **29**: 25-49

[4] Bernheim B. D. [1986]: "On the Voluntary and Involuntary Provision of Public Goods" *American Economic Review* **76**: 789-793

[5] Bernheim B. D. and M. D. Whinston [1998]: "Incomplete Contracts and Strategic Ambiguity" *American Economic Review* **88**: 902-32.

[6] Brandenburger A. Dekel E. and J. Geanakoplos [1992]: "Correlated Equilibrium with Generalized Information Structures" *Games and Economic Behavior* **4**: 182-201

[7] Dekel E. and F. Gul [1997]: "Rationality and Knowledge in Game Theory" in Kreps D. M. and F. K. Wallis (Eds.) *Advances in Economic Theory, Seventh World Congress*, Cambridge, Cambridge University Press

[8] Fienberg S. [1989]: *The Evolving Role of Statistical Assessments as Evidence in the Courts,* Springer, New York

[9] Fudenberg D. and J. Tirole [1991]: *Game Theory* Cambridge MA, MIT press

[10] Geanakoplos J. [1989]: "Game Theory without Partitions, and Applications to Speculations and Consensus" *Cowles Foundation for Research at Yale University* mimeo

[11] Grossman S. and O. Hart [1986]: "The Costs and Benefits of Ownership: a Theory of Vertical and Lateral Integration" *Journal of Political Economy* **94**: 691-719

[12] Hart O. [1995]: *Firms, Contracts and Financial Structure* Oxford University Press

[13] Hart O. and J. Moore [1988]: "Incomplete Contracts and Renegotiation" *Econometrica* **4**: 755-785

[14] Hart O. and J. Moore [1990]: "Property Rights and the Nature of the Firm" *Journal of Political Economy* **98**, 1119-1158

[15] Hart O. and J. Moore [1999]: "Foundations of Incomplete Contracts" *Review of Economic Studies* **66**: 115-138

[16] Hellwig M. [1996]: "Notes on the Theory of Incomplete Contracts" mimeo, University of Mannheim

[17] Holmstrom B. [1982]: "Moral Hazard in Teams" *The Bell Journal of Economics* **13**: 324-340

[18] Holmstrom B. and J. Tirole [1989]: "The Theory of the Firm" *Handbook of Industrial Organization,* Vol. 1, North Holland

[19] Laffont J. J. and E. Maskin [1982]: "The Theory of Incentives: An Overview" in *Advances of Economic Theory*, ed. W. Hildebrand, Cambridge University Press

[20] Legros P. and S. Matthews [1991]: "Efficient and Nearly-Efficient Partnerships" *Review of Economic Studies* **61**: 109-129

[21] Maskin E. and J. Tirole [1999a]: "Unforeseen Contingencies and Incomplete Contracts" *Review of Economic Studies* **66**: 83-114

[22] Maskin E. and J. Tirole [1999b]: "Two Remarks on the Property-Rights Literature" *Review of Economic Studies* **66**: 139-150

[23] Milgrom P. and J. Roberts [1986] "Relying on the Information of Interested Parties" *RAND Journal of Economics* **17**: 18-32

[24] Monderer D. and D. Samet [1989]: "Approximating Common Knowledge with Common Beliefs" *Games-and-Economic-Behavior* **1**: 170-90

[25] Moore J. and R. Repullo [1998]: "Subgame Perfect Implementation" *Econometrica* **56**: 1191-1220

[26] Noldeke G. and K. Schmidt [1995]: "Option Contracts and Renegotiation: a Solution of the Hold-Up Problem" *Rand Journal of Economics* **26**: 163-179

[27] Okuno-Fujiwara M, A. Postlewaite and K. Suzumura [1990]: "Strategic Information Revelation" *Review of Economic Studies* **57**: 25-47

[28] Palfrey T. R. and H. Rosenthal [1984]: "Participation and the provision of discrete public goods: a Strategic Analysis" *Journal of Public Economics* **24**: 171-193

[29] Rubinfeld D. and D. Sappington [1987]: "Efficient Awards and Standards of Proof in Judicial Proceedings" *Rand Journal of Economics* **18**: 308-15

[30] Segal I. [1999]: "Complexity and Renegotiation: A Foundation for Incomplete Contracts" *Review of Economic Studies* **66**: 57-82

[31] Seidmann D. [1997]: "Can the Right of Silence Benefit Innocent Suspects who tell the Truth?" *University of Newcastle-upon-Tyne* mimeo

[32] Stole L. [1997]: "Lectures on the Theory of Contracts and Organizations" Chapters 1 and 2, *University of Chicago Graduate School of Business,* unpublished notes

[33] Tirole J. [1986]: "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations" *Journal of Law Economics and Organization* **2:** 181-212

[34] Tirole J. [1992]: "Collusion and the Theory of Organizations" *Advances of Economic Theory: Sixth World Congress* (ed. J. J. Laffont): 151-206

[35] Tirole J. [1999]: "Incomplete Contracts, Where do we Stand?" *Econometrica* **67**: 741-782