RATIONALITY
& COMPETITION
CRC TRR 190

Principled Mechanism Design with Evidence

Sebastian Schweighofer-Kodritsch (HU Berlin)
Roland Strausz (HU Berlin)

Discussion Paper No. 504

May 16, 2024

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Principled Mechanism Design with Evidence*

Sebastian Schweighofer-Kodritsch and Roland Strausz†

January 18, 2024

Abstract

We cast mechanism design with evidence in the framework of Myerson (1982), whereby his generalized revelation principle directly applies and yields standard notions of incentive compatible direct mechanisms. Their specific nature depends on whether the agent's (verifiable) presentation of evidence is contractually controllable, however. For deterministic implementation, we show that, in general, such control has value, and we offer two independent conditions under which this value vanishes, one on evidence (WET) and another on preferences (TIWO). Allowing for fully stochastic mechanisms, we also show how randomization generally has value and clarify to what extent this value vanishes under the common assumption of evidentiary normality (NOR). While, in general, the value of control extends to stochastic implementation, neither control nor randomization have any value if NOR holds together with WET or TIWO.

JEL Classification: D82

Keywords: Mechanism Design, Revelation Principle, Evidence, Verifiable Information, Value of Control, Value of Randomization

*This paper supersedes the mimeo "Mechanism Design with Partially Verifiable Information" written by Roland Strausz at the Cowles Foundation at Yale University in the spring of 2016. We thank Andreas Asseyer, Ian Ball, Dirk Bergemann, Jesse Bull, Françoise Forges, Tibor Heumann, Marina Halac, Johannes Hörner, Navin Kartik, Deniz Kattwinkel, Frédéric Koessler, Daniel Krähmer, Matthias Lang, Barton Lipman, Vincent Meisner, Roger Myerson, Mallesh Pai, Philip Reny, Vasiliki Skreta, Philipp Strack, Yiman Sun, Juuso Toikka, Juuso Välimäki, Joel Watson, Alex Wolitzky, and Jiawei Zhang for very helpful discussions and comments on earlier drafts. We also thank seminar participants at WU Vienna, (virtual) CMID 2020 Klagenfurt, (virtual joint) CUHK–HKU–HKUST, (virtual) ESNAWM 2021, (virtual) SAET Conference 2021 Seoul, PSE, Yale, Toronto, Penn State, ESSET Gerzensee, Cambridge (UK), Zürich, Bristol, Bielefeld, Chicago and Northwestern. Funded by the European Union (ERC, PRIVDIMA, 101096682). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. Support by the *Deutsche Forschungsgemeinschaft* through CRC TRR 190 (project number 280092119) is gratefully acknowledged.

†School of Business and Economics, Humboldt-Universität zu Berlin; sebastian.kodritsch@hu-berlin.de and roland.strausz@hu-berlin.de.

1 Introduction

We develop a Myersonian approach to mechanism design with evidence. By casting the problem in the general framework of Myerson (1982), the presentation of evidence is captured as an *action* rather than as a form of *communication*. As we demonstrate, this avoids and also clarifies a number of fundamental issues that have been raised under existing approaches (e.g., Green and Laffont (1986), Bull and Watson (2007), Deneckere and Severinov (2008)). Most fundamentally, the general formulation of the revelation principle of Myerson (1982) directly applies and simplifies the design problem with standard notions of direct mechanisms and incentive compatibility, thus unifying mechanism design with and without evidence – both conceptually and technically. With regards to applications, our framework enables us to identify as well as relax the implicit assumptions concerning the contractibility of evidence that underlie existing approaches, and we characterize their economic significance in terms of general properties of the design problem.

In particular, due to the validity of the revelation principle, the framework permits a general characterization of optimal mechanisms, without resorting to restrictions on the underlying evidence structure such as a "nested range condition" (NRC, Green and Laffont (1986)), or "normality" (NOR, Bull and Watson (2007)).[1] Hence, our approach allows to dispense with these restrictions for applications, as for instance in Hart, Kremer, and Perry (2017), or Ben-Porath, Dekel, and Lipman (2019). For the same reason, when these evidentiary conditions are not met, it also allows to dispense with strong simplifying assumptions on preferences, as for instance in Singh and Wittman (2001), or Glazer and Rubinstein (2006). However, as we show, combining the evidentiary condition of NOR with either of two further conditions that we introduce below—WET, another evidentiary condition, or TIWO, a condition on preferences—renders applications with evidence especially tractable in two important ways. First, it eliminates the need for stochastic mechanisms, so that a focus on deterministic implementation is without loss. Second, it ensures that the extent to which the presentation of evidence is contractible becomes inconsequential for implementability.

Indeed, our Myersonian approach immediately implies that the definition of incentive compatible direct mechanisms crucially depends on the contractibility of evidence – or, in the language of Myerson (1982), on its *controllability*. Loosely speaking, this refers to whether the principal may offer contracts that bind the agent to presenting specific evidence or instead faces moral hazard. Consequently, we obtain two different specifications

[1] NOR allows to pin down the evidence on which mechanisms can focus without loss, to "maximal" evidence. It is equivalent to the "full reports condition" proposed in Lipman and Seppi (1995), as well as the "minimal closure condition" in Forges and Koessler (2005). NRC is stronger than NOR, however (see our Appendix C for details).

of mechanism design with evidence, one for *controllable evidence* and one for *non-controllable evidence*.

Applying the resulting framework to the case of a single privately informed agent, we show how it allows to fully characterize the set of implementable social choice functions, first for deterministic and then for general stochastic implementation. Moreover, in each case, we characterize and compare implementability under controllable and non-controllable evidence.

With regard to deterministic implementation, direct mechanisms for controllable evidence correspond to simple menus of contracts from which the privately informed agent gets to select one. Each contract specifies both the principal's allocation and the agent's evidence, so the agent effectively gets to choose his favorite allocation among those that his evidence gives him access to. By contrast, direct mechanisms for non-controllable evidence are more complex due to evidentiary moral hazard. In particular, they involve recommendations to the agent for what evidence he should present, and incentive compatibility requires that the agent obediently follows these recommendations (in addition to honestly reporting his type). We show, however, that these mechanisms permit a representation that coincides with the contracts that are studied in the literature's alternative approaches to evidence; in particular, we thereby clarify the notion of inalienability as used in Bull and Watson (2007).

The notion of controllability is economically meaningful, in the sense that control has value even though, without control, only a problem of "verifiable moral hazard" obtains (as evidence is verifiable by definition): We show with an explicit example (Example 1) that controllable evidence supports the implementation of social choice functions that are not implementable with non-controllable evidence.[2]

We then identify two independent conditions each of which is sufficient (and also weakly necessary) for implementability to coincide under both contracting regimes, so that assuming controllability of evidence to analytically simplify the problem is without loss. Underlying both conditions is the conceptual insight (Corollary 1) that evidentiary control benefits the principal only if she benefits from restricting the pieces of evidence that the agent can use. The first condition, which we call WET, is novel in the literature and concerns only the problem's evidence structure. It requires that any evidence is "maximal" with respect to some agent-type, in the sense of proving all that the agent could possibly prove if that were truly his type.[3] This means there is a weakest evidence type (WET) for every piece

[2]This value of control then also implies that in settings with evidence, the "delegation principle" generally fails, because despite verifiability, evidentiary moral hazard is costly; see Rochet (1985), who introduces it as "taxation principle."

[3]The notion of "maximal evidence" is well-known in the existing literature on evidence as the defining feature of NOR (see Bull and Watson, 2007, p. 79). Following our terminology, NOR requires that any agent-type has evidence that is maximal with respect to his true type. NOR and WET are two independent conditions, however; an evidence structure may satisfy none, both, or only one of them.

of evidence, who cannot prove more than any other type that has this evidence, so that a principal can effectively exclude it by assigning to it the same outcome as for the associated weakest evidence type. The second condition, which we call TIWO, concerns only the problem's preference structure and requires that some allocation is worst for the agent regardless of his type. This means a principal can effectively exclude any evidence by assigning to it this type-independent worst outcome (TIWO). TIWO is satisfied in particular whenever the agent's ordinal preferences over outcomes are type-independent, as is assumed in much of the literature on evidence games (see, e.g., Glazer and Rubinstein (2006), Sher (2011, 2014), or Hart et al. (2017)). Notably, TIWO naturally arises in settings with monetary transfers, which clarifies why "verifiable moral hazard" would usually be an oxymoron.

Turning to stochastic implementation of deterministic social choice functions, we show that the value of control extends, and we provide explicit examples to establish that also randomization has value, both for controllable evidence (Example 2) and for non-controllable evidence (Example 3). These examples illustrate also the general insights that: (i) when evidence is controllable, stochastic contracts that randomize over evidence allow to weaken incentive constraints considerably, in that any agent-type need only be prevented from mimicking types for whom he has *all* the evidence (i.e., in the absence of moral hazard randomization allows to "maximize" evidence), and (ii) when evidence is non-controllable, implementation of a deterministic objective can require randomization that perfectly correlates the stochastic allocation rule with the stochastic evidence recommendation. We further prove that, while NOR is sufficient (and also weakly necessary) for randomization not to have any value with controllable evidence, this is not generally true when evidence is non-controllable: Although deterministic evidence recommendations are then without loss, there can still be value to randomizing allocations "off-path" to mitigate evidentiary moral hazard (Example 4); this latter result clarifies the restriction to "simple type dependence" in Ben-Porath et al. (2019) to establish that randomization has no value in a setting with NOR and non-controllable evidence. Hence, with controllable evidence, the only purpose of randomization is to overcome the physical limits of evidence presentation when some agent-types are unable to present maximal evidence for their type (i.e., when NOR is violated); with non-controllable evidence, the additional moral hazard problem means that randomization may additionally serve the purpose of mitigating evidentiary moral hazard, as an imperfect substitute of control, while also being less powerful in "maximizing" evidence.

Putting our results together, when NOR is satisfied jointly with WET or TIWO, control has again no value, so also randomization then has no value, even for non-controllable evidence. Notably, the evidence structure introduced by Dye (1985), which is commonly assumed in applications (e.g., Shin (2003); Acharya, DeMarzo, and Kremer (2011)), satisfies both NOR and WET. Hence, our results clarify the tractability this assumption affords and suggest directions for generalizing existing results. Similarly, the aforementioned theoretical analyses of evidence games usually assume type-independent preferences over outcomes for the sender, implying TIWO, so this shows the strength of NOR in such settings for mechanism design analysis – not only is randomization then unnecessary but the analysis can be carried out as if evidence were controllable, for simplicity. A final example (Example 5) highlights a complementarity of control and randomization, by showing that, without NOR, control can have value for stochastic implementation even when both WET and TIWO hold jointly.

The rest of our paper is structured as follows: In the next section, we introduce the evidentiary implementation model, and we explain how we operationalize the mechanism design framework and revelation principle of Myerson (1982). Throughout we apply the revelation principle thus derived for each contractual regime—controllable evidence and non-controllable evidence—to characterize implementability of deterministic social choice functions. In Section 3, we first restrict attention to deterministic implementation, for which we establish and characterize the value of control, introducing WET and TIWO. In Section 4, we consider fully stochastic implementation (of the deterministic objective) and the value of randomization, where NOR plays a key role in simplifying the design problem, and we also revisit the value of control. The final Section 5 discusses the related literature (in addition to more specific remarks in the results sections) and straightforward extensions as well as limitations of our analysis, concerning in particular settings with multiple agents and sequential evidence presentation.

2 Model and Mechanism Design Set-up

Setting and Problem. A principal (she) has to decide about an allocation $x \in X$ that affects an agent (he). The agent has private information captured by his type $\theta \in \Theta$. The agent's type identifies *both* his preferences over allocations $x \in X$, according to the utility function $u : X \times \Theta \to \mathbb{R}$, and the evidence $e \in E$ that he has available, according to the mapping $\mathcal{E} : \Theta \to 2^E \setminus \{\emptyset\}$. Hence, the agent's type θ captures all of his private information, and any type report $\hat{\theta}$ by the agent implies a claim both to his preferences as well as to the evidence he has available, namely $u(\cdot|\hat{\theta})$ and $\mathcal{E}(\hat{\theta})$.

Adopting the general formulation of evidence by Bull and Watson (2007),[4] E is an abstract set describing all potentially available evidence, in the sense that $E = \cup_{\theta\in\Theta}\mathcal{E}(\theta)$,

[4]This formulation generalizes both the "type-reports-as-evidence" model introduced in the pioneering work by Green and Laffont (1986) and the "reduced-form" model recently studied by Ben-Porath et al. (2019). It also covers the evidence formulation of Deneckere and Severinov (2008), which makes explicit all "atomic" pieces of evidence.

while $\mathcal{E}(\theta) \subseteq E$ represents the evidence available to a given type θ.[5] By convention, type θ can present *only one* element $e \in \mathcal{E}(\theta)$.[6] Evidence $e \in E$ proves that the agent's type is one in $\{\theta \in \Theta : e \in \mathcal{E}(\theta)\}$. Thus, a given type θ can partially verify his private information by presenting some evidence $e \in \mathcal{E}(\theta)$ available to him. Note that the possibility of "no evidence" can be captured by some $e_0 \in E$ such that $e_0 \in \mathcal{E}(\theta)$ holds true for all types $\theta \in \Theta$ of the agent, whereby e_0 proves nothing.

We refer to the pair (X, u) as the problem's *preference structure* and the pair $(E, \mathcal{E})$ as the problem's *evidence structure*. To circumvent measure-theoretical complications, we assume that X, Θ and E are finite sets. Consequently, we let $p \in \Delta(\Theta)$ denote the common prior probability distribution over types, and we assume it has full support; i.e., p is a mapping from Θ to $[0, 1]$ with $\sum_{\theta\in\Theta} p(\theta) = 1$ and $p(\theta) > 0$ for all $\theta \in \Theta$.

Summarizing, the collection $(\Theta, p, X, u, E, \mathcal{E})$ represents the primitives of our *evidentiary implementation model*, describing the model's type structure (Θ, p), its preference structure (X, u), and its evidence structure $(E, \mathcal{E})$. Our main focus is on the question of what (deterministic) social choice functions $f_X : \Theta \to X$ the principal is able to implement by relying on evidence as her only means to provide incentives. We refer to this as the *evidentiary implementation problem*. It means we analyze the fundamental constraint that the agent's private information imposes on *any* objective of the principal, which we therefore leave unspecified.[7] Moreover, and in line with the literature studying such problems, we abstract from any participation decision by the agent, so the agent must participate in the mechanism chosen by the principal. This is without loss if the agent always (i.e., for any type) prefers any allocation in X over his outside option.

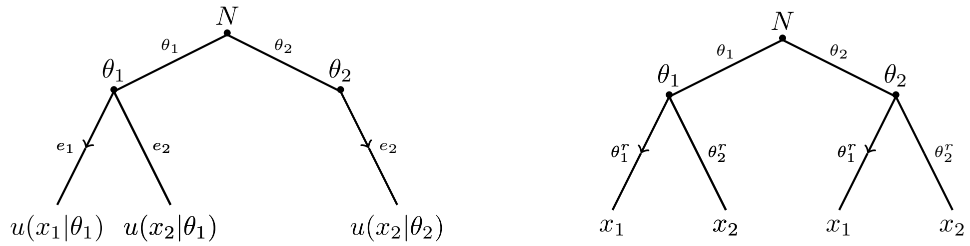
Mechanism Design and Bayesian Games. Our first goal is to clarify the extent to which we can employ the standard tools of mechanism design to evidentiary implementation problems, in particular standard formulations of the revelation principle (Myerson, 1979, 1982). We therefore briefly illustrate here why, with evidence, this is not immediate. The complication arises from the defining feature of evidence. Any mechanism that calls upon the agent to present evidence induces a game with *type-dependent strategy spaces*. Mechanism design, however, is built on the theory of Bayesian games as introduced by Harsanyi (1967), which presumes that strategy spaces are type-*independent*.

To see that such type-independent strategy spaces are essential for the general applicability of the revelation principle, consider the following example. The principal has to decide between two allocations $X = \{x_1, x_2\}$, and the agent has two possible types $\Theta = \{\theta_1, \theta_2\}$.

[5] We use $\subseteq$ to denote weak set inclusion and $\subset$ to denote strict set inclusion.

[6] That the agent can only present one element $e \in E$ is without loss, as any combination of pieces of evidence that the agent could present altogether can be represented by a specific $e \in E$.

[7] Consequently, beyond its support Θ, the prior distribution p over types will play no role in our results.



Game Γ^d with type-dependent strategy spaces

Direct mechanism for Γ^d

Figure 1: Failure of the revelation principle with type-dependent strategy spaces. The left panel represents the game with type-dependent strategy spaces. The right panel represents the corresponding direct mechanism, which fails to replicate the equilibrium outcome.

While both types prefer x_1 over x_2, type θ_1 has evidence $\mathcal{E}(\theta_1) = E = \{e_1, e_2\}$ and type θ_2 has evidence $\mathcal{E}(\theta_2) = \{e_2\}$. Suppose the principal offers an "evidence matching mechanism" such that, if the agent presents evidence e_i then allocation x_i is selected, $i \in \{1, 2\}$. This mechanism induces a game with type-dependent strategy spaces, as illustrated in the left panel of Figure 1. Since type θ_2 can only present evidence e_2, whereas type θ_1 can and also will present e_1 in order to obtain his preferred allocation x_1, this mechanism implements the social choice function f_X such that $f_X(\theta_i) = x_i$, $i \in \{1, 2\}$.

The revelation principle states that it is possible to replicate any equilibrium outcome via honest type reports that "directly" lead to the corresponding allocation. This claim is not true when strategies in the original mechanism are type-dependent. As the right panel of Figure 1 illustrates, the issue is that the direct mechanism that corresponds to the indirect mechanism in the left panel gives type θ_2 access to the preferred allocation x_1 by simply reporting his type as θ_1.[8]

Harsanyi (1967) already explicitly addresses the possibility of type-dependent strategy spaces and argues that the assumption of type-independent strategy spaces is without loss, because any game with type-dependent strategy spaces has a strategically equivalent representation by a game with type-independent strategy spaces: "the assumption that a given strategy $s_i = s_i^0$ is not *available* to player i is equivalent, from a game-theoretical point of view, to the assumption that player i will never actually *use* the strategy s_i^0 (even though it would be physically available to him) because by using s_i^0 he would always obtain some extremely low (i.e. high negative) payoffs" (Harsanyi, 1967, p. 1809; emphasis and parentheses in original).

In our context, we obtain this strategically equivalent representation by defining the

[8]Indeed, when the agent's true type is θ_2, it is also *physically* impossible to replicate type θ_1's presentation of e_1, since this evidence is then not available. Relatedly, see Myerson (1991, Chapter 6) for a discussion of the contract-theoretical foundation of the revelation principle's replication argument.

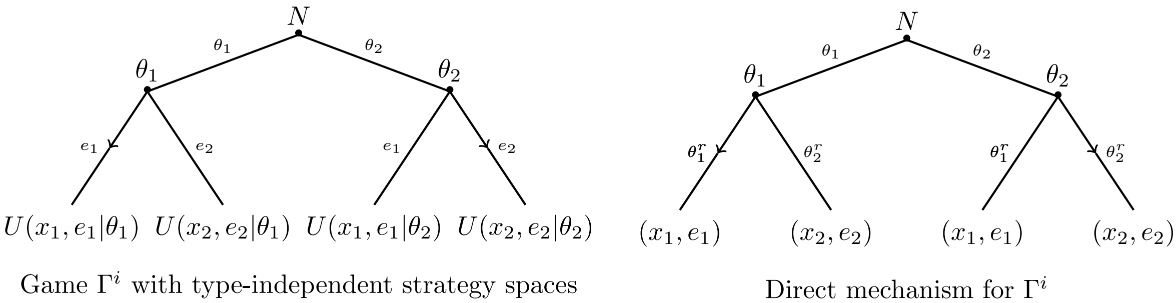


Figure 2: The revelation principle restored. The left panel represents the game under the equivalent "as-if" interpretation with type-independent strategy spaces and evidence-extended allocations/utility. The right panel represents the corresponding direct mechanism, which now replicates the equilibrium outcome.

following *(evidence-)extended utility function*:

$$U\left(x, e \,|\theta\right) = \begin{cases} u\left(x \,|\theta\right) & \text{if } e \in \mathcal{E}\left(\theta\right), \\ -c & \text{otherwise,} \end{cases} \tag{1}$$

with c sufficiently large.[9] Evidence-extended utility functions translate the type-dependence of the agent's evidence into a property of his payoffs, so we can interpret any mechanism *as-if* inducing a game in which any type of agent θ has *any* evidence $e \in E$ available, but his payoff depends on the *(evidence-) extended allocation* (x, e) according to $U(x, e|\theta)$. Introducing this extension (of utility functions and allocations) thus ensures that the principal's mechanism induces a Bayesian game in the strict sense of Harsanyi (1967) and is a first necessary step for operationalizing standard tools of mechanism design to evidentiary implementation problems.[10]

Applying Harsanyi's procedure to our specific example, Figure 2 confirms that it restores the validity of the revelation principle. The left panel of Figure 2 illustrates the game with type-independent strategy spaces that obtains from using the extended utility functions; the right panel shows its direct representation does indeed replicate its equilibrium outcome.

The example also serves to illustrate a related point: In contrast to communication, evidence is a primitive of the setting and problem, hence cannot be arbitrarily designed.

[9]We can make "sufficiently large" precise by defining $\overline{u} \equiv \max\left\{u\left(x\,|\theta\right) : (x,\theta) \in X \times \Theta\right\}$ and $\underline{u} \equiv \min\left\{u\left(x\,|\theta\right) : (x,\theta) \in X \times \Theta\right\}$. When restricting attention to deterministic mechanisms, as we do in Section 3, $c > -\underline{u}$ is sufficient. When considering stochastic mechanisms, as we do in Section 4, $c > -\underline{u} + (|E| - 1) \cdot (\overline{u} - \underline{u})$ is sufficient, where $|E|$ is the cardinality of the (finite) set E.

[10]It is already at this level where our approach diverges from the existing literature on mechanism design with evidence, which has disregarded the issue of type-dependent strategy sets but nevertheless applied the usual notions developed for Bayesian games. For instance, Bull and Watson (2007, p. 80) speak of "a Bayesian game with type-contingent restrictions on actions," and although in their footnote 7, they explicitly mention the use of evidence-extended utility functions as an "alternative way", they do not pursue this modeling alternative. By contrast, and in the context of full implementation with evidence, Kartik and Tercieux (2012) use an approach similar to ours.

We will therefore conceptualize the presentation of evidence as an *action* by the agent. This enables us to cast the evidentiary implementation problem in the framework of Myerson (1982), where—in line with evidence-extended utility—the factual presentation of evidence is to be considered a part of the outcome implemented by any mechanism.[11]

Bayes-Nash Implementability. Having ensured that any mechanism induces a proper Bayesian game, we next apply the usual definition of (Bayes-Nash) implementability in mechanism design. That is, we say that a mechanism *implements* the social choice function f_X, if the (proper) Bayesian game induced by the mechanism has a Bayes-Nash equilibrium (BNE) whose outcome is such that an agent of type θ receives allocation $f_X(\theta)$. Moreover, we say that a social choice function f_X is *implementable*, if there exists a mechanism that implements it.

To explicitly keep track of how evidence supports implementation, we introduce *evidence extensions* $f_E : \Theta \to E$ such that $f_E(\theta) \in \mathcal{E}(\theta)$ for all types θ, which describe for each type θ the evidence $e \in \mathcal{E}(\theta)$ that he presents. We call a pair $f = (f_X, f_E)$, consisting of a social choice function f_X and an evidence extension f_E, an *(evidence-) extended social choice function*, and we denote the space of all such extended social choice functions by $\mathcal{F}$. We say that a mechanism implements the extended social choice function $f = (f_X, f_E) \in \mathcal{F}$, if it implements the social choice function f_X in a BNE whose outcome is such that an agent of type θ presents evidence $f_E(\theta)$. Likewise, we say that an extended social choice function f is implementable, if there exists a mechanism that implements it.

While we follow the literature in focusing on the implementability of *deterministic* social choice functions, a well-known complication is that even the implementation of such a deterministic objective may require randomization (see, e.g., Deneckere and Severinov, 2008, Example 5). It will be instructive, however, to first analyze deterministic implementation in Section 3 and postpone the discussion of general (i.e., possibly stochastic) implementation of deterministic social choice functions to Section 4.

Revelation Principle. For studying the limits of implementability under asymmetric information, the revelation principle plays a crucial role, because it yields a tractable characterization of a subset of feasible mechanisms by which any implementable allocation can be attained. This subset of mechanisms is therefore often referred to as *canonical*. The most general formulation of the revelation principle for Bayes-Nash implementability is that of Myerson (1982). This work extends results based on the traditional domain of pure adverse selection problems (see, e.g., Myerson, 1979) to domains including problems of moral hazard

[11]See Forges and Koessler (2005) for a related discussion, when taking as given the type-dependent "verifying messages" (evidence) available to players and allowing for any type-independent messages (communication).

(see, e.g., Holmström, 1979), the traditional domain of contract theory. Our evidence extensions enable us to cast the evidentiary implementation problem in this general framework and thereby operationalize this generalized revelation principle.[12]

What makes the framework of Myerson (1982) especially insightful for analyzing evidentiary implementation is that it requires an explicit assumption about whether actions by the agent—here the presentation of evidence—are (contractually) controllable or not. This determines whether an action is subject to moral hazard, where controllability is stronger than the concept of verifiability in contract theory. That is, all controllable actions are verifiable, but a non-controllable action is verifiable if the principal can condition her controllable actions on it and is non-verifiable if she cannot. Hence, a non-controllable verifiable action underlies *verifiable moral hazard*.

While in many economic contexts, the concepts of controllablity and verifiability yield identical results in terms of implementability, we will show that this is not the case for evidentiary implementation. Indeed, while throughout this paper we follow the literature's assumption that, by its very nature, what evidence the agent presents is verifiable, we obtain differences in implementability, depending on whether this action is controllable or not. These differences are reflected by the revelation principle, in that the concrete structure of direct mechanisms and the associated notion of incentive compatibility differ depending on whether evidence is controllable or non-controllable. It is useful to first illustrate these differences in the context of deterministic implementation.

3 Deterministic Mechanism Design with Evidence

In this section, we consider evidentiary implementation problems under the restriction to deterministic mechanisms. Accordingly, we say that a mechanism *d-implements* the social choice function f_X if it induces a Bayesian game in which there are no moves of nature—apart from determining the agent's type according to the common prior—and this game has a *pure* BNE whose outcome is such that an agent of type θ receives allocation $f_X(\theta)$. We say that a social choice function f_X is *d-implementable* if there exists a mechanism that d-implements it. By using the notions of evidence-extended utility and social choice functions, we derive a generalized revelation principle in the full spirit of Myerson (1982) for d-implementability (see Appendix B, in particular Theorem B): Any d-implementable social choice function f_X is implementable by a deterministic mechanism that is direct and incentive compatible

[12]Because BNE is a static solution concept, Myerson (1982) concerns static/one-shot principal-agent problems. Hence, introducing the presentation of evidence as an action implies that there is a single stage at which this is decided. While this is without loss with a single agent, where it is merely a matter of appropriately specifying the evidence structure (see footnote 4), we discuss the extension to potentially multiple stages of evidence presentation in Section 5, where we also consider multiple agents.

in the sense of Myerson (1982). We first explicitly derive the implied representations of incentive compatible direct mechanisms, as they depend on whether evidence is controllable or non-controllable, and then analyze the economic significance of this formal distinction.

3.1 Controllable Evidence

With controllable evidence, the mechanism design problem does not involve any moral hazard. As a result, the revelation principle reduces to its traditional version for pure adverse selection problems: A direct mechanism asks the agent to report a type and, depending on this report, selects a contract to be executed; it is incentive compatible if the agent always finds it optimal to honestly report his type.

Formally, we can represent deterministic *direct mechanisms* with controllable evidence as functions $\gamma^C : \Theta \to X \times E$. We will often express these component-wise, as a pair of functions $\left(\gamma_X^C, \gamma_E^C\right)$ with $\gamma_X^C : \Theta \to X$ and $\gamma_E^C : \Theta \to E$, specifying for a type report $\theta^r \in \Theta$, the allocation $\gamma_X^C(\theta^r)$ that the principal must provide, and the evidence $\gamma_E^C(\theta^r)$ that the agent must present. Denoting the set of all such direct mechanisms by $\Gamma^C \equiv (X \times E)^\Theta$, observe that the set of extended social choice functions is a subset of the set of all direct mechanisms, $\mathcal{F} \subseteq \Gamma^C$.

A direct mechanism γ^C is *incentive compatible* if it provides any type of agent with an incentive to report honestly; i.e., if it satisfies:

$$IC^C: \; U\left(\gamma^C(\theta)\,|\theta\right) \geq U\left(\gamma^C(\theta')\,|\theta\right), \; \forall (\theta, \theta') \in \Theta \times \Theta.$$

The revelation principle for d-implementability implies the following characterization with controllable evidence.

Theorem 1. *Suppose evidence is controllable. Then, a social choice function f_X is d-implementable if and only if there exists an evidence extension f_E such that the direct mechanism given by $(f_X, f_E) \in \Gamma^C$ satisfies IC^C.*

Intuitively, the d-implementability of a social choice function f_X is all about whether its allocations $f_X(\theta)$ can be matched with evidence $f_E(\theta) \in \mathcal{E}(\theta)$ such that

$$f_E(\theta') \in \mathcal{E}(\theta) \Rightarrow u(f_X(\theta)\,|\theta) \geq u(f_X(\theta')\,|\theta), \; \forall (\theta, \theta') \in \Theta \times \Theta. \tag{2}$$

This characterization in terms of type-dependent allocation preferences and evidence follows immediately from IC^C. By contrast, any *extended* social choice function $f = (f_X, f_E)$ already fully pins down the direct mechanism that would have to d-implement it under honesty, so IC^C immediately characterizes its d-implementability.

We can represent any of the mechanisms identified by Theorem 1 as a menu of *evidence-extended allocation contracts* $M = \{(x(\theta), e(\theta))\}_{\theta \in \Theta}$, from which the agent has to choose one. The interpretation is that if the agent picks a pair $(x, e) \in M$, the principal must provide the allocation x and the agent must present the evidence e. Since options in M for which the agent does not possess the evidence are prohibitively costly, effectively only options for which he possesses the required evidence are open to the agent. This reveals how the principal can use evidence to regulate the agent's "access" to allocations: A type θ with evidence $\mathcal{E}(\theta)$ gets to choose his most preferred allocation from the subset of allocations that his evidence gives him access to, i.e., from $X(\theta) = \{x \in X : \exists e \in \mathcal{E}(\theta), (x, e) \in M\}$.

Remarks: The assumption of controllable evidence finds its analogue in mechanism design settings with transfers, because these models also assume that transfers are controllable in the sense of Myerson (1982). Indeed, in an auction, the auction rules do not only specify who obtains the object as a function of messages (or bids) but they also specify how much each player *must* pay given these messages. Similarly in monopolistic screening models with a single buyer such as those of Mussa and Rosen (1978), the optimal mechanism is a menu of options from which the buyer can freely pick, but each of these options does not only specify the quantity (or quality) that the seller has to deliver, but it also explicitly specifies the price that the buyer *must* pay. Hence, the analogue of a menu of evidence-extended allocation contracts in our context of controllable evidence, is a menu of "transfer-extended allocation contracts" in the context of a monopolist using second-degree price discrimination. This then also implies that with controllable evidence, evidence plays a role that is structurally similar to the role of transfers in monopolistic screening models.

3.2 Non-Controllable Evidence

With non-controllable evidence, the mechanism design problem involves both adverse selection and (verifiable) moral hazard. Applying our deterministic revelation principle to this set-up yields a class of canonical mechanisms that are "direct" in the dual sense that they induce the following two-stage Bayesian game. In the first stage, after learning his type, the agent *directly* reports some type. In the second stage, the agent receives a *direct* (report-contingent) recommendation concerning his non-controllable action upon which the agent chooses his action freely and thus determines the allocation according to a (report-contingent) allocation rule. Hence, a direct mechanism with non-controllable evidence specifies both how it picks a recommendation and how it picks a final allocation depending on the evidence presented. Incentive compatibility is accordingly twofold. It requires that the agent finds it optimal to first honestly report his type and then obey the recommendation.

Formally, we can represent deterministic *direct mechanisms* in this setting as functions $\gamma^N : \Theta \to X^E \times E$. We will often express these component-wise, as a pair of functions

$\left(\gamma_X^N, \gamma_E^N\right)$ with $\gamma_X^N : \Theta \to X^E$ specifying an evidence-contingent allocation rule (which we write as $\gamma_X^N(\cdot|\theta) : E \to X$ for any given type report θ) and $\gamma_E^N : \Theta \to E$ specifying an evidence recommendation, for any type report. Denoting the set of direct mechanisms by $\Gamma^N \equiv \left(X^E \times E\right)^\Theta$, observe that such a mechanism is *incentive compatible*, ensuring "honesty-plus-obedience," if and only if it satisfies the constraint

$$IC^N: \; U\left(\gamma_X^N\left(\gamma_E^N(\theta)|\theta\right), \gamma_E^N(\theta)|\theta\right) \geq U\left(\gamma_X^N(e'|\theta'), e'|\theta\right), \; \forall(\theta, \theta', e') \in \Theta \times \Theta \times E.$$

Theorem 2. *Suppose evidence is non-controllable. Then, a social choice function f_X is d-implementable if and only if there exist an evidence extension f_E and a rule $\gamma_X^N : \Theta \to X^E$ such that (i) $\gamma_X^N(f_E(\theta)|\theta) = f_X(\theta)$ for all types θ and (ii) the direct mechanism given by $\left(\gamma_X^N, f_E\right) \in \Gamma^N$ satisfies IC^N.*

Given a direct mechanism $\gamma^N \in \Gamma^N$ that satisfies requirement (i) for some evidence extension, the incentive compatibility requirement (ii) reduces to

$$u(f_X(\theta)|\theta) \geq u(\gamma_X^N(e'|\theta')|\theta), \; \forall(\theta, \theta', e') \in \Theta \times \Theta \times E : e' \in \mathcal{E}(\theta). \tag{3}$$

With non-controllable evidence, the agent's presentation of evidence is subject to moral hazard. Consequently, even a given *extended* social choice function $f = (f_X, f_E)$ only partially pins down the direct mechanism $\gamma^N = (\gamma_X^N, \gamma_E^N)$ that would have to implement it. It does so only "on-path," meaning $\gamma_E^N(\theta) = f_E(\theta)$ and $\gamma_X^N(f_E(\theta)|\theta) = f_X(\theta)$ must hold for all types θ.

With but a single agent, recommendations turn out redundant when focusing on deterministic implementation: Recommendations that deterministically depend on the agent's type report cannot disseminate any information to the agent in excess of what the agent can already deduce from the mechanism itself. One may therefore reduce any mechanism in Γ^N that satisfies IC^N to (an indirect) one in which communication and evidence presentation are collapsed into a single stage, without any recommendation/communication from the mechanism. We can represent such mechanisms as functions $g^N : \Theta \times E \to X$, specifying allocations x conditional on the agent's type report and evidence (θ, e). While obedience loses its literal meaning in the absence of a recommendation, the honesty requirement on type reports imposes the following incentive constraint:

$$\begin{aligned} IC_G^N : \; & \exists f_E : \Theta \to E \text{ with } f_E(\theta) \in \mathcal{E}(\theta), \; \forall \theta \in \Theta, \text{ such that} \\ & U\left(g^N(\theta, f_E(\theta)), f_E(\theta)|\theta\right) \geq U\left(g^N(\theta', e'), e'|\theta\right), \; \forall(\theta, \theta', e') \in \Theta \times \Theta \times E. \end{aligned}$$

The next proposition shows that, with non-controllable evidence, mechanisms $g^N \in G^N$ subject to IC_G^N are without loss for d-implementability.

Proposition 1. *Suppose evidence is non-controllable. Then, a social choice function f_X is d-implementable if and only if there exists a mechanism $g^N : \Theta \times E \to X$ that satisfies IC_G^N and d-implements f_X.*

Proof. Only necessity requires proof, so suppose the social choice function f_X is d-implementable. By Theorem 2, there exists an evidence extension f_E and a rule $\gamma_X^N : \Theta \to X^E$ such that (i) $\gamma_X^N(f_E(\theta)\,|\theta) = f_X(\theta)$ for all types θ and (ii) the direct mechanism given by $\left(\gamma_X^N, f_E\right) \in \Gamma^N$ satisfies IC^N.

Define then mechanism $g^N : \Theta \times E \to X$ such that $g^N(\theta, e) \equiv \gamma_X^N(e|\theta)$ for all $(\theta, e) \in \Theta \times E$. By (i), $g^N(\theta, f_E(\theta)) = f_X(\theta)$ for all $\theta \in \Theta$, and by (ii), using (3), for all types $\theta \in \Theta$,

$$\begin{aligned} U\left(g^N(\theta, f_E(\theta)), f_E(\theta)|\theta\right) &= \max_{(\theta',e')\in\Theta\times\mathcal{E}(\theta)} u\left(\gamma_X^N(e'|\theta')|\theta\right) \\ &= \max_{(\theta',e')\in\Theta\times\mathcal{E}(\theta)} u\left(g^N(\theta', e')|\theta\right) \quad = \max_{(\theta',e')\in\Theta\times E} U\left(g^N(\theta', e'), e'|\theta\right). \end{aligned}$$

□

Remarks: Proposition 1 establishes precisely how our Myersonian approach relates to the existing revelation results in the literature on evidence. Both Bull and Watson (2007, Theorems 1, 2 and 5) and Deneckere and Severinov (2008, Theorem 1) derive the sufficiency of the canonical revelation mechanisms $g^N : \Theta \times E \to X$ for d-implementability of social choice functions f_X. Our approach derives these from a standard Myersonian revelation principle for settings with non-controllable evidence, hence with both adverse selection and moral hazard. Thus, we formally clarify, on the one hand, the implicit underlying assumption on the controllability of evidence in the existing literature, which is that evidence presentation is non–controllable (and to which Bull and Watson (2007) refer as "inalienability"), and, on the other hand, the otherwise hidden notion of incentive compatibility with respect to evidence presentation, which is obedience. It is worthwhile emphasizing here that the redundancy of recommendations is a consequence of the determinism assumption, as we will show in Section 4.

3.3 The Value of Control

Distinguishing the presentation of evidence according to its controllability results in two structurally different classes of canonical mechanisms and associated notions of incentive compatibility: (Γ^C, IC^C) for controllable and (Γ^N, IC^N) for non-controllable evidence. We now analyze the *economic* significance of this distinction, i.e., the implications of controllability for what the principal can implement.

We begin by showing that a principal's contractual control over the agent's evidence presentation could never hurt her. This confirms the appropriateness of the notion of controllability, as (weakly) enlarging the set of d-implementable social choice functions.

Proposition 2. *A social choice function f_X is d-implementable with non-controllable evidence only if f_X is d-implementable with controllable evidence.*

Proof. Suppose the social choice function f_X is d-implementable with non-controllable evidence. By Theorem 2, there exists an evidence extension f_E and a rule $\gamma_X^N : \Theta \to X^E$ such that (i) $\gamma_X^N(f_E(\theta)\,|\theta) = f_X(\theta)$ for all types θ and (ii) the direct mechanism given by $\left(\gamma_X^N, f_E\right) \in \Gamma^N$ satisfies IC^N.

Now let $f \equiv (f_X, f_E) \in \mathcal{F}$ and recall that this extended social choice function fully pins down the direct mechanism $f \in \Gamma^C$ that would have to d-implement it subject to IC^C with controllable evidence. Using (i) and (ii), for any $(\theta, \theta') \in \Theta \times \Theta$,

$$U(f(\theta)|\theta) = U(\gamma_X^N(f_E(\theta)|\theta), f_E(\theta)|\theta) \geq U(\gamma_X^N(f_E(\theta')|\theta), f_E(\theta')|\theta) = U(f(\theta')|\theta),$$

i.e., f satisfies IC^C, whereby f_X is d-implementable with controllable evidence. □

The intuition behind the proposition is simple. The principal faces fewer incentive constraints when evidence is controllable. Specificially, due to the absence of moral hazard, there are no obedience constraints. Hence, with controllable evidence, the principal can d-implement any social choice function that is d-implementable with non-controllable evidence.

Based on this result, given any type structure and associated evidence and preference structures, we say that *control has value*, if there exists a social choice function that is d-implementable when evidence is controllable but not when evidence is non-controllable; and we say that *control has no value*, if the sets of d-implementable social choice functions in the two cases coincide.

The following simple example shows that, in general, control indeed has value; equivalently, evidentiary moral hazard is generally costly to the principal, despite its verifiability.

Example 1. The principal has to decide between two allocations $X = \{x_1, x_2\}$, and the agent has two possible types $\Theta = \{\theta_1, \theta_2\}$. The agent's type-dependent preferences over allocations are as in the following table:

$u(x\|\theta)$	θ_1	θ_2
x_1	0	1
x_2	1	0

The evidence structure consists of three possible elements $E = \{e_0, e_1, e_2\}$ distributed over types according to

$$\mathcal{E}(\theta_1) = \{e_0, e_1\} \text{ and } \mathcal{E}(\theta_2) = \{e_0, e_2\}.$$

This means that each agent-type θ_i can either prove himself by presenting evidence e_i, or prove nothing by presenting "null-evidence" e_0. We may also interpret this as any agent-type having "hard" information to prove himself, which he may release or withhold.

Consider then the social choice function $f_X : \{\theta_1, \theta_2\} \to \{x_1, x_2\}$ such that $f_X(\theta_i) = x_i$ for both $i \in \{1, 2\}$. According to f_X, the agent should always receive his less preferred allocation.

First, note that f_X is not d-implementable when evidence is non-controllable. Consider any $\gamma^N = (\gamma_X^N, \gamma_E^N) \in \Gamma^N$. Either $\gamma_X^N(e_0|\theta_1) = x_1$, in which case type θ_2 can and therefore must obtain x_1, or $\gamma_X^N(e_0|\theta_1) = x_2$, in which case type θ_1 can and therefore must obtain x_2. Simply put, when evidence is non-controllable, any mechanism must specify some allocation in response to the provision of evidence e_0, and since both types can present it, one of them must obtain his preferred allocation.

Second, when evidence is controllable, the following mechanism $\gamma^C \in \Gamma^C$ satisfies IC^C and implements f_X: Let $\gamma^C(\theta_i) \equiv (x_i, e_i)$ for both $i \in \{1, 2\}$. Each type of agent reveals himself, because the alternative of lying means he would have to prove his lie, which is impossible. □

Jointly, Proposition 2 and Example 1 prove that, in general, control has value. The example suggests that a value of control arises from the availability of shared evidence that pools otherwise separable types. When evidence is non-controllable every such evidence has to be followed by some allocation. This, in turn, enlarges the set of "accessible" allocations for several types, thus interfering with incentives.

We next provide two independent conditions that guarantee that control has no value. The first condition concerns the evidence structure; it identifies what constitutes potentially "costly" evidence to the principal and rules out its existence. The second condition concerns the preference structure; it identifies when the presentation of potentially "costly" evidence can be deterred by the principal. Each of the two conditions is also necessary for there to be no value of control, in a weak sense.

The significance of these two conditions is twofold. First, if one of them is satisfied, then any assumptions concerning the controllability of evidence are irrelevant for deriving optimal mechanisms (assuming d-implementability). However, because the structure of direct mechanisms and incentive constraints is considerably simpler with controllable evidence, this allows to simplify the design problem by analyzing it "as if" evidence were controllable.

Second, the controllability of evidence relates to which evidentiary contracts are enforced by the court system; because in many practical situations, this institutional aspect constitutes a political "meta-decision" and is therefore endogenous itself (as opposed to being dictated by technological constraints), our two conditions also contribute to identifying the settings where this higher-level institutional design matters, and to understanding how it affects the welfare of the parties involved.

Providing the groundwork for these conditions is the following result that clarifies the essence of controllability of evidence – namely, as the principal's ability to effectively rule out some evidence. Consider the class of (indirect) mechanisms G^C defined by all functions $g_R^C : \Theta \times R \to X$ such that R is any non-empty subset of E (and a "parameter" of the mechanism). This class generalizes the class G^N of indirect mechanisms we analyzed with non-controllable evidence, which obtains for $R = E$. The interpretation is therefore similar, except that the agent is restricted to presenting only evidence from the subset R.

In line with this generalization, say that a mechanism g_R^C is *R-incentive compatible*, if it satisfies the following honesty constraint:

$$
\begin{aligned}
IC_G^C : \exists f_R : \Theta \to R \text{ with } f_R(\theta) \in \mathcal{E}(\theta),\ \forall \theta \in \Theta,\ \text{such that} \\
U\left(g_R^C\left(\theta, f_R(\theta)\right), f_R(\theta)|\theta\right) \geq\ U\left(g_R^C\left(\theta', e'\right), e'|\theta\right),\ \forall(\theta, \theta', e') \in \Theta \times \Theta \times R.
\end{aligned}
$$

Proposition 3. *Suppose evidence is controllable. Then, a social choice function f_X is d-implementable if and only if there is a subset $R \subseteq E$ and a mechanism $g_R^C : \Theta \times R \to X$ such that g_R^C satisfies IC_G^C and d-implements f_X.*

Proof. With the presentation of evidence as an action, mechanisms g_R^C with $R \neq E$ are technically not covered by the revelation principle of Theorem 1 for controllable evidence. We therefore prove both necessity and sufficiency.

For necessity, suppose the social choice function f_X is d-implementable (with controllable evidence). By Theorem 1, there exists an evidence extension f_E such that the direct mechanism given by extended social choice function $f = (f_X, f_E) \in \Gamma^C$ satisfies IC^C. Construct then $R \subseteq E$ and a mechanism $g_R^C \in G^C$ as follows:

$$
\begin{aligned}
R\ &\equiv \{e \in E\ :\ \exists \theta \in \Theta,\ f_E(\theta) = e\}, \\
g_R^C(\theta, e)\ &\equiv \begin{cases} f_X(\theta), & \text{if } e = f_E(\theta), \\ f_X(t(e)), & \text{otherwise}, \end{cases} \\
&\text{for all } (\theta, e) \in \Theta \times E, \\
&\text{for some } t : R \to \Theta \text{ with } f_E(t(e)) = e \text{ for all } e \in R.
\end{aligned}
$$

Note that, given the specification of R, a mapping t exists and is well-defined. We now show that this mechanism satisfies IC_G^C with f_R such that $f_R(\theta) = f_E(\theta) \in R$ for all types θ, which immediately implies that it d-implements f_X.

Take any type $\theta \in \Theta$ and observe that, by IC^C, for any $\theta' \in \Theta$, we have

$$\begin{aligned} U\left(g_R^C\left(\theta, f_E(\theta)\right), f_E(\theta)|\theta\right) &= U\left(f(\theta)|\theta\right) \\ &\geq U\left(f(\theta')|\theta\right) \\ &= U\left(g_R^C\left(\theta', f_E(\theta')\right), f_E(\theta')|\theta\right). \end{aligned}$$

Hence, among all choices $(\theta', e) \in \Theta \times R$ where report and evidence are consistent in the sense that $e = f_E(\theta')$, an honest type report together with the consistent evidence is optimal. However, by construction, any inconsistent choice $(\theta', e) \in \Theta \times R$ where $e \neq f_E(\theta')$ yields the same as some consistent one, because

$$U\left(g_R^C\left(\theta', e\right), e|\theta\right) = U\left(f(t(e))|\theta\right).$$

Hence, IC_G^C holds true, establishing the necessity part.

For sufficiency, suppose $R \subseteq E$ and the mechanism $g_R^C \in G^C$ satisfies IC_G^C with $f_R : \Theta \to R$ and d-implements f_X. Since f_R is an evidence extension, in terms of extended social choice functions, the mechanism d-implements $f = (f_X, f_R) \in \mathcal{F}$. We will now establish the result by showing that f satisfies IC^C. This follows immediately from IC_G^C, as, for all $(\theta, \theta') \in \Theta \times \Theta$,

$$\begin{aligned} U\left(f(\theta)|\theta\right) &= U\left(g_R^C\left(\theta, f_R(\theta)\right), f_R(\theta)|\theta\right) \\ &\geq U\left(g_R^C\left(\theta', f_R(\theta')\right), f_R(\theta')|\theta\right) \\ &= U\left(f(\theta')|\theta\right). \end{aligned}$$

□

Proposition 3 is analogous to Proposition 1, which considers the non-controllable case. The two results allow to directly relate implementability in the controllable and non-controllable case.

Corollary 1. *A social choice function f_X is d-implementable with controllable evidence in the evidentiary implementation model $(\Theta, p, X, u, E, \mathcal{E})$ if and only if there exists a subset of evidence $R \subseteq E$ such that f_X is d-implementable with non-controllable evidence in the evidentiary implementation model $(\Theta, p, X, u, R, \mathcal{R})$ with $\mathcal{R}(\theta) \equiv \mathcal{E}(\theta) \cap R \neq \emptyset$ for all $\theta \in \Theta$.*

The corollary clarifies that any value of contractual control over evidence arises from the ability to prevent the agent from presenting certain kinds of evidence.

In addition, Propositions 1 and 3 jointly offer an alternative proof that contractual control could only ever be beneficial to the principal (Proposition 2), since all mechanisms in G^N satisfying IC_G^N are also mechanisms in G^C. The proof of Proposition 2 is however more direct.

Some authors interpret communication subject to certain "lying" constraints as evidence (see Green and Laffont, 1986; Forges and Koessler, 2005; Glazer and Rubinstein, 2006; Deneckere and Severinov, 2008). In view of the value of control established here, the comparison of Propositions 1 and 3 shows that the ability to design this type of communication—specifically, to restrict it—is valuable.[13]

Finally, Proposition 3 very directly points to how control would also be valuable once we endogenized the agent's participation in the mechanism. It would enable the principal to exclude certain types of agent from her allocation, should she want to do so. Suppose the agent had a type-independent outside option of u_0, while always preferring any allocation over it (i.e., $\min_{(x,\theta)\in X\times\Theta} u(x|\theta) > u_0$). If the principal designed $R \subseteq E$ such that $\mathcal{E}(\theta)\cap R = \emptyset$ for some type of agent θ, then this type would be bound to violate the contract due to lack of evidence and hence not participate in the mechanism. Note that this means the mechanism would not implement a social choice function $f_X : \Theta \to X$ since $f_X(\theta) \notin X$. By contrast, every type of agent will participate in *any* mechanism that is feasible with non-controllable evidence, all of which implement a social choice function $f_X : \Theta \to X$.[14]

3.3.1 Weakest Evidence Types (WET)

We first develop a condition that guarantees that none of the evidence ever available to the agent is costly, in the above sense that ruling it out would be beneficial to the principal. The following notion of maximality of evidence will be key here.

Definition 1. Evidence $e \in E$ is **maximal with respect to** type $\theta \in \Theta$, if, for any type $\theta' \in \Theta$, $e \in \mathcal{E}(\theta')$ implies $\mathcal{E}(\theta) \subseteq \mathcal{E}(\theta')$. Evidence $e \in E$ is **maximal evidence** if it is maximal with respect to some type $\theta \in \Theta$.

If evidence e is maximal with respect to type θ, then it proves all that could possibly be proven if the agent's true type were θ. This is so, because e establishes that the agent has all of the evidence $\mathcal{E}(\theta)$ that type θ has. Equivalently, whenever the agent has evidence e

[13]Studying games extended by pre-play communication with type-dependent message spaces, Forges and Koessler (2005) observe that restricting these message spaces is potentially valuable. Proposition 3 identifies that the reason for this is that the availability of such restrictions effectively transforms a setting with non-controllable evidence into one with controllable evidence.

[14]Relatedly, with controllable evidence, the cost c we introduce with extended utility in equation (1) corresponds to the cost imposed on the agent by the court system in case he violates the contract chosen (footnote 9 explicitly states what this "fine" has to satisfy for our results on controllable evidence to apply). This is different from c's technical role with non-controllable evidence (where "$c = \infty$" is appropriate).

available, $\mathcal{E}(\theta)$ defines a lower bound on what he could prove (in this sense, θ is a "weakest evidence type" for evidence e).

Clearly, evidence that is exclusive to a type is maximal with respect to this type, as this evidence then provides proof of all of his private information. However, the definition is much more permissive. Evidence e can be maximal with respect to a type θ of agent without being exclusive to this type, and even without this type's possessing e.[15]

In case type θ does have evidence that is maximal with respect to himself, however, i.e., if $e \in \mathcal{E}(\theta)$, he has a way of proving everything he could possibly prove, so presenting any other (non-maximal) evidence is tantamount to withholding evidence. Observe now that in Example 1's evidence structure, evidence e_0 is not only shared by both types, but it is also not maximal evidence. Thus, it violates the following condition.

Definition 2. An evidence structure $(E, \mathcal{E})$ satisfies the **weakest evidence types condition (WET)**, if every evidence $e \in E$ is maximal evidence.

WET still allows for "null" evidence that proves nothing, i.e., some evidence e_0 such that $e_0 \in \cap_{\theta \in \Theta} \mathcal{E}(\theta)$; e.g., adding a third type θ_0 to Example 1, with evidence $\mathcal{E}(\theta_0) = \{e_0\}$, evidence e_0 still proves nothing but is then maximal with respect to this additional type, whereby the evidence structure satisfies WET. Indeed, what WET ensures is that any evidence that is shared by multiple types is still maximal with respect to some type. As the following result formally establishes, under this condition evidentiary moral hazard is costless, and the set of d-implementable allocations is independent of whether evidence is controllable.

Proposition 4. *Fix the type structure (Θ, p) and evidence structure $(E, \mathcal{E})$ of any evidentiary implementation model. Then, the evidence structure $(E, \mathcal{E})$ satisfies WET if and only if control has no value for* any *preference structure (X, u).*

The sufficiency of WET for control not to have value follows from a straightforward revealed preference argument: If evidence e is maximal with respect to some type θ, then any other type that has e could *always* mimic θ, irrespective of whether e is ruled out. Hence, if every evidence is maximal evidence, i.e., WET holds, then control has no value.

The proposition also expresses the sense in which WET is actually necessary for control not to have value: Whenever the evidence structure of an evidentiary implementation model violates WET, we can modify the preference structure such that control allows to implement strictly more. Indeed, WET is not *generally* necessary for control to not have value. That

[15] In fact, e can be maximal with respect to type θ when this type does not even possess *any* evidence that would be maximal with respect to himself. For instance, consider the following evidence structure for four types: $\mathcal{E}(\theta_1) = \{e_1\}$, $\mathcal{E}(\theta_2) = \{e_2\}$, $\mathcal{E}(\theta_3) = \{e_1, e_2, e_3\}$, and $\mathcal{E}(\theta_4) = \{e_1, e_2\}$. Neither e_1 nor e_2 are maximal with respect to θ_4, yet e_3 is maximal with respect to *every* type, hence in particular θ_4, because it proves θ_3, whereas no other evidence would disprove this type.

is, there exist evidentiary implementation models where WET is violated, yet control has no value (see the TIWO condition below).

The necessity proof is based on graph-theoretical arguments. We first show that non-maximality of some evidence $\tilde{e}$ implies a subset $\tilde{\Theta} \subset \Theta$ of at least two types that all have this evidence $\tilde{e} \in E$, such that every type θ (in the full set of types Θ, hence including also $\tilde{\Theta}$ itself) has some evidence that some (other) type in $\tilde{\Theta}$ does not have. Based on this insight, we construct a preference structure and social choice function generalizing Example 1 where control has value: There are as many different allocations in X as types in $\tilde{\Theta}$, such that each of these types desires exactly one of them (all others are indifferent over all allocations), and control allows to ensure none of them gets their desired allocation, by always requiring evidence for it that the corresponding type does not have. This is impossible without control, because then $\tilde{e}$ has to be followed by one such type's desired allocation.

Remarks: WET is a novel condition in the literature. Its key underlying notion of maximal evidence, however, also plays an important role as part of the common assumption of "normality" (see Bull and Watson, 2007), another condition on the evidence structure.[16] Using our Definition 1, we can define normality as follows, which immediately relates it to WET.

Definition 3. An evidence structure $(E, \mathcal{E})$ satisfies **normality (NOR)**, if every type $\theta \in \Theta$ has evidence $e \in \mathcal{E}(\theta)$ that is maximal with respect to θ.

While WET requires that any evidence is maximal with respect to some type, NOR requires that any type has some evidence that is maximal with respect to himself. Under NOR, the agent can always provide maximal evidence for his type, whereby not doing so means he is withholding evidence. This permits a further simplification of the canonical mechanisms identified via the revelation principle here: For controllable evidence, it is then without loss to consider direct mechanisms $\gamma^C = (\gamma^C_X, \gamma^C_E)$ where every $\gamma^C_E(\theta)$ is some evidence that is maximal with respect to type θ. In analogy to honest type reporting, the mechanism therefore requires incentive compatibility only for *maximal* evidence presentation. For non-controllable evidence, NOR implies that it is without loss to consider direct mechanisms $\gamma^N = (\gamma^N_X, \gamma^N_E)$ with an evidence recommendation γ^N_E that recommends each type to present maximal evidence for his type. Hence, NOR directly allows to pin down γ^N_E (see also Forges and Koessler, 2005; Bull and Watson, 2007; Deneckere and Severinov, 2008). We return to this implication and further clarify the role of NOR in mechanism design with evidence below, in Section 4, where we extend the analysis to stochastic mechanisms.

Even though related, WET and NOR are independent properties in that neither implies the other. Example 1's evidence structure satisfies NOR, because $e_i \in \mathcal{E}(\theta_i)$ is maximal

[16] See also the "full reports condition" by Lipman and Seppi (1995) and the "minimal closure condition" by Forges and Koessler (2005).

with respect to θ_i, for each $i \in \{1,2\}$, while it violates WET, because e_0 is not maximal evidence. To see that WET does not imply NOR, consider the evidence structure for three types $\Theta = \{\theta_1, \theta_2, \theta_3\}$ given by $\mathcal{E}(\theta_1) = \{e_1\}$, $\mathcal{E}(\theta_2) = \{e_2\}$, and $\mathcal{E}(\theta_3) = \{e_1, e_2\}$: Each e_i is maximal evidence because it is maximal with respect to type θ_i, $i \in \{1,2\}$, satisfying WET, but type θ_3 does not have evidence that is maximal with respect to himself, because each of his evidence is available to some other type while he also has strictly more evidence than any other type, violating NOR.

At the same time, WET and NOR are not mutually exclusive. For instance, the nested-range condition (NRC, see Green and Laffont (1986)) implies both WET and NOR, whereby they both weaken NRC, but in distinct ways.[17] The evidence structure introduced by Dye (1985), which has proven highly useful in applications (e.g., see Shin, 2003; Acharya et al., 2011; Ben-Porath et al., 2018; Shishkin, 2023, among many others) satisfies both WET and NOR. Hence, our results regarding the value of control offer an additional perspective on why such evidence structures afford tractability; specifically, optimal mechanisms do not depend on the controllability of evidence presentation.[18]

3.3.2 Type-Independent Worst Options (TIWO)

Example 1 shows that there is value to contractually controlling the presentation of evidence by ruling out what is otherwise costly evidence. WET identifies costly evidence as non-maximal evidence and removes the value of control by directly requiring that no such evidence exists. If the set of allocations includes one that is sure to be the least preferred by the agent, however, there is also an indirect way for the principal to control evidence, even if, formally, evidence is non-controllable. We explore this idea next.

Definition 4. A preference structure (X, u) satisfies the **type-independent worst options condition (TIWO)**, if there exists an allocation $x_w \in X$ such that

$$u(x, \theta) \geq u(x_w, \theta), \ \forall (x, \theta) \in X \times \Theta.$$

Intuitively, if the principal controls an allocation that is the least preferred for any type,

[17]More precisely, jointly requiring "WET+NOR" is a weaker condition than NRC. We prove this claim in Appendix C, where we also restate NRC using our Definition 1, which allows us to shed new light on the results of Green and Laffont, complementing related treatments in Bull and Watson (2007) and Ball and Kattwinkel (2023).

[18]For works introducing "Dye evidence" see also Farrell (1985) and Jung and Kwon (1988). It takes the following form: The set of types is $\Theta = \{\theta_{i,0}, \theta_{i,1}\}_{i=1}^n$ for some $n \geq 1$, where $\theta_{i,0}$ and $\theta_{i,1}$ have the same private information "θ_i", except that $\theta_{i,0}$ has no evidence whereas $\theta_{i,1}$ could perfectly prove it, i.e., $\mathcal{E}(\theta_{i,0}) = \{e_0\}$ and $\mathcal{E}(\theta_{i,1}) = \{e_0, e_i\}$, $i \in \{1, \ldots, n\}$. This satisfies both WET and NOR, because each e_i is maximal with respect to type $\theta_{i,1}$ whereas e_0 is maximal with respect to all the other types $\{\theta_{i,0}\}_{i=1}^n$. In recent work, Ben-Porath et al. (2019) demonstrate a close connection between Dye evidence and costly verification, and Asseyer and Weksler (2022) show how Dye evidence also arises endogenously when quality certification is offered by a profit-maximizing third party.

then the principal can simply deter the agent from presenting any evidence she wants to deter, by imposing this worst allocation in the event that the agent were to present it. Example 1 serves to illustrate this point. In this example, each allocation is preferred over the other for some type, whereby it violates TIWO. Suppose now that the principal additionally had a third allocation $x_3 \in X$ at her discretion, and $u(x_3, \theta_i) = -1$ for both $i = \{1, 2\}$, so that TIWO is satisfied. Then, the social choice function f_X considered there would be implementable also with non-controllable evidence; simply setting $\gamma_X^N(e_0|\theta_i) = x_3$ for both $i = \{1, 2\}$ deters evidence e_0, and together with $\gamma_X^N(e_i|\theta_i) = x_{3-i}$ and $\gamma_E^N(\theta_i) = e_i$ defines a mechanism in Γ^N that satisfies IC^N and implements f_X. The following proposition generalizes this observation.

Proposition 5. *Fix the type structure (Θ, p) and preference structure (X, u) of any evidentiary implementation model. Then, the preference structure (X, u) satisfies TIWO if and only if control has no value for any evidence structure $(E, \mathcal{E})$.*

The proposition establishes that, analogous to WET, TIWO is not only sufficient but also necessary for control to have no value, when we require this to hold for any evidence structure. We prove this result in Appendix A.2. Sufficiency is straightforward based on the deterrence argument above. We prove necessity by showing that, given a failure of TIWO, there is an evidence structure such that each type only has evidence required for his least preferred allocations, plus some null evidence e_0 that every type has (which is non-maximal evidence because every type has a least preferred allocation and therefore some other evidence as well, violating WET). Thus, a social choice function such that every type of agent receives his worst allocation is implementable with controllable evidence, whereas this is impossible with non-controllable evidence, as no matter which allocation follows evidence e_0, there is always a type for whom it is not least preferred. (This construction again generalizes Example 1.)

Remarks: Deneckere and Severinov (2008) introduce TIWO in their analysis of evidence structures violating NOR. Their results suggest a role similar to NOR, though they are concerned with non-controllable evidence only and consider the question of when mechanisms with a single stage of evidence presentation are sufficient. We show here that, with regards to the value of control for d-implementability, NOR and TIWO have different implications – TIWO guarantees that there is no such value, while NOR does not.

TIWO is trivially satisfied in settings where the agent's (ordinal) preferences are type-independent, hence known to the principal. This is a defining feature of the "persuasion" settings analyzed by Glazer and Rubinstein (2004, 2006); Sher (2011, 2014), as well as the general class of "evidence games" analyzed by Hart et al. (2017). While this literature studies non-controllable evidence, our result that TIWO is sufficient for control not to have value shows that one can significantly simplify the mechanism design part of such analyses, in that it suffices to study the structurally simpler mechanisms γ^C with controllable evidence (see also Ben-Porath et al., 2019). This literature permits randomization by the receiver, however, which we cover—for the case of commitment—in our extension to stochastic mechanisms below.

Looking beyond mechanism design with evidence, Proposition 5 explains why, in the agency and contracting literature, "verifiable moral hazard" is considered an oxymoron, and the controllability of verifiable actions has not received serious attention. Agency theory examines the extent to which monetary transfers can provide incentives, and these monetary transfers naturally lead to type-independent worst options; specifically, the principal can always deter any undesirable verifiable action simply by requiring a sufficiently large payment from the agent in the event the agent were to take this action.[19]

Indeed, the result links contractual control with (off-path) deterrence, similar to the shadow of the law perspective in contract theory, where contractual control obtains because large penalties by the court system deter parties from not abiding by the contract. In the context of evidence, controllability means that if the agent presents any evidence other than that stipulated in the contract, he is punished by the court system. A type-independent worst option allows the principal to achieve the same deterrence herself also when evidence is non-controllable.

4 Stochastic Mechanism Design with Evidence

The generalized revelation principle of Myerson (1982) explicitly allows for stochastic mechanisms and stochastic implementation. Our evidence-extended utility and social choice functions render this principle directly applicable to evidentiary implementation problems (see Appendix B, in particular Theorem A). More specifically, and in line with our analysis of canonical mechanisms under d-implementability, this principle yields canonical mechanisms of the form $\tilde{\gamma}^C : \Theta \to \Delta(X \times E)$ subject to honest reporting for controllable evidence, whereas for non-controllable evidence, they take the form $\tilde{\gamma}^N : \Theta \to \Delta(X^E \times E)$ subject to honest reporting and obedient evidence presentation.[20]

[19] For analyses of mechanism design with evidence as well as transfers, see, e.g., Singh and Wittman (2001), Sher and Vohra (2015), Koessler and Perez-Richet (2019), Ali, Lewis, and Vasserman (2023), Pram (2023), Lang (2020), or Dasgupta, Krasikov, and Lamba (2022).

[20] As is standard, we denote by $\Delta(S)$ the space of all probability distributions over the (finite) set S.

4.1 The Value of Randomization

Maintaining our focus on implementing *deterministic* social choice functions $f_X : \Theta \to X$, the question that we explore in this section is the extent to which stochastic mechanisms are helpful and even indispensable for implementing such social choice functions. In particular, we address whether and when there is a *value of randomization* (allowing to implement deterministic social choice functions that are not d-implementable), how this depends on the controllability of evidence presentation, and how results on the value of control that we have established for d-implementability extend when allowing for stochastic mechanisms.[21]

While the revelation principle does not generally guarantee that the restriction to d-implementability is without loss, deterministic mechanisms often turn out optimal in popular applications of mechanism design, including settings with evidence.[22] With regard to the implementability of *deterministic* social choice functions, this may suggest that randomization has no value, so considering only deterministic mechanisms is without loss. The following two examples show that such a view would be incorrect, both with controllable evidence and with non-controllable evidence. Moreover, they also indicate how the role of randomization differs depending on whether evidence is controllable or not.

Example 2. (Controllable evidence) The principal has to decide between two allocations $X = \{x_b, x_g\}$, and the agent has three possible types $\Theta = \{\theta_1, \theta_2, \theta_3\}$. The agent's preferences over allocations are type-independent, such that $u(x_g|\theta) = 1$ and $u(x_b|\theta) = 0$ for all types θ. The evidence structure consists of three possible elements $E = \{e_0, e_1, e_2\}$ distributed over types according to

$$\mathcal{E}(\theta_1) = \{e_0, e_1\},\ \mathcal{E}(\theta_2) = \{e_0, e_2\},\ \text{and } \mathcal{E}(\theta_3) = \{e_0, e_1, e_2\}.$$

Note that $\mathcal{E}(\theta_1) \cup \mathcal{E}(\theta_2) = \mathcal{E}(\theta_3) = E$, so any evidence that type θ_3 might present is evidence that also another type might present.

Consider now the social choice function $f_X : \Theta \to X$ with $f_X(\theta_1) = f_X(\theta_2) = x_b$ and $f_X(\theta_3) = x_g$. The first claim is that f_X is not d-implementable. Following Theorem 1, this would require that it is implemented by an incentive compatible direct mechanism $\gamma^C : \Theta \to X \times E$, where, in particular, $\gamma^C(\theta_3) = (x_g, e_i)$ for some $e_i \in \mathcal{E}(\theta_3) = E$. However, any evidence $e_i \in E$ can be presented also by θ_1 or θ_2, whereby regardless of which

[21] Since deterministic mechanisms are contained in the class of stochastic mechanisms, the analogue of Proposition 2—which shows that control could only ever allow the principal to implement more—is immediate for the value of randomization: d-implementability implies implementability.

[22] For instance, Strausz (2003) shows that the revelation principle fails when restricting attention to deterministic direct mechanisms even with a single agent, while Strausz (2006) shows that in the "regular" versions of the classic monopolistic screening problem the optimal direct mechanism is deterministic. For mechanism design settings with evidence and verifiability, where randomization has no value, see Glazer and Rubinstein (2004, 2006), and Ben-Porath et al. (2019).

evidence we specify, the mechanism cannot be incentive compatible with the requirement that *both* of these other types shall obtain the worse allocation, i.e., $f_X(\theta_1) = f_X(\theta_2) = x_b$: $\gamma^C(\theta_3) = (x_g, e_1)$ would enable type θ_1 to obtain x_g, whereas $\gamma^C(\theta_3) = (x_g, e_2)$ would enable type θ_2 to obtain x_g, and $\gamma^C(\theta_3) = (x_g, e_0)$ would enable both types to obtain x_g.

The second claim is that f_X is implementable, once we allow for randomization. Consider the stochastic "menu" mechanism giving the agent the choice of either the contract (x_b, e_0), or the stochastic contract of a 50:50 randomization between (x_g, e_1) and (x_g, e_2). Choosing the latter means that the principal must provide allocation x_g, while the agent must present either evidence e_1 or evidence e_2, depending on the result of a coin flip.[23] While type θ_3 is then certain to be able to present the evidence required, each of the other two types would face a positive probability of violating the contract due to lacking the required evidence. Given arbitrarily large costs of doing so, this menu implements f_X. □

Example 2 shows that implementation of a deterministic social choice function with controllable evidence may require randomization. The randomization pertains here only to the evidence that the agent must present. As we show below, this is a general feature for the implementability of deterministic social choice functions with controllable evidence: It then suffices to consider stochastic direct mechanisms that randomize only the evidence component, i.e., mechanisms of the form $\tilde{\gamma}^C : \Theta \to X \times \Delta(E)$.

The next example illustrates that, with non-controllable evidence, it is not sufficient to consider stochastic mechanisms that randomize only in the evidence component, however. In particular, the example shows that the implementability of a deterministic social choice function may then require randomization also with respect to the allocation rule.

Example 3. (Non-controllable evidence) The principal has to decide between three allocations $X = \{x_w, x_b, x_g\}$, and the agent has three possible types $\Theta = \{\theta_1, \theta_2, \theta_3\}$. The agent's preferences over allocations are type-independent, such that

$$u(x_w|\theta) = 0, \; u(x_b|\theta) = 3, \; \text{and } u(x_g|\theta) = 4,$$

for any type θ. The evidence structure is the same as in Example 2.

Consider now the social choice function $f_X : \Theta \to X$ with $f_X(\theta_1) = f_X(\theta_2) = x_b$ and $f_X(\theta_3) = x_g$. As in the previous example, f_X is not d-implementable, because any deterministic mechanism would allow one (or both) of the types $\{\theta_1, \theta_2\}$ to mimic type θ_3 and thus obtain the preferred allocation x_g that f_X would reserve for this type.

[23] This menu corresponds to a 50:50 randomization over the two deterministic direct mechanisms (f_X, f_E^1) and (f_X, f_E^2), where f_E^i has $f_E^i(\theta_1) = f_E^i(\theta_2) = e_0$ and $f_E^i(\theta_3) = e_i$, for $i \in \{1, 2\}$. Choosing contract (x_b, e_0) in the menu mechanism thus corresponds to reporting θ_1 or θ_2 in the stochastic direct mechanism, while choosing the stochastic contract, which we may write as $(x_g, (\frac{1}{2} : e_1, \frac{1}{2} : e_2))$, corresponds to reporting θ_3.

In particular, fix any $k \in \{1, 2\}$, and take the deterministic direct mechanism $\gamma^{N,k} = (\gamma_X^{N,k}, \gamma_E^{N,k}) \in \Gamma^N$ with

$$\gamma_X^{N,k}(e|\theta_1) = \gamma_X^{N,k}(e|\theta_2) = \begin{cases} x_b, & \text{if } e = e_0, \\ x_w, & \text{if } e \neq e_0; \end{cases}, \text{ and } \gamma_X^{N,k}(e|\theta_3) = \begin{cases} x_g, & \text{if } e = e_k, \\ x_w, & \text{if } e \neq e_k. \end{cases}$$
$$\gamma_E^{N,k}(\theta_1) = \gamma_E^{N,k}(\theta_2) = e_0, \text{ and } \gamma_E^{N,k}(\theta_3) = e_k.$$

Note that the mechanism depends on k only in terms of which evidence is recommended as well as rewarded with x_g after type report θ_3. In any case, being deterministic, neither $\gamma^{N,1}$ nor $\gamma^{N,2}$ is incentive compatible. Indeed, given mechanism $\gamma^{N,k}$, type θ_k would mimic type θ_3, for any $k \in \{1, 2\}$.

However, the stochastic direct mechanism that randomizes over $\gamma^{N,1}$ and $\gamma^{N,2}$ with equal probability implements f_X, satisfying honesty and obedience also for both types θ_1 and θ_2. To see this, note that disobeying a recommendation results in the worst allocation x_w. If they are honest, they face no randomness: They are sure to receive recommendation e_0, which they will obey, to obtain the better allocation x_b rather than x_w, yielding them utility 3. If they dishonestly report to be of type θ_3, however, they are subject to randomness: Either they will receive a recommendation they can and will obey (for θ_i, $i \in \{1, 2\}$, this would be e_i), to obtain the best allocation x_g, yielding utility 4; or they will receive a recommendation they cannot possibly obey (for θ_i, $i \in \{1, 2\}$, this would be e_{3-i}), resulting in the worst allocation x_w, yielding utility 0. Each of these two outcomes has probability one half so that dishonestly reporting type θ_3 yields either type an expected utility of 2. Hence, any type of agent maximizes his expected utility by being honest and obedient here. The stochastic direct mechanism is incentive compatible, and it implements the deterministic social choice function f_X, when it is not d-implementable.

Observe that both $\gamma_E^{N,1} \neq \gamma_E^{N,2}$ and $\gamma_X^{N,1} \neq \gamma_X^{N,2}$. Hence, the randomization is with respect to both the recommendation and the allocation rule, and it perfectly correlates the two. Indeed, if the evidence-contingent allocation rule were independent of the recommendation, the latter would be irrelevant, and even if the allocation rule were stochastic in a way that depended on the agent's type report, f_X could not be implemented here, because the agent's preferences are type-independent and θ_3 would be mimicked by another type. □

Deneckere and Severinov (2008, Example 5) present an example similar to Example 3. Our explicit formulation uncovers its special property that, in order to preserve its deterministic allocation, it requires a perfectly correlated randomization between the evidence that certain types are to provide, as represented by γ_E^N, and the way in which evidence translates into an allocation, as represented by γ_X^N. This allows us to connect this feature of (non-controllable) evidence to the contracting literature that studies moral hazard, where this type of correlated randomization has also been identified as beneficial. In particular, Rahman and Obara (2010) show the power of such correlation in the context of moral hazard in teams, referring to such stochastic contracts as "mediated contracts." Strausz (2012) shows that these mediated contracts indirectly implement the required correlated randomization of direct mechanisms in the framework of Myerson (1982). As we show below, Example 3 expresses the fact that the obedience constraint under non-controllable evidence does not allow us to further simplify the structure of canonical stochastic mechanisms, even when restricting to deterministic social choice functions.

Finally, observe that the evidence structure common to both examples violates NOR. Indeed, as we show below, NOR is both closely and delicately related to the value of randomization. While it guarantees that deterministic *recommendations* are without loss, it does not generally guarantee the optimality of deterministic mechanisms, but does so only for controllable evidence. Moreover, NOR is also a sufficient condition for extending our result on d-implementability, that with WET or TIWO control has no value, to general implementability.

4.2 Additional Notation and Definitions

The following natural notation will be used. Let S_0, S_1 and S_2 be any finite sets. First, whenever we have a mapping $\sigma : S_0 \to \Delta(S_1)$, we write $\sigma(s_1 | s_0)$ for the probability of $s_1 \in S_1$ conditional on $s_0 \in S_0$. Second, whenever we have three mappings $\sigma : S_0 \to \Delta(S_1 \times S_2)$, $\sigma_1 : S_0 \to S_1$ and $\sigma_2 : S_0 \to \Delta(S_2)$ such that

$$\sigma(s_1, s_2 | s_0) = \mathbb{I}(s_1 = \sigma_1(s_0)) \cdot \sigma_2(s_2 | s_0)$$

for any $(s_0, s_1, s_2) \in S_0 \times S_1 \times S_2$, we write σ also as $\sigma = (\sigma_1, \sigma_2) : S_0 \to S_1 \times \Delta(S_2)$.[24]

In line with this notation, we define a *stochastic evidence extension* as any mapping $\tilde{f}_E : \Theta \to \Delta(E)$ such that, for all $(\theta, e) \in \Theta \times E$, $\tilde{f}_E(e|\theta) > 0$ implies $e \in \mathcal{E}(\theta)$, i.e., as any randomization over evidence extensions. A natural special case, which will be useful in our analysis of controllable evidence, is what we call *uniform evidence extension* and denote by $\tilde{f}_E^U$; it is defined for all $(e, \theta) \in E \times \Theta$ as

$$\tilde{f}_E^U(e|\theta) \equiv \begin{cases} \frac{1}{|\mathcal{E}(\theta)|}, & \text{if } e \in \mathcal{E}(\theta), \\ 0, & \text{if } e \notin \mathcal{E}(\theta), \end{cases}$$

where $|\mathcal{E}(\theta)|$ is the cardinality of the (finite) set $\mathcal{E}(\theta)$. It will also be useful to define a *maximal evidence extension* f_E^M as any (deterministic) evidence extension f_E such that

[24] $\mathbb{I}(a = b)$ denotes the indicator function, which is equal to 1 if $a = b$ holds true, and 0 otherwise.

$f_E(\theta)$ is maximal with respect to type θ, for all $\theta \in \Theta$. Recalling that, by definition, any evidence extension satisfies $f_E(\theta) \in \mathcal{E}(\theta)$ for all $\theta \in \Theta$, maximal evidence extensions exist if and only if the evidence structure satisfies NOR, and under this assumption, they will be useful in our analysis of both controllable and non-controllable evidence.

We call any pair $\tilde{f} = (f_X, \tilde{f}_E)$, consisting of a social choice function f_X and a stochastic evidence extension $\tilde{f}_E$, a *(stochastic-evidence-) extended social choice function*, and we denote the space of all such extended social choice functions by $\tilde{\mathcal{F}}$. Extending our earlier definitions of implementability by explicitly allowing stochastic evidence extensions, we say that a mechanism implements the extended social choice function $\tilde{f} = (f_X, \tilde{f}_E) \in \tilde{\mathcal{F}}$, if it implements the social choice function f_X in a BNE whose outcome is such that an agent of type θ presents evidence $e \in \mathcal{E}(\theta)$ with probability $\tilde{f}_E(e|\theta)$. Finally, we say that an extended social choice function $\tilde{f} \in \tilde{\mathcal{F}}$ is implementable, if there exists a mechanism that implements it.

4.3 Controllable Evidence

In the case of controllable evidence, stochastic *direct mechanisms* take the form $\tilde{\gamma}^C : \Theta \to \Delta(X \times E)$. The agent reports a type θ, and, conditional on this report, the mechanism randomly chooses pair (x, e) with probability $\tilde{\gamma}^C(x, e|\theta)$; the realization (x, e) means the principal must provide allocation x and the agent must present evidence e. We denote the set of all such mechanisms by $\tilde{\Gamma}^C$. Observe that $\tilde{\mathcal{F}} \subseteq \tilde{\Gamma}^C$.

A mechanism $\tilde{\gamma}^C \in \tilde{\Gamma}^C$ is *incentive compatible* if it provides any type of agent with an incentive to report honestly, i.e.,

$$\tilde{IC}^C : \sum_{(x,e)\in X\times E} \tilde{\gamma}^C(x, e\,|\theta) \cdot U(x, e\,|\theta) \geq \sum_{(x,e)\in X\times E} \tilde{\gamma}^C(x, e\,|\theta') \cdot U(x, e\,|\theta)\,,\ \forall (\theta, \theta') \in \Theta \times \Theta.$$

The following characterization of implementability shows that, with controllable evidence, the need for randomization in direct mechanisms concerns only the evidence that the agent must present and can without loss be taken to be uniform.

Theorem 3. *Suppose evidence is controllable. Then, a social choice function f_X is implementable if and only if the stochastic direct mechanism given by $(f_X, \tilde{f}_E^U) : \Theta \to X \times \Delta(E)$ satisfies $\tilde{IC}^C$.*

Proof. Only necessity requires proof, so suppose the stochastic direct mechanism $\tilde{\gamma}^C : \Theta \to \Delta(X \times E)$ satisfies $\tilde{IC}^C$ and implements f_X. Clearly, the following must then hold true:

$$\mathcal{E}(\theta') \subseteq \mathcal{E}(\theta) \Rightarrow u(f_X(\theta)|\theta) \geq u(f_X(\theta')|\theta). \tag{4}$$

Consider now the stochastic direct mechanism $(f_X, \tilde{f}_E^U)$. Since (4) is a property of f_X only, this mechanism satisfies it. For it to satisfy $\tilde{IC}^C$, it is then sufficient that, in addition to (4), the mechanism $(f_X, \tilde{f}_E^U)$ also satisfies that, for all $(\theta, \theta') \in \Theta \times \Theta$ such that $\mathcal{E}(\theta') \not\subseteq \mathcal{E}(\theta)$,

$$u(f_X(\theta)|\theta) \geq \left(1 - \frac{1}{|E|}\right) \cdot u(f_X(\theta')|\theta) + \frac{1}{|E|} \cdot (-c). \tag{5}$$

While, for c large enough, it is clear that $(f_X, \tilde{f}_E^U)$ implies (5) whenever $\mathcal{E}(\theta') \not\subseteq \mathcal{E}(\theta)$, we next derive an explicit lower bound on c such that this is guaranteed.

Letting $\overline{u} \equiv \max\{u(x|\theta) : (x, \theta) \in X \times \Theta\}$ and $\underline{u} \equiv \min\{u(x|\theta) : (x, \theta) \in X \times \Theta\}$, it suffices for (5) to hold that

$$\underline{u} > \left(1 - \frac{1}{|E|}\right) \cdot \overline{u} + \frac{1}{|E|} \cdot (-c),$$

which is equivalent to

$$c > \underline{c} \equiv (|E| - 1) \cdot (\overline{u} - \underline{u}) - \underline{u}.$$

Hence, for any c exceeding the lower bound $\underline{c}$, the stochastic direct mechanism $(f_X, \tilde{f}_E^U)$ satisfies $\tilde{IC}^C$.[25] □

The intuition behind the theorem is straightforward. Stochastic direct mechanisms that put positive probability on every evidence that the agent claims to possess are extremely powerful; given sufficiently high costs to violating the contract ex post, they effectively create evidence corresponding to the entire evidence set $\mathcal{E}(\theta)$ for every type of agent θ. As a consequence, incentive compatibility reduces to (4), which is as weak as possible, since type θ can perfectly mimic any type θ' such that $\mathcal{E}(\theta') \subseteq \mathcal{E}(\theta)$ under any conceivable mechanism. Thus, (the proof of) Theorem 3 delivers a simple characterization result of implementability, which we state explicitly as a corollary.

Corollary 2. *Suppose evidence is controllable. Then, a social choice function f_X is implementable if and only if f_X satisfies (4).*

In particular, this dispenses with any need for correlation between evidence and allocation.[26] The specification in Theorem 3, which pins down the direct mechanisms by extending f_X with the uniform evidence extension $\tilde{f}_E^U$, is used only for convenience, though it also appears natural, a priori.

Since, in terms of the model's primitives, evidence does not directly affect the agent's utility, it is natural to consider economic allocations only at the level of (non-extended) social

[25] This derivation underlies the explicit lower bound on c stated in footnote 9.

[26] Indeed, even if we were interested in implementing a stochastic social choice function $\tilde{f}_X : \Theta \to \Delta(X)$, it would be sufficient to consider incentive compatibility of the stochastic direct mechanism $(\tilde{f}_X, \tilde{f}_E^U) : \Theta \to \Delta(X) \times \Delta(E)$.

choice functions f_X, and the existing literature on evidence has done so exclusively. Theorem 3 shows that any randomization required for implementing such social choice functions could concern only the evidence that must be presented by the agent, and Example 2 shows that such randomization has value, meaning some f_X is implementable but not d-implementable. However, as is straightforward from Theorem 3, if we restricted attention to deterministic *extended* social choice functions $f = (f_X, f_E) \in \mathcal{F}$, implementability of f would be equivalent to d-implementability of f—i.e., randomization would have no value—with controllable evidence. As we show below, this is different with non-controllable evidence.

Corollary 3. *Suppose evidence is controllable. Then, a deterministic evidence-extended social choice function f is implementable if and only if f is d-implementable.*

Proof. Only necessity requires proof. Simply observe that any stochastic direct mechanism $(f_X, \tilde{f}_E) : \Theta \to X \times \Delta(E)$ satisfying $\tilde{IC}^C$ implements the stochastic-evidence-extended social choice function $(f_X, \tilde{f}_E)$. □

This raises the question of whether there are natural conditions under which randomization concerning evidence can be dispensed with also for implementing (non-extended) social choice functions f_X. We next show that NOR is indeed such a condition, and that it is also necessary if we require that randomization have no value for *any* preference structure (similar to how WET and TIWO are necessary for there to be no value of control with respect to d-implementability). As a first step towards this result, we establish the following lemma.

Lemma 1. *Suppose evidence is controllable and the evidence structure satisfies NOR. Then, a social choice function f_X is implementable if and only if, for any maximal evidence extension f_E^M, the deterministic direct mechanism $(f_X, f_E^M) \in \Gamma^C$ satisfies IC^C.*

Proof. Only necessity requires proof. From Theorem 3, f_X is implementable if and only if it satisfies (4). From Theorem 1, f_X is d-implementable if and only if it satisfies (2). Suppose now that the evidence structure $(E, \mathcal{E})$ satisfies NOR, and consider any maximal evidence extension f_E^M, which exists by NOR. Now observe that, for any pair of types $(\theta, \theta') \in \Theta \times \Theta$, $\mathcal{E}(\theta') \subseteq \mathcal{E}(\theta)$ implies $f_E^M(\theta') \in \mathcal{E}(\theta)$, whereby (4) implies (2). □

Proposition 6. *Suppose evidence is controllable. Fix the type structure (Θ, p) and evidence structure $(E, \mathcal{E})$ of any evidentiary implementation model. Then, the evidence structure $(E, \mathcal{E})$ satisfies NOR if and only if randomization has no value for* any *preference structure (X, u).*

Proof. Sufficiency is immediate from Lemma 1. For necessity, suppose the evidence structure $(E, \mathcal{E})$ violates NOR. Let then $\hat{\theta}$ be a type such that, for every $e \in \mathcal{E}(\hat{\theta})$, there exists another type θ' such that $e \in \mathcal{E}(\theta')$ and $\mathcal{E}(\hat{\theta}) \nsubseteq \mathcal{E}(\theta')$. The violation of NOR implies existence of $\hat{\theta}$.

Consider then the preference structure (X, u) given by $X = \{x_1, x_2\}$ and $u : X \times \Theta \to \mathbb{R}$ such that $u(x_1|\theta) > u(x_2|\theta)$ for all types θ, together with the social choice function f_X such that

$$f_X(\theta) = \begin{cases} x_1, & \text{if } \mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta), \\ x_2, & \text{if } \mathcal{E}(\hat{\theta}) \nsubseteq \mathcal{E}(\theta). \end{cases}$$

Recalling the characterization of implementability by (4), f_X is clearly implementable. Recalling the characterization of d-implementability by (2), take any (deterministic) evidence extension f_E and observe that, by construction via $\hat{\theta}$, there exists a type θ' such that $f_E(\hat{\theta}) \in \mathcal{E}(\theta')$ and $\mathcal{E}(\hat{\theta}) \nsubseteq \mathcal{E}(\theta')$. This implies $u(f_X(\theta')|\theta') = u(x_2|\theta') < u(x_1|\theta') = u(f_X(\hat{\theta})|\theta')$ while $f_E(\hat{\theta}) \in \mathcal{E}(\theta')$, violating (2), whereby f_X is not d-implementable. □

NOR means that every type of agent θ can actually prove his full evidence set $\mathcal{E}(\theta)$ by presenting some evidence $e \in \mathcal{E}(\theta)$ that is maximal with respect to θ. Recalling that the value of randomization with controllable evidence arises from allowing to effectively create such evidence, it becomes clear why this value vanishes under NOR: This property renders redundant the randomization that requires types to present any of their evidence with a strictly positive probability. Thus, implementability of a social choice function f_X then implies its d-implementability.

Lemma 1 additionally establishes that, under NOR, the characterization of implementability reduces substantially: To determine whether a social choice function f_X is implementable it is sufficient to check IC^C only for a single evidence extension, where we may pick any maximal one f_E^M. Hence, under NOR and controllable evidence, the revelation principle does not only state which type the agent is to reveal—his actual type—but also which evidence he is to present—his maximal evidence.

4.4 Non-controllable Evidence

In the case of non-controllable evidence, stochastic *direct mechanisms* take the form $\tilde{\gamma}^N : \Theta \to \Delta(X^E \times E)$, and we denote the space of such mechanisms by $\tilde{\Gamma}^N$. The agent reports a type $\hat{\theta}$, and, conditional on this report, the mechanism randomly chooses pair $(d_0, e) \in X^E \times E$ with probability $\tilde{\gamma}^N(d_0, e|\hat{\theta})$, informing the agent of e (the realized recommendation) but not of d_0 (the realized allocation rule), upon which the agent may present any evidence $\hat{e} \in E$, and the principal then must provide allocation $d_0(\hat{e})$, as under the realized allocation rule d_0.[27,28]

[27]Following Myerson (1982), when the agent's presentation of evidence is a non-controllable action by the agent, the principal's decision is a mapping from evidence presented by the agent (recall its verifiability) into allocations provided by herself, and we adopt Myerson's notation d_0 for a particular such decision; see also Appendix B.

[28]The agent should not observe the *realized* d_0 *before* choosing his evidence $\hat{e}$, because revealing d_0 only makes incentive compatibility more difficult to meet. By contrast, the mechanism might as well publicly

The agent's strategy in such a mechanism specifies, for any type θ, a pair $(\hat{\theta}, \hat{\delta}) \in \Theta \times E^E$, which consists of a type report $\hat{\theta} \in \Theta$ and rule $\hat{\delta} \in E^E$ for presenting evidence as a function of the evidence recommended (i.e., $\hat{\delta} : E \to E$). The strategy is honest and obedient for type θ, if it specifies $(\hat{\theta}, \hat{\delta}) \in \Theta \times E^E$ such that $\hat{\theta} = \theta$ and $\hat{\delta}(e) = e$ for all $e \in E$ with $\sum_{d_0 \in X^E} \tilde{\gamma}^N(d_0, e|\theta) > 0$. A mechanism $\tilde{\gamma}^N \in \tilde{\Gamma}^N$ is *incentive compatible* if it provides any type of agent with an incentive to be honest and obedient, i.e.,

$$\tilde{IC}^N : \sum_{(d_0,e) \in X^E \times E} \tilde{\gamma}^N(d_0, e|\theta) \cdot U(d_0(e), e|\theta) \geq \sum_{(d_0,e) \in X^E \times E} \tilde{\gamma}^N(d_0, e|\hat{\theta}) \cdot U(d_0(\hat{\delta}(e)), \hat{\delta}(e)|\theta),$$
$$\forall (\theta, \hat{\theta}, \hat{\delta}) \in \Theta \times \Theta \times E^E.$$

Note that for checking $\tilde{IC}^N$, we may ignore any strategy that, for some type θ and some recommendation e that occurs with positive probability $\sum_{d_0 \in X^E} \tilde{\gamma}^N(d_0, e|\hat{\theta}) > 0$, specifies $(\hat{\theta}, \hat{\delta}) \in \Theta \times E^E$ such that $\hat{\delta}(e) \notin \mathcal{E}(\theta)$: Replacing $\hat{\delta}(e)$ by any $\hat{e} \in \mathcal{E}(\theta)$ yields a greater expected utility.[29] Hence, incentive compatibility requires that the mechanism only recommends evidence that would have to be available to the agent *if he were honest*, i.e., that $\sum_{d_0 \in X^E} \tilde{\gamma}^N(d_0, e|\theta) > 0$ only if $e \in \mathcal{E}(\theta)$, for any type $\theta \in \Theta$.

Example 3 already shows that, in contrast to the case of controllable evidence, the determinism of social choice functions affords no further structural simplification of canonical mechanisms within $\tilde{\Gamma}^N$ for characterizing implementability with non-controllable evidence, in general. A social choice function f_X is implementable with non-controllable evidence if and only if there exists an incentive compatible stochastic direct mechanism $\tilde{\gamma}^N \in \tilde{\Gamma}^N$ that implements it.

Theorem 4. *Suppose evidence is non-controllable. Then, a social choice function f_X is implementable if and only if there exists a stochastic direct mechanism $\tilde{\gamma}^N \in \tilde{\Gamma}^N$ such that (i) $\tilde{\gamma}^N(d_0, e|\theta) > 0$ implies $d_0(e) = f_X(\theta)$ for all $(\theta, e) \in \Theta \times E$, and (ii) $\tilde{\gamma}^N$ satisfies $\tilde{IC}^N$.*

In light of Proposition 6, a natural conjecture might be that NOR would remove any potential need for randomization, as it does with controllable evidence. Intuitively, unlike in Example 3, which violates NOR, it should then be clear what evidence to recommend to the agent, namely evidence that is maximal with respect to the type he reports. Given a deterministic allocation objective, randomization would therefore seem pointless under NOR also with non-controllable evidence. However, the following example shows that this is not true. Indeed, despite NOR, randomization here has value even for the implementability of

reveal the realized d_0 *after* the agent has presented his evidence $\hat{e}$.

[29] While we may simply consider the limiting case $c \to \infty$ in using evidence-extended utility U, which would imply that the expected utility of type θ under any such strategy approaches $-\infty$, see footnote 9 for explicit lower bounds on c. Under non-controllable evidence, $c > -\underline{u}$ is actually sufficient even when allowing for stochastic mechanisms.

deterministic *extended* social choice functions $f = (f_X, f_E) \in \mathcal{F}$, in contrast to controllable evidence (Corollary 3).

Example 4. The principal has to decide between three allocations $X = \{x_1, x_2, x_3\}$, and the agent has three possible types $\Theta = \{\theta_1, \theta_2, \theta_3\}$. The agent's type-dependent preferences over allocations are as in the following table:

$u(x\|\theta)$	θ_1	θ_2	θ_3
x_1	2	3	0
x_2	0	2	3
x_3	3	0	2

The evidence structure is given by $E = \{e_0, e_1, e_2, e_3\}$, such that for each $i \in \{1, 2, 3\}$, $\mathcal{E}(\theta_i) = \{e_0, e_i\}$; i.e., the agent can always either prove his type or prove nothing. This evidence structure clearly satisfies NOR.

Consider now the social choice function f_X such that $f_X(\theta_i) = x_i$, for all $i \in \{1, 2, 3\}$. We will now show that f_X is not d-implementable, which is most readily established using Proposition 1. Take then any (deterministic) mechanism $g^N : \Theta \times E \to X$, and consider its allocation in the event that the agent presents the "null-evidence" e_0, where it will suffice to consider the allocation $g^N(\theta_1, e_0)$ for concreteness: Either $g^N(\theta_1, e_0) = x_1$, in which case type θ_2 could obtain allocation x_1, which θ_2 prefers over $f_X(\theta_2) = x_2$; or $g^N(\theta_1, e_0) = x_2$, in which case type θ_3 could obtain allocation x_2, which θ_3 prefers over $f_X(\theta_3) = x_3$; or $g^N(\theta_1, e_0) = x_3$, in which case type θ_1 could obtain allocation x_3, which θ_1 prefers over $f_X(\theta_1) = x_1$. Hence, no such mechanism could implement f_X, and, by Proposition 1, f_X is therefore not d-implementable.

Let then, for all $i, j \in \{1, 2, 3\}$, the evidence-contingent allocation rule $d_{0,i,j} : E \to X$ be given by

$$d_{0,i,j}(e) \equiv \begin{cases} x_i, & \text{if } e = e_i, \\ x_j, & \text{if } e \neq e_i, \end{cases}$$

and consider the stochastic direct mechanism $\tilde{\gamma}^N = (\tilde{\gamma}^N_X, \gamma^N_E) : \Theta \to \Delta(X^E) \times E$ such that, for all $i \in \{1, 2, 3\}$, $\gamma^N_E(\theta_i) \equiv e_i$ and $\tilde{\gamma}^N_X(d_{0,i,j}|\theta_i) \equiv \frac{1}{3}$ for all $j \in \{1, 2, 3\}$. From the agent's perspective, this means that upon report θ_i he is sure to receive recommendation e_i, and then also sure to obtain allocation x_i if he is obedient, while any disobedience results in the stochastic allocation $(\frac{1}{3} : x_1, \frac{1}{3} : x_2, \frac{1}{3} : x_3)$. If this mechanism is incentive compatible—i.e., the agent is always honest and obedient—it implements the (deterministic) *extended* social choice function $f = (f_X, f_E)$ that has $f_E(\theta_i) = e_i$, for all $i \in \{1, 2, 3\}$. Now observe that the agent can only obey the recommendation after an honest type report, because $e_j \in \mathcal{E}(\theta_i)$ if and only if $i = j$, for any pair (i, j) with $i, j \in \{1, 2, 3\}$. Since any disobedience results in the stochastic allocation $(\frac{1}{3}:x_1, \frac{1}{3}:x_2, \frac{1}{3}:x_3)$, with expected utility $\frac{5}{3}$ for any type of agent, while honesty and obedience yields type θ_i the sure allocation x_i, with utility 2, for any $i \in \{1,2,3\}$, the mechanism $\tilde{\gamma}^N$ is indeed incentive compatible. It therefore implements the (deterministic) extended social choice function f, and, in particular, the social choice function f_X is thus implementable even though it is not d-implementable. □

Despite satisfying NOR, Example 4 exhibits a value of randomization. This randomization is only with respect to the allocation rule, however, and not the evidence recommendation. Moreover, the randomization only occurs off-path, following disobedience. Its value arises from the agent's type-dependent preferences, where randomization allows to create an allocation lottery that every type of agent finds worse than his allocation according to the social choice function. Thus, the example establishes that, even under NOR, additional assumptions are generally required for randomization to not have any value; revisiting the value of control with stochastic mechanisms will allow us to derive two general sufficient conditions for this to be the case, see Corollary 5 in the next section.[30]

At the same time, NOR is indeed sufficient for determinism with respect to evidence recommendations; i.e., restricting direct mechanisms to mappings $\tilde{\gamma}^N : \Theta \to \Delta(X^E) \times E$ is then without loss, where we say that *randomization with respect to recommendations has no value*. Moreover, NOR is also necessary for this property to hold for *any* preference structure. To show this, we first establish a useful lemma, analogous to Lemma 1 for controllable evidence. The intuition is similar, so we relegate the proof to Appendix A.3.

Lemma 2. *Suppose evidence is non-controllable and the evidence structure satisfies NOR. Then, a social choice function f_X is implementable if and only if there exists a mapping $\tilde{\gamma}^N_X : \Theta \to \Delta(X^E)$ such that, for every maximal evidence extension f^M_E, (i) $\tilde{\gamma}^N_X(d_0|\theta) > 0$ implies $d_0(f^M_E(\theta)) = f_X(\theta)$ for all $\theta \in \Theta$, and (ii) the stochastic direct mechanism $(\tilde{\gamma}^N_X, f^M_E) \in \tilde{\Gamma}^N$ satisfies $\tilde{IC}^N$.*

Proposition 7. *Suppose evidence is non-controllable. Fix the type structure (Θ, p) and evidence structure $(E, \mathcal{E})$ of any evidentiary implementation model. Then, the evidence structure $(E, \mathcal{E})$ satisfies NOR if and only if randomization with respect to recommendations has no value for* any *preference structure (X, u).*

We relegate the proof of this proposition to Appendix A.4. Lemma 2 establishes sufficiency. Concerning necessity, the construction of a preference structure and social choice

[30] Relatedly, Ben-Porath et al. (2019) obtain this for a class of optimal mechanism design problems by assuming NOR together with "simple type dependence" of agents' preferences. This generalizes the commonly used assumption of type-*independence* of preferences, which would imply TIWO and also yield that randomization has no value (given NOR, Corollary 5). It is easy to check that the preferences in Example 4 indeed violate not only TIWO but also simple type dependence.

function such that deterministic recommendations are with loss whenever NOR is violated generalizes the logic of Example 3.

When recommendations are deterministic, the (generally stochastic) allocation rule is independent from the recommendation, for any type report. Thus, whenever NOR is satisfied, evidence recommendations actually become redundant. This suggests that we can then generalize Proposition 1 for d-implementability to (general) implementability, though subject to the important caveat from Example 4 above that we cannot generally dispense with randomization with respect to allocations.

To do so, consider therefore (indirect) stochastic mechanisms of the form $\tilde{g}^N : \Theta \times E \to \Delta(X)$, which specify stochastic allocations conditional on the agent's type report and evidence (θ, e), where $\tilde{g}^N(x|\theta, e)$ denotes the corresponding probability of any allocation x; $\tilde{G}^N$ will denote the set of all such mechanisms. Again, in the absence of recommendations, obedience loses its literal meaning, while honesty requires that

$$\tilde{IC}_G^N: \ \exists f_E : \Theta \to E \text{ with } f_E(\theta) \in \mathcal{E}(\theta), \ \forall \theta \in \Theta, \text{ such that}$$
$$\sum_{x \in X} \tilde{g}^N\left(x|\theta, f_E(\theta)\right) \cdot U\left(x, f_E(\theta)|\theta\right) \geq \sum_{x \in X} \tilde{g}^N\left(x|\theta', e'\right) \cdot U\left(x, e'|\theta\right), \ \forall \left(\theta, \theta', e'\right) \in \Theta \times \Theta \times E.$$

We are now ready to state the generalization of Proposition 1 for d-implementability to general implementability under NOR.

Proposition 8. *Suppose evidence is non-controllable and the evidence structure satisfies NOR. Then, a social choice function f_X is implementable if and only if there exists a mechanism $\tilde{g}^N : \Theta \times E \to \Delta(X)$ such that, for every maximal evidence extension f_E^M, (i) $\tilde{g}^N(x|\theta, f_E^M(\theta)) = \mathbb{I}(x = f_X(\theta))$, for all $(\theta, x) \in \Theta \times X$, and (ii) $\tilde{g}^N$ satisfies $\tilde{IC}_G^N$ for f_E^M.*

Proof. Only necessity requires proof, so suppose the social choice function f_X is implementable. In view of Lemma 2, let mapping $\tilde{\gamma}_X^N : \Theta \to \Delta(X^E)$ be such that, for every maximal evidence extension f_E^M, (i) $\tilde{\gamma}_X^N(d_0|\theta) > 0$ implies $d_0(f_E^M(\theta)) = f_X(\theta)$ for all $\theta \in \Theta$, and (ii) the stochastic direct mechanism $(\tilde{\gamma}_X^N, f_E^M) \in \Gamma^N$ satisfies $\tilde{IC}^N$.

Define then the stochastic indirect mechanism $\tilde{g}^N : \Theta \times E \to \Delta(X)$ such that, for any $(\theta, e, x) \in \Theta \times E \times X$,

$$\tilde{g}^N(x|\theta, e) \equiv \sum_{d_0 \in X^E} \tilde{\gamma}_X^N(d_0|\theta) \cdot \mathbb{I}(d_0(e) = x).$$

This construction is well-defined: $\tilde{g}^N(x|\theta, e) \in [0, 1]$ for any $(x, \theta, e) \in X \times \Theta \times E$, and $\sum_{x \in X} \tilde{g}^N(x|\theta, e) = 1$ for any $(\theta, e) \in \Theta \times E$. By (i), for every maximal evidence extension f_E^M, $\tilde{g}^N(x|\theta, f_E^M(\theta)) = \mathbb{I}(x = f_X(\theta))$ for all $(\theta, x) \in \Theta \times X$, so that it remains only to establish that $\tilde{g}^N$ satisfies $\tilde{IC}_G^N$ for every f_E^M.

Take then any type θ and consider any strategy specification $(\hat{\theta}, \hat{e}) \in \Theta \times E$ for this type under the mechanism $\tilde{g}^N$. By construction, it yields (expected) utility:

$$\sum_{x \in X} \tilde{g}^N(x|\hat{\theta}, \hat{e}) \cdot U(x, \hat{e}|\theta) = \sum_{x \in X} \left(\sum_{d_0 \in X^E} \tilde{\gamma}_X^N(d_0|\theta) \cdot \mathbb{I}(d_0(\hat{e}) = x) \right) \cdot U(x, \hat{e}|\theta)$$
$$= \sum_{d_0 \in X^E} \tilde{\gamma}_X^N(d_0|\theta) \cdot U(d_0(\hat{e}), \hat{e}|\theta).$$

For every maximal evidence extension f_E^M, using the above as well as (ii), we then have that, for any type θ and any $(\hat{\theta}, \hat{e}) \in \Theta \times E$,

$$\sum_{x \in X} \tilde{g}^N(x|\theta, f_E^M(\theta)) \cdot U(x, f_E^M(\theta)|\theta) = u(f_X(\theta)|\theta) = \sum_{d_0 \in X^E} \tilde{\gamma}_X^N(d_0|\theta) \cdot U(d_0(f_E^M(\theta)), f_E^M(\theta)|\theta)$$
$$\geq \sum_{d_0 \in X^E} \tilde{\gamma}_X^N(d_0|\hat{\theta}) \cdot U(d_0(\hat{e}), \hat{e}|\theta) = \sum_{x \in X} \tilde{g}^N(x|\hat{\theta}, \hat{e}) \cdot U(x, \hat{e}|\theta),$$

whereby $\tilde{g}^N$ satisfies $\tilde{IC}_G^N$ for f_E^M. □

4.5 The Value of Control

We now revisit the value of contractual control over the agent's evidence presentation, allowing for stochastic mechanisms. First, we generally confirm the notion of controllability as (at least in a weak sense) enhancing the principal's ability to implement various social choice functions: Implementability with controllable evidence is necessary for implementability with non-controllable evidence.

Proposition 9. *A social choice function f_X is implementable with non-controllable evidence only if f_X is implementable with controllable evidence.*

Proof. By Corollary 2, implementability with controllable evidence is characterized by (4), which is the weakest possible condition for implementability, since a type θ having all the evidence of another type θ' can perfectly mimic the latter under any mechanism. □

In view of this result, for any given setting, we will say that *control has value* (resp., *control has no value*), if there exists a social choice function f_X that is implementable with controllable evidence but not with non-controllable evidence.

Recall then Example 1 and reconsider the case of non-controllable evidence while now allowing for stochastic mechanisms. Since the evidence structure satisfies NOR, Lemma 2 tells us that we only need to consider randomization with respect to allocations. However, there are only two allocations, and no randomization over the two allocations when the agent presents the "null-evidence" e_0 could simultaneously deter both types from doing so in order to increase their chance of the preferred allocation. While d-implementable with controllable evidence, f_X is therefore not implementable with non-controllable evidence; in particular, control has value.

Relatedly, Example 4 does satisfy NOR, and implementation of the social choice function considered there requires a stochastic mechanism with non-controllable evidence, while it is d-implementable with controllable evidence. This raises the question whether Proposition 9 can be strengthened, and d-implementability with controllable evidence, though not sufficient, is generally necessary for implementability with non-controllable evidence. Combining Propositions 9 and 6 immediately yields the following result.

Corollary 4. *Suppose the evidence structure satisfies NOR. Then, a social choice function f_X is implementable with non-controllable evidence only if f_X is d-implementable with controllable evidence.*

This result does not generalize beyond settings whose evidence structure satisfies NOR. Example 3 violates NOR, and the social choice function considered there is implementable with non-controllable evidence, whereas it is easy to see that it is not d-implementable with controllable evidence, since any evidence that type θ_3 has is evidence that also another type has.

We are left with the question whether the two conditions WET and TIWO identified in Section 3 remain sufficient for control to have no value, even when stochastic mechanisms are available. Again, in any setting satisfying NOR, we readily obtain these generalizations. In fact, we obtain the following stronger result.

Corollary 5. *Suppose the evidence structure satisfies NOR. If the evidence structure additionally satisfies WET or the preference structure satisfies TIWO, then neither control nor randomization have value.*

Proof. Suppose the evidence structure satisfies NOR and take any social choice function f_X that is implementable with non-controllable evidence. By Proposition 9, f_X is implementable with controllable evidence. By Proposition 6, given NOR, f_X is d-implementable with controllable evidence.

If the evidence structure additionally satisfies WET, then, by Proposition 4, f_X is d-implementable with non-controllable evidence. Since d-implementability implies implementability, all implications established are equivalences. The same obtains if we impose TIWO instead of WET and invoke Proposition 5 instead of Proposition 4. □

The main additional insight from this corollary is that, given NOR, either WET or TIWO is sufficient for there to be no value of randomization with non-controllable evidence. To understand this result, recall Example 4, which shows that under NOR randomization with non-controllable evidence could be required only off-path, where it would help mitigate the evidentiary moral hazard problem. Both WET and TIWO achieve this perfectly, however, as WET means there is no "costly" evidence, and TIWO allows to deter any "costly" evidence. When neither of the two holds true, whereas NOR is satisfied, randomizing allocations works as an imperfect substitute for the control achieved under WET or TIWO.

A final example, which satisfies both WET and TIWO but violates NOR, shows that Corollary 5 does not hold without NOR.

Example 5. The principal has to decide between two allocations $X = \{x_b, x_g\}$, and the agent has three possible types $\Theta = \{\theta_1, \theta_2, \theta_3\}$. His preferences over allocations are type-independent, such that $u(x_g|\theta) \equiv u_g > u_b \equiv u(x_b|\theta)$ for all types θ, and the evidence structure is given as follows:

$$\mathcal{E}(\theta_1) = \{e_1\},\ \mathcal{E}(\theta_2) = \{e_2\},\ \text{and } \mathcal{E}(\theta_3) = \{e_1, e_2\} = E.$$

This setting satisfies WET, since each evidence e_i is maximal with respect to type θ_i, for both $i \in \{1, 2\}$, and it satisfies TIWO, since preferences are type-independent. Yet, it violates NOR, since neither evidence e_1 nor evidence e_2 is maximal with respect to type θ_3.

Consider now the social choice function f_X such that $f_X(\theta_1) = f_X(\theta_2) = x_b$ and $f_X(\theta_3) = x_g$. With controllable evidence, f_X is implementable, by Corollary 2, as it satisfies (4), which here imposes only that type θ_3 must weakly prefer $f_X(\theta_3)$ over either of $f_X(\theta_1)$ and $f_X(\theta_2)$.

With non-controllable evidence, and despite WET and TIWO, f_X is not implementable, however. Given Theorem 4, it suffices to show that there is no incentive compatible direct mechanism $\tilde{\gamma}^N \in \tilde{\Gamma}^N$ that implements it. Suppose, by way of contradiction, that $\tilde{\gamma}^N$ is such a mechanism, satisfying (i) $\tilde{\gamma}^N(d_0, e|\theta) > 0$ implies $d_0(e) = f_X(\theta)$ for all $(\theta, e) \in \Theta \times E$, and (ii) $\tilde{IC}^N$. Denote by $\tilde{f}_E$ the stochastic evidence extension such that $\tilde{\gamma}^N$ implements the stochastic-evidence-extended social choice function $(f_X, \tilde{f}_E) \in \mathcal{F}$, which is given by

$$\tilde{f}_E\left(e|\theta\right) = \sum_{d_0 \in X^E} \tilde{\gamma}^N\left(d_0, e|\theta\right),\ \forall(\theta, e) \in \Theta \times E.$$

Let then $\tilde{f}_E(e_1|\theta_3) \equiv p$, implying $\tilde{f}_E(e_2|\theta_3) \equiv 1 - p$. Incentive compatibility for types θ_1 and θ_2 then requires, respectively, that

$$u_b \geq p \cdot u_g + (1 - p) \cdot u_b \text{ and } u_b \geq (1 - p) \cdot u_g + p \cdot u_b,$$

which is equivalent to $p \leq 0$ and $p \geq 1$, respectively, and cannot simultaneously hold true, a contradiction. □

5 Concluding Discussion

We conclude by summarizing how our work relates to the existing literature, and by discussing the extent to which our insights generalize to settings with multiple agents and to sequential evidence presentation.

5.1 Related Literature

The literature on mechanism design with evidence originates with Green and Laffont (1986), who identified a failure of the revelation principle in settings where the evidence structure violates their nested range condition. Their model fits our Myersonian approach for non-controllable evidence, and their failure of the revelation principle arises from a severe implicit restriction on mechanisms, which is to disallow any communication of private information beyond the presentation of evidence.[31]

The most closely related works are Bull and Watson (2007), and Deneckere and Severinov (2008). Both characterize the implementatibility of deterministic social choice functions via mechanism design with evidence, and both already clarify the restrictive assumption in Green and Laffont (1986) (also closely related is Forges and Koessler (2005), see the discussion below on sequential evidence presentation). Our analysis contains their canonical mechanisms—the "special three-stage mechanism" in Bull and Watson (2007, Theorem 6), and "revelation mechanism R" in Deneckere and Severinov (2008, Theorem 3)—as incentive compatible direct mechanisms from the generalized revelation principle of Myerson (1982) for non-controllable evidence and thus provides an interpretation in terms of standard incentive compatibility (honesty and obedience).[32] Bull and Watson (2007) then focus mainly on the question of when recommendations are unnecessary and establish NOR as a sufficient condition (see their Theorem 5). Deneckere and Severinov (2008) allow for a larger class of mechanisms, as their focus is different (sequential evidence presentation, see discussion below) but obtain a similar result (see their Theorem 1). Both of these results concerning the implications of NOR for simplifying the mechanism design problem restrict attention to d-implementability, however (see their respective proofs). Besides uncovering the contractibility assumptions underlying these analyses and deriving standard notions of incentive compatible direct mechanisms, our Myersonian approach leads us to extend the analysis to when *any* randomization is redundant, by allowing for fully stochastic mechanisms in our setting: We show that randomization can still have value even under NOR unless further

[31]Their motivation for doing so relates to lying aversion, such that agent-types might never report certain other types; for applications in this vein see, e.g., Alger and Ma (2003), and Alger and Renault (2006, 2007).

[32]Both of these works' results allow for multiple agents, whereas we focus on a single agent. Concerning this basic point, the difference is purely expositional, however; see also the discussion below for further detail.

restrictions are imposed (Example 4), we add a necessity result for NOR regarding the redundancy of recommendations (Proposition 7), and we offer sufficient conditions in addition to NOR for there to indeed be no value of *any* kind of randomization (WET or TIWO, Corollary 5).

Generally, all existing derivations of canonical mechanisms are for non-controllable evidence. Our Myersonian approach points out that this constitutes an implicit assumption regarding the contractibility of evidence, and, in this respect, we contribute the novel analysis of controllable evidence, including the value of control.[33] Apart from raising this contractibility issue for applied work, we find the comparison between these contractual regimes also conceptually helpful for understanding the informational problem and incentives. Maybe most importantly, even for analyses of non-controllable evidence we contribute a significant analytical simplification device, whenever the setting satisfies WET or TIWO (see the characterization results in Section 3.3), though NOR may additionally be required when allowing for fully stochastic mechanisms (see Section 4.5).

5.2 Extensions

Multiple Agents. Important applications of mechanism design with evidence involve multiple agents (see, e.g., Ben-Porath et al. (2019)). Since the framework of Myerson (1982) for which he establishes the generalized revelation principle is for any number $n \in \mathbb{N}$ of agents, nothing fundamental stands in the way of extending our analysis. For Bayes-Nash implementation, incentive compatibility for each agent-type takes as given incentive compatibility for every other agent-type, and various conditions—WET, TIWO, NOR, as well as (non)-controllability of evidence—would simply need to be applied at the level of every agent.

Apart from noting that all of Examples 1 through 5 immediately generalize to insights for any $n \geq 1$, we point out here only two issues regarding d-implementability with non-controllable evidence that arise when agents' types are not independently distributed but such that some type profile has zero probability even though each individual type in the profile has positive (marginal) probability (all of Forges and Koessler (2005), Bull and Watson (2007), and Deneckere and Severinov (2008) allow for this, while assuming evidence is non-controllable): First, the direct mechanisms for non-controllable evidence can then not generally dispense with recommendations—i.e., Proposition 1 does not generalize—because

[33] Kartik and Tercieux (2012) pursue a formally similar idea to controllable evidence in their analysis of full implementation with evidence (see also Ben-Porath and Lipman (2012)), though with the important difference that, for full implementation, the revelation principle is not helpful to begin with. Also, Corollary 4 of Deneckere and Severinov (2008) considers a setting in which agents can be subjected to "infinite" punishment upon disobeying the (possibly stochastic) recommendation, which is tantamount to contractual control over evidence, see our Theorem 3.

it may be valuable to tailor the evidence recommendation for one agent to the report of another in order to incentivize the latter's honesty; see the example by Bull and Watson (2007, pp. 86–87), who also show that recommendations are, however, generally redundant in the n-agent case under NOR. Second, and relatedly, it is not clear when Bayes-Nash equilibrium is the appropriate notion in such settings. This is because direct mechanisms have a dynamic structure, and an agent may then learn from the recommendation he receives that another agent was dishonest, in which case Bayes-Nash equilibrium allows for irrational (non-credible) behavior to support honesty; see Gerardi (2004), and Gerardi and Myerson (2007) on this issue.

Sequential Evidence Presentation. We operationalize Myerson (1982) by modeling evidence presentation as an action in this framework. Consequently, we formally allow but a single stage where the agent chooses what evidence to present, after any communication has taken place. One is naturally led to ask when this restriction is without loss of generality upon allowing the principal to design the extensive form for *both* evidence presentation and communication, as opposed to only communication, subject to the restriction that evidence is exogenously given (in contrast to message spaces for communication).

First of all, with a single agent, it clearly is, both for controllable evidence and non-controllable evidence, once we interpret (or redefine) the evidence structure as specifying various combinations of atomic evidence that any agent-type could overall present (relatedly, see Sugaya and Wolitzky (2021) for the revelation principle in multistage games, and Sher (2014) for optimal design of possibly dynamic persuasion rules in the sender–receiver model of Glazer and Rubinstein (2006)). Furthermore, for controllable evidence, a similar generalization argument applies with any number of agents, since their evidence presentation is fully governed by the principal's contracts. This reduces the question to problems with non-controllable evidence and more than one agent, where note that, by the argument just given, an analysis for controllable evidence with a single stage of evidence presentation anyways provides an upper bound on what is implementable in this case.

Both Forges and Koessler (2005), and Deneckere and Severinov (2008) consider mechanisms for multiple agents and non-controllable evidence with (finite) sequential evidence presentation alongside communication.[34] The former authors are interested in characterizing (Bayes-Nash) equilibrium outcomes of a given game when allowing for any finite number of pre-play communication stages, each involving also verifying messages from fixed type-dependent message spaces. The latter authors are interested in the closely related question

[34]Bull and Watson (2007) restrict their mechanisms to extensive forms such that along every path every agent gets to present evidence only once. They establish the rather intuitive result from an incentive point of view that having all communication of private information before a single stage of simultaneous evidence presentation, as in the revelation principle derived here, is without loss in this larger class of mechanisms.

of whether and when a principal might benefit from designing evidence presentation as sequential. Of course, multiple rounds of evidence presentation may mechanically expand what evidence is overall feasible to present. However, under the restriction that any evidence that could be presented overall along any path could also be presented in a single stage, these works show that each of NOR and TIWO is a sufficient condition for a single stage (after any "soft" communication) to be without loss, while if both are violated, then there exist settings where sequential design of evidence presentation has value.[35]

The results of this paper naturally lead us to conjecture that WET, meaning every evidence is maximal evidence, would be another such condition. A general analysis of mechanism design with evidence where the extensive form of evidence presentation as well as communication is subject to design is beyond the scope of this paper, however. We only note that, conceiving of evidence presentation as an action rather than communication that can be arbitrarily designed, as we do here, immediately indicates why this is a challenging endeavour (see, however, Doval and Ely (2020), and Sugaya and Wolitzky (2021) for recent related advances).

References

Acharya, V. V., P. DeMarzo, and I. Kremer (2011). Endogenous information flows and the clustering of announcements. *The American Economic Review 101*(7), 2955–2979.

Alger, I. and C.-t. A. Ma (2003). Moral hazard, insurance, and some collusion. *Journal of Economic Behavior and Organization 50*(2), 225–247.

Alger, I. and R. Renault (2006). Screening ethics when honest agents care about fairness. *International Economic Review 47*(1), 59–85.

Alger, I. and R. Renault (2007). Screening ethics when honest agents keep their word. *Economic Theory 30*, 291–311.

Ali, S. N., G. Lewis, and S. Vasserman (2023). Voluntary disclosure and personalized pricing. *The Review of Economic Studies 90*(2), 538–571.

Asseyer, A. and R. Weksler (2022, September). Certification design with common values. Available at SSRN: http://dx.doi.org/10.2139/ssrn.4232034.

Ball, I. and D. Kattwinkel (2023, April). Probabilistic verification in mechanism design. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3189941.

Ben-Porath, E., E. Dekel, and B. L. Lipman (2018). Disclosure and choice. *The Review of Economic Studies 85*(3), 1471–1501.

Ben-Porath, E., E. Dekel, and B. L. Lipman (2019). Mechanisms with evidence: Commitment and robustness. *Econometrica 87*(2), 529–566.

Ben-Porath, E. and B. L. Lipman (2012). Implementation with partial provability. *Journal of Economic Theory 147*(5), 1689–1724.

Bull, J. and J. Watson (2007). Hard evidence and mechanism design. *Games and Economic Behavior 58*(1), 75–93.

[35] Deneckere and Severinov (2008) interpret this is as a "debate," see also Glazer and Rubinstein (2001).

Dasgupta, S., I. Krasikov, and R. Lamba (2022, July). Hard information design. Available at SSRN: http://dx.doi.org/10.2139/ssrn.4158934.

Deneckere, R. and S. Severinov (2008). Mechanism design with partial state verifiability. *Games and Economic Behavior 64*(2), 487–513.

Doval, L. and J. C. Ely (2020). Sequential information design. *Econometrica 88*(6), 2575–2608.

Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of Accounting Research 23*(1), 123–145.

Farrell, J. (1985, May). Voluntary disclosure: Robustness of the unraveling result, and comments on its importance. MIT Dept. of Economics Working Paper No. 374.

Forges, F. and F. Koessler (2005). Communication equilibria with partially verifiable types. *Journal of Mathematical Economics 41*, 793–811.

Gerardi, D. (2004). Unmediated communication in games with complete and incomplete information. *Journal of Economic Theory 114*(1), 104–131.

Gerardi, D. and R. B. Myerson (2007). Sequential equilibria in Bayesian games with communication. *Games and Economic Behavior 60*(1), 104–134.

Glazer, J. and A. Rubinstein (2001). Debates and decisions: On a rationale of argumentation rules. *Games and Economic Behavior 36*(2), 158–173.

Glazer, J. and A. Rubinstein (2004). On optimal rules of persuasion. *Econometrica 72*(6), 1715–1736.

Glazer, J. and A. Rubinstein (2006). A study in the pragmatics of persuasion: A game theoretical approach. *Theoretical Economics 1*(4), 395–410.

Green, J. R. and J.-J. Laffont (1986). Partially verifiable information and mechanism design. *The Review of Economic Studies 53*(3), 447–456.

Harsanyi, J. C. (1967). Games with incomplete information played by "Bayesian" players: Part I. The basic model. *Management Science 14*(3), 159–182.

Hart, S., I. Kremer, and M. Perry (2017). Evidence games: Truth and commitment. *American Economic Review 107*(3), 690–713.

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics 10*(1), 74–91.

Jung, W.-O. and Y. K. Kwon (1988). Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research 26*(1), 146–153.

Kartik, N. and O. Tercieux (2012). Implementation with evidence. *Theoretical Economics 7*(2), 323–355.

Koessler, F. and E. Perez-Richet (2019). Evidence reading mechanisms. *Social Choice and Welfare 53*, 375–397.

Lang, M. (2020, October). Mechanism design with narratives. CESifo Working Paper No. 8502.

Lipman, B. L. and D. J. Seppi (1995). Robust inference in communication games with partial provability. *Journal of Economic Theory 66*, 370–405.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic Theory 18*(2), 301–317.

Myerson, R. B. (1979). Incentive compatibility and the bargaining problem. *Econometrica 47*(1), 61–73.

Myerson, R. B. (1982). Optimal coordination mechanisms in generalized principal-agent problems. *Journal of Mathematical Economics 10*(1), 67–81.

Myerson, R. B. (1991). *Game Theory: Analysis of Conflict*. Harvard University Press.

Pram, K. (2023). Learning and evidence in insurance markets. *International Economic Review 64*(4), 1685–1714.

Rahman, D. and I. Obara (2010). Mediated partnerships. *Econometrica 78*(4), 285–308.

Rochet, J. C. (1985). The taxation principle and multi-time Hamilton-Jacobi equations. *Journal of Mathematical Economics 14*, 113–128.

Sher, I. (2011). Credibility and determinism in a game of persuasion. *Games and Economic Behavior 71*(2), 409–419.

Sher, I. (2014). Persuasion and dynamic communication. *Theoretical Economics 9*(1), 99–136.

Sher, I. and R. Vohra (2015). Price discrimination through communication. *Theoretical Economics 10*(2), 597–648.

Shin, H. S. (2003). Disclosures and asset returns. *Econometrica 71*(1), 105–133.

Shishkin, D. (2023, April). Evidence acquisition and voluntary disclosure. Retrieved on November 22, 2023, from https://denisshishkin.com/papers/evidence_acquisition.pdf.

Singh, N. and D. Wittman (2001). Implementation with partial verification. *Review of Economic Design 6*, 63–84.

Strausz, R. (2003). Deterministic mechanisms and the revelation principle. *Economics Letters 79*, 333–37.

Strausz, R. (2006). Deterministic versus stochastic mechanisms principal-agent models. *Journal of Economic Theory 128*, 306–314.

Strausz, R. (2012). Mediated contracts and mechanism design. *Journal of Economic Theory 147*(3), 1280–1290.

Sugaya, T. and A. Wolitzky (2021). The revelation principle in multistage games. *The Review of Economic Studies 88*(3), 1503–1540.

(Online) Appendix

A Proofs Omitted from the Main Text

A.1 Proposition 4

By proving the following proposition, we first establish the sufficiency of WET for control to have no value.

Proposition 10. *If the evidence structure satisfies WET, then control has no value.*

Proof. In view of Proposition 2, we only need to show that any social choice function that is d-implementable with controllable evidence is d-implementable with non-controllable evidence. Suppose then that the evidence structure $(E, \mathcal{E})$ satisfies WET and the social choice function f_X is d-implementable with controllable evidence. In view of Proposition 3, let $g_R^C : \Theta \times R \to X$ satisfy IC_G^C for $f_R : \Theta \to R$ such that $f_R(\theta) \in \mathcal{E}(\theta)$, and d-implement f_X, where $R \subseteq E$. Moreover, for any $e \in E \setminus R$, let $t(e) \in \Theta$ denote a type such that e is maximal with respect to $t(e)$, which exists by WET.

In view of Proposition 1, construct a mechanism $g^N \in G^N$ as follows: For all $(\theta, e) \in \Theta \times E$,

$$g^N(\theta, e) \equiv \begin{cases} g_R^C(\theta, e), & \text{if } e \in R, \\ f_X(t(e)), & \text{otherwise.} \end{cases}$$

We will now establish the claim by showing that g^N satisfies IC_G^N for the evidence extension f_R, from which it is immediate that g^N d-implements f_X.

Take then any type $\theta \in \Theta$. First, observe that, for any $(\theta', e) \in \Theta \times R$,

$$\begin{aligned} U\left(g^N(\theta, f_R(\theta)), f_R(\theta)|\theta\right) &= U\left(g_R^C(\theta, f_R(\theta)), f_R(\theta)|\theta\right) \\ &\geq U\left(g_R^C(\theta', e), e|\theta\right), \end{aligned}$$

since g_R^C satisfies IC_G^C for f_R. Second, for any $(\theta', e) \in \Theta \times E \setminus R$,

$$\begin{aligned} U\left(g^N(\theta', e), e|\theta\right) &= U\left(f_X(t(e)), e|\theta\right) \\ &= \begin{cases} u\left(f_X(t(e))|\theta\right), & \text{if } e \in \mathcal{E}(\theta), \\ -c, & \text{otherwise.} \end{cases} \end{aligned}$$

To see that $U(g^N(\theta, f_R(\theta)), f_R(\theta)|\theta) \geq u\left(f_X(t(e))|\theta\right)$ in case $e \in \mathcal{E}(\theta)$, note that because e is maximal with respect to $t(e)$, $e \in \mathcal{E}(\theta)$ implies that $\mathcal{E}(t(e)) \subseteq \mathcal{E}(\theta)$, so in particular $f_R(t(e)) \in \mathcal{E}(\theta)$. Then, by the first observation and the fact that g_R^C d-implements f_X with evidence extension f_R, it follows that $U\left(g^N(\theta, f_R(\theta)), f_R(\theta)|\theta\right) \geq U(g_R^C(t(e), f_R(t(e))), f_R(t(e))|\theta) = u\left(f_X(t(e))|\theta\right)$, completing the proof. □

By proving the following proposition, we establish the sense in which WET is also necessary for control to have no value.

Proposition 11. *Given any type structure* (Θ, p) *and associated evidence structure* $(E, \mathcal{E})$*, if* $(E, \mathcal{E})$ *violates WET, then there exists a preference structure* (X, u) *such that control has value.*

The proof of this proposition exploits properties of two binary relations that we now introduce. First, for any pair of types $(\theta, \theta') \in \Theta \times \Theta$, let $\theta \perp^e \theta'$ if and only if $\mathcal{E}(\theta) \not\subseteq \mathcal{E}(\theta')$ and $\mathcal{E}(\theta') \not\subseteq \mathcal{E}(\theta)$, so the binary relation $\perp^e$ expresses that each type has some evidence that the other does not have. Note that, while symmetric and irreflexive, $\perp^e$ is not transitive, in general. Second, for any subset of types $\tilde{\Theta} \subseteq \Theta$, we say that a pair of types from this subset, $(\theta, \theta') \in \tilde{\Theta} \times \tilde{\Theta}$, is $\perp^e$-connected in $\tilde{\Theta}$, written as $\theta \overset{\perp^e}{\leftrightarrow}_{\tilde{\Theta}} \theta'$, if and only if there exists a finite sequence $(\theta_i)_{i=1}^k \in \tilde{\Theta}^k$, $k \geq 1$, such that $\theta_1 = \theta$, $\theta_i \perp^e \theta_{i+1}$ for all $i \in \{1, \ldots, k-1\}$, and $\theta_k = \theta'$. Note that $\overset{\perp^e}{\leftrightarrow}_{\tilde{\Theta}}$ defines a binary relation on $\tilde{\Theta}$ that is reflexive, symmetric and transitive.[36] We now present two useful lemmas.

Lemma 3. *Fix any finite sequence* $(\theta_i)_{i=1}^k \in \Theta^k$*,* $k \geq 2$*, such that* $\theta_i \perp^e \theta_{i+1}$ *for all* $i \in \{1, \ldots, k-1\}$ *and let type* $\hat{\theta} \in \Theta$ *be such that* $\hat{\theta} \not\perp^e \theta_i$ *for all* $i \in \{1, \ldots, k\}$*. Then, either (i)* $\mathcal{E}(\theta_i) \subset \mathcal{E}(\hat{\theta})$ *for all* $i \in \{1, \ldots, k\}$*, or (ii)* $\mathcal{E}(\hat{\theta}) \subset \mathcal{E}(\theta_i)$ *for all* $i \in \{1, \ldots, k\}$*.*

Proof. We prove the lemma by induction. First, we show that either (i) $\mathcal{E}(\theta_1) \subset \mathcal{E}(\hat{\theta})$ or (ii) $\mathcal{E}(\hat{\theta}) \subset \mathcal{E}(\theta_1)$: By definition, $\hat{\theta} \not\perp^e \theta_1$ says that $\mathcal{E}(\theta_1) \subseteq \mathcal{E}(\hat{\theta})$ or $\mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta_1)$; since $\hat{\theta} \not\perp^e \theta_2$, however, $\mathcal{E}(\theta_1) = \mathcal{E}(\hat{\theta})$ would imply $\theta_1 \not\perp^e \theta_2$, in contradiction to $\theta_1 \perp^e \theta_2$.

Second, we show the induction step that, for any $i \in \{1, \ldots, k-1\}$, (i) $\mathcal{E}(\theta_i) \subset \mathcal{E}(\hat{\theta})$ implies $\mathcal{E}(\theta_{i+1}) \subset \mathcal{E}(\hat{\theta})$, and (ii) $\mathcal{E}(\hat{\theta}) \subset \mathcal{E}(\theta_i)$ implies $\mathcal{E}(\hat{\theta}) \subset \mathcal{E}(\theta_{i+1})$: By definition, $\hat{\theta} \not\perp^e \theta_{i+1}$ says that $\mathcal{E}(\theta_{i+1}) \subseteq \mathcal{E}(\hat{\theta})$ or $\mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta_{i+1})$; if (i) $\mathcal{E}(\theta_i) \subset \mathcal{E}(\hat{\theta})$, then $\mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta_{i+1})$ would imply $\mathcal{E}(\theta_i) \subset \mathcal{E}(\theta_{i+1})$, in contradiction to $\theta_i \perp^e \theta_{i+1}$, so we must have $\mathcal{E}(\theta_{i+1}) \subset \mathcal{E}(\hat{\theta})$; if (ii) $\mathcal{E}(\hat{\theta}) \subset \mathcal{E}(\theta_i)$, then $\mathcal{E}(\theta_{i+1}) \subseteq \mathcal{E}(\hat{\theta})$ would imply $\mathcal{E}(\theta_{i+1}) \subset \mathcal{E}(\theta_i)$, in contradiction to $\theta_i \perp^e \theta_{i+1}$, so we must have $\mathcal{E}(\hat{\theta}) \subset \mathcal{E}(\theta_{i+1})$. □

Lemma 4. *Take any subset of types* $\tilde{\Theta} \subseteq \Theta$ *and let*

$$P \equiv \{\theta \in \tilde{\Theta} : \exists \theta' \in \tilde{\Theta},\ \theta \perp^e \theta'\} \text{ and } Q \equiv \tilde{\Theta} \setminus P.$$

Then the following holds true:

[36] Reflexivity follows from the fact that we allow for $k = 1$. Symmetry follows from the symmetry of $\perp^e$, so we can "reverse" any sequence. Transitivity follows from the fact that we can concatenate any two sequences that end and begin, respectively, with the same type.

(i) If Q is non-empty, it contains a type $\theta \in Q$ such that, for any type $\theta' \in Q$, $\mathcal{E}(\theta) \subseteq \mathcal{E}(\theta')$.

(ii) If P is non-empty, it contains a subset $P^* \subseteq P$ that has at least two types $\{\theta, \theta'\} \subseteq P^*$ with $\mathcal{E}(\theta) \neq \mathcal{E}(\theta')$ and is such that, for any pair of types $(\theta, \theta') \in P^* \times P^*$, we have $\theta \overset{\perp^e}{\leftrightarrow}_{\tilde{\Theta}} \theta'$, and for any pair of types $(\theta, \theta') \in P^* \times P \setminus P^*$, we have $\mathcal{E}(\theta) \subset \mathcal{E}(\theta')$.

Proof. For (i), note that, it holds trivially if Q is a singleton. If it is not a singleton, note that for any two types $\{\theta, \theta'\} \subseteq Q$, we have that $\theta \not\perp^e \theta'$, implying $\mathcal{E}(\theta) \subseteq \mathcal{E}(\theta')$ or $\mathcal{E}(\theta') \subseteq \mathcal{E}(\theta)$. Hence, the weak set inclusion $\subseteq$ is complete on Q, and its finiteness implies the claim.

For (ii), note that, whenever P is non-empty, the irreflexivity and symmetry of $\perp^e$ imply that it contains at least two elements, and the symmetry of $\overset{\perp^e}{\leftrightarrow}_{\tilde{\Theta}}$ implies that $(P, \overset{\perp^e}{\leftrightarrow}_{\tilde{\Theta}})$ defines an undirected graph. If this graph is complete, then let $P^* = P$ and the claim follows immediately.

Otherwise, because $\overset{\perp^e}{\leftrightarrow}_{\tilde{\Theta}}$ is transitive, the maximal cliques of P in graph $(P, \overset{\perp^e}{\leftrightarrow}_{\tilde{\Theta}})$ partition P into $k \geq 2$ non-empty subsets $\{P_1, \ldots, P_k\}$. Moreover, note that each such subset P_i contains at least two elements, because $\theta \in P_i$ implies existence of some $\theta' \in P$ such that $\theta \perp^e \theta'$, which implies both $\mathcal{E}(\theta) \neq \mathcal{E}(\theta')$ and $\theta \overset{\perp^e}{\leftrightarrow}_{\tilde{\Theta}} \theta'$, so that $\theta' \neq \theta$ and $\theta' \in P_i$. Lemma 3 implies that we can order/relabel the subsets $\{P_1, \ldots, P_k\}$ such that, for any $i \in \{1, \ldots, k-1\}$ we have $\mathcal{E}(\theta) \subset \mathcal{E}(\theta')$ for any pair of types $(\theta, \theta') \in P_i \times P_{i+1}$. Letting $P^* = P_1$, the claim follows from transitivity of strict set inclusion $\subset$. $\square$

With the help of the previous two lemmas, we next prove Proposition 11 in two steps. In a first step, we show that a failure of WET implies that there is a non-empty set of types, which—in view of Lemma 4—we denote by P^*, such that each type in P^* has the non-maximal evidence $\tilde{e}$, while *every* possible type $\theta \in \Theta$, hence including also the types in P^*, has some evidence that some type in P^* does not have. In a second step, we show that this feature allows us to construct a preference structure (X, u) such that control has value.

Proof. Fix any type structure (Θ, p) and associated evidence structure $(E, \mathcal{E})$. Suppose now that $(E, \mathcal{E})$ violates WET, implying that there is some non-maximal evidence, say $\tilde{e}$, in E. Let

$$\tilde{\Theta} \equiv \{\theta \in \Theta : \tilde{e} \in \mathcal{E}(\theta)\}$$

denote the (non-empty) set of all types that have the non-maximal evidence $\tilde{e} \in E$. Observe that non-maximality of $\tilde{e}$ means that for any type $\theta \in \Theta$, there exists a type $\theta' \in \tilde{\Theta}$ such that $\mathcal{E}(\theta) \not\subseteq \mathcal{E}(\theta')$.

With reference to Lemma 4, consider the partition of set $\tilde{\Theta}$ of all types that have evidence $\tilde{e}$ into the two sets $\{P, Q\}$ such that

$$P \equiv \{\theta \in \tilde{\Theta} : \exists \theta' \in \tilde{\Theta},\ \theta \perp^e \theta'\} \text{ and } Q \equiv \tilde{\Theta} \setminus P.$$

Non-maximality of $\tilde{e}$ implies that P is non-empty. If P were empty, then $Q = \tilde{\Theta}$ so that Q is non-empty and, hence, Lemma 4 (i) would imply that there exists a type $\theta \in \tilde{\Theta}$ such that $\mathcal{E}(\theta) \subseteq \mathcal{E}(\theta')$ for all $\theta' \in \tilde{\Theta}$. Evidence $\tilde{e}$ would then be maximal with respect to any such type θ, a contradiction.

Given P is non-empty, Lemma 4 (ii) implies that it contains a subset $P^* \subseteq P$ with at least two types that have different evidence sets and is such that, for any pair of types $(\theta, \theta') \in P^* \times P^*$, we have $\theta \overset{\perp e}{\leftrightarrow}_{\tilde{\Theta}} \theta'$, and for any pair of types $(\theta, \theta') \in P^* \times P \setminus P^*$, we have $\mathcal{E}(\theta) \subset \mathcal{E}(\theta')$.

Non-maximality of $\tilde{e}$ implies that for any pair of types $(\theta, \theta') \in P^* \times Q$, we have that $\mathcal{E}(\theta) \subset \mathcal{E}(\theta')$. This is vacuously true if Q is empty. Otherwise, by Lemma 4 (i), there is a type $\hat{\theta} \in Q$ such that $\mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta')$ for any type $\theta' \in Q$. By Lemma 3, we have that either $\mathcal{E}(\theta) \subset \mathcal{E}(\hat{\theta})$ for all $\theta \in P^*$ or $\mathcal{E}(\hat{\theta}) \subset \mathcal{E}(\theta)$ for all $\theta \in P^*$. If the latter was true, however, this together with Lemma 4 (ii) would imply that $\mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta)$ for all $\theta \in P^* \cup (P \setminus P^*) = P$ and hence for all $\theta \in \tilde{\Theta}$, whereby $\tilde{e}$ would be maximal with respect to type $\hat{\theta}$, a contradiction. Hence, we must have $\mathcal{E}(\theta) \subset \mathcal{E}(\hat{\theta})$ for all $\theta \in P^*$, whereby $\mathcal{E}(\theta) \subset \mathcal{E}(\theta')$ for any pair of types $(\theta, \theta') \in P^* \times Q$.

Finally, non-maximality of $\tilde{e}$ then further implies that, for any type $\theta \notin \tilde{\Theta}$, there exists some type $\theta' \in P^*$ such that $\mathcal{E}(\theta) \nsubseteq \mathcal{E}(\theta')$. Recall that non-maximality of $\tilde{e}$ directly implies $\mathcal{E}(\theta) \nsubseteq \mathcal{E}(\tilde{\theta})$ for some type $\tilde{\theta} \in \tilde{\Theta}$. Fixing $\tilde{\theta}$, the result trivially follows if $\tilde{\theta} \in P^*$. Hence, consider the non-trivial case that $\tilde{\theta} \in \tilde{\Theta} \setminus P^*$, implying either $\tilde{\theta} \in Q$ or $\tilde{\theta} \in P \setminus P^*$. Then, as shown above, in either case we have $\mathcal{E}(\theta') \subset \mathcal{E}(\tilde{\theta})$ for all types $\theta' \in P^*$, and combining this with $\mathcal{E}(\theta) \nsubseteq \mathcal{E}(\tilde{\theta})$ yields that $\mathcal{E}(\theta) \nsubseteq \mathcal{E}(\theta')$ holds true for any $\theta' \in P^*$.

This completes our first step, as we have shown that non-maximality of $\tilde{e}$ implies that there is a non-empty set P^* of types that all have evidence $\tilde{e}$ and is such that *every* type $\theta \in \Theta$ has some evidence that some type in P^* does not have.

We are now ready to take our second step, which is to construct a particular preference structure (X, u) such that control has value. First, partition P^* into evidentiary equivalence classes $\{\tilde{\Theta}^*_1, \ldots, \tilde{\Theta}^*_k\}$ such that, for any $\{i, j\} \subseteq \{1, \ldots, k\}$ and any pair of types $(\theta, \theta') \in \tilde{\Theta}^*_i \times \tilde{\Theta}^*_j$, we have that $\mathcal{E}(\theta) = \mathcal{E}(\theta')$ if and only if $i = j$; recall that, because of Lemma 4 (ii), $k \geq 2$. Also, let $\{E_1, \ldots, E_k\}$ denote the evidence sets corresponding to these equivalence classes, so that, for each $i \in \{1, \ldots, k\}$ and any $\theta \in \tilde{\Theta}^*_i$ we have $E_i = \mathcal{E}(\theta)$. Now take (X, u) to be as follows:

$$X \equiv \{x_1, \ldots, x_k\}, \text{ and } u(x_i|\theta) \equiv \begin{cases} 1, & \text{if } \theta \in \tilde{\Theta}^*_i, \\ 0, & \text{otherwise.} \end{cases}$$

There are as many allocations as there are types with different evidence sets in P^*, and each type in P^* most prefers a unique allocation x_i corresponding to his evidence set E_i, whilst being indifferent over all others. Types not in P^* are indifferent among all allocations.

Given this preference structure, there is no mechanism with non-controllable evidence that could implement a social choice function f_X such that no type in P^* obtains the allocation they most prefer, i.e., that $f_X(\theta) \neq x_i$ holds for every $i \in \{1, \ldots, k\}$ and every $\theta \in \tilde{\Theta}_i^*$. For, if there were, then, by Proposition 1, an indirect mechanism of the form $g : \Theta \times E \to X$ would implement such a social choice function; however, fixing any report $\theta \in \Theta$, whichever $i \in \{1, \ldots, k\}$ we choose, $g(\theta, \tilde{e}) = x_i$ implies that a type in $\tilde{\Theta}_i^* \subset P^*$ obtains his preferred allocation, a contradiction.

We complete the proof by finally showing that, with controllable evidence, there is a mechanism that implements such a social choice function f_X. Consider the (indirect) mechanism such that the agent has to choose an element from the menu $\{\{(x_i, e)\}_{e \in E \setminus E_i}\}_{i=1}^k$. For each $i \in \{1, \ldots, k\}$, this menu associates the preferred allocation x_i of all types in $\tilde{\Theta}_i^*$ with all the evidence that these types do not have, which is $E \setminus E_i$. Note that $\tilde{e} \notin \bigcup_{i=1}^k E \setminus E_i$.

Recall that for every type θ (in the full type set Θ) there exists some type θ' in P^* such that $\mathcal{E}(\theta) \not\subseteq \mathcal{E}(\theta')$, i.e., there exists some $i \in \{1, \ldots, k\}$ such that $\mathcal{E}(\theta) \not\subseteq \mathcal{E}(\theta')$ for all θ' in $\tilde{\Theta}_i^*$. Hence, $\mathcal{E}(\theta) \cap \bigcup_{i=1}^k E \setminus E_i$ is non-empty, and this menu mechanism implements some social choice function. By construction, this social choice function satisfies $f_X(\theta) \neq x_i$ for any $\theta \in \tilde{\Theta}_i^*$, for all $i \in \{1, \ldots, k\}$, i.e., no type in P^* obtains his preferred allocation, because they have evidence E_i while x_i requires evidence $E \setminus E_i$. □

A.2 Proposition 5

By proving the following proposition, we first prove the sufficiency of TIWO for control to have no value.

Proposition 12. *If the preference structure satisfies TIWO, then control has no value.*

Proof. In view of Proposition 2, only necessity requires proof, i.e., that any social choice function that is d-implementable with controllable evidence is d-implementable with non-controllable evidence. Suppose then that the preference structure (X, u) satisfies TIWO and the social choice function f_X is d-implementable with controllable evidence. TIWO implies that there is a type-independent worst option $x_w \in X$. By Proposition 3, d-implementability of f_X with controllable evidence implies that there there is some $R \subseteq E$ which together with $g_R^C : \Theta \times R \to X$ satisfies IC_G^C for some $f_R : \Theta \to R$ with $f_R(\theta) \in \mathcal{E}(\theta)$, and d-implement f_X.

In view of Proposition 1, construct a mechanism $g^N \in G^N$ as follows: For all $(\theta, e) \in \Theta \times E$,

$$g^N(\theta, e) \equiv \begin{cases} g_R^C(\theta, e), & \text{if } e \in R, \\ x_w, & \text{otherwise.} \end{cases}$$

We will now establish the claim by showing that g^N satisfies IC_G^N for the evidence extension f_R, from which it is immediate that g^N d-implements f_X.

Take then any type $\theta \in \Theta$. First, observe that, for any $(\theta', e) \in \Theta \times R$,

$$\begin{aligned} U\left(g^N(\theta, f_R(\theta)), f_R(\theta))|\theta\right) &= U\left(g_R^C(\theta, f_R(\theta)), f_R(\theta)|\theta\right) \\ &\geq U\left(g_R^C(\theta', e), e|\theta\right), \end{aligned}$$

since g_R^C satisfies IC_G^C for f_R. Second, for any $(\theta', e) \in \Theta \times (E \setminus R)$,

$$\begin{aligned} U\left(g^N(\theta', e), e|\theta\right) &= U\left(x_w, e|\theta\right) \\ &= \begin{cases} u\left(x_w|\theta\right), & \text{if } e \in \mathcal{E}(\theta), \\ -c, & \text{otherwise;} \end{cases} \end{aligned}$$

so that $U\left(g^N(\theta, f_R(\theta)), f_R(\theta))|\theta\right) = u\left(f_X(\theta)|\theta\right) \geq u\left(x_w|\theta\right) \geq U\left(g^N(\theta', e), e|\theta\right)$ follows from the fact that g_R^C d-implements f_X and x_w is a type-independent worst option. Hence, g^N satisfies IC_G^N for the evidence extension f_R, which completes the proof. □

By proving the following proposition, we establish the sense in which TIWO is also necessary for control to have no value.

Proposition 13. *Given any type structure (Θ, p) and associated preference structure (X, u), if (X, u) violates TIWO, then there exists an evidence structure $(\mathcal{E}, E)$ such that control has value.*

Proof. Fix any type structure (Θ, p) and associated preference structure (X, u) that violates TIWO. For each allocation $x \in X$, let

$$P(x) \equiv \{\theta \in \Theta : x \in \arg\min_{x' \in X} u(x'|\theta)\}, \text{ and } Q(x) \equiv \Theta \setminus P(x).$$

$P(x)$ denotes the subset of types for whom x is a worst option, and $Q(x)$ its complement, containing all the types for whom x is not a worst option. Note that we have $\cup_{x \in X} P(x) = \Theta$, because every type has some worst option.

A violation of TIWO implies that $P(x) \neq \Theta$ and hence $Q(x) \neq \emptyset$ for every $x \in X$. Because $\cup_{x \in X} P(x) = \Theta$, this implies that the collection $\{P(x)\}_{x \in X}$ has at least two (different) non-empty sets.

Consequently, we can find two natural numbers $l \geq k \geq 2$ and express X as

$$X = \{x_1, \ldots, x_{k-1}, x_k, \ldots, x_l\}$$

such that $P(x_i) \neq \emptyset$ (i.e., allocation x_i is worst for some type) if and only if $i \in \{1, \ldots, k\}$.

Construct now the evidence structure $(E, \mathcal{E})$ such that $E = \{e_0, e_1, \ldots, e_k\}$ and

$$e_i \in \mathcal{E}(\theta) \Leftrightarrow i = 0 \text{ or } \theta \in P(x_i).$$

That is, each type θ has the evidence e_0 and, additionally, any evidence e_i with $i > 0$ where i is such that x_i is a worst option for type θ, i.e., $\theta \in P(x_i)$. Because $\cup_{i=1}^k P(x_i) = \Theta$, for every type θ, there exists an $i > 0$ such that $e_i \in \mathcal{E}(\theta)$, i.e., every $\mathcal{E}(\theta)$ contains at least one other evidence besides e_0. Moreover, because $P(x_i) \neq \emptyset$, every evidence e_i is held by some type θ, i.e., $E = \cup_{\theta \in \Theta} \mathcal{E}(\theta)$.

Given preference structure (X, u) and the defined evidence structure $(E, \mathcal{E})$, consider the (indirect) mechanism with controllable evidence given by the menu $\{(x_i, e_i)\}_{i=1}^k$. This mechanism implements a social choice function $f_X : \Theta \to X$ such that every type receives an allocation that is a worst option for him, i.e., $\theta \in P(f_X(\theta))$ for every $\theta \in \Theta$.

By contrast, there is no mechanism with non-controllable evidence that could implement such a social choice function. For, if there were, then, by Proposition 1, there would in particular be an indirect mechanism $g : \Theta \times E \to X$ that implements it. However, fixing any type report $\theta \in \Theta$, whichever $i \in \{1, \ldots, l\}$ we choose for $g(\theta, e_0) = x_i$, any type $\theta' \in Q(x_i)$, who always exists, can obtain an allocation that is not a worst option by reporting θ along with presenting e_0, a contradiction. □

A.3 Lemma 2

Proof. Only necessity requires proof, so suppose the stochastic direct mechanism $\tilde{\gamma}^N \in \tilde{\Gamma}^N$ is such that (i) $\tilde{\gamma}^N(d_0, e|\theta) > 0$ implies $d_0(e) = f_X(\theta)$ for all $(\theta, e) \in \Theta \times E$, and (ii) $\tilde{\gamma}^N$ satisfies $\tilde{IC}^N$; i.e., $\tilde{\gamma}^N$ is incentive compatible and implements f_X.

Construct now a mapping $\tilde{\gamma}_X^N : \Theta \to \Delta(X^E)$ as follows: For any type θ, and any allocation rule $d_0 \in X^E$, let $\alpha_{d_0} \in X^E$ be the allocation rule such that, for any $e \in E$,

$$\alpha_{d_0}(e) \equiv \begin{cases} f_X(\theta), & \text{if } e = f_E^M(\theta) \text{ for some maximal } f_E^M, \\ d_0(e), & \text{if } e \neq f_E^M(\theta) \text{ for any maximal } f_E^M, \end{cases}$$

and let then, for any allocation rule $d_0' \in X^E$,

$$\tilde{\gamma}_X^N(d_0'|\theta) \equiv \sum_{d_0 \in X^E : \alpha_{d_0} = d_0'} \left(\sum_{e \in E} \tilde{\gamma}^N(d_0, e|\theta) \right).$$

This construction is well-defined: $\tilde{\gamma}_X^N(d_0'|\theta) \in [0, 1]$ for any $d_0' \in X^E$, and $\sum_{d_0' \in X^E} \tilde{\gamma}_X^N(d_0'|\theta) = 1$. Moreover, we clearly have that $\tilde{\gamma}_X^N(d_0'|\theta) > 0$ only if $d_0'(f_E^M(\theta)) = f_X(\theta)$ for any type θ and any maximal evidence extension f_E^M.

Now fix any maximal evidence extension $\hat{f}_E^M$ and consider the stochastic direct mechanism $(\tilde{\gamma}_X^N, \hat{f}_E^M) \in \Gamma^N$. To prove the lemma, it only remains to show that this mechanism satisfies $\tilde{IC}^N$, because, by the above, it then implements f_X, and in particular, randomization with respect to recommendations thus has no value.

Take then any type $\theta \in \Theta$ and consider any strategy specification $(\hat{\theta}, \hat{\delta})$ for this type. Since recommendations under the mechanism $(\tilde{\gamma}_X^N, \hat{f}_E^M) \in \Gamma^N$ are deterministic, following the strategy with $(\hat{\theta}, \hat{\delta})$ results in recommendation $\hat{f}_E^M(\hat{\theta})$ with probability one, so only $\hat{\delta}(\hat{f}_E^M(\hat{\theta}))$ is outcome-relevant about $\hat{\delta}$. Let then $\hat{e}$ denote $\hat{\delta}(\hat{f}_E^M(\hat{\theta}))$, so that any strategy specification for type θ is without loss described by a pair $(\hat{\theta}, \hat{e})$. By construction of $\tilde{\gamma}_X^N$, if $\hat{e} = f_E^M(\hat{\theta})$ for some maximal evidence extension f_E^M, $(\hat{\theta}, \hat{e})$ results in sure allocation $f_X(\hat{\theta})$ and hence utility $U(f_X(\hat{\theta}), \hat{e}|\theta)$; otherwise, it yields the following (expected) utility, where we use that we can set $\alpha_{d_0} = d_0$ in this case:

$$\begin{aligned}
\sum_{d_0' \in X^E} \tilde{\gamma}_X^N(d_0'|\hat{\theta}) \cdot U(d_0'(\hat{e}), \hat{e}|\theta) &= \sum_{d_0' \in X^E} \left(\sum_{d_0 \in X^E : \alpha_{d_0} = d_0'} \left(\sum_{e \in E} \tilde{\gamma}^N(d_0, e|\theta) \right) \right) \cdot U(d_0'(\hat{e}), \hat{e}|\theta) \\
&= \sum_{d_0' \in X^E} \left(\sum_{e \in E} \tilde{\gamma}^N(d_0', e|\hat{\theta}) \right) \cdot U(d_0'(\hat{e}), \hat{e}|\theta) \\
&= \sum_{(d_0, e) \in X^E \times E} \tilde{\gamma}^N(d_0, e|\hat{\theta}) \cdot U(d_0(\hat{e}), \hat{e}|\theta).
\end{aligned}$$

Suppose then that there exists a strategy in mechanism $(\tilde{\gamma}_X^N, \hat{f}_E^M) \in \Gamma^N$ that yields type θ an expected utility greater than this type's utility from being honest and obedient, i.e., greater than $u(f_X(\theta)|\theta)$. If this strategy specifies a pair $(\hat{\theta}, \hat{e})$ such that $\hat{e} = f_E^M(\hat{\theta})$ for some maximal evidence extension f_E^M, then type θ obtains utility $U(f_X(\hat{\theta}), \hat{e}|\theta) > u(f_X(\theta)|\theta)$, implying $\hat{e} \in \mathcal{E}(\theta)$; since $\hat{e}$ is maximal with respect to $\hat{\theta}$, we have that $\mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta)$, whereby type θ could perfectly mimic type $\hat{\theta}$ under mechanism $\tilde{\gamma}^N$, resulting in a contradiction to this mechanism's satisfying $\tilde{IC}^N$. Otherwise, i.e., if this strategy specifies a pair $(\hat{\theta}, \hat{e})$ such that $\hat{e} \neq f_E^M(\hat{\theta})$ for any maximal evidence extension f_E^M, then—by the above—type θ obtains expected utility

$$\sum_{(d_0, e) \in X^E \times E} \tilde{\gamma}^N(d_0, e|\hat{\theta}) \cdot U(d_0(\hat{e}), \hat{e}|\theta),$$

which he could also obtain under mechanism $\tilde{\gamma}^N$, by first reporting type $\hat{\theta}$ and then presenting evidence $\hat{e}$ regardless of the recommendation. Hence, we again have a contradiction to this mechanism's satisfying $\tilde{IC}^N$. □

A.4 Proposition 7

Proof. Sufficiency is immediate from Lemma 1. For necessity, suppose the evidence structure $(E, \mathcal{E})$ violates NOR. Let the type $\hat{\theta}$ be such that, for all $e \in \mathcal{E}(\hat{\theta})$, there exists a type $\theta \in \Theta$ such that $e \in \mathcal{E}(\theta)$ and $\mathcal{E}(\hat{\theta}) \not\subseteq \mathcal{E}(\theta)$. The violation of NOR implies that $\hat{\theta}$ exists.

Consider then a preference structure (X, u) such that there are three allocations $X = \{x_1, x_2, x_3\}$, and the agent has type-independent preferences such that

$$u(x_1|\theta) = 1 > u(x_2|\theta) = 0 > u(x_3|\theta) = -k$$

for all types θ and any $k \geq (|\mathcal{E}(\hat{\theta})| - 1)/|\mathcal{E}(\hat{\theta})|$, together with the social function f_X such that

$$f_X(\theta) = \begin{cases} x_1, & \text{if } \mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta), \\ x_2, & \text{if } \mathcal{E}(\hat{\theta}) \not\subseteq \mathcal{E}(\theta). \end{cases}$$

This social choice function would discriminate between those types of agent that have all the evidence that type $\hat{\theta}$ has, who would receive the best allocation x_1, and those types that do not, who would receive the intermediate allocation x_2.

The violation of NOR implies existence of a type θ such that $e' \in \mathcal{E}(\theta)$ and $\mathcal{E}(\hat{\theta}) \not\subseteq \mathcal{E}(\theta)$. Letting $\rho(\theta) = \theta''$ and $\delta(\gamma_E^N(\theta'')) = e''$ be the report made and evidence presented by type θ on the path of the supposed BNE, we must have that $\tilde{\gamma}_X^N(d_0|\theta'') > 0$ implies $d_0(e'') = f_X(\theta) = x_2$, where also $e'' \in \mathcal{E}(\theta)$. However, since $e' \in \mathcal{E}(\theta)$,

$$\sum_{d_0 \in X^E} \tilde{\gamma}_X^N(d_0|\theta')U\left(d_0(e'), e'|\theta\right) = u(x_1|\theta) = 1 > 0 = u(x_2|\theta) = \sum_{d_0 \in X^E} \tilde{\gamma}_X^N(d_0|\theta'')U\left(d_0(e''), e''|\theta\right),$$

contradicting that $\sigma = (\rho, \delta)$ constitutes a BNE of the game induced by mechanism $\tilde{\gamma}^N = (\tilde{\gamma}_X^N, \gamma_E^N)$.

The following stochastic direct mechanism $\tilde{\gamma}^N \in \tilde{\Gamma}^N$ has stochastic recommendations and implements f_X for k sufficiently large: The marginal distribution over evidence recommendations is

$$\sum_{d_0 \in X^E} \tilde{\gamma}^N(d_0, e|\theta) \equiv \begin{cases} \frac{1}{|\mathcal{E}(\hat{\theta})|}, & \text{if } \mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta) \text{ and } e \in \mathcal{E}(\hat{\theta}), \\ \frac{1}{|\mathcal{E}(\theta)|}, & \text{if } \mathcal{E}(\hat{\theta}) \not\subseteq \mathcal{E}(\theta) \text{ and } e \in \mathcal{E}(\theta), \\ 0, & \text{otherwise;} \end{cases}$$

and the (conditional) allocation rules are given by

$$\tilde{\gamma}^N(d_0, e|\theta) > 0 \Rightarrow d_0(e') = \begin{cases} x_1, & \text{if } \mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta) \text{ and } e' = e, \\ x_2, & \text{if } \mathcal{E}(\hat{\theta}) \not\subseteq \mathcal{E}(\theta) \text{ and } e' = e, \\ x_3, & \text{if } e' \neq e. \end{cases}$$

The marginal distribution over evidence recommendations conforms to a uniform evidence extension, except that any type report θ such that $\mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta)$ is treated as if it was report $\hat{\theta}$. Obedient evidence presentation then yields the allocation according to f_X, so that if the agent is honest and obedient the mechanism implements this social choice function. Disobedience yields the worst allocation for ex-post utility $-k$. Since x_1 is the best allocation, incentive compatibility need only be ensured for those types that under honesty and obedience would receive allocation x_2, for a utility of zero (as opposed to one, if they received x_1). Consider then any such type θ, for whom $\mathcal{E}(\hat{\theta}) \not\subseteq \mathcal{E}(\theta)$: The only potentially profitable deviation is to report a type θ' such that $\mathcal{E}(\hat{\theta}) \subseteq \mathcal{E}(\theta')$, in which case the probability that θ could not obey the recommendation is at least $1/|\mathcal{E}(\hat{\theta})|$, since type $\hat{\theta}$ has at least some evidence that type θ does not have. This yields the following upper bound on the expected utility under any deviation from being honest and obedient:

$$\left(1 - \frac{1}{|\mathcal{E}(\hat{\theta})|}\right) - \frac{1}{|\mathcal{E}(\hat{\theta})|}k.$$

The assumed lower lower bound on k implies this is non-positive, establishing incentive compatibility of the mechanism, and hence that f_X is implementable.

□

B The Revelation Principle

Here, we restate the revelation principle from Myerson (1982) and show how it applies to our evidentiary implementation problem. First, we do so for the case where mechanisms are allowed to be stochastic, as in the original. Subsequently, we derive the revelation principle for the case of only deterministic mechanisms, as we use it in Section 3. Recall here the notation introduced in the main text's Section 4.2 for stochastic mechanisms, with the exception that $\tilde{F}$ will here denote the set of all "fully stochastic" evidence-extended social choice functions, i.e., all mappings $\tilde{f} : \Theta \to \Delta(X \times E)$ such that $\tilde{f}(x, e|\theta) > 0$ implies $e \in \mathcal{E}(\theta)$.

B.1 The Revelation Principle with Stochastic Mechanisms

Myerson (1982) formulates the economic problem in abstract terms, namely via a given decision structure (D_0, D_1), such that the principal and the agent simultaneously decide over $d_0 \in D_0$ and $d_1 \in D_1$, respectively. He defines preferences over decision profiles (d_0, d_1) denoted by $d \in D = D_0 \times D_1$. Moreover, his revelation principle concerns the implementability in pure Bayes-Nash equilibrium (BNE) of what we will call here stochastic social *decision* functions $\phi_D : \Theta \to \Delta(D)$.[37] It can be stated as follows (see Myerson, 1982, Proposition 2, p. 73):[38]

Theorem A. *A stochastic social decision function ϕ_D is implementable by some mechanism if and only if ϕ_D is implementable by an incentive compatible direct mechanism.*

For the formal definitions of the space of all feasible mechanisms, direct mechanisms and incentive compatibility of the latter, see the original work. Informally, the space of all feasible mechanisms is that of game forms which enrich the basic decision problem by arbitrary communication via the mechanism; direct mechanisms reduce this communication to having the agent first privately report a type $\theta \in \Theta$ to the mechanism and then privately receive a decision recommendation $d_1 \in D_1$ in return; incentive compatibility is the requirement that every type of every agent shall find it optimal to first honestly report his type and then obediently take the recommended decision.

Theorem A allows for stochastic mechanisms. Doing so as well, we only need to specify the decision sets for the present setting, derive the implied preferences over decision profiles, and characterize the implementability of social choice functions in terms of the implementability of social decision functions.

With controllable evidence, this is straightforward. The principal's decision set is $D_0 = X \times E$, and the decision set D_1 of the agent is a trivial singleton. We will abuse notation by identifying the sets D and D_0. The agent's preferences over decision profiles are immediate from his extended utility function: Denoting his utility under various decision profiles by $U^D : D \times \Theta \to \mathbb{R}$, we simply have $U^D(x, e \,|\theta) = U(x, e\,|\theta)$. Moreover, note that the space of stochastic social decision functions ϕ_D is identical to that of stochastic (evidence-extended) social choice functions $\tilde{f}$.

With non-controllable evidence, the principal's decision set is $D_0 = X^E$, and the decision set of the agent is $D_1 = E$. Again denoting the agent's utility from various decision profiles by $U^D : D \times \Theta \to \mathbb{R}$, we have $U^D(d_0, e\,|\theta) = U(d_0(e), e\,|\theta)$. In this case, the stochastic extended social choice function $\tilde{f} : \Theta \to \Delta(X \times E)$ is implementable if and only if there

[37] Myerson uses $t \in T$ rather than $\theta \in \Theta$ for the agent's types, and he formally defines neither social decision nor social choice functions.

[38] While the original statement is different, this is what is proven.

exists an implementable stochastic social decision function $\phi_D : \Theta \to \Delta(D_0 \times E)$ such that, for any $(x, e) \in X \times E$,

$$\tilde{f}(x, e|\theta) = \sum_{d_0 \in D_0} \phi_D(d_0, e|\theta) \cdot \mathbb{I}(d_0(e) = x).$$

B.2 The Revelation Principle with Deterministic Mechanisms

We now restrict the design to deterministic mechanisms. We prove here a general revelation principle for this case, from which Theorems 1 and 2 directly follow.[39] Our proofs rely on the formalism introduced in Myerson (1982) but are otherwise straightforward.

Say that a social decision function ϕ_D is *d-implementable* if there exists a deterministic mechanism that implements it. (This corresponds to the notion of implementability we use in Section 3.) Clearly, d-implementability implies implementability. The following lemma makes the basic observation that deterministic mechanisms cannot implement social decision functions that are not deterministic.

Lemma 5. *A social decision function ϕ_D is d-implementable only if ϕ_D is deterministic.*

Proof. Fix a mechanism $((R, M), \pi)$ and a strategy $\sigma = (\rho, \delta)$ of the agent in (the game induced by) this mechanism. Then, for any $(d_1, \theta) \in D_1 \times \Theta$,

$$\delta_\sigma^{-1}(d_1|\,\theta) = \{m \in M : \delta(m, \theta) = d_1\} \tag{6}$$

is the set of messages such that the agent's decision is d_1 when his type is θ (this set may be empty, of course).[40] Using this definition, the distribution over decision profiles under strategy σ in mechanism $((R, M), \pi)$ is given by

$$\pi_\sigma^*(d_0, d_1|\,\theta) = \sum_{m \in \delta_\sigma^{-1}(d_1|\theta)} \pi(d_0, m\,|\rho(\theta)) \tag{7}$$

(if $\delta_\sigma^{-1}(d_1|\theta) = \emptyset$, then $\pi_\sigma^*(d_0, d_1|\theta) = 0$ for any $d_0 \in D_0$, of course). A mechanism $((R, M), \pi)$ is deterministic if there exists a mapping $\tilde{\pi} : R \to D_0 \times M$ such that, for any possible reporting strategy $r \in R$ by the agent,

$$\pi(d_0, m\,|r) = \mathbb{I}((d_0, m) = \tilde{\pi}(r)).$$

[39] We maintain here the restriction to *pure* BNE. Otherwise, given their determinism, direct mechanisms would generally be insufficient (see Strausz, 2003).

[40] In his definition, Myerson (1982, p. 74) uses entire decision profiles d instead of just the agent's decision d_1, as we do here. Since the principal's decision component d_0 never affects the set that is defined, this makes no difference.

Suppose now that $((R, M), \pi)$ is deterministic and implements the social decision function ϕ_D. Then, there exists a BNE $\sigma = (\rho, \delta)$ such that, for any $(d, \theta) \in D \times \Theta$, $\pi^*_\sigma(d|\theta) = \phi_D(d|\theta)$. Since $((R, M), \pi)$ is deterministic, there exists a mapping $\tilde{\pi} : R \to D_0 \times M$ such that

$$\pi \left(d_0, m \left| \rho \left(\theta \right) \right. \right) = \mathbb{I} \left(\left(d_0, m \right) = \tilde{\pi} \left(\rho \left(\theta \right) \right) \right),$$

whereby, using (6) and (7),

$$\phi_D \left(d \left| \theta \right. \right) = \pi^*_\sigma \left(d \left| \theta \right. \right) = \sum_{m \in \delta_\sigma^{-1}(d, \theta)} \mathbb{I} \left(\left(d_0, m \right) = \tilde{\pi} \left(\rho \left(\theta \right) \right) \right),$$

which equals either zero or one. □

We are now ready to state the deterministic version of the revelation principle.

Theorem B. *A social decision function ϕ_D is d-implementable if and only if there exists a deterministic incentive compatible direct mechanism that implements ϕ_D.*

Proof. Only necessity requires proof. Suppose then that the social decision function ϕ_D is d-implementable and the deterministic mechanism $((R, M), \pi)$ implements it in BNE $\sigma = (\rho, \delta)$. Let $\tilde{\pi} : R \to D_0 \times M$ be such that, for any possible reporting strategy $r \in R$ by the agent,

$$\pi \left(d_0, m \left| r \right. \right) = \mathbb{I} \left(\left(d_0, m \right) = \tilde{\pi} \left(r \right) \right).$$

Construct the direct mechanism $((\Theta, D_1), \pi^*_\sigma)$ and note that it is deterministic, since $\pi^*_\sigma(d|\theta) = \phi_D(\theta)$ and ϕ_D is deterministic by Lemma 5. Consider any type θ of the agent. Suppose he could obtain greater utility by reporting θ' and following decision rule $\delta' : D_1 \to D_1$ than by reporting honestly and deciding obediently; i.e.,

$$\sum_{d \in D} \pi^*_\sigma \left(d \left| \theta \right. \right) U^D \left(d \left| \theta \right. \right) < \sum_{d \in D} \pi^*_\sigma \left(d \left| \theta' \right. \right) U^D \left(\delta' \left(d_1 \right) \left| \, \theta \right. \right).$$

Note that the left-hand side equals the utility under BNE σ in the original mechanism, and consider the following deviation in that mechanism:

$$\rho'' \left(\theta_i \right) = \rho \left(\theta' \right) \text{ and } \delta'' \left(m, \theta \right) = \delta' \left(\delta \left(m, \theta' \right) \right).$$

This deviation yields the same utility as the right-hand side in the above inequality and hence the same gain, contradicting BNE. □

Using similar arguments as in the case of stochastic mechanisms makes this result directly applicable to our setting and yields Theorems 1 and 2.

C On Green and Laffont's (1986) Model and NRC

Green and Laffont (1986) study a special case of evidence structures, relative to those considered in this paper. Specifically, for a given type structure (Θ, p) with $\Theta = \{\theta_1, \dots, \theta_n\}$, they consider evidence structures $(E, \mathcal{E})$ satisfying $E = \{e_1, \dots, e_n\}$ and $e_i \in \mathcal{E}(\theta_i)$ for all $i \in \{1, \dots, n\}$. In fact, Green and Laffont would identify every evidence e_i with the type report θ_i. Restricting attention to mechanisms of the form $g_E : E \to X$, they essentially show that restricting to "truthful" such mechanisms in the sense that type θ_i will "report" e_i is without loss if (and, in a weak sense, also only if) the evidence structure satisfies the so-called nested-range condition (NRC). Using our Definition 1 of maximal evidence, NRC can be usefully (re-)stated as follows.

Definition 5. An evidence structure $(E, \mathcal{E})$ in the Green and Laffont (1986) model satisfies the **nested-range condition (NRC)**, if evidence e_i is maximal with respect to type θ_i, for every $i \in \{1, \dots, n\}$.

NRC immediately implies that every evidence is maximal evidence (i.e., WET). Moreover, since $e_i \in \mathcal{E}(\theta_i)$ for every $i \in \{1, \dots, n\}$, it also implies that every type has evidence that is maximal with respect to himself (i.e., NOR).

The following example shows that neither WET nor NOR generally implies NRC, by showing the stronger result that they can even be jointly satisfied while NRC is violated: Take $n = 4$, where $\mathcal{E}(\theta_1) = \mathcal{E}(\theta_2) = \{e_1, e_2, e_4\}$, $\mathcal{E}(\theta_3) = \{e_1, e_3, e_4\}$ and $\mathcal{E}(\theta_4) = \{e_1, e_4\}$. Both of evidence e_1 and e_4 are maximal with respect to type θ_4, while e_2 is maximal with respect to both of types θ_1 and θ_2, and e_3 is maximal with respect to type θ_3. Hence, every evidence is maximal evidence (i.e., WET), and every type has evidence that is maximal with respect to himself (i.e., NOR), but evidence e_1 is not maximal with respect to type θ_1, violating NRC.

We add two further examples, which show that also within the Green and Laffont (1986) model, none of WET and NOR implies the other. (Given NRC implies both, the two examples also re-iterate the above point that neither WET nor NOR implies NRC.) First, take $n = 3$, where $\mathcal{E}(\theta_1) = \{e_1\}$, $\mathcal{E}(\theta_2) = \{e_2, e_3\}$, and $\mathcal{E}(\theta_3) = \{e_1, e_2, e_3\}$. Evidence e_1 is then maximal with respect to type θ_1, while both of evidence e_2 and e_3 are maximal with respect to type θ_2, and none is maximal with respect to type θ_3. Hence, every evidence is maximal evidence (i.e., WET), but not every type has evidence that is maximal with respect to himself (i.e., not NOR). Second, take again $n = 3$, but where $\mathcal{E}(\theta_1) = \mathcal{E}(\theta_2) = \{e_1, e_2\}$, and $\mathcal{E}(\theta_3) = \{e_1, e_3\}$. Evidence e_2 is maximal with respect to both of types θ_1 and θ_2, while e_3 is maximal with respect to type θ_3, and e_1 is not maximal evidence. Hence, every type has evidence that is maximal with respect to himself (i.e., NOR), but not every evidence is maximal evidence (i.e., not WET).

Finally, observe that a violation of NRC means that for some $i \in \{1, \dots, n\}$, evidence e_i is not maximal with respect to type θ_i. Given Green and Laffont (1986) identify evidence e_i with type report θ_i, this means that an honest report by type θ_i would not prove all this type could possibly prove. Hence, while a literal interpretation of "report" e_i by type θ_i would call it truthful, this has no factual foundation (see Bull and Watson (2007) for a closely related treatment of Green and Laffont's model, and also its discussion in Ball and Kattwinkel (2023)). Two observations serve to clarify this point. First, recall Proposition 1: It shows that, by extending the mechanisms $g_E : E \to X$ considered in Green and Laffont (1986) to mechanisms $g^N : \Theta \times E \to X$, where the agent also makes an "unrestricted" type report, one recovers the property that truthfulness in the sense of honest (unrestricted) type reporting is without loss. Second, consider the following example.

Example 6. The principal has to decide between two allocations $X = \{x_b, x_g\}$, and the agent has three possible types $\Theta = \{\theta_1, \theta_2, \theta_3\}$. His preferences over the two allocations are type-independent, such that he (strictly) prefers x_g over x_b. (We will not consider any randomization, so this is sufficient information.)

The evidence structure consists of three possible elements $E = \{e_1, e_2, e_3\}$ distributed over types according to

$$\mathcal{E}(\theta_1) = \mathcal{E}(\theta_2) = \{e_1, e_2\} \text{ and } \mathcal{E}(\theta_3) = \{e_1, e_3\}.$$

This corresponds to the example above, showing that an evidence structure in Green and Laffont (1986) may satsify NOR while violating WET, hence also violating NRC. Evidence e_1 is not maximal evidence—in particular not maximal with respect to type θ_1—while e_2 is maximal with respect to and possessed by both of types θ_1 and θ_2, and e_3 is maximal with respect to and possessed by type θ_3.

Note that the agent's private information effectively concerns only what he can prove, his preferences are common knowledge. Consider then the social choice function $f_X : \{\theta_1, \theta_2, \theta_3\} \to \{x_b, x_g\}$ such that $f_X(\theta_1) = f_X(\theta_2) = x_g$ and $f_X(\theta_3) = x_b$. There is no mechanism $g_E : E \to X$ that implements f_X such that every type θ_i presents evidence e_i: It would have to specify $g_E(e_1) = g_E(e_2) = x_g$ and $g_E(e_3) = x_b$, but then type θ_3 would deviate to presenting e_1. However, the mechanism g_E given by $g_E(e_2) = x_g$ and $g_E(e_1) = g_E(e_3) = x_b$ clearly implements f_X such that every type θ_i presents evidence that is maximal with respect to himself. Finally, note that extending this mechanism g_E to a mechanism $g^N : \Theta \times E \to X$ such that $g^N(\theta, e) = g_E(e)$ for all $(\theta, e) \in \Theta \times E$, we have g^N implement f_X such that every type both honestly reports his type and presents such maximal evidence. $\square$