Probabilistic Verification in Mechanism Design

Ian Ball*
Deniz Kattwinkel**

*Department of Economics, Yale University, ian.ball@yale.edu
**Department of Economics, University of Bonn, denizkattwinkel@gmail.com

Probabilistic Verification in Mechanism Design[*]

Ian Ball[†]   Deniz Kattwinkel[‡]

September 3, 2019

Abstract

We introduce a model of probabilistic verification in a mechanism design setting. The principal verifies the agent's claims with statistical tests. The agent's probability of passing each test depends on his type. In our framework, the revelation principle holds. We characterize whether each type has an associated test that best screens out all the other types. In that case, the testing technology can be represented in a tractable reduced form. In a quasilinear environment, we solve for the revenue-maximizing mechanism by introducing a new expression for the virtual value that encodes the effect of testing.

Keywords: probabilistic verification, testing, revelation principle, ordering tests, evidence.
JEL Codes: D82, D86.

[†]Department of Economics, Yale University, `ian.ball@yale.edu`.

[‡]Department of Economics, University of Bonn, `denizkattwinkel@gmail.com`.

1 Introduction

In the standard mechanism design paradigm, the principal can commit to an arbitrary mechanism, which induces a game among the agents. There are no explicit constraints on the mechanism, but there is an implicit assumption that the outcome of the induced game does not depend directly on the agents' types. Thus, each type can freely mimic every other type. The principal learns an agent's type only if the mechanism makes it optimal for that agent to reveal it.

In practice, claims about private information are often verified. If an employee applies for disability benefits, the provider performs a medical exam to assess the employee's condition. If a driver makes an insurance claim after a car accident, the insurer checks the claim against a police report. If a consumer reports his income to a lender, the lender requests a monthly pay stub for confirmation. In these examples, verification is noisy—medical tests are imperfect, witnesses fallible, and pay stubs incomplete.

The seminal paper of Green and Laffont (1986) first incorporates *partial verification* into mechanism design. In a principal–agent setting, they illustrate how verification relaxes incentive compatibility and hence makes more social choice functions implementable. In their model, mechanisms are direct and each type faces an exogenous restriction on which reports he can send to the principal. The following interpretation is suggested. As a function of the agent's type, certain reports are always detected as false while other reports never are. The punishment for detection is prohibitively costly, so the agent chooses among the reports he can send without being detected.

Partial verification cannot capture the *noisiness* of real verification. Under partial verification, the agent knows with certainty which reports will be detected as false. But in many applications, each report is detected with an associated probability. Thus, agents trade off the benefits of successful misreporting against the risk of detection.

In this paper, we model *probabilistic verification* by endowing a principal with a stochastic testing technology. We consider a principal–agent setting. The agent has a private type. The principal has full commitment power and controls decisions (which may or may not include transfers). To this standard setting, we add a family of *pass–fail tests*. The principal elicits a message from the agent and then conducts a test. The agent sees the test and privately chooses whether to exert effort, which is costless. If he exerts effort, then his passage probability depends on his type and on the test; if he does not exert effort, then he fails with certainty. The principal observes the result of the test—but not the agent's effort—and then takes a decision.

We analyze which social choice functions can be implemented with a given testing technology. We reduce the class of mechanisms in two steps.

First, we simplify communication. Since testing does not intrude into the communication stage, we get a version of the revelation principle (Theorem 1): There is no loss in restricting attention to direct mechanisms that induce the agent to report truthfully and to exert effort on every test. We contrast our result with the failure of the revelation principle in the setting of Green and Laffont (1986).

Second, we simplify the choice of tests. In general, which test is best for verifying a particular type report depends on which types the principal would like to screen away. We introduce for each type θ an associated order over tests. One test is more θ-*discerning* than another if it can better screen all other types away from type θ. This family of orders is the appropriate analogue of Blackwell's (1953) informativeness order for our testing setting. These discernment orders provide a unified generalization of various conditions imposed in the deterministic verification literature, such as nested range (Green and Laffont, 1986), full reports (Lipman and Seppi, 1995), and normality (Bull and Watson, 2007).

A function assigning to each type θ a most θ-discerning test is a *most-discerning* testing function. The sufficiency part of our main implementation result (Theorem 2) states: If there exists a most-discerning testing function, then every implementable social choice function can be implemented using that testing function. In this case, the testing technology induces an *authentication rate*, which specifies the probabilities with which each type can pass the test assigned to each other type. The principal's problem reduces to an optimization over decision rules, subject to incentive constraints involving the authentication probabilities.

The reduction from a testing technology to an authentication rate can be inverted. We provide a necessary and sufficient condition for an authentication rate to be induced by a most-discerning testing function. In applications, we can directly specify an authentication rate that satisfies our condition—testing need not be modeled explicitly. If the authentication rate takes values 0 and 1 only, then our condition reduces to the nested range condition that Green and Laffont (1986) use to recover the revelation principle.

We are the first to analyze verification with the first-order approach. Partial verification is not amenable to this approach because the authentication probability jumps discontinuously from 0 to 1. Under probabilistic verification, the authentication rate can depend continuously on the agent's report, so each local constraint is loosened but not eliminated. In a quasilinear environment, we aggregate these loosened local constraints to derive a *virtual value* that encodes the testing technology.

We use this virtual value to solve for revenue-maximizing mechanisms in three classical settings: nonlinear pricing, selling a single good, and auctions. With verification, the revenue-maximizing allocations have their usual expressions, except that our virtual value appears in place of the classical virtual value. The associated transfers are higher in the

3

presence of verification. If the tests are completely uninformative, then our virtual value equals the classical virtual value. As the tests become more precise, our virtual value increases toward the agent's true value, and the revenue-maximizing allocation becomes more efficient. When selling a single good, a posted price is not optimal. Instead, the price depends on the agent's report. To study auction settings, we extend the model to allow for competing agents who submit reports and are tested separately. A virtual value is defined for each agent. The revenue-maximizing auction allocates the good to the agent whose virtual value is highest.

Finally, we consider tests with more than two results. Our discernment orders naturally extend. As in the baseline model, if there exists a most-discerning testing function, then there is no loss in using that testing function only.

The rest of the paper is organized as follows. Section 2 presents our model of testing. Section 3 establishes the revelation principle. Section 4 introduces the discernment orders and states the most-discerning implementation result. Section 5 reduces the verification technology to an authentication rate. Section 6 considers applications to revenue-maximization. We extend the model to allow for multiple agents in Section 7 and nonbinary tests in Section 8. Section 9 connects our model to previous models of verification in economics and computer science; other relevant literature is referenced throughout the text. Section 10 concludes. Measure-theoretic definitions are in Appendix A. Proofs are in Appendices B and C.

2 Model

2.1 Setting

There are two players: a principal (she) and an agent (he). The agent draws a private type $\theta \in \Theta$ from a commonly known distribution. The principal takes a decision $x \in X$.[1] Preferences depend on the decision x and on the agent's type θ. The Bernoulli utility functions for the agent and the principal are

$$u \colon X \times \Theta \to \mathbf{R} \quad \text{and} \quad v \colon X \times \Theta \to \mathbf{R}.$$

A social choice function, denoted

$$f \colon \Theta \to \Delta(X),$$

assigns a decision lottery to each type.[2]

---

[1]Decisions are completely abstract; they may or may not include transfers.

[2]We make the following standing technical assumptions. Each set is a Polish space endowed with its Borel σ-algebra. The space of Borel probability measures on a Polish space Z is denoted $\Delta(Z)$. All

2.2 Verification technology

To the principal–agent setting we add a verification technology. There is a set T of *tests*, with generic element τ. Each test generates a binary result—pass or fail, denoted 1 or 0.[3] The distribution of test results is determined by the *passage rate*

$$\pi \colon T \times \Theta \to [0,1],$$

which assigns to each pair (τ, θ) the probability with which type θ can pass test τ. The passage rate π is common knowledge, as is the rest of the setting, except the agent's private type.

The principal can conduct one test from the set T. The procedure is as follows. First the principal selects a test τ. Next the agent observes τ and chooses to exert effort or not, denoted $\bar{e}$ or $\underline{e}$. Effort is costless. If the agent exerts effort, nature draws the test result 1 with probability $\pi(\tau|\theta)$ and the test result 0 otherwise. If the agent does not exert effort, the test result is 0 with certainty. The principal observes the test result, but not the agent's effort choice.[4]

Effort is an inalienable choice of the agent, as in models of evidence in which the agent chooses whether to produce the evidence he possesses (Bull and Watson, 2007). Indeed, if the passage rate takes values 0 and 1 only, then each test can be interpreted as a request for a particular piece of hard evidence; see Example 2. In general, the passage rate takes interior values, so the agent is unsure whether he will pass each test. The agent can, however, intentionally fail each test by exerting no effort. The active role played by the agent and the resulting asymmetry between passage and failure are what distinguish tests from statistical experiments.

2.3 Mechanisms and strategies

The principal commits to a mechanism, which induces an extensive-form game that proceeds as follows. The agent sends a message to the principal. Based on the message, the principal selects a test. The agent observes the test and chooses whether to exert effort. Nature draws the test result, as prescribed by the passage rate and the agent's effort. The principal observes this result and takes a decision.

In view of this timing, we define mechanisms and strategies.

---

primitive functions are measurable. Social choice functions and other measure-valued maps satisfy a weaker measurability condition called universal measurability. The details are in Appendix A.

[3]We consider nonbinary tests in Section 8; our main results go through.

[4]The same procedure is used in Deb and Stewart (2018). They allow the principal to conduct tests (termed *tasks*) sequentially before making a binary classification. In DeMarzo et al. (2019), a seller of an asset can conduct a test of the asset's quality. Each test has a *null result*, which the seller can always claim to have received. If there is only one non-null result, then this technology is equivalent to ours.
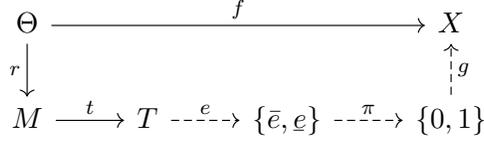
$$\Theta \xrightarrow{\quad\quad\quad\quad f \quad\quad\quad\quad} X$$

$$\begin{array}{ccc} \Theta & \xrightarrow{\quad f \quad} & X \\ {\scriptstyle r}\downarrow & & \uparrow{\scriptstyle g} \\ M & \xrightarrow{\;t\;} T \dashrightarrow{\;e\;} \{\bar{e},\underline{e}\} \dashrightarrow{\;\pi\;} \{0,1\} \end{array}$$

Figure 1. Social choice function induced by a profile

Definition 1 (Mechanism). A *mechanism* (M, t, g) consists of a message space M together with a *testing rule* $t\colon M \to \Delta(T)$ and a *decision rule*

$$g\colon M \times T \times \{0,1\} \to \Delta(X).$$

Once the principal commits to a mechanism, the agent faces a dynamic decision problem. First he sends a message. Then, after observing the selected test, he chooses effort.

Definition 2 (Strategy). A *strategy* (r, e) for the agent consists of a *reporting strategy* $r\colon \Theta \to \Delta(M)$ and an *effort strategy*

$$e\colon \Theta \times M \times T \to [0,1],$$

specifying the probability of exerting effort.

A mechanism (M, t, g) and a strategy (r, e) together constitute a *profile*, which induces a social choice function by composition, as indicated in Figure 1. The functions e, π, and g are represented by dashed arrows as a reminder that these functions depend on histories, not just their source sets in the diagram. Using the notation for composition from Markov processes—think of a probability row vector right-multiplied by stochastic matrices—the induced social choice function f is $(r \times t \times (e\pi))g$. The map $r \times t \times (e\pi)$ from Θ to $\Delta(M \times T \times \{0,1\})$ is applied before the map g. For measure-theoretic definitions of these operations, see Appendix A.1.

The players use expected utility to evaluate lotteries over decisions. A profile $(M, t, g; r, e)$ is *incentive compatible* if the strategy (r, e) is a *best response* to the mechanism (M, t, g), i.e., the strategy (r, e) maximizes the agent's utility over all strategies in the game induced by the mechanism (M, t, g). A profile $(M, t, g; r, e)$ *implements* a social choice function f if $(M, t, g; r, e)$ is incentive compatible and $f = (r \times t \times (e\pi))g$. A social choice function is *implementable* if there exists a profile that implements it.
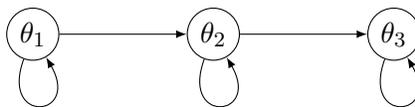
6

Figure 2. Feasible reports

3   Revelation principle

We now begin our analysis of which social choice functions can be implemented, given a testing technology. The space of incentive-compatible profiles is large. We reduce this space by establishing a version of the revelation principle. First we revisit the failure of the revelation principle in Green and Laffont's (1986) model of partial verification.

Example 1 (Partial verification, Green and Laffont, 1986)**.** The agent has one of three types, labeled θ_1, θ_2, and θ_3. Type θ_1 can report θ_1 or θ_2; type θ_2 can report θ_2 or θ_3; and type θ_3 can report only θ_3. This correspondence is represented as a directed graph in Figure 2. Each type is a node, and edges connect each type to each of his feasible reports.

The principal chooses whether to allocate a single good to the agent, who prefers to receive it, no matter his type. Consider the social choice function that allocates the good if and only if the agent's type is θ_2 or θ_3. To truthfully implement this allocation, the principal must allocate the good if the agent reports θ_2. But then type θ_1 can report θ_2 in order to get the good. Therefore, this allocation cannot be implemented truthfully. It can, however, be implemented *untruthfully*: The principal allocates the good if and only if the agent reports θ_3. Types θ_2 and θ_3 both report θ_3, but type θ_1 cannot.

What goes wrong in this example? In the partial verification model, reports are not cheap-talk messages. Instead, each report serves as a test that certain types can pass. Truthful implementation implicitly assigns to each type θ the test "report θ", regardless of which other types can pass that test. By contrast, our testing framework separates communication from verification. Reports retain their usual meaning and thus the revelation principle holds.

In the standard mechanism design setting, the revelation principle states that there is no loss in restricting to direct and truthful profiles. In our testing framework, we demand also that the agent exert effort on every test.

Definition 3 (Canonical)**.** A profile $(M, t, g; r, e)$ is *canonical* if (i) $M = \Theta$, (ii) $r = \mathrm{id}$, and (iii) $e(\theta, \theta, \tau) = 1$ for all $\theta \in \Theta$ and $\tau \in T$.

Here the identity, id, maps each type θ to the point mass δ_θ in $\Delta(\Theta)$.[5] Condition (i) says that the mechanism is direct. Condition (ii) says that the agent reports truthfully.

---

[5]We sometimes identify a map $y \mapsto f(y)$ from Y to Z with the map $y \mapsto \delta_{f(y)}$ from Y to $\Delta(Z)$.

Condition (iii), which is specific to the verification setting, says that the agent exerts effort on every test.[6] A social choice function is *canonically implementable* if it can be implemented by a canonical profile. Our revelation principle states that there is no loss in restricting to canonical implementation.

Theorem 1 (Revelation principle)
Every implementable social choice function is canonically implementable.

The structure of the proof is similar to that of the standard revelation principle. Given an arbitrary profile $(M, t, g; r, e)$ that implements a social choice function f, we construct a canonical profile that also implements f. Play proceeds as follows.[7] The agent truthfully reports his type θ. The principal feeds this report θ into the reporting strategy r to draw a message m, which is then passed to the testing rule t to draw a test τ. The agent exerts effort, so nature draws the test result 1 with probability $\pi(\tau|\theta)$ and the test result 0 otherwise. If the test result is 0, the principal feeds 0 into the decision rule g. If the test result is 1, the principal feeds into g the result 1 with probability $e(\theta, m, \tau)$ and the result 0 otherwise. Therefore, the input to g is 1 with probability $\pi(\tau|\theta)e(\theta, m, \tau)$.[8]

This canonical profile induces f. To check that this profile is incentive compatible, we show that for any deviation in the canonical mechanism, there is a corresponding deviation in the original mechanism that induces the same (stochastic) decision. The original profile is incentive compatible, so this deviation cannot be profitable.

Suppose that the agent reports θ' and then follows the strategy of exerting effort with probability $\tilde{e}(\tau)$ on each test τ. The agent can get the same decision in the original mechanism by playing as follows. Feed θ' into the strategy r to draw a message m to send to the principal. If test τ is conducted, exert effort with probability $\tilde{e}(\tau)e(\theta', m, \tau)$.

By separating communication from testing, we recovered the revelation principle. Conceptually, our testing framework elucidates the role of verification in eliciting private information. Computationally, our progress is less clear. The complexity of untruthful reporting has been replaced by the complexity of testing rules. Reducing this class of testing rules is what we turn to next.

---

[6]The agent's off-path behavior can be ignored because the best-response condition is in strategic form. Alternatively, condition (iii) could be strengthened to require that $e(\theta, \theta', \tau) = 1$ for all $\theta, \theta' \in \Theta$ and $\tau \in T$. The results would not change.

[7]In the proof, we use a more general procedure that works also for nonbinary tests.

[8]Technically, this is not a mechanism but a *generalized* mechanism (Mertens et al., 2015, Exercise 10, p. 70) because the distribution of the randomization device—the message the principal draws—depends on the agent's type report. In the proof, we eliminate this device by applying the disintegration theorem.

4 Ordering tests

In this section, we identify a smaller class of testing rules that suffices for implementation. To define this class, we introduce a family of orders over tests.

4.1 Discernment orders

The basic question is whether one test can *always* be used in place of another. More precisely, fix a type θ and tests τ and ψ. Suppose that the principal conducts test ψ on type θ. Is it *always* possible to replace test ψ with test τ and then adjust the decision rule so that (i) the decision for type θ is preserved, and (ii) no new deviations are introduced? The key is to *convert* each score[9] on test τ into an equivalent score on test ψ. The principal feeds this converted score into the original decision rule.

A score conversion is a Markov transition k on $\{0, 1\}$, which associates to each score s in $\{0, 1\}$ a measure k_s in $\Delta(\{0, 1\})$. A transition k on $\{0, 1\}$ is *monotone* if k_1 first-order stochastically dominates k_0, denoted $k_1 \geq_{\mathrm{SD}} k_0$. For each test τ and type θ, denote by $\pi_{\tau|\theta}$ the measure on $\{0, 1\}$ that puts probability $\pi(\tau|\theta)$ on 1. When a test result is drawn from the measure $\pi_{\tau|\theta}$ and a Markov transition k is applied, the resulting distribution is denoted $\pi_{\tau|\theta}k$, which can be viewed as the product of a row vector and a stochastic matrix.

Definition 4 (Discernment order). Fix a type θ. A test τ is *more θ-discerning* than a test ψ, denoted $\tau \succeq_\theta \psi$, if there exists a monotone Markov transition k on $\{0, 1\}$ satisfying:

(i) $\pi_{\tau|\theta}k = \pi_{\psi|\theta}$;

(ii) $\pi_{\tau|\theta'}k \leq_{\mathrm{SD}} \pi_{\psi|\theta'}$ for all types θ' with $\theta' \neq \theta$.

Conditions (i) and (ii) correspond to parts (i) and (ii) of the motivating question above. Each condition compares two testing procedures. On the left side, the agent exerts effort on test τ and his score is converted by k into a score on test ψ. On the right side, the agent exerts effort on test ψ and his score is drawn. Condition (i) says that for type θ these two procedures give the same score distribution. Condition (ii) says that for all other types the converted score distribution is first-order stochastically dominated by the unconverted score distribution on test ψ. The conversion k is required to be monotone so that effort weakly improves the distribution of the converted score. In short, the definition ensures that the conversion k from τ-scores to ψ-scores is fair for type θ but (weakly) unfavorable for all other types.

The relation $\succeq_\theta$ is reflexive and transitive. For reflexivity, take k to be the identity transition. For transitivity, compose the score conversions and note that monotone transitions preserve first-order stochastic dominance and are closed under composition; see

---

[9]We use *score* synonymously with *result*. Tests are pass–fail, but the language of scoring is more intuitive and our constructions immediately extend to the nonbinary case.

Appendix A.2. Two tests can be incomparable under $\succeq_\theta$. We use the notation $\sim_\theta$ for equivalence under $\succeq_\theta$, and $\succ_\theta$ for the strict part of $\succeq_\theta$.

The θ-discernment order resembles Blackwell's (1953) informativeness order between experiments. Given a state space Ω and a signal space S, an *experiment* is a map from Ω to $\Delta(S)$. In an experiment, the signal realizations are drawn exogenously by nature. A *garbling* is a Markov transition on S. An experiment τ is *more Blackwell informative* than an experiment ψ if there exists a garbling g such that

$$\tau g = \psi. \tag{1}$$

To bring out the connection with the discernment orders, set $\Omega = \Theta$, and denote by $\pi_{\tau|\theta'}$ the distribution of signals from experiment τ in state θ'. Then (1) can be expressed as

$$\pi_{\tau|\theta'} g = \pi_{\psi|\theta'} \quad \text{for all } \theta' \in \Theta. \tag{2}$$

The Blackwell order is concerned with information, not incentives. The garbled signal from experiment τ must have the same distribution as the ungarbled signal from experiment ψ, in every state of the world. No state is privileged, and no structure on the signal space is required. In contrast, the discernment orders reflect the agent's incentives to report truthfully and to exert effort. There is a family of discernment orders, one associated with each type θ. For the distinguished type θ, the converted score on test τ must have the same distribution as the unconverted score on test ψ. For all other types—the potential deviators—the converted score on test τ need only be stochastically dominated by the score on test ψ. A conversion, unlike a garbling, is required to be monotone so that effort weakly improves the distribution of the converted score. For dominance and monotonicity to make sense, scores must be totally ordered.

4.2 Implementation with most-discerning testing functions

We are interested in maximum tests with respect to the discernment orders.

Definition 5 (Most discerning). A test τ is *most θ-discerning* if $\tau \succeq_\theta \psi$ for all $\psi \in T$. A function $t\colon \Theta \to T$ is *most discerning* if, for each type θ, the test $t(\theta)$ is most θ-discerning.

To state the implementation result, we need a few definitions. Given a type space Θ, a *testing environment* consists of a test set T and a passage rate $\pi\colon T \times \Theta \to [0, 1]$. A *decision environment* consists of a decision set X and a utility function $u\colon X \times \Theta \to \mathbf{R}$ for the agent. Given a testing rule $\hat{t}\colon \Theta \to \Delta(T)$, a social choice function f is *canonically implementable with $\hat{t}$* if there exists a decision rule g such that the mechanism $(\hat{t}, g)$ canonically implements f.

Theorem 2 (Most-discerning implementation)

Fix a type space Θ and a testing environment (T, π). For a measurable testing function $\hat{t} \colon \Theta \to T$, the following are equivalent.

1. *$\hat{t}$ is most discerning.*
2. *In every decision environment (X, u), every implementable social choice function is canonically implementable with $\hat{t}$.*

The implication from condition 1 to condition 2 means that a single most-discerning testing function suffices for implementation. The proof formalizes the test replacement that motivated our definition of the discernment orders. By the revelation principle (Theorem 1), there is no loss in considering only canonical profiles. Suppose that in a canonical profile, some report θ is assigned a test ψ with $\psi \neq \hat{t}(\theta)$. Since $\hat{t}(\theta) \succeq_\theta \psi$, the principal can use a score conversion to replace test ψ with test $\hat{t}(\theta)$, without introducing any new deviations. We perform this replacement simultaneously for every type to construct an incentive-compatible canonical profile with testing rule $\hat{t}$.

The implication from condition 2 to condition 1 means that the most-discerning property is necessary. If a testing function $\hat{t}$ is not most discerning, then in some decision environment there is an implementable social choice function that cannot be canonically implemented with testing rule $\hat{t}$. The construction is as follows. Since $\hat{t}$ is not most-discerning, there exists some type θ and some test τ such that $\hat{t}(\theta) \not\succeq_\theta \tau$. We start with a social choice function that can be implemented by assigning test τ to type θ. To replace test τ with test $\hat{t}(\theta)$, we would need a score conversion from τ to $\hat{t}(\theta)$ that preserves the decision for type θ. But any such conversion is either nonmonotone or violates the dominance condition (ii). If the conversion is nonmonotone, then type θ can improve his passage probability by not exerting effort. If (ii) is violated, then there is some other type θ' whose score distribution after reporting θ is improved by the conversion. In the proof, we construct a decision environment in which these deviations are profitable.

The most-discerning property takes a simple form when tests are deterministic.

Example 2 (Deterministic tests and evidence)**.** If the passage rate π is $\{0, 1\}$-valued, then our testing framework reduces to Bull and Watson's (2007) model of hard evidence. Each test can be interpreted as a request for a piece of evidence. Define a correspondence $E \colon \Theta \twoheadrightarrow T$ by

$$E(\theta) = \{\tau \in T : \pi(\tau|\theta) = 1\}.$$

Type θ can provide the evidence requested by test τ if and only if τ is in $E(\theta)$. It can be

shown that[10] a test τ in $E(\theta)$ is most θ-discerning if and only if, for every test ψ in $E(\theta)$,

$$E^{-1}(\tau) \subseteq E^{-1}(\psi).$$

That is, test τ is the hardest test that type θ can pass. The existence of a most-discerning testing function is equivalent to Bull and Watson's (2007) evidentiary normality condition, which in turn is equivalent to Lipman and Seppi's (1995) full reports condition.

If a most-discerning testing function does not exist, we can still reduce the class of testing rules that we need to consider.

Definition 6 (Most-discerning correspondence). A subset T_0 of T is *most θ-discerning* if for each test $\psi \in T$ there exists a test $\tau \in T_0$ such that $\tau \succeq_\theta \psi$. A correspondence $\hat{T} : \Theta \twoheadrightarrow T$ is *most discerning* if for each type θ the set $\hat{T}(\theta)$ is most θ-discerning.

A testing function $\hat{t} : \Theta \to T$ is a *selection* from a correspondence $\hat{T} : \Theta \twoheadrightarrow T$ if $\hat{t}(\theta) \in \hat{T}(\theta)$ for each $\theta \in \Theta$. We extend this notion to stochastic testing rules. A testing rule $\hat{t} : \Theta \to \Delta(T)$ is *supported on* a correspondence $\hat{T} : \Theta \twoheadrightarrow T$ if $\text{supp}\,\hat{t}_\theta \subset \hat{T}(\theta)$ for each $\theta \in \Theta$. The next result says that if a correspondence is most discerning, then we can restrict attention to the testing rules supported on that correspondence. To avoid measurability problems, we impose additional regularity conditions.

Theorem 3 (Implementation with a most-discerning correspondence)
Suppose that the passage rate π is continuous. Let $\hat{T}$ be a correspondence from Θ to T with closed values and a measurable graph. If $\hat{T}$ is most discerning, then for every implementable social choice function f, there exists a testing rule $\hat{t}$ supported on $\hat{T}$ such that f is canonically implementable with $\hat{t}$.

If $\hat{T}$ is singleton-valued, then exactly one testing rule is supported on $\hat{T}$, and Theorem 3 reduces to the sufficiency part of Theorem 2.[11] In general, Theorem 3 allow us to restrict attention to testing rules supported on a fixed correspondence $\hat{T}$. For example, suppose each type θ in some subset Θ_0 of Θ has a most θ-discerning test $\hat{t}(\theta)$. Take $\hat{T}(\theta) = \{\hat{t}(\theta)\}$ for $\theta \in \Theta_0$ and $\hat{T}(\theta) = T$ for $\theta \notin \Theta_0$.

We apply Theorem 3 to the the Green–Laffont example, which we reformulate with tests.

---

[10] Argue directly from the definition using deterministic score conversions. Alternatively, apply Proposition 2, stated below.

[11] In this case, the continuity assumption on π is unnecessary. We use the regularity assumptions only to apply a measurable selection theorem, which ensures that there is a measurable way to assign each pair (θ, ψ) a test τ in $\hat{T}(\theta)$ satisfying $\tau \succeq_\theta \psi$. If it can be shown independently that such an assignment exists, then the conclusion of Theorem 3 follows without any assumptions on π or $\hat{T}$. When $\hat{T}$ is singleton-valued, there is at most one such assignment.
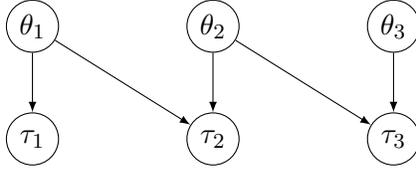
Figure 3. Passage correspondence

Example 3 (Green–Laffont with tests). Set $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and $T = \{\tau_1, \tau_2, \tau_3\}$. Type θ_i can pass test τ_j if and only if, in Example 1, type θ_i can report θ_j. This $\{0, 1\}$-valued passage rate is represented in Figure 3 as a directed graph on $\Theta \cup T$.[12] Edges connect each type to each of the tests he can pass. The discernment relations are given by

$$\tau_1 \succ_{\theta_1} \tau_2 \succ_{\theta_1} \tau_3, \qquad \tau_2, \tau_3 \succ_{\theta_2} \tau_1, \qquad \tau_3 \succ_{\theta_3} \tau_2 \sim_{\theta_3} \tau_1.$$

Tests τ_2 and τ_3 are incomparable under $\succeq_{\theta_2}$ because τ_2 screens away type θ_3 (but not θ_1) and τ_3 screens away θ_1 (but not θ_3). The following correspondence $\hat{T}$ is most-discerning:

$$\hat{T}(\theta_1) = \{\tau_1\}, \qquad \hat{T}(\theta_2) = \{\tau_2, \tau_3\}, \qquad \hat{T}(\theta_3) = \{\tau_3\}.$$

By Theorem 3, there is no loss in restricting to testing rules supported on $\hat{T}$. But we cannot assume that type θ_2 is assigned test τ_2—this is the crux of the original counterexample.

4.3 Characterizing the discernment orders

We provide a more practical characterization of each discernment order. Fix a type θ and tests τ and ψ. To determine whether τ is more θ-discerning than ψ, we first characterize the monotone Markov transitions k satisfying $\pi_{\tau|\theta} k = \pi_{\psi|\theta}$. We parameterize these transitions as convex combinations of two extreme points.

The first extreme point is obtained by matching quantiles, much like scores are converted between the SAT and the ACT. Figure 4 shows this Markov transition, separated into two cases. In the left panel, $\pi(\tau|\theta) \geq \pi(\psi|\theta)$, so a score of 0 on τ is never converted to a 1 on ψ. In the right panel, $\pi(\tau|\theta) < \pi(\psi|\theta)$, so a score of 1 on τ is never converted to a 0 on ψ.

To formally define this transition, we construct Markov transitions that are analogues of distribution and quantile functions. Given a type θ and a test τ, the associated *cumulative distribution transition*

$$\tilde{F}_{\tau|\theta} \colon \{0, 1\} \to \Delta([0, 1])$$

maps 0 and 1 to the uniform distributions over $[0, 1 - \pi(\tau|\theta)]$ and $[1 - \pi(\tau|\theta), 1]$, respectively.

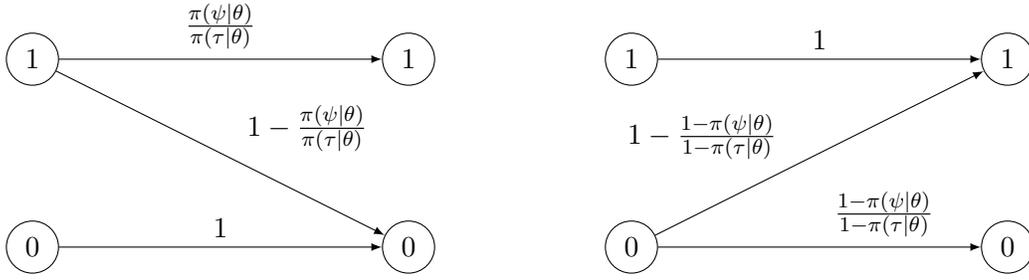[12]Compare this graph on $\Theta \cup T$ with the graph on Θ in Figure 2.

Figure 4. Conversion from test τ to test ψ for type θ

The associated *quantile transition*

$$\tilde{Q}_{\tau|\theta} \colon [0,1] \to \Delta(\{0,1\})$$

maps points in $[0, 1 - \pi(\tau|\theta)]$ to δ_0 and points in $(1 - \pi(\tau|\theta), 1]$ to δ_1.[13] The quantile-matching transition is the composition $\tilde{F}_{\tau|\theta}\tilde{Q}_{\psi|\theta}$.

The second extreme point is the constant transition that maps both scores 0 and 1 to the measure $\pi_{\psi|\theta}$. This transition is denoted $\pi_{\psi|\theta}$ as well.

Proposition 1 (Score conversion characterization)
Fix a type θ and tests τ and ψ. For a Markov transition k on $\{0,1\}$, the following are equivalent.

1. *k is monotone and $\pi_{\tau|\theta}k = \pi_{\psi|\theta}$.*
2. *$k = \lambda\tilde{F}_{\tau|\theta}\tilde{Q}_{\psi|\theta} + (1 - \lambda)\pi_{\psi|\theta}$ for some $\lambda \in [0,1]$.*

We characterize the discernment order in terms of this parameter $\lambda \in [0,1]$. For any passage rate π, define the associated failure rate $\bar{\pi}$ by $\bar{\pi} = 1 - \pi$.

Proposition 2 (Discernment order characterization)
Fix a type θ and tests τ and ψ.

1. *Suppose $\pi(\tau|\theta) \geq \pi(\psi|\theta)$. We have $\tau \succeq_\theta \psi$ if and only if there exists $\lambda \in [0,1]$ such that, for all types θ',*

$$[\lambda\pi(\tau|\theta') + (1 - \lambda)\pi(\tau|\theta)]\pi(\psi|\theta) \leq \pi(\psi|\theta')\pi(\tau|\theta). \tag{3}$$

2. *Suppose $\pi(\tau|\theta) < \pi(\psi|\theta)$. We have $\tau \succeq_\theta \psi$ if and only if there exists $\lambda \in [0,1]$ such that, for all types θ',*

$$[\lambda\bar{\pi}(\tau|\theta') + (1 - \lambda)\bar{\pi}(\tau|\theta)]\bar{\pi}(\psi|\theta) \geq \bar{\pi}(\psi|\theta')\bar{\pi}(\tau|\theta). \tag{4}$$

---

[13]To handle an edge case in Proposition 1, we redefine $\tilde{Q}_{\tau|\theta}$ to map 1 to δ_1 even if $\pi(\tau|\theta) = 0$. For more general definitions of these transitions and for some of their properties, see Appendix A.2.
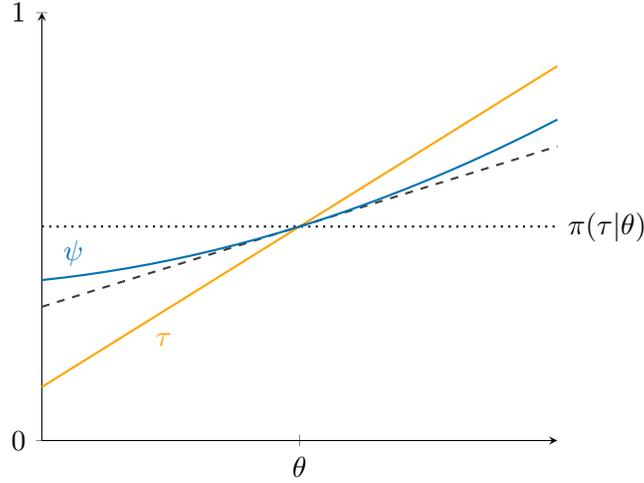
Figure 5. More θ-discerning with $\lambda \neq 1$

Remark. If $\pi(\tau|\theta) \geq \pi(\tau|\theta')$ for all θ', then (3) and (4) are each weakest when $\lambda = 1$, so we can equivalently require $\lambda = 1$ in the statement of Proposition 2. If, in addition, the passage rates are interior, then (3) and (4) can be expressed as

$$\frac{\pi(\tau|\theta)}{\pi(\tau|\theta')} \geq \frac{\pi(\psi|\theta)}{\pi(\psi|\theta')} \quad \text{and} \quad \frac{\bar{\pi}(\tau|\theta)}{\bar{\pi}(\tau|\theta')} \leq \frac{\bar{\pi}(\psi|\theta)}{\bar{\pi}(\psi|\theta')}.$$

For the relation $\tau \succeq_\theta \psi$, the passage (failure) rate ratio is what matters if type θ is more likely to pass (fail) test τ than test ψ.

Example 4 (More θ-discerning with $\lambda \neq 1$)**.** For simplicity, we consider a type θ and two tests τ and ψ such that $\pi(\tau|\theta)$ and $\pi(\psi|\theta)$ are equal and nonzero. In this case, test τ is more θ-discerning than test ψ if and only if there exists $\lambda \in [0,1]$ such that

$$\lambda \pi(\tau|\theta') + (1-\lambda)\pi(\tau|\theta) \leq \pi(\psi|\theta') \quad \text{for all } \theta' \in \Theta. \tag{5}$$

The passage rates for these tests are plotted in Figure 5.[14] The type space is an interval, plotted on the horizontal axis. For test τ, the passage rate is an increasing affine function; for test ψ, the passage rate is increasing and convex. The dotted line takes the constant value $\pi(\tau|\theta)$, and the dashed line is the average of the passage rate $\pi(\tau|\cdot)$ and the constant $\pi(\tau|\theta)$. From the graph we see that (5) is satisfied with $\lambda = 1/2$, so $\tau \succeq_\theta \psi$. Moreover, the tangency at the point $(\theta, \pi(\tau|\theta))$ shows that $1/2$ is the only value of λ for which (5) holds.

Example 5 (Relative performance and scaling tests)**.** Consider a type θ and tests τ, ψ_1,

---

[14]Algebraically, $\Theta = [1/4, 3/4]$ and $\theta = 1/2$. The passage rates are $\pi(\tau|\theta') = 1/2 + (3/4)(\theta' - 1/2)$ and $\pi(\psi|\theta') = \theta'(\theta' - 1/4) + 3/8$.
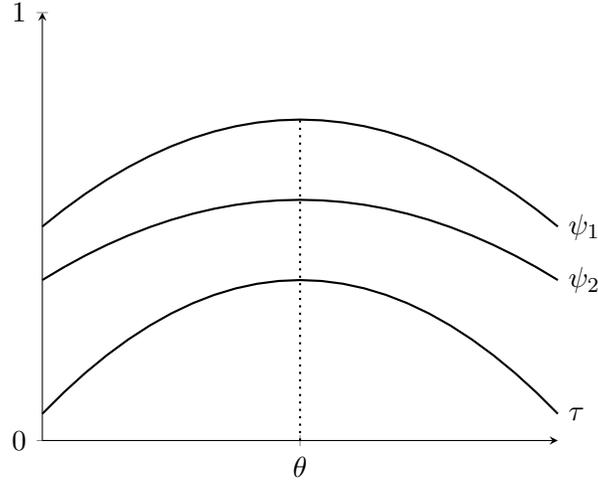
15

Figure 6. Relative performance and scaling

and ψ_2 whose passage rates are plotted in Figure 6.[15] Test τ is the test that type θ is least likely to pass, but τ is more θ-discerning than ψ_1 and ψ_2. This is possible because test τ is difficult for every type. The performance of type θ *relative to the other types* is better on test τ than on tests ψ_1 and ψ_2.

This example illustrates also the subtle effect of scaling the passage rate. The passage rate on ψ_2 is a scaling of the passage rate on ψ_1: $\pi(\psi_2|\cdot) = 0.75\pi(\psi_1|\cdot)$. Since $\pi(\psi_1|\theta) > \pi(\psi_2|\theta)$, the relation $\psi_1 \succeq_\theta \psi_2$ depends on the relative passage rates; it is satisfied. The reverse relation depends on the relative failure rates; it is not satisfied. Thus, $\psi_1 \succ_\theta \psi_2$.

Lastly, we study equivalence with respect to each discernment order. Two tests are *θ-equivalent* if they are equivalent with respect to $\succeq_\theta$. Two tests are *equal* if their passage rates are equal. A type θ is *minimal* on a test τ if $\pi(\tau|\theta) \leq \pi(\tau|\theta')$ for all $\theta' \in \Theta$.

Proposition 3 (θ-discernment equivalence)
Fix a type θ. Tests τ_1 and τ_2 are θ-equivalent if and only if (i) τ_1 and τ_2 are equal, or (ii) θ is minimal on τ_1 and τ_2.

If two unequal tests are θ-equivalent, then these tests have no power to screen other types away from type θ. Whichever of the two tests is used, every decision that is feasible for type θ is also feasible for all other types. Proposition 3 considers θ-equivalence for a fixed type θ. If two tests are θ-equivalent for every type θ, then their passage rates are constant, but the constants are not necessarily equal.

---

[15] Again, $\Theta = [1/4, 3/4]$ and $\theta = 1/2$. The passage rates are $\pi(\psi_1|\theta') = 0.75 - 4(\theta' - 0.5)^2$; $\pi(\psi_2|\theta') = 0.75(0.75 - 4(\theta' - 0.5)^2)$; and $\pi(\tau|\theta') = 0.375 - 5(\theta' - 0.5)^2$.

16

5 Testing in reduced form

If there exists a most-discerning testing function, then our framework takes a reduced form in which tests do not appear explicitly.

5.1 Induced authentication rates

When the agent makes a type report, his passage probability depends on the test that the principal conducts after that report. If there is a most-discerning testing function $\hat{t}$, then by Theorem 2 we can restrict attention to mechanisms in which each report θ' is assigned test $\hat{t}(\theta')$. In this case, the passage probabilities for each report are pinned down for each type.

Definition 7 (Induced authentication rate)**.** Given a most-discerning testing function $\hat{t}\colon \Theta \to T$, the *authentication rate induced by* $\hat{t}$ is the function $\alpha\colon \Theta \times \Theta \to [0,1]$ given by

$$\alpha(\theta'|\theta) = \pi(\hat{t}(\theta')|\theta).$$

There can be multiple most-discerning testing functions. Each induces a different authentication rate. But Theorem 2 guarantees that every most-discerning testing function can be used to implement the same set of social choice functions. There is a corresponding equivalence between the authentication rates induced by different most-discerning testing functions. We need a few definitions. A testing environment is *most discerning* if it admits a most-discerning testing function. A type θ is *minimal* for an authentication rate α if $\alpha(\theta|\theta) \leq \alpha(\theta|\theta')$ for all $\theta' \in \Theta$. Authentication rates α_1 and α_2 are *essentially equal* if (i) α_1 and α_2 have the same minimal types, and (ii) $\alpha_1(\theta|\cdot) = \alpha_2(\theta|\cdot)$ for all types θ that are not minimal for α_1 and α_2. Essential equality is an equivalence relation.

Proposition 4 (Essential uniqueness)
In a most-discerning testing environment, the authentication rates induced by most-discerning testing functions are all essentially equal.

Hereafter, we identify authentication rates that are essentially equal, so we speak of *the* authentication rate induced by a most-discerning testing environment.

5.2 Incentive compatibility

Given a most-discerning testing environment (T, π), we reformulate the principal's problem in terms of the authentication rate α induced by (T, π). When the agent reports θ', he is *authenticated* if he passes the associated most θ'-discerning test. A *reduced-form mechanism* consists of functions g_0 and g_1 from Θ to $\Delta(X)$. When the agent reports θ', the principal

takes the decision $g_1(\theta')$ if the agent is authenticated and the decision $g_0(\theta')$ otherwise. If type θ reports θ' and exerts effort on the associated test, his interim utility $u(\theta'|\theta)$ is given by

$$u(\theta'|\theta) = \alpha(\theta'|\theta)u(g_1(\theta'), \theta) + (1 - \alpha(\theta'|\theta))u(g_0(\theta'), \theta).$$

On the right side, the function u is extended linearly from X to $\Delta(X)$. Even in this reduced form, the cost of lying is determined endogenously by the mechanism g, in contrast to models of lying costs.[16]

The incentive-compatibility constraint becomes

$$u(\theta|\theta) \geq u(\theta'|\theta) \vee u(g_0(\theta'), \theta) \quad \text{for all } \theta, \theta' \in \Theta. \tag{IC}$$

The right side is the interim utility for type θ if he reports θ' and then chooses effort to maximize his utility. The principal selects a reduced-form mechanism to solve

$$\text{maximize} \quad \mathbf{E}[\alpha(\theta|\theta)v(g_1(\theta), \theta) + (1 - \alpha(\theta|\theta))v(g_0(\theta), \theta)]$$
$$\text{subject to} \quad \text{(IC)}.$$

In the applications below, we also impose participation constraints.

5.3 Primitive authentication rates

We showed that a most-discerning testing environment has a simpler representation as an authentication rate. Can we start with the authentication rate as a primitive? To retain the testing interpretation, a primitive authentication rate must be induced by a most-discerning testing environment. We characterize when this is the case.

An authentication rate α implicitly associates to each report θ' a test $\hat{t}(\theta')$ with passage rate $\pi(\hat{t}(\theta')|\cdot) = \alpha(\theta'|\cdot)$. To check whether this testing function $\hat{t}$ is most discerning, we must specify which other tests are in the test set. We claim that there is no less in choosing the minimal test set $\hat{t}(\Theta) = \{\hat{t}(\theta') : \theta' \in \Theta\}$. If $\hat{t}$ is not most-discerning with this test set, then it cannot be most-discerning with any larger test set because adding tests adds constraints to Definition 4. We translate this condition—that the testing function $\hat{t}$ is most discerning with the test set $\hat{t}(\Theta)$—into a condition on the authentication rate α. For an authentication rate α, let $\bar{\alpha} = 1 - \alpha$.

---

[16]In models of lying costs, the agent's utility is the difference between his consumption utility and an exogenous lying cost, which depends on the agent's true type and the agent's report. Lying costs relax the incentive constraints. See, for example, Lacker and Weinberg (1989), Maggi and Rodriguéz-Clare (1995), Crocker and Morgan (1998), Kartik et al. (2007), Kartik (2009), and Deneckere and Severinov (2017). In computer science, Kephart and Conitzer (2016) show that, with lying costs, the revelation principle holds if the lying cost function satisfies the triangle inequality.

Definition 8 (Most-discerning authentication)**.** An authentication rate α is *most discerning* if the following hold for all types θ_2 and θ_3.

1. If $\alpha(\theta_2|\theta_2) \geq \alpha(\theta_2|\theta_3)$, then there exists $\lambda \in [0,1]$ such that, for all types θ_1,

$$[\lambda\alpha(\theta_2|\theta_1) + (1-\lambda)\alpha(\theta_2|\theta_2)]\alpha(\theta_3|\theta_2) \leq \alpha(\theta_3|\theta_1)\alpha(\theta_2|\theta_2).$$

2. If $\alpha(\theta_2|\theta_2) < \alpha(\theta_2|\theta_3)$, then there exists $\lambda \in [0,1]$ such that, for all types θ_1,

$$[\lambda\bar{\alpha}(\theta_2|\theta_1) + (1-\lambda)\bar{\alpha}(\theta_2|\theta_2)]\bar{\alpha}(\theta_3|\theta_2) \geq \bar{\alpha}(\theta_3|\theta_1)\bar{\alpha}(\theta_2|\theta_2).$$

From our characterization of the discernment order (Proposition 2), we get the following characterization for authentication rates.

Theorem 4 (Authentication rate characterization)
An authentication rate α is induced by some most-discerning testing environment if and only if α is most discerning.

Remark. If $\alpha(\theta|\theta) \geq \alpha(\theta|\theta')$ for all types θ and θ', then α is most discerning if and only if

$$\alpha(\theta_3|\theta_2)\alpha(\theta_2|\theta_1) \leq \alpha(\theta_3|\theta_1)\alpha(\theta_2|\theta_2), \tag{6}$$

for all $\theta_1, \theta_2, \theta_3 \in \Theta$.

If α is $\{0,1\}$-valued and $\alpha(\theta|\theta) = 1$ for all θ, then α induces a message correspondence $M \colon \Theta \twoheadrightarrow \Theta$ defined by

$$M(\theta) = \{\theta' : \alpha(\theta'|\theta) = 1\}.$$

This correspondence M satisfies $\theta \in M(\theta)$ for each θ, as in Green and Laffont (1986). In terms of M, (6) becomes

$$\theta_3 \in M(\theta_2) \quad \& \quad \theta_2 \in M(\theta_1) \implies \theta_3 \in M(\theta_1),$$

which is exactly Green and Laffont's (1986) nested range condition.

6 Applications to revenue-maximization

We solve for revenue-maximizing mechanisms with the local first-order approach. The solutions use a new expression for the virtual value.

6.1 Quasilinear setting with verification

Consider the nonlinear pricing setting from Mussa and Rosen (1978). The agent's type $\theta \in \Theta = [\underline{\theta}, \bar{\theta}]$ is drawn from a distribution function F with strictly positive density f. The principal allocates a quantity $q \in \mathbf{R}_+$ and receives a transfer $t \in \mathbf{R}$.[17] Utilities for the agent and the principal are

$$u(q, t, \theta) = \theta q - t \quad \text{and} \quad v(q, t) = t - c(q).$$

Here, c is the cost of production. Assume that $c(0) = c'(0) = 0$ and that the marginal cost c' is strictly increasing and unbounded.

The verification technology is represented by a measurable most-discerning authentication rate $\alpha \colon \Theta \times \Theta \to [0, 1]$ that satisfies the following conditions.

(i) $\alpha(\theta|\theta) = 1$ for all types θ.

(ii) For each $\theta' \in \Theta$, the function $\theta \mapsto \alpha(\theta'|\theta)$ is absolutely continuous.

(iii) For each $\theta \in \Theta$, the right and left partial derivatives $D_{2+}\alpha(\theta|\theta)$ and $D_{2-}\alpha(\theta|\theta)$ exist, and the functions $\theta \mapsto D_{2+}\alpha(\theta|\theta)$ and $\theta \mapsto D_{2-}\alpha(\theta|\theta)$ are integrable.[18]

Condition (i) ensures that the agent is authenticated if he reports truthfully. The regularity conditions (ii) and (iii) allow us to apply the envelope theorem. Since α is most discerning, (i) implies that $\alpha(\theta_3|\theta_2)\alpha(\theta_2|\theta_1) \leq \alpha(\theta_3|\theta_1)$ for all $\theta_1, \theta_2, \theta_3 \in \Theta$. Figure 7 plots an authentication rate that satisfies our assumptions. The agent's type is on the horizontal axis. Each curve corresponds to a fixed report. In this example, the authentication probability decays exponentially in the absolute difference between the agent's type and the agent's report.

We assume that the agent is free to walk away at any time, so we impose an ex post participation constraint. Whether or not the agent is authenticated, his utility must be nonnegative.[19] Without these constraints, the principal could apply severe punishments to effectively prohibit the agent from making any report that is not authenticated with certainty. In that case, the model reduces to partial verification, as in Caragiannis et al. (2012).

We work with reduced-form mechanisms. Since the agent is always authenticated on-path, there is no loss in holding him to his outside option if he is not authenticated. We take $g_0(\theta) = (0, 0)$ for all θ, and we optimize over the decision rule g_1. Without loss, we

---

[17] The pair (q, t) is the decision x in the general model. In applications, t always denotes transfers, *not* testing. Since we work directly with authentication rates, we make no reference to tests.

[18] These derivatives are defined by

$$D_{2+}\alpha(\theta'|\theta) = \lim_{h\downarrow 0} \frac{\alpha(\theta'|\theta + h) - \alpha(\theta'|\theta)}{h}, \qquad D_{2-}\alpha(\theta'|\theta) = \lim_{h\downarrow 0} \frac{\alpha(\theta'|\theta) - \alpha(\theta'|\theta - h)}{h}.$$

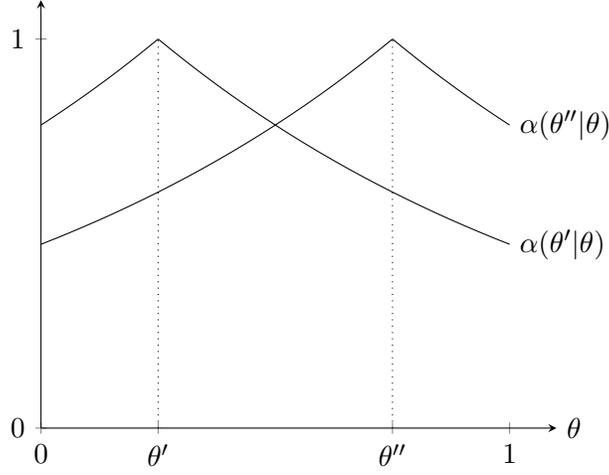[19] In particular, we rule out upfront payments like those used in Border and Sobel (1987).

Figure 7. Exponential authentication rate

restrict g_1 to be deterministic. The component functions of g_1 are denoted q and t.

The principal selects a quantity function $q\colon \Theta \to \mathbf{R}_+$ and a transfer function $t\colon \Theta \to \mathbf{R}$ to solve

$$
\begin{aligned}
\text{maximize} \quad & \int_{\underline{\theta}}^{\bar{\theta}} [t(\theta) - c(q(\theta))] f(\theta) \, \mathrm{d}\theta \\
\text{subject to} \quad & \theta q(\theta) - t(\theta) \geq \alpha(\theta'|\theta)[\theta q(\theta') - t(\theta')], \quad \theta, \theta' \in \Theta \\
& \theta q(\theta) - t(\theta) \geq 0, \quad \theta \in \Theta.
\end{aligned}
\tag{7}
$$

The first constraint is incentive compatibility. The second is ex post participation, conditional upon being authenticated. The utility $u(\theta'|\theta)$ takes a simple form because the agent gets zero utility if he is not authenticated. The maximum operation from (IC) is dropped because it is subsumed by the participation constraint.

6.2 Virtual value

To motivate our new expression for the virtual value, we use the envelope theorem to compute the agent's equilibrium utility function U, defined by

$$
U(\theta) = u(\theta|\theta) = \max_{\theta' \in \Theta} u(\theta'|\theta).
$$

In the classical setting without verification, $u(\theta'|\theta) = \theta q(\theta') - t(\theta')$. By the envelope theorem, $U'(\theta) = D_2 u(\theta|\theta) = q(\theta)$. Integrating gives

$$
U(\theta) = \int_{\underline{\theta}}^{\theta} q(z) \, \mathrm{d}z.
\tag{8}
$$

With verification, the equilibrium utility U takes a different form. We sketch the derivation here. From (7), the interim utility is given by

$$u(\theta'|\theta) = \alpha(\theta'|\theta)[\theta q(\theta') - t(\theta)].$$

The envelope theorem gives the bounds

$$q(\theta) + D_{2+}\alpha(\theta|\theta)U(\theta) \leq U'(\theta) \leq q(\theta) + D_{2-}\alpha(\theta|\theta)U(\theta).$$

Since $\alpha(\theta|\theta) = 1$, we have $D_{2+}\alpha(\theta|\theta) \leq 0 \leq D_{2-}\alpha(\theta|\theta)$. If α has a cusp, as in the example in Figure 7, these inequalities are strict and the derivative of U is not pinned down. To maximize the principal's revenue, we set $U'(\theta)$ equal to the lower bound. Define the precision function $\lambda \colon \Theta \to \mathbf{R}_+$ by

$$\lambda(\theta) = -D_{2+}\alpha(\theta|\theta).$$

The larger is $\lambda(\theta)$, the steeper is the function $\alpha(\theta|\cdot)$ to the right of θ. For $\theta' \leq \theta$, let

$$\Lambda(\theta'|\theta) = \exp\left(\int_{\theta'}^{\theta} -\lambda(w)\,\mathrm{d}w\right).$$

The minimum equilibrium utility U is given by

$$U(\theta) = \int_{\underline{\theta}}^{\theta} \Lambda(z|\theta)q(z)\,\mathrm{d}z. \tag{9}$$

With this expression for equilibrium utility, we define the virtual value. Recall Myerson's (1981) *virtual value*

$$\varphi^M(\theta) = \theta - \frac{1 - F(\theta)}{f(\theta)} = \theta - \frac{1}{f(\theta)}\int_{\theta}^{\bar{\theta}} f(z)\,\mathrm{d}z. \tag{10}$$

The virtual value of type θ is the marginal expected revenue with respect to the quantity $q(\theta)$. There are two parts. First, the principal can extract the additional consumption utility from type θ, so the marginal revenue from type θ equals θ. Second, the quantity $q(\theta)$ pushes up the equilibrium utility according to (8), so the marginal revenue from each type z above θ is -1; this effect is integrated against the relative density $f(z)/f(\theta)$. Verification does not change the marginal revenue from type θ, but the marginal revenue from each higher type z becomes $-\Lambda(\theta|z)$, by (9). We define the *virtual value* by

$$\varphi(\theta) = \theta - \frac{1}{f(\theta)}\int_{\theta}^{\bar{\theta}} \Lambda(\theta|z)f(z)\,\mathrm{d}z. \tag{11}$$
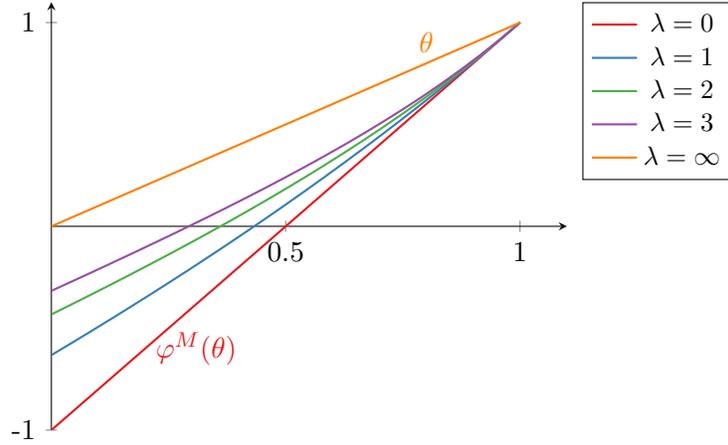
Figure 8. Virtual value for different precision functions λ

Comparing (10) and (11) gives the inequality

$$\varphi^M(\theta) \leq \varphi(\theta) \leq \theta.$$

The virtual value $\varphi(\theta)$ tends towards these bounds in limiting cases.

Proposition 5 (Testing precision)
As λ converges to 0 pointwise, $\varphi(\theta)$ converges to $\varphi^M(\theta)$ for each type θ. As λ converges to ∞ pointwise, $\varphi(\theta)$ converges to θ for each type θ.

Figure 8 illustrates these limits in a simple example, where the agent's type is uniformly distributed on the unit interval and the precision function λ is constant.

Remark. If $\lambda(\theta) = 0$ for all θ, then $\Lambda(\theta|z) = 1$ for $\theta \leq z$. Therefore, our virtual value coincides with the classical virtual value, and by the results below, the revenue-maximizing mechanism is unaffected by verification. This holds in particular if α has no kink on the diagonal, e.g., if $\alpha(\theta'|\theta) = 1 - |\theta' - \theta|^\sigma$ with $\sigma > 1$.

6.3 Nonlinear pricing

The virtual value is derived from the envelope representation of the equilibrium utility, which uses only local incentive constraints. We assume that the virtual value is increasing. But because the interim utility is not linear in the agent's type (due to the authentication rate α), we need further assumptions to ensure that the local incentive constraints imply the global incentive constraints. This implication holds in particular for the *exponential*

authentication rates, given by

$$\alpha(\theta'|\theta) = \exp\left(-\int_{\theta'}^{\theta} \lambda(z)\,\mathrm{d}z\right),$$

for integrable functions $\lambda\colon \Theta \to \mathbf{R}_+$. We permit a larger class of authentication rates.

We impose a global condition on the relative values of α and Λ. The function Λ is determined by the behavior of α in a neighborhood of the diagonal in $\Theta \times \Theta$. Because α is most discerning, it follows that Λ is a global lower bound for α.

Proposition 6 (Lower bound)
For all types θ' and θ with $\theta' \leq \theta$, we have $\alpha(\theta'|\theta) \geq \Lambda(\theta'|\theta)$.

For the exponential authentication rates, this inequality holds with equality. We require that α not be much greater than Λ. The precise condition depends on the optimal quantity function $q^\star$, which will be defined in the theorem statement. The *global upper bound* states that, for all types θ' and θ with $\theta' \leq \theta$,

$$\alpha(\theta'|\theta) \leq \Lambda(\theta'|\theta)A(\theta'|\theta),$$

where

$$A(\theta'|\theta) = \int_{\underline{\theta}}^{\theta} \Lambda(z|\bar{\theta})q^\star(z)\,\mathrm{d}z \Big/ \int_{\underline{\theta}}^{\theta} \Lambda(z \wedge \theta'|\bar{\theta})q^\star(z \wedge \theta')\,\mathrm{d}z.$$

For the quantity function $q^\star$ in the theorem statement, $A(\theta'|\theta) \geq 1$ for $\theta' \leq \theta$.

Proposition 7 (Optimal nonlinear pricing)
Suppose that the virtual value φ is increasing. The optimal quantity function $q^\star$ and transfer function $t^\star$ are unique and given by

$$c'(q^\star(\theta)) = \varphi(\theta)_+, \qquad t^\star(\theta) = \theta q^\star(\theta) - \int_{\underline{\theta}}^{\theta} \Lambda(z|\theta)q^\star(z)\,\mathrm{d}z,$$

provided that the global upper bound is satisfied.

In the optimal mechanism, type θ receives the quantity that is efficient for type $\varphi(\theta)_+$, just as in Mussa and Rosen (1978), except that φ is our new virtual value. Transfers are pinned down by the equilibrium utility U from (9). The faster the virtual value φ increases, the faster the optimal quantity function $q^\star$ increases and the more permissive is the global upper bound.

6.4 Selling a single good

Suppose that the principal is selling a single indivisible good, which she does not value. The agent's type is his valuation for the good. The principal allocates the good with probability $q \in [0, 1]$ and receives a transfer $t \in \mathbf{R}$. Utilities are

$$u(q, t, \theta) = \theta q - t \quad \text{and} \quad v(q, t) = t.$$

Without verification, the revenue-maximizing mechanisms is a posted price (Riley and Zeckhauser, 1983). With verification, the price my depend on the agent's report.

Proposition 8 (Optimal sale of a single good)
Suppose that the virtual value φ is increasing. The optimal quantity and transfer functions are unique and given as follows, provided that the global upper bound is satisfied. Let $\theta^\star = \inf\{\theta : \varphi(\theta) \geq 0\}$. Each type below $\theta^\star$ receives nothing and pays nothing. Each type θ above $\theta^\star$ receives the good and pays

$$t^\star(\theta) = \theta - \int_{\theta^\star}^\theta \Lambda(z|\theta) \, \mathrm{d}z.$$

As in the no-verification solution, there is a cutoff type $\theta^\star$ who receives the good and pays his valuation. Each type below the cutoff is excluded; each type above the cutoff receives the good and pays less than his valuation. The allocation probability takes values 0 and 1 only—there is no randomization. Verification increases the virtual value relative to the classical virtual value, so the cutoff type is lower and more types receive the good. The price is (weakly) increasing in the agent's report, and strictly increasing if λ is strictly positive. Nevertheless, types above the cutoff cannot profit by misreporting downward—the benefit of a lower price is outweighed by the risk of not being authenticated.

7 Testing multiple agents

7.1 Testing and implementation

We extend our model to allow for n agents, labeled $i = 1, \ldots, n$. Each agent i independently draws his type $\theta_i \in \Theta_i$ from a commonly known distribution $\mu_i \in \Delta(\Theta_i)$. Set $\Theta = \prod_{i=1}^n \Theta_i$. The decision set is denoted by X, as before. Each agent i has utility function $u_i \colon X \times \Theta \to \mathbf{R}$; the principal has utility function $v \colon X \times \Theta \to \mathbf{R}$.

For each agent i, there is a set T_i of tests and a passage rate

$$\pi_i \colon T_i \times \Theta_i \to [0, 1],$$

where $\pi_i(\tau_i|\theta_i)$ is the probability with which type θ_i can pass test τ_i. Set $T = \prod_{i=1}^n T_i$. Each agent sees his own test—but not the tests of the other agents—and then chooses whether to exert effort. Nature draws the test result for each agent independently. A mechanism specifies a message set M_i for each agent i. Set $M = \prod_{i=1}^n M_i$. The rest of the mechanism consists of a testing rule $t\colon M \to \Delta(T)$ and a decision rule $g\colon M \times T \times \{0,1\}^n \to \Delta(X)$. The test conducted on each agent can depend on the messages sent by other agents. For each agent i, a strategy consists of a reporting strategy $r_i\colon \Theta_i \to \Delta(M_i)$ and an effort strategy $e_i\colon \Theta_i \times M_i \times T_i \to [0,1]$. The equilibrium concept is Bayes–Nash equilibrium.

In this multi-agent setting, the revelation principle (Theorem 1) holds. The discernment orders extend. For each agent i and each type θ_i in Θ_i, the θ_i-discernment order $\succeq_{\theta_i}$ over T_i is defined as in the baseline model with π_i in place of π. Given testing functions $t_i\colon \Theta_i \to T_i$ for each i, define the product testing function $\otimes_i t_i\colon \Theta \to T$ by $t(\theta_1, \ldots, \theta_n) = (t_1(\theta_1), \ldots, t_n(\theta_n))$. If for each agent i there is a most-discerning testing function, then the product of these testing functions suffices for implementation. In particular, the test for agent i depends only on agent i's report.

Theorem 5 (Most-discerning implementation with multiple players)
Fix a type space Θ and a testing environment (T, π). For a testing function $\hat{t} = \otimes_i \hat{t}_i$, the following are equivalent.

1. *$\hat{t}_i$ is most discerning for all i.*
2. *In every decision environment (X, u), every implementable social choice function is canonically implementable with $\hat{t}$.*

For applications, we make the same assumptions as in the single-agent case. Each agent is free to walk away, so we impose ex post participation constraints. For each agent i, the testing environment is represented by a measurable most-discerning authentication rate $\alpha_i\colon \Theta_i \times \Theta_i \to [0,1]$ that satisfies assumptions (i)–(iii). For each i, define λ_i and Λ_i by putting α_i in place of α in the definitions of λ and Λ.

7.2 Auctions

Consider an auction for a single indivisible good. Each agent i independently draws his type $\theta_i \in \Theta_i = [\underline{\theta}_i, \bar{\theta}_i]$ from a distribution function F_i with positive density f_i. The principal allocates the good to each agent i with probability $q_i \in [0,1]$, where $q_1 + \cdots + q_n \leq 1$; the principal receives transfers $t_1, \ldots, t_n \in \mathbf{R}$. Set $q = (q_1, \ldots, q_n)$ and $t = (t_1, \ldots, t_n)$. For simplicity, we assume that the principal does not value the good. Utilities are given by

$$u_i(q, t, \theta) = \theta_i q_i - t_i \quad \text{and} \quad v(q, t) = \sum_{i=1}^n t_i.$$

Let $f_{-i}(\theta_{-i})$ denote $\prod_{j \neq i} f_j(\theta_j)$. For quantity functions $q_i \colon \Theta \to \mathbf{R}$, interim expectations are denoted with capital letters:

$$Q_i(\theta_i) = \int_{\Theta_{-i}} q_i(\theta_i, \theta_{-i}) f_{-i}(\theta_{-i}) \, \mathrm{d}\theta_{-i}.$$

As in the single-agent case, we impose a condition that depends on the optimal quantity function $q^\star$, which is defined in Proposition 9. For each agent i, the *global upper bound* states that, for all $\theta_i, \theta_i' \in \Theta_i$ with $\theta_i' \leq \theta_i$, we have

$$\alpha_i(\theta_i'|\theta_i) \leq A_i(\theta_i'|\theta_i)\Lambda_i(\theta_i'|\theta_i),$$

where

$$A_i(\theta_i'|\theta_i) = \int_{\underline{\theta}_i}^{\theta} \Lambda_i(z_i|\bar{\theta}_i)Q_i^\star(z_i) \, \mathrm{d}z_i \Big/ \int_{\underline{\theta}_i}^{\theta} \Lambda_i(z_i \wedge \theta_i'|\bar{\theta}_i)Q_i^\star(z_i \wedge \theta_i') \, \mathrm{d}z_i.$$

For each i and all types $\theta_i, \theta_i' \in \Theta_i$ with $\theta_i' \leq \theta_i$, we have $A_i(\theta_i'|\theta_i) \geq 1$.

Proposition 9 (Optimal auctions)
Suppose that each virtual value φ_i is increasing. The seller's maximum revenue is achieved by the allocation function $q^\star$ and transfer function $t^\star$ given by

$$q_i^\star(\theta) = \begin{cases} 1 & \text{if } \varphi_i(\theta_i) > 0 \vee \max_{j \neq i} \varphi_j(\theta_j), \\ 0 & \text{otherwise,} \end{cases}$$

and

$$t_i^\star(\theta) = q_i^\star(\theta)\left[\theta_i - \int_{\underline{\theta}_i}^{\theta_i} \Lambda_i(z_i|\theta_i)\frac{Q_i^\star(z_i)}{Q_i^\star(\theta_i)} \, \mathrm{d}z_i\right],$$

provided that the global upper bound is satisfied for each agent i.

The transfers are chosen so that each agent pays only if he receives the good, thus ensuring that the the ex post participation constraints are satisfied. The allocation coincides with Myerson's (1981) solution, except φ is our new virtual value. In Myerson's (1981) solution, the allocation rule disadvantages bidders whose valuation distributions are greater in the sense of hazard rate dominance. In our solution, the allocation rule also advantages bidders who can be verified more precisely in the sense that their local precision functions are pointwise greater.

8   Nonbinary tests

In the main model, we consider pass–fail tests because of their natural connection with partial verification and evidence. Here we extend the baseline principal–agent model to allow for tests with more than two results. There is a family T of tests. Each test generates scores in a finite subset S of $\mathbf{R}$. The passage rate

$$\pi \colon T \times \Theta \to \Delta(S)$$

assigns to each pair (τ, θ) the score distribution $\pi_{\tau|\theta}$ in $\Delta(S)$ for type θ on test τ.

   We generalize the agent's effort choice. On a pass–fail test τ, when type θ exerts effort with probability e, he passes with probability $e \cdot \pi(\tau|\theta)$. Thus, there is a correspondence between effort probabilities in $[0,1]$ and passage probabilities in $[0, \pi(\tau|\theta)]$. We can equivalently model the agent as choosing the passage probability directly, subject to a stochastic dominance constraint. This alternative definition of a strategy extends immediately to nonbinary tests. A *performance strategy* is a map

$$p \colon \Theta \times M \times T \to \Delta(S),$$

satisfying $p_{\theta,m,\tau} \leq_{\mathrm{SD}} \pi_{\tau|\theta}$ for all $(\theta, m, \tau) \in \Theta \times M \times T$. The other components of a profile are defined as before with S in place of $\{0,1\}$.

   The revelation principle (Theorem 1) is proved with performance strategies, so it applies to nonbinary tests. To define the discernment orders, say that a Markov transition k on S is *monotone* if $k_r \geq_{\mathrm{SD}} k_s$ whenever $r > s$. Put S in place of $\{0,1\}$ in the definition of most-discerning. As before, a single most-discerning testing function suffices for implementation.

Theorem 6 (Most-discerning implementation with nonbinary tests)
If a testing function $\hat{t} \colon \Theta \to T$ is most discerning, then every implementable social choice function is canonically implementable with $\hat{t}$.

9   Review of verification models

Verification has been modeled in many ways, in both economics and computer science. We organize our discussion around the taxonomy in Table 1, which focuses on the primitives in each model.

   Green and Laffont (1986) introduce **partial verification**. They restrict their analysis to direct mechanisms. Verification is represented as a correspondence $M \colon \Theta \twoheadrightarrow \Theta$ satisfying $\theta \in M(\theta)$ for all $\theta \in \Theta$. Each type θ can report any type θ' in $M(\theta)$. In particular, each

	Reduced form	**Microfoundation**
partial	message correspondence $M : \Theta \twoheadrightarrow \Theta$ Green and Laffont (1986)	evidence correspondence $E : \Theta \twoheadrightarrow \mathcal{E}$ Bull and Watson (2007)
probabilistic	authentication rate $\alpha : \Theta \times \Theta \to [0, 1]$ Caragiannis et al. (2012)	passage rate $\pi : T \times \Theta \to [0, 1]$ our paper

Table 1. Taxonomy of verification models

type can report truthfully. In this framework the revelation principle does not hold,[20] as Green and Laffont (1986) illustrate with a three-type counterexample, which we adapt in Example 1. The revelation principle does hold, however, if the correspondence M satisfies the *nested range condition*, which requires that the relation associated to M is transitive. Without the revelation principle, it is generally difficult to determine whether a particular social choice function is implementable (Nisan and Ronen, 2001; Singh and Wittman, 2001; Fotakis and Zampetakis, 2015; Auletta et al., 2011; Yu, 2011; Rochet, 1987; Vohra, 2011).

Bull and Watson (2004, 2007) and Lipman and Seppi (1995) model verification with **hard evidence**.[21] They introduce an evidence set $\mathcal{E}$ and an evidence correspondence $E : \Theta \twoheadrightarrow \mathcal{E}$. Type θ possesses the evidence in $E(\theta)$; he can present one piece of evidence from $E(\theta)$ to the principal. The evidence environment is *normal* if each type θ has a piece of evidence $e(\theta)$ in $E(\theta)$ that is maximal in the following sense: Every other type θ' who has $e(\theta)$ also has every other piece of evidence in $E(\theta)$. Therefore, type θ' can mimic type θ if and only if $E(\theta')$ contains $e(\theta)$. A normal evidence environment induces an abstract mimicking correspondence that satisfies the nested range condition. Normality is a special case of our most-discerning condition; see Example 2.

In computer science, Caragiannis et al. (2012) and Ferraioli and Ventre (2018) study a reduced-form model of **probabilistic verification** in mechanism design. They restrict their analysis to direct mechanisms, and they specify the probabilities with which each type can successfully mimic each other type. Dziuda and Salas (2018) and Balbuzanov (2019) study a setting without commitment in which these probabilities are constant.[22] Our testing framework microfounds these models of partial verification, provided that the primitive authentication rate is most discerning. If the primitive authentication rate is not

---

[20]Strausz (2016) recovers the revelation principle by modeling verification as a component of the outcome.
[21]Evidence was introduced in games (without commitment) by Milgrom (1981) and Grossman (1981); for recent work on evidence games, see Hart et al. (2017), Ben-Porath et al. (2017), and Koessler and Perez-Richet (2017).
[22]For some fixed $p \in (0, 1)$, the authentication probability $\alpha(\theta' | \theta)$ equals p if $\theta' \neq \theta$ and 1 if $\theta' = \theta$. This authentication satisfies (6) and hence is most discerning.

most-discerning, then it cannot be microfounded by our testing framework.

The computer science literature on probabilistic verification has a different focus than us. Caragiannis et al. (2012) allow the principal to use arbitrarily severe punishments to deter the agent from misreporting.[23] If one type cannot mimic another type perfectly, then he risks being detected and facing a prohibitive fine. Therefore, this setting reduces to partial verification. In our model, the agent can walk away at any time, so punishment is limited.

If the environment is most discerning, tests can also be interpreted as **stochastic evidence**. Each test τ in T corresponds to a request for a particular piece of evidence. The agent is asked to send a cheap talk message to the principal after he learns his payoff type but *before* he learns which evidence is available to him. With probability $\pi(\tau|\theta)$, type θ will have the evidence requested by test τ. Deneckere and Severinov (2008) study a different kind of stochastic evidence. In their model the agent simultaneously learns his payoff type and his set of feasible evidence messages.

In economics, "verification" traditionally means that the principal can learn the agent's type perfectly by taking some action, e.g., paying a fee or allocating a good. This literature began with Townsend (1979) who studied **costly verification** in debt contracts. Ben-Porath et al. (2019) connect costly verification and evidence. When monetary transfers are infeasible, costly verification is often used as a substitute; see Ben-Porath et al. (2014); Erlanson and Kleiner (2015); Halac and Yared (2017); Li (2017); Mylovanov and Zapechelnyuk (2017).

10   Conclusion

We model probabilistic verification as a family of stochastic tests available to the principal. Our testing framework provides a unified generalization of previous verification models. Because verification is noisy, our framework is amenable to the local first-order approach. We illustrate this approach in a few classical revenue-maximization problems. We believe this local approach will make verification tractable in other settings as well.

As the precision of the verification technology varies, our setting continuously interpolates between private information and complete information. Thus we can quantify the value to the principal of a particular verification technology. This is the first step toward analyzing a richer setting in which the principal decides how much to invest in verification.

---

[23]Another difference is that Caragiannis et al. (2012) investigate which allocation rules can be supported by some transfer rule. We view transfers as part of the outcome and we study revenue-maximization.

A   Measure theory

A.1   Markov transitions

This section introduces Markov transitions, which are continuous generalizations of stochastic matrices. For further details, see Kallenberg (2017, Chapter 1).

Definition 9 (Markov transition). Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be measurable spaces. A *Markov transition*[24] from $(X, \mathcal{X})$ to $(Y, \mathcal{Y})$ is a function $k\colon X \times \mathcal{Y} \to [0, 1]$ satisfying:

  (i) for each $x \in X$, the map $B \mapsto k(x, B)$ is a probability measure on $(Y, \mathcal{Y})$;
  (ii) for each $B \in \mathcal{Y}$, the map $x \mapsto k(x, B)$ is a measurable function on $(X, \mathcal{X})$.

A Markov transition k from $(X, \mathcal{X})$ to $(Y, \mathcal{Y})$ is sometimes written as $k\colon X \to \Delta(Y)$. Each measure $k(x, \cdot)$ on Y is denoted k_x, and we write $k_x(B)$ for $k(x, B)$. A Markov transition from $(X, \mathcal{X})$ to $(X, \mathcal{X})$ is called a *Markov transition on* $(X, \mathcal{X})$. When the σ-algebras are clear, we will speak of Markov transitions betweens sets.

We introduce three operations between Markov transitions—composition, products, and outer products.

$$(\mu k)(B) = \int_X \mu(\mathrm{d}x) k(x, B),$$

for all $B \in \mathcal{Z}$.

First we define composition. Let k be a Markov transition from $(X, \mathcal{X})$ to $(Y, \mathcal{Y})$ and ℓ a Markov transition from $(Y, \mathcal{Y})$ to $(Z, \mathcal{Z})$. The *composition* of k and ℓ, denoted $k\ell$, is the Markov transition from $(X, \mathcal{X})$ to $(Z, \mathcal{Z})$ defined by

$$(k\ell)(x, C) = \int_Y k(x, \mathrm{d}y)\, \ell(y, C),$$

for all $x \in X$ and $C \in \mathcal{Z}$. Here the function $y \mapsto \ell(y, C)$ is integrated over Y with respect to the measure k_x. Inside the integral, it is standard to place the measure before the integrand so that the sequencing of the variables mirrors the timing of the process.

Next we define products. Let k be a Markov transition from $(X, \mathcal{X})$ to $(Y, \mathcal{Y})$ as before, and let m be a Markov transition from $(X \times Y, \mathcal{X} \otimes \mathcal{Y})$ to $(Z, \mathcal{Z})$. Here $\mathcal{X} \otimes \mathcal{Y}$ is the product σ-algebra generated by the measurable rectangles $A \times B$ for $A \in \mathcal{X}$ and $B \in \mathcal{Y}$. The *product* of k and m, denoted $k \times m$, is the unique Markov transition from $(X, \mathcal{X})$ to $(Y \times Z, \mathcal{Y} \otimes \mathcal{Z})$ satisfying

$$(k \otimes m)(x, B \times C) = \int_B k(x, \mathrm{d}y)\, m((x, y), C),$$

for all $x \in X$, $B \in \mathcal{Y}$, and $C \in \mathcal{Z}$. Here the function $y \mapsto m((x, y), C)$ is integrated over the set B with respect to the measure k_x. If m is a Markov transition from $(Y, \mathcal{Y})$ to $(Z, \mathcal{Z})$, we use the same notation, with the understanding that the integrand is $m(y, C)$ rather than $m((x, y), C)$.

Finally, we define outer products. Let k_1 be a Markov transition from $(X_1, \mathcal{X}_1)$ to $(Y, \mathcal{Y}_1)$ and k_2 a Markov transition from $(X_2, \mathcal{X}_2)$ to $(Y_2, \mathcal{Y}_2)$. The *outer product* of k_1

---

[24]Markov transitions are also called kernels or Markov/stochastic/probability/transition kernels.

and k_2, denoted $k_1 \otimes k_2$, is the unique Markov transition from $(X_1 \times X_2, \mathcal{X}_1 \otimes \mathcal{X}_2)$ to $(Y_1 \times Y_2, \mathcal{Y}_1 \otimes \mathcal{Y}_2)$ satisfying

$$(k_1 \otimes k_2)((x_1, x_2), B_1 \times B_2) = k_1(x_1, B_1) \cdot k_2(x_2, B_2),$$

for all $x_1 \in X_1$, $x_2 \in X_2$, $B_1 \in \mathcal{Y}_1$, and $B_2 \in \mathcal{Y}_2$. All three operations are associative. This holds trivially for outer products; for composition and products, see Kallenberg (2017, Lemma 1.17, p. 33). We drop parentheses when there is no ambiguity.

A.2 Markov transitions on the real line

The real line $\mathbf{R}$ is endowed with its usual Borel σ-algebra.

First, we define Markov transitions that are analogues of the cumulative distribution and quantile functions. Let μ be a measure on $\mathbf{R}$, and let $F_\mu \colon \mathbf{R} \to [0,1]$ be the associated right-continuous cumulative distribution function. Suppose μ has compact support S. Define the left-continuous quantile function $Q_\mu \colon [0,1] \to S$ by

$$Q_\mu(p) = \inf\{s \in S : F_\mu(s) \geq p\}.$$

The *cumulative distribution transition* associated to μ, denoted $\tilde{F}_\mu$, is the Markov transition from $\mathbf{R}$ to $[0,1]$ that assigns to each point s in $\mathbf{R}$ the uniform measure on $[F_\mu(s-), F_\mu(s)]$, where $F_\mu(s-)$ is the left-limit of F_μ at s. In particular, if F_μ is continuous at s, then $F_\mu(s-) = F_\mu(s)$ and this uniform measure is the Dirac measure $\delta_{F_\mu(s)}$.

The *quantile transition* associated to μ, denoted $\tilde{Q}_\mu$, is the Markov transition from $[0,1]$ to $\mathbf{R}$ that assigns to each number p in $[0,1]$ the Dirac measure $\delta_{Q_\mu(p)}$.

These Markov transitions extend the usual properties of distribution and quantile functions to nonatomic distributions. Let $U_{[0,1]}$ denote the uniform measure on $[0,1]$.

Lemma 1 (Distribution and quantile transitions)**.** For measures μ and ν on $\mathbf{R}$ with compact support, the following hold:
  (i) $\mu \tilde{F}_\mu = U_{[0,1]}$;
  (ii) $U_{[0,1]} \tilde{Q}_\nu = \nu$;
  (iii) $\mu \tilde{F}_\mu \tilde{Q}_\nu = \nu$.

For measures μ and ν on the real line, μ *first-order stochastically dominates* ν, denoted $\mu \geq_{\mathrm{SD}} \nu$, if $F_\mu(x) \leq F_\nu(x)$ for all real x. In particular, first-order stochastic dominance is reflexive.

To state the next results, we make the standing assumption that S is a compact subset of $\mathbf{R}$, endowed with the restriction of the Borel σ-algebra. A Markov transition d on S is *downward* if $d(s, (-\infty, s] \cap S) = 1$ for all $s \in S$.

Lemma 2 (Downward transitions)**.** For measures μ and ν on S, the following are equivalent:
  (i) $\mu \geq_{\mathrm{SD}} \nu$;
  (ii) $\tilde{F}_\mu \tilde{Q}_\nu$ is downward;
  (iii) $\mu d = \nu$ for some downward transition d.

A Markov transition m on S is *monotone* if $s > t$ implies $m_s \geq_{\mathrm{SD}} m_t$, for all $s, t \in S$.

Lemma 3 (Monotone transitions).

   (i) A Markov transition m on S is monotone if and only if $\mu m \geq_{\mathrm{SD}} \nu m$ for all measures μ and ν on S satisfying $\mu \geq_{\mathrm{SD}} \nu$.

  (ii) The composition of monotone Markov transitions is monotone.

A.3   Measurability and universal completions

Our definition of θ-discernment includes an inequality for every type. To ensure that we can select score conversions in a measurable way, we enlarge the Borel σ-algebra to its universal completion, which we introduce here.

    Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be measurable spaces. A function f from X to Y is $\mathcal{X}/\mathcal{Y}$-measurable if $f^{-1}(B)$ is in $\mathcal{X}$ for all B in $\mathcal{Y}$. This condition is written more compactly as $f^{-1}(\mathcal{Y}) \subset \mathcal{X}$. If the σ-algebra $\mathcal{Y}$ is understood, we say f is $\mathcal{X}$-measurable, and if both σ-algebras are understood, we say that f is measurable.

    Let $(X, \mathcal{X}, \mu)$ be a probability space. A set A in $\mathcal{X}$ is a μ-*null set* if $\mu(A) = 0$. The μ-*completion* of σ-algebra $\mathcal{X}$, denoted $\overline{\mathcal{X}}_\mu$, is the smallest σ-algebra that contains every set in $\mathcal{X}$ and every subset of every μ-null set. It is straightforward to check that a subset A of X is a member of $\overline{\mathcal{X}}_\mu$ if and only if there are sets A_1 and A_2 in $\mathcal{X}$ such that $A_1 \subset A \subset A_2$ and $\mu(A_2 \setminus A_1) = 0$. The *universal completion* $\overline{\mathcal{X}}$ of $\mathcal{X}$ is the σ-algebra on X defined by

$$\overline{\mathcal{X}} = \cap_\mu \overline{\mathcal{X}}_\mu,$$

where the intersection is taken over all probability measures on $(X, \mathcal{X})$.

    It is convenient to work with the universal completion because of the following measurable projection theorem (Cohn, 2013, Proposition 8.4.4, p. 264).

Theorem 7 (Measurable projection)
Let $(X, \mathcal{X})$ be a measurable space, Y a Polish space, and C a set in the product σ-algebra $\mathcal{X} \otimes \mathcal{B}(Y)$. Then the projection of C on X belongs to $\overline{\mathcal{X}}$.

    By taking universal completions, we do not lose any Markov transitions.

Lemma 4 (Completing transitions). A Markov transition k from $(X, \mathcal{X})$ to $(Y, \mathcal{Y})$ can be uniquely extended to a Markov transition $\bar{k}$ from $(X, \overline{\mathcal{X}})$ to $(Y, \overline{\mathcal{Y}})$.

    Next we consider the universal completion of a product σ-algebra.

Lemma 5 (Product spaces). For measurable spaces $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$,

$$\overline{\mathcal{X}} \otimes \overline{\mathcal{Y}} \subset \overline{\mathcal{X} \otimes \mathcal{Y}} = \overline{\overline{\mathcal{X}} \otimes \overline{\mathcal{Y}}}.$$

    We adopt the convention that when a Markov transition is defined between Polish spaces, both the domain and codomain are endowed with the universal completions of their Borel σ-algebras. In particular, when we take a product of Markov transitions, we extend this transition. By Lemma 5, it does not matter whether the component transitions are extended first.

    We conclude with one note of caution about universal completions. Let k be a Markov transition k from $(X \times Y, \overline{\mathcal{X} \otimes \mathcal{Y}}$ to $(Z, \mathcal{Z})$. Since $\overline{\mathcal{X} \otimes \mathcal{Y}}$ is generally larger than $\overline{\mathcal{X}} \otimes \overline{\mathcal{Y}}$,

the section k_x may not be a Markov transition. For fixed $x \in X$ and $B \in \mathcal{Z}$, the map $y \mapsto k(x, y, B)$ is defined, but may not be $\mathcal{Y}$-measurable. But if Y is countable, and $\mathcal{Y} = 2^Y$, then measurability is automatic.

A.4 Defining mechanisms and strategies

In the model, we make the following technical assumptions. The sets Θ, T, and X are all Polish spaces. The finite signal space S is endowed with the discrete topology. The Markov transition π is from $(\Theta \times T, \mathcal{B}(\Theta \times T))$ to $(S, \mathcal{B}(S))$. Recall that for Polish spaces Y and Z, we have $\mathcal{B}(Y \times Z) = \mathcal{B}(Y) \otimes \mathcal{B}(Z)$. In a mechanism, the message space M is Polish. Testing rules, decision rules, reporting strategies, and passage strategies are all Markov transitions, with the domain and codomain endowed with the universal completions of their Borel (product) σ-algebras.

B Proofs

B.1 Proof of Theorem 1

We begin with new notation. To avoid confusion in the commutative diagrams below, we denote the message set in a direct mechanism by Θ'. This set is a copy of Θ, but it is helpful to keep the sets Θ and Θ' distinct.

We work with performance strategies, as defined in Section 8. Let f be an implementable social choice function. Select a profile $(M, t, g; r, p)$ that implements f. We construct a direct mechanism $(\hat{t}, \hat{g})$ that canonically implements f.

Let $\hat{t}$ be the composition rt, which is a Markov transition from Θ' to T. By disintegration of Markov transitions (Kallenberg, 2017, Theorem 1.25, p. 39), there is a Markov transition h from $\Theta' \times T$ to M such that $r \times t = \hat{t} \otimes h$.[25]

Define a Markov transition d from $\Theta' \times M \times T \times S$ to S as follows. For each $(\theta', m, \tau) \in \Theta' \times M \times T$, set

$$d_{\theta', m, \tau} = \tilde{F}_{\tau|\theta'} \tilde{Q}_{\theta', m, \tau},$$

where $\tilde{F}_{\tau|\theta'}$ is the distribution Markov transition corresponding to $\pi_{\tau|\theta'}$ and $\tilde{Q}_{\theta', m, \tau}$ is the quantile Markov transition corresponding to $p_{\theta', m, \tau}$; see Appendix A.2 for the definitions. By Lemma 2,

$$\pi_{\tau|\theta'} d_{\theta', m, \tau} = p_{\theta', m, \tau}, \tag{12}$$

and $d_{\theta', m, \tau}$ is downward because $p_{\theta', m, \tau} \leq_{\mathrm{SD}} \pi_{\tau|\theta'}$.

Define the direct decision rule $\hat{g}$ as the composition shown in the following commutative diagram, where j denotes the identity transition from Θ to Θ'.[26]

---

[25]We cannot directly apply the result to the universal completions of the Borel σ-algebras. Argue as follows. First, restrict $r \times t$ to a Markov transition from $(\Theta, \overline{\mathcal{B}(\Theta')})$ to $(M \times T, \mathcal{B}(M) \otimes \mathcal{B}(T))$ and $\hat{t}$ to a Markov transition from $(\Theta', \overline{\mathcal{B}(\Theta')})$ to $(T, \mathcal{B}(T))$. By Kallenberg (2017, Theorem 1.25, p. 39), there exists a Markov transition h from $(\Theta' \times T, \overline{\mathcal{B}(\Theta')} \otimes \mathcal{B}(T))$ to $(M, \mathcal{B}(M))$ that satisfies the desired equality for the restricted transitions from $(\Theta', \mathcal{B}(\Theta'))$ to $(M \times T, \mathcal{B}(M) \otimes \mathcal{B}(T))$. By Lemma 4, we can extend h to a Markov transition from $(\Theta \times T, \overline{\mathcal{B}(\Theta)} \otimes \mathcal{B}(T))$ to $(M, \overline{\mathcal{B}(M)})$.

[26]To keep the diagrams uncluttered, we adopt the following conventions. The labeled Markov transition maps a subproduct of the source space into a subproduct of the target space. The other sets in the target

---
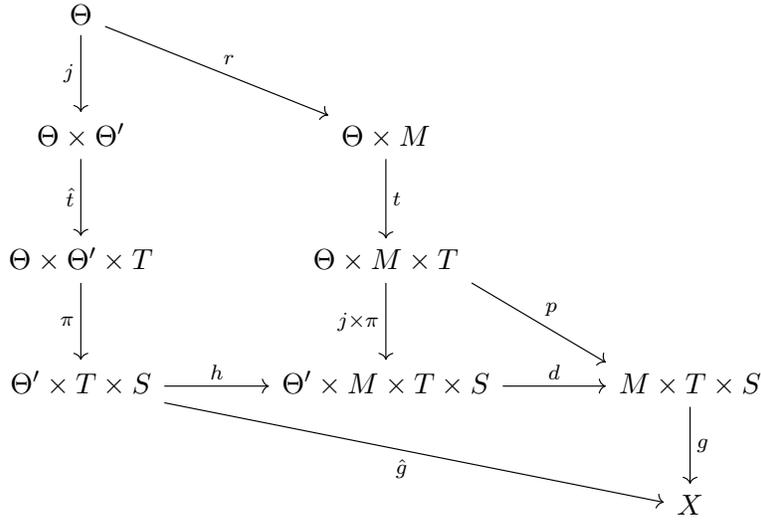
By construction, this diagram commutes, so the direct mechanism $(\hat{t}, \hat{g})$ induces f. To prove incentive compatibility, we show that for any strategy $(\hat{z}, \hat{q})$ in the direct mechanism, there is a strategy (z, q) in the original mechanism that induces the same social choice function. Given $(\hat{z}, \hat{q})$, define (z, q) by the following commutative diagrams.

$$(13)$$

Since each Markov transition $d_{\theta', m, \tau}$ is downward, it follows that q is feasible. These deviations induce the same social choice function because the following diagram commutes.

35
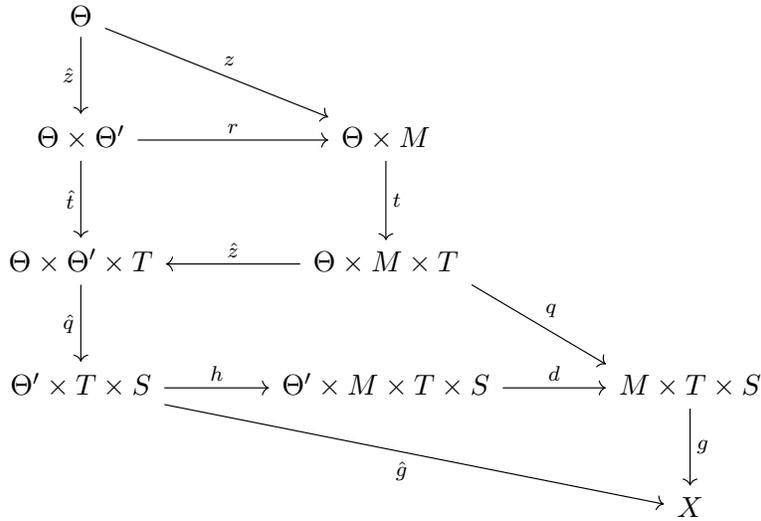
Extension to multiple agents Following the notation in Section 7, let $\Theta = \prod_i \Theta_i$, $T = \prod_i T_i$, and $S = \prod_i S_i$. Set $\pi = \otimes_i \pi$. With this notation, π is a Markov transition from $T \times \Theta$ to S, as in the single-agent case. After setting $M = \prod_i M_i$, a mechanism in the multi-agent setting maps between the same sets. Similarly for the agent's strategies, set $r = \otimes_i r_i$ and $p = \otimes_i p_i$. With these new definitions, construct h as before. Define d_i separately for each agent i and then set $d = \otimes_i d_i$. The proof is exactly as before. To consider a deviation by a fixed player j, take $\hat{z} = \hat{z}_j \otimes r_{-j}$ and $\hat{q} = \hat{q}_j \otimes p_{-j}$. Then the deviation (z, q) constructed in (13) will have the form $z = z_j \otimes r_{-j}$ and $q = q_j \otimes p_{-j}$.

B.2 Proof of Theorem 2

We begin with some notation. As in the proof of Theorem 1, denote by Θ' the message set in a direct mechanism. Thus, Θ' is a copy of Θ. Similarly, denote by T' a copy of T that will be the codomain of a most-discerning testing function. Keeping these copies separate will be helpful for the commutative diagrams below.

The proof is organized as follows. First, we select score conversions in a measurable way. Next, we prove sufficiency and then necessity.

Selecting score conversions Let $\mathcal{K}$ denote the space $\Delta(S)^S$ of Markov transitions on S, viewed as a subset of $\mathbf{R}^{S \times S}$, with the usual Euclidean topology and inner product $\langle \cdot, \cdot \rangle$. For $k \in \mathcal{K}$, denote by $k(s, s')$ the transition probability from s to s'.

Define the domain

$$D = \{(\theta, \tau, \psi) \in \Theta' \times T' \times T : \tau \succeq_\theta \psi\}.$$

Define the correspondence $K \colon D \twoheadrightarrow \mathcal{K}$ by putting $K(\theta, \tau, \psi)$ equal to the set of monotone Markov transitions k in $\mathcal{K}$ satisfying (i) $\pi_{\tau|\theta} k = \pi_{\psi|\theta}$, and (ii) $\pi_{\tau|\theta'} k \leq_{\mathrm{SD}} \pi_{\psi|\theta'}$ for all types θ'. By the choice of domain D, the correspondence K is nonempty-valued.

Endow D with the restriction of the σ-algebra $\overline{\mathcal{B}(\Theta' \times T' \times T)}$. To prove that there exists a measurable selection $\hat{k}$ from K, we apply the Kuratowski–Ryll-Nardzewski selection theorem (Aliprantis and Border, 2006, 18.13, p. 600). The correspondence K has compact convex values, so it suffices to check that associated support functions for K are measurable (Aliprantis and Border, 2006, 18.31, p. 611).

Fix $\ell \in \mathbf{R}^{S \times S}$. Define the map $C \colon D \to \mathbf{R}$ by

$$C(\theta, \tau, \psi) = \max_{k \in K(\theta, \tau, \psi)} \langle k, \ell \rangle.$$

It suffices to show that C is $\overline{\mathcal{B}(\Theta' \times T' \times T)}$-measurable. Define a sequence of auxiliary functions $C_m \colon D \times (\Theta')^m \to \mathbf{R}$ as follows. Let $C_m(\theta, \tau, \psi, \theta'_1, \ldots, \theta'_m)$ be the value of the

program

$$
\begin{aligned}
\text{maximize} \quad & \langle k, \ell \rangle \\
\text{subject to} \quad & k \in \mathcal{K} \\
& k \text{ is monotone} \\
& \pi_{\tau|\theta} k = \pi_{\psi|\theta} \\
& \pi_{\tau|\theta'_j} k \leq_{\text{SD}} \pi_{\psi|\theta'_j}, \quad j = 1, \ldots, m.
\end{aligned}
$$

This is a standard linear programming problem with a compact feasible set. By Berge's theorem (Aliprantis and Border, 2006, 17.30, p. 569), the value of the linear program is upper semicontinuous (and hence Borel) as a function of the coefficients appearing in the constraints. Since π is Borel, so is each function C_m. By the measurable projection theorem (Theorem 7), each map

$$
(\theta, \tau, \psi) \mapsto \inf_{\boldsymbol{\theta'} \in (\Theta')^m} C_m(\theta, \tau, \psi, \boldsymbol{\theta'})
$$

is $\overline{\mathcal{B}(\Theta' \times T' \times T)}$-measurable. A compactness argument shows that[27]

$$
C(\theta, \tau, \psi) = \inf_m \inf_{\boldsymbol{\theta'} \in (\Theta')^m} C_m(\theta, \tau, \psi, \boldsymbol{\theta'}),
$$

so C is also $\overline{\mathcal{B}(\Theta' \times T' \times T)}$-measurable.

Sufficiency Fix a decision environment (X, u) and let f be an implementable social choice function. By the revelation principle (Theorem 1), there is a direct mechanism (t, g) that canonically implements f. We now construct a decision rule $\hat{g}$ such that the direct mechanism $(\hat{t}, \hat{g})$ canonically implements f. Define $\hat{g}$ by the following commutative diagram,

---

[27]We claim that for each positive ε there exists a natural number m and a vector $\boldsymbol{\theta'} \in (\Theta')^m$ such that $C_m(\theta, \tau, \psi, \boldsymbol{\theta'}) < C(\theta, \tau, \psi) + \varepsilon$. Suppose not. For each $\theta' \in \Theta'$, let $K_{\theta'}$ be the compact set of monotone Markov transitions $k \in \mathcal{K}$ satisfying (i) $\pi_{\tau|\theta} k = \psi_{\psi|\theta}$, (ii) $\pi_{\tau|\theta'} k \leq_{\text{SD}} \pi_{\psi|\theta'}$, and (iii) $\langle k, \ell \rangle \geq C(\theta, \tau, \psi) + \varepsilon$. This family has the finite intersection property, but the intersection over all $\theta' \in \Theta'$ is empty, which is a contradiction.

where j denotes the identity transition from Θ to Θ'.

$$
\begin{array}{c}
\Theta \\
\downarrow \text{\scriptsize id} \\
\Theta \times \Theta' \\
\downarrow \hat{t} \qquad \searrow t \\
\Theta \times \Theta' \times T' \xleftarrow{\ \hat{t}\ } \Theta \times \Theta' \times T \\
\downarrow \pi \qquad\qquad \searrow \pi \\
\Theta' \times T' \times S \xrightarrow{\ t\ } \Theta' \times T' \times T \times S \xrightarrow{\ \hat{k}\ } \Theta' \times T \times S \\
\qquad\qquad \searrow \hat{g} \qquad\qquad \downarrow g \\
\qquad\qquad\qquad X
\end{array}
$$
(14)

Since the diagram commutes, the direct mechanism $(\hat{t}, \hat{g})$ induces f. For incentive compatibility, we show that for any strategy $(\hat{z}, \hat{q})$ in the new mechanism, there is a strategy (z, q) in the old mechanism inducing the same social choice function. Given $(\hat{z}, \hat{q})$, set $z = \hat{z}$ and define q by the following commutative diagram.

$$
\begin{array}{ccc}
\Theta \times \Theta' \times T' \times T & \xleftarrow{\ \hat{t}\ } & \Theta \times \Theta' \times T \\
\downarrow \hat{q} & & \downarrow q \\
\Theta' \times T' \times T \times S & \xrightarrow{\ \hat{k}\ } & S
\end{array}
$$
(15)

Since each transition $\hat{k}_{\theta,\tau,\psi}$ is downward, it follows that q is feasible. These deviations induce the same social choice function because the following diagram commutes.

$$
\begin{array}{c}
\Theta \\
\downarrow \hat{z} \\
\Theta \times \Theta' \\
\downarrow \hat{t} \qquad \searrow t \\
\Theta \times \Theta' \times T' \xleftarrow{\ \hat{t}\ } \Theta \times \Theta' \times T \\
\downarrow \hat{q} \qquad\qquad \searrow q \\
\Theta' \times T' \times S \xrightarrow{\ t\ } \Theta' \times T' \times T \times S \xrightarrow{\ \hat{k}\ } \Theta' \times T \times S \\
\qquad\qquad \searrow \hat{g} \qquad\qquad \downarrow g \\
\qquad\qquad\qquad X
\end{array}
$$
(16)

Necessity If $\hat{t}$ is not most-discerning, then there exists a fixed type θ such that $\hat{t}(\theta)$ is not most θ-discerning. Set $\tau = \hat{t}(\theta)$. Select a test ψ such that $\tau \not\succeq_\theta \psi$.

Define the decision set

$$X = \{x_{\theta'} : \theta' \in \Theta\} \cup \{\underline{x}\},$$

where $\underline{x}$, $x_{\theta'}$, and $x_{\theta''}$ are distinct for all distinct types θ' and θ''.

Define utilities as follows. Type θ gets utility 1 from x_θ and utility 0 from every other decision in X. For $\theta' \neq \theta$, type θ' gets utility 1 from decision x_θ, utility $\pi(\psi|\theta')$ from $x_{\theta'}$, and utility 0 from all other decisions.

Let f be the social choice function that assigns to each type θ' with $\theta' \neq \theta$ the decision $x_{\theta'}$ with certainty, and assigns to type θ decision x_θ with probability $\pi(\psi|\theta)$ and decision $\underline{x}$ with probability $1 - \pi(\psi|\theta)$. Then f can be canonically implemented by the mechanism $(t, \hat{g})$, where t is any function satisfying $t(\theta) = \psi$ and $\hat{g}$ is the decision rule specified as follows. If the agent reports $\theta' \neq \theta$, assign $x_{\theta'}$ no matter the test result; if the agent reports θ, select x_θ if the agent passes test ψ and $\underline{x}$ if the agent fails test ψ.

We claim that f cannot be canonically implemented with the testing function $\hat{t}$. Suppose for a contradiction that there is a decision rule $\hat{g}$ such that f is canonically implemented by the mechanism $(\hat{t}, \hat{g})$. We separate into two cases.

First suppose $\pi(\tau|\theta) = 0$. Then every type can get the good with probability $\pi(\psi|\theta)$. Since $\tau \not\succeq_\theta \psi$, there is some type $\theta' \neq \theta$ such that $\pi(\psi|\theta') < \pi(\psi|\theta)$, so type θ' has a profitable deivation.

Next suppose $\pi(\tau|\theta) > 0$. Define a Markov transition k on S, represented by a vector in $[0,1]^2$ by letting $k(s)$ be the probability that the measure $g_{\theta,\tau,s}$ places on decision x_θ. Then $\pi_{\tau|\theta} k = \pi_{\psi|\theta}$. Since $\tau \not\succeq_\theta \psi$, either k is not monotone, in which case type θ can profitably deviate by reporting type θ and failing the test, or there is some type $\theta' \neq \theta$ such that $\pi_{\tau|\theta'} k >_{\text{SD}} \pi_{\psi|\theta'}$. In this case, type θ' can profitably deviate by reporting θ' and exerting effort.

B.3 Proof of Theorem 3

First we use the regularity assumptions to prove that there exists a measurable test selection. Then we follow the proof of the sufficiency part of Theorem 2.

Measurable test selection We prove that there exists a measurable function $\bar{t}$ from $(\Theta' \times T, \overline{\mathcal{B}(\Theta' \times T)})$ to $(T', \mathcal{B}(T'))$ such that the test $\bar{t}(\theta, \psi)$ is in $\hat{T}(\theta)$ and satisfies $\bar{t}(\theta, \psi) \succeq_\theta \psi$, for each $\theta \in \Theta$ and $\psi \in T$. Define a correspondence $H \colon \Theta' \times T \twoheadrightarrow T'$ by

$$H(\theta, \psi) = \{\tau \in T : \tau \succeq_\theta \psi\}.$$

Since π is continuous, the graph of H is closed in $\Theta' \times T \times T'$.[28] Since the graph of $\hat{T}$ is Borel, so is the set $\{(\theta, \psi, \tau) : \tau \in \hat{T}(\theta)\}$ and also the intersection

$$\{(\theta, \psi, \tau) : \tau \succeq_\theta \psi \text{ and } \tau \in \hat{T}(\theta)\}.$$

By the measurable projection theorem (Theorem 7), the associated correspondence from $(\Theta' \times T, \overline{\mathcal{B}(\Theta' \times T)})$ to T' is measurable. Moreover, this correspondence has closed values, so we can apply the Kuratowski–Ryll-Nardzewksi selection theorem (Aliprantis and Border, 2006, 18.13, p. 600) to obtain the desired function $\bar{t}$.

Sufficiency With $\bar{t}$ in hand, the proof is almost the same as the proof of Theorem 2 in Appendix B.2, but we are not given $\hat{t}$. Instead, $\hat{t}$ is defined by the commutative diagram. The second and third rows of (14) and (16) become

$$
\begin{array}{ccc}
\Theta \times \Theta' & & \\
\hat{t} \downarrow & \searrow{t} & \\
\Theta \times \Theta' \times T' & \xleftarrow{\bar{t}} & \Theta \times \Theta' \times T
\end{array}
$$

In (15), put $\bar{t}$ in place of $\hat{t}$. The rest of the proof is completed as above.

B.4   Proof of Proposition 1

A Markov transition k on $\{0, 1\}$ can be represented as a vector

$$(k(0), k(1)) \in [0, 1]^2,$$

where $k(s)$ is the probability of transitioning from $s \in \{0, 1\}$ to 1. A Markov transition k is monotone and satisfies $\pi_{\tau|\theta} k = \pi_{\psi|\theta}$ if and only if the vector $(k(0), k(1))$ satisfies

$$k(1) \geq k(0) \quad \text{and} \quad \pi(\tau|\theta)k(1) + (1 - \pi(\tau|\theta))k(0) = \pi(\psi|\theta).$$

We separate the solution into cases.

   1. If $\pi(\tau|\theta) \geq \pi(\psi|\theta)$, then the solutions are given by

$$
\begin{bmatrix} k(0) \\ k(1) \end{bmatrix} = \lambda \begin{bmatrix} 0 \\ \pi(\psi|\theta)/\pi(\tau|\theta) \end{bmatrix} + (1 - \lambda) \begin{bmatrix} \pi(\psi|\theta) \\ \pi(\psi|\theta) \end{bmatrix},
$$

     for $\lambda \in [0, 1]$, provided that we use the convention that $0/0$ equals 1.

   2. If $\pi(\tau|\theta) < \pi(\psi|\theta)$, then the solutions are given by

$$
\begin{bmatrix} k(0) \\ k(1) \end{bmatrix} = \lambda \begin{bmatrix} \bar{\pi}(\psi|\theta)/\bar{\pi}(\tau|\theta) \\ 1 \end{bmatrix} + (1 - \lambda) \begin{bmatrix} \pi(\psi|\theta) \\ \pi(\psi|\theta) \end{bmatrix},
$$

---

[28]Take a sequence $(\theta_n, \psi_n, \tau_n)$ in $\mathrm{gr}\,H$ converging to a limit (θ, ψ, τ) in $\Theta' \times T \times T'$. For each n, there is a monotone Markov transition k_n on S such that (i) $\pi_{\tau_n|\theta_n} k_n = \pi_{\psi_n|\theta_n}$, and (ii) $\pi_{\tau_n|\theta'} k_n \leq_{\text{SD}} \pi_{\psi_n|\theta'}$ for all $\theta' \in \Theta$. The space of Markov transitions on S is compact, so after passing to a subsequence, we may assume that k_n converges to a limit k, which must be monotone. Since π is continuous, taking limits gives (i) $\pi_{\tau|\theta} k = \pi_{\psi|\theta}$, and (ii) $\pi_{\tau|\theta'} k \leq_{\text{SD}} \pi_{\psi|\theta'}$ for each $\theta' \in \Theta$. Therefore, (θ, ψ, τ) is in $\mathrm{gr}\,H$.

for $\lambda \in [0, 1]$.

In each case, the left term is the vector representation of $\tilde{F}_{\tau|\theta}\tilde{Q}_{\tau|\theta}$ and the right term is the vector representation of the constant Markov transition $\pi_{\psi|\theta}$.

B.5 Proof of Proposition 2

Fix a type θ and tests τ and ψ. If $k = \lambda\tilde{F}_{\tau|\theta}\tilde{Q}_{\tau|\theta} + (1-\lambda)\pi_{\psi|\theta}$, then by Lemma 1,

$$\pi_{\tau|\theta'}k = \lambda\pi_{\tau|\theta'}\tilde{F}_{\tau|\theta}\tilde{Q}_{\psi|\theta} + (1-\lambda)\pi_{\psi|\theta} = \left(\lambda\pi_{\tau|\theta'} + (1-\lambda)\pi_{\tau|\theta}\right)\tilde{F}_{\tau|\theta}\tilde{Q}_{\psi|\theta}.$$

Therefore, by Proposition 1, we have $\tau \succeq_\theta \psi$ if and only if there exists $\lambda \in [0, 1]$ such that

$$\left(\lambda\pi_{\tau|\theta'} + (1-\lambda)\pi_{\tau|\theta}\right)\tilde{F}_{\tau|\theta}\tilde{Q}_{\psi|\theta} \leq_{\text{SD}} \pi_{\psi|\theta'}, \tag{17}$$

for all $\theta' \in \Theta$. Now work with cases.

1. If $\pi(\tau|\theta) = \pi(\psi|\theta) = 0$, the result is clear.
2. If $\pi(\tau|\theta) \geq \pi(\psi|\theta)$ and $\pi(\tau|\theta) > 0$, we compare the probability of passage. The inequality holds if and only if

$$[\lambda\pi(\tau|\theta') + (1-\lambda)\pi(\tau|\theta)]\frac{\pi(\psi|\theta)}{\pi(\tau|\theta)} \leq \pi(\psi|\theta'),$$

   This reduces to the desired inequality.
3. If $\pi(\tau|\theta) < \pi(\psi|\theta)$, we compare the probability of failure. The inequality holds if and only if

$$[\lambda\bar{\pi}(\tau|\theta') + (1-\lambda)\bar{\pi}(\tau|\theta)]\frac{\bar{\pi}(\psi|\theta)}{\bar{\pi}(\tau|\theta)} \geq \bar{\pi}(\psi|\theta').$$

B.6 Proof of Proposition 3

Fix a type θ and tests τ_1 and τ_2.

One direction is clear. If τ_1 and τ_2 are equal, then τ_1 and τ_2 are clearly θ-equivalent. If type θ is minimal on τ_1 and τ_2, take $\lambda = 1$ in Proposition 2 to see that τ_1 and τ_2 are θ-equivalent.

For the other direction, suppose τ_1 and τ_2 are θ-equivalent. Without loss, we may assume $\pi(\tau|\theta) \geq \pi(\psi|\theta)$. We separate into cases according to whether θ is minimal on τ.

First suppose θ is minimal on τ_1. Since $\tau_1 \succeq_\theta \tau_2$, there are probabilities $k(0)$ and $k(1)$ such that for all types θ',

$$\begin{aligned}
\pi(\tau_2|\theta) &= k(0) + (k(1) - k(0))\pi(\tau_1|\theta) \\
&\leq k(0) + (k(1) - k(0))\pi(\tau_1|\theta') \\
&\leq \pi(\tau_2|\theta').
\end{aligned}$$

Now suppose θ is not minimal for τ, so there exists some type θ'' such that $\pi(\tau|\theta'') < \pi(\tau|\theta)$. In particular, $\pi(\tau|\theta) > 0$.

Claim. $\pi(\tau|\theta) = \pi(\psi|\theta)$.

Suppose for a contradiction that $\pi(\tau_1|\theta) > \pi(\tau_2|\theta)$. By Proposition 2, there are constants λ_1 and λ_2 in $[0,1]$ such that for all types θ',

$$[\lambda_1\pi(\tau_1|\theta') + (1-\lambda_1)\pi(\tau_1|\theta)]\pi(\tau_2|\theta) \le \pi(\tau_2|\theta')\pi(\tau_1|\theta),$$
$$[\lambda_2\bar{\pi}(\tau_2|\theta') + (1-\lambda_2)\bar{\pi}(\tau_2|\theta)]\bar{\pi}(\tau_1|\theta) \ge \bar{\pi}(\tau_1|\theta')\bar{\pi}(\tau_2|\theta).$$

With $\theta' = \theta''$ the first inequality is weakest when $\lambda_1 = 1$, so

$$\pi(\tau_1|\theta'')\pi(\tau_2|\theta) \le \pi(\tau_2|\theta'')\pi(\tau_1|\theta). \tag{18}$$

Taking $\lambda_2 = 0$ in the second inequality, and noting that $\bar{\pi}(\tau_2|\theta) = 1 - \pi(\tau_2|\theta) > 0$, yields the contradiction $\bar{\pi}(\tau_1|\theta) \ge \bar{\pi}(\tau_1|\theta'')$. Therefore, the inequality must hold with $\lambda_2 = 1$, so

$$\bar{\pi}(\tau_2|\theta'')\bar{\pi}(\tau_2|\theta) \ge \bar{\pi}(\tau_2|\theta)\bar{\pi}(\tau_1|\theta''). \tag{19}$$

We show that (18) and (19) are incompatible. In (19), subtract $\bar{\pi}(\tau_1|\theta)\bar{\pi}(\tau_2|\theta)$ from both sides to get

$$[\bar{\pi}(\tau_2|\theta'') - \bar{\pi}(\tau_2|\theta)]\bar{\pi}(\tau_1|\theta) \ge [\bar{\pi}(\tau_1|\theta'') - \bar{\pi}(\tau_1|\theta)]\bar{\pi}(\tau_2|\theta),$$

which is equivalently,

$$[\pi(\tau_2|\theta) - \pi(\tau_2|\theta'')]\bar{\pi}(\tau_1|\theta) \ge [\pi(\tau_1|\theta) - \pi(\tau_1|\theta'')]\bar{\pi}(\tau_2|\theta).$$

The right side is strictly positive and $\bar{\pi}(\tau_2|\theta) < \bar{\pi}(\tau_1|\theta)$, so

$$\pi(\tau_2|\theta) - \pi(\tau_2|\theta'') > \pi(\tau_1|\theta) - \pi(\tau_1|\theta''). \tag{20}$$

Now negate (18) and add $\pi(\tau_1|\theta)\pi(\tau_2|\theta)$ to both sides to obtain

$$[\pi(\tau_1|\theta) - \pi(\tau_1|\theta'')]\pi(\tau_2|\theta) \ge [\pi(\tau_2|\theta) - \pi(\tau_2|\theta'')]\pi(\tau_1|\theta).$$

But $\pi(\tau_2|\theta) < \pi(\tau_1|\theta)$, so (20) gives the opposite inequality.

With the claim established, we now complete the proof. By Proposition 2 there are constants λ_1 and λ_2 in $[0,1]$ such that for all types θ',

$$[\lambda_1\pi(\tau_1|\theta') + (1-\lambda_1)\pi(\tau_1|\theta)]\pi(\tau_2|\theta) \le \pi(\tau_2|\theta')\pi(\tau_1|\theta),$$
$$[\lambda_2\pi(\tau_2|\theta') + (1-\lambda_2)\pi(\tau_2|\theta)]\pi(\tau_1|\theta) \le \pi(\tau_1|\theta')\pi(\tau_2|\theta).$$

After cancelling the common value of $\pi(\tau_1|\theta)$ and $\pi(\tau_2|\theta)$, which is nonzero by assumption, we have

$$\lambda_1\pi(\tau_1|\theta') + (1-\lambda_1)\pi(\tau_1|\theta) \le \pi(\tau_2|\theta'), \tag{21}$$
$$\lambda_2\pi(\tau_2|\theta') + (1-\lambda_2)\pi(\tau_2|\theta) \le \pi(\tau_1|\theta'). \tag{22}$$

It suffices to show that $\lambda_1 = \lambda_2 = 1$, for then $\pi(\tau_1|\theta') = \pi(\tau_2|\theta')$ for all types θ'. Take $\theta' = \theta''$ in both inequalities. We have $\pi(\tau_2|\theta) = \pi(\tau_1|\theta) > \pi(\tau_1|\theta'')$, so (22) implies that $\pi(\tau_2|\theta'') \le \pi(\tau_1|\theta'')$. Substituting this into (21), we get $\lambda_1 = 1$. Then $\pi(\tau_1|\theta'') \le \pi(\tau_2|\theta'')$,

so (22) implies $\lambda_2 = 1$.

B.7   Proof of Proposition 4

We simply translate Proposition 3 into the language of authentication rates. Suppose $\hat{t}_1$ and $\hat{t}_2$ are most-discerning testing functions, and let α_1 and α_2 be the induced authentication rates. For each type θ, we know $\hat{t}_1(\theta)$ and $\hat{t}_2(\theta)$ are most θ-discerning tests. Apply Proposition 3 and translate the conclusion into the language of authentication rates.

B.8   Proof of Theorem 4

Let α be an authentication rate. First, suppose that α is most discerning. Let $T = \{\tau_{\theta'} : \theta' \in \Theta\}$, and define the passage rate π by $\pi(\tau_{\theta'}|\theta) = \alpha(\theta'|\theta)$ for all types θ and θ'. Combining Definition 8 and Proposition 2, we see that the testing function $\theta \mapsto \tau_\theta$ is most discerning. By construction, this testing function induces α.

Now suppose α is induced by a most-discerning testing function $\hat{t}$ in a testing environment (T, π). Substitute the equality $\alpha(\theta'|\theta) = \pi(\hat{t}(\theta'), \theta)$ into Proposition 2 to conclude that α is most discerning.

B.9   Proof of Proposition 5

We apply the dominated convergence theorem. As λ converges to 0 pointwise, $\Lambda(z|\theta)$ converges to 1 for all z and θ with $z \leq \theta$. Hence $\varphi(\theta)$ converges to $\varphi^M(\theta)$, for each θ. Likewise, as λ converges to ∞ pointwise, $\Lambda(z|\theta)$ converges to 0 for all z and θ with $z < \theta$. Hence $\varphi(\theta)$ converges to θ, for each θ.

B.10   Proof of Proposition 6

We will prove a stronger result that we use below. Define functions λ_+ and λ_- from Θ to $[0, \infty)$ by
$$\lambda_+(\theta) = -D_{2+}\alpha(\theta|\theta), \qquad \lambda_-(\theta) = D_{2-}\alpha(\theta|\theta).$$
In the main text, we only work with λ_+, which is denoted λ. Extend the function Λ to $\Lambda \colon \Theta \times \Theta \to [0, 1]$ by

$$\Lambda(\theta'|\theta) = \begin{cases} \exp\left(-\int_{\theta'}^{\theta} \lambda_+(s)\,\mathrm{d}s\right) & \text{if } \theta \geq \theta', \\[2ex] \exp\left(-\int_{\theta}^{\theta'} \lambda_-(s)\,\mathrm{d}s\right) & \text{if } \theta < \theta'. \end{cases}$$

With these definitions, we now prove that $\alpha(\theta'|\theta) \geq \Lambda(\theta'|\theta)$ all types θ' and θ. Fix θ and θ'. For each h, transitivity gives

$$\alpha(\theta'|\theta + h) \geq \alpha(\theta'|\theta)\alpha(\theta|\theta + h).$$

Subtract $\alpha(\theta'|\theta)$ from each side to get

$$
\begin{aligned}
\alpha(\theta'|\theta + h) - \alpha(\theta'|\theta) &\geq \alpha(\theta'|\theta)(\alpha(\theta|\theta + h) - 1)) \\
&= \alpha(\theta'|\theta)[\alpha(\theta + h, \theta) - \alpha(\theta|\theta)].
\end{aligned}
$$

Dividing by h and passing to the limit as $h \downarrow 0$ and $h \uparrow 0$ gives

$$
D_{2+}\alpha(\theta'|\theta) \geq -\lambda_+(\theta)\alpha(\theta'|\theta) \quad \text{and} \quad D_{2-}\alpha(\theta'|\theta) \leq \lambda_-(\theta)\alpha(\theta'|\theta).
$$

Now we use absolute continuity to convert these local bounds into global bounds. Fix a report θ'. Define the function Δ on $[\underline{\theta}, \bar{\theta}]$ by

$$
\Delta(\theta) = \frac{\alpha(\theta'|\theta)}{\Lambda(\theta'|\theta)}.
$$

By construction, $\Delta(\theta') = 1$. We will argue that $\Delta(\theta) \geq 1$ for all θ. For $\theta' < \theta$, if $D_{2+}\Lambda(\theta'|\theta)$ exists, then

$$
D_+\Delta(\theta) = \frac{1}{\Lambda(\theta'|\theta)} \left(D_+\alpha(\theta'|\theta) + \lambda_+(\theta)\alpha(\theta'|\theta) \right) \geq 0.
$$

For $\theta' > \theta$, if $D_{2-}\Lambda(\theta'|\theta)$ exists, then

$$
D_-\Delta(\theta) = \frac{1}{\Lambda(\theta'|\theta)} \left(D_-\alpha(\theta'|\theta) - \lambda_-(\theta)\alpha(\theta'|\theta) \right) \leq 0,
$$

Since $\Lambda(\theta'|\cdot)$ is absolutely continuous, these inequalities hold almost surely. Moreover, the product of absolutely continuous functions on a compact set is absolutely continuous, so Δ is absolutely continuous, and hence the fundamental theorem of calculus gives $\Delta(\theta) \geq 1$, as desired.

B.11 Proof of Proposition 7

First we introduce notation. For a given quantity function q, there is a one-to-one correspondence between the transfer function t and the utility function U, given by $U(\theta) = \theta q(\theta) - t(\theta)$. We will interchangeably refer to such a mechanism as (q, t) or (q, U). Let

$$
u(\theta'|\theta) = \alpha(\theta'|\theta)[\theta q(\theta') - t(\theta')], \qquad U(\theta) = u(\theta|\theta) = \max_{\theta' \in \Theta} u(\theta'|\theta).
$$

Lemma 6 (Utility bound). Let q be a bounded quantity function, and let U be a utility function. If (q, U) is incentive compatible, then for every type θ, we have

$$
U(\theta) \geq \int_{\underline{\theta}}^{\theta} \Lambda(z|\theta)q(z)\, dz. \tag{23}
$$

If the function $\theta \mapsto \Lambda(\theta|\bar{\theta})q(\theta)$ is increasing and the global upper bound is satisfied, it is incentive compatible for (23) to hold with equality.

Lemma 6 is proved in Appendix B.12. Here we prove the theorem, taking Lemma 6

as given. There is no loss in restricting attention to bounded quantity functions.[29] Pick a bounded quantity function $q\colon \Theta \to \mathbf{R}_+$. The principal's objective function can be decomposed as the difference between the total surplus and the agent's rents:

$$\int_{\underline{\theta}}^{\bar{\theta}} [\theta q(\theta) - c(q(\theta))] f(\theta) \, \mathrm{d}\theta - \int_{\underline{\theta}}^{\bar{\theta}} U(\theta) \, \mathrm{d}\theta.$$

Plug in the bound from Lemma 6 and switch the order of integration to obtain the following upper bound on the principal's objective:

$$V(q) = \int_{\underline{\theta}}^{\bar{\theta}} [\varphi(\theta) q(\theta) - c(q(\theta))] f(\theta) \, \mathrm{d}\theta.$$

The quantity function $q^\star$ from the theorem statement maximizes the expression in brackets pointwise, and $t^\star$ is the corresponding transfer that achieves the utility bound. Since φ is increasing and the global upper bound is satisfied, this mechanism is incentive compatible. Since c' is strictly increasing, the pointwise maximizer is unique, so this quantity function $q^\star$ is unique almost everywhere.

B.12   Proof of Lemma 6

Let (q, t) be a bounded incentive compatible mechanism. The first step is showing that the equilibrium utility function U is absolutely continuous. Fix types θ and θ'. We have

$$\begin{aligned}
U(\theta') &\geq \alpha(\theta'|\theta) \left(\theta' q(\theta) - t(\theta)\right)_+ \\
&\geq \Lambda(\theta'|\theta) \left(\theta' q(\theta) - t(\theta)\right)_+ \\
&\geq \Lambda(\theta'|\theta) \left(\theta' q(\theta) - t(\theta)\right),
\end{aligned}$$

where the first inequality uses individual rationality and incentive compatibility, and the second uses the inequality between α and Λ established in Appendix B.10. Therefore,

$$\begin{aligned}
U(\theta) - U(\theta') &\leq \theta q(\theta) - t(\theta) - \Lambda(\theta'|\theta) \left(\theta' q(\theta) - t(\theta)\right) \\
&= \theta q(\theta) - t(\theta) - \Lambda(\theta'|\theta) \left((\theta' - \theta) q(\theta) + \theta q(\theta) - t(\theta)\right) \\
&= (1 - \Lambda(\theta'|\theta))(\theta q(\theta) - t(\theta)) - \Lambda(\theta'|\theta)(\theta' - \theta) q(\theta) \\
&\leq (1 - \Lambda(\theta'|\theta))(\bar{\theta} \|q\|_\infty + \|t\|_\infty) + |\theta' - \theta| \|q\|_\infty.
\end{aligned}$$

To avoid separating into cases according to the relative sizes of θ and θ', set $\hat{\lambda} = \lambda_+ \vee \lambda_-$. Since $\hat{\lambda} \leq \lambda_+ + \lambda_-$, we know $\hat{\lambda}$ is integrable. Using the inequality $e^x \geq 1 + x$, we have

$$1 - \Lambda(\theta'|\theta) \leq 1 - \exp\left(-\int_{\theta' \wedge \theta}^{\theta' \vee \theta} \hat{\lambda}(z) \, \mathrm{d}z\right) \leq \int_{\theta' \wedge \theta}^{\theta' \vee \theta} \hat{\lambda}(z) \, \mathrm{d}z.$$

---

[29]Pick a quantity $\bar{q}$ such that $\bar{\theta}\bar{q} = c(\bar{q})$. Offering more than $\bar{q}$ will always result in weakly negative profits, so we can remove those offerings from the menu and increase the principal's revenue. Therefore, there is no loss in focusing on quantity functions that are bounded above by $\bar{q}$.

Plug this into the inequality above to get

$$U(\theta) - U(\theta') \leq C \int_{\theta' \wedge \theta}^{\theta' \vee \theta} (\hat{\lambda}(z) + 1) \, \mathrm{d}z,$$

where $C = \max\{1, \bar{\theta}\} \|q\|_\infty + \|t\|_\infty$. Switching the roles of θ and θ' gives the same inequality, so we conclude that

$$|U(\theta) - U(\theta')| \leq C \int_{\theta' \wedge \theta}^{\theta' \vee \theta} (\hat{\lambda}(z) + 1) \, \mathrm{d}z,$$

which proves the desired absolute continuity.

Now we use the absolute continuity of U to establish the bound. Define the auxiliary function Δ on $[\underline{\theta}, \bar{\theta}]$ by

$$\Delta(\theta) = \Lambda(\theta|\bar{\theta})\left(U(\theta) - \int_{\underline{\theta}}^{\theta} \Lambda(z|\theta)q(z) \, \mathrm{d}z\right) = \Lambda(\theta|\bar{\theta})U(\theta) - \int_{\underline{\theta}}^{\theta} \Lambda(z|\bar{\theta})q(z) \, \mathrm{d}z.$$

The function Δ is absolutely continuous since it is the product of absolutely continuous functions on a compact set. By Theorem 1 in [Milgrom and Segal (2002)](#), whenever U is differentiable, we have

$$q(s) - \lambda_+(s)U(s) = D_{2+}u(\theta|\theta) \leq U'(s) \leq D_{2-}u(\theta, \theta) = q(s) + \lambda_-(s)U(s).$$

At each point θ where all the functions involved are differentiable, which holds almost surely, we have

$$\begin{aligned}
\Delta'(\theta) &= \lambda_+(\theta)\Lambda(\theta|\bar{\theta})U(\theta) + \Lambda(\theta|\bar{\theta})U'(\theta) - \Lambda(\theta|\bar{\theta})q(\theta) \\
&= \Lambda(\bar{\theta}, \theta)\left[U'(\theta) - (q(\theta) - \lambda_+(\theta)U(\theta))\right] \\
&\geq 0.
\end{aligned}$$

Since $\Delta(\underline{\theta}) = 0$, the fundamental theorem of calculus implies that $\Delta(\theta) \geq 0$ for all θ, as desired.

It remains to check that the global incentive constraints are satisfied, provided that q is monotone. Expressing incentive-compatibility in terms of U, we need to show that for all types θ and θ',

$$U(\theta) \geq \alpha(\theta'|\theta)(U(\theta') + (\theta - \theta')q(\theta')). \tag{24}$$

We consider upward and downward deviations separately. First suppose $\theta' > \theta$. Write out (24) as

$$U(\theta) + \alpha(\theta'|\theta)(\theta' - \theta)q(\theta') \geq \alpha(\theta'|\theta)U(\theta'),$$

or equivalently,

$$\int_{\underline{\theta}}^{\theta} \Lambda(z|\theta)q(z) \, \mathrm{d}z + \alpha(\theta'|\theta) \int_{\theta}^{\theta'} q(\theta') \, \mathrm{d}z$$

$$\geq \alpha(\theta'|\theta) \int_{\underline{\theta}}^{\theta} \Lambda(z|\theta')q(z) \, \mathrm{d}z + \alpha(\theta'|\theta) \int_{\theta}^{\theta'} \Lambda(z|\theta')q(z) \, \mathrm{d}z.$$

46

Compare the corresponding terms on each side. Since $\Lambda(z|\theta) \geq \Lambda(z|\theta')$ for $z \leq \theta \leq \theta'$, we get the inequality between the first terms. For the inequality between the second terms, multiply by $\Lambda(\theta'|\bar{\theta})/\alpha(\theta'|\theta)$ and use the fact that $\Lambda(z|\bar{\theta})q(z)$ is increasing in z.

Now suppose $\theta' < \theta$. Express (24) as

$$\int_{\underline{\theta}}^{\theta'} \Lambda(z|\theta)q(z)\,\mathrm{d}z + \int_{\theta'}^{\theta} \Lambda(z|\theta)q(z)\,\mathrm{d}z$$

$$\geq \alpha(\theta'|\theta) \left[\int_{\underline{\theta}}^{\theta'} \Lambda(z|\theta')q(z)\,\mathrm{d}z + \int_{\theta'}^{\theta} q(\theta')\,\mathrm{d}z \right]$$

$$= \frac{\alpha(\theta'|\theta)}{\Lambda(\theta'|\theta)} \left[\int_{\underline{\theta}}^{\theta'} \Lambda(z|\theta)q(z)\,\mathrm{d}z + \int_{\theta'}^{\theta} \Lambda(\theta'|\theta)q(\theta')\,\mathrm{d}z \right].$$

Multiply both sides by $\Lambda(\theta|\bar{\theta})$ and then rearrange to get the equivalent inequality

$$\alpha(\theta'|\theta) \leq \Lambda(\theta'|\theta) \frac{\int_{\underline{\theta}}^{\theta'} \Lambda(z|\bar{\theta})q(z)\,\mathrm{d}z + \int_{\theta'}^{\theta} \Lambda(z|\bar{\theta})q(z)\,\mathrm{d}z}{\int_{\underline{\theta}}^{\theta'} \Lambda(z|\bar{\theta})q(z)\,\mathrm{d}z + \int_{\theta'}^{\theta} \Lambda(\theta'|\bar{\theta})q(\theta')\,\mathrm{d}z}$$

$$= \Lambda(\theta'|\theta) \int_{\underline{\theta}}^{\theta} \Lambda(z|\bar{\theta})q(z)\,\mathrm{d}z \bigg/ \int_{\underline{\theta}}^{\theta} \Lambda(z \wedge \theta'|\bar{\theta})q(z \wedge \theta')\,\mathrm{d}z.$$

B.13   Proof of Proposition 8

The same argument as in the proof of Proposition 7 (Appendix B.11) shows that the principal's value as a function of q can be written as

$$V(q) = \int_{\underline{\theta}}^{\bar{\theta}} \varphi(\theta)q(\theta)\,\mathrm{d}\theta.$$

The quantity function $q^\star$ from the theorem statement maximizes this quantity pointwise, and $t^\star$ is the corresponding transfer function. Since $q^\star$ is monotone, this mechanism is incentive compatible and hence optimal. Except at points θ where $\varphi(\theta) = 0$, the pointwise maximizer is unique, and hence the mechanism is unique almost everywhere outside the set $\varphi^{-1}(0)$.

B.14   Proof of Theorem 5

The proof of Theorem 2 (Appendix B.2) can be extended to allow for multiple agents by taking products like in the proof of Theorem 1 (Appendix B.1). For the sufficiency, define $\hat{k}_i$ for each agent i and set $\hat{k} = \otimes_i \hat{k}_i$. For necessity, if there is some j such that $\hat{t}_j$ is not most-discerning, apply the construction above on agent j, assuming every type of every other agent is indifferent over all decisions.

B.15 Proof of Proposition 9

Applying the same argument player by player gives

$$V(Q) = \int_\Theta \Big(\sum_{i=1}^n \varphi_i(\theta_i) q_i(\theta_i) \Big) f(\theta) \, d\theta.$$

This is maximized by the interim quantity functions $q^\star$ in the theorem statement, which induces monotone interim quantity functions. For each agent i, the transfer function is pinned down by the envelope expression for U_i. We then choose a transfer function $t^\star$ consistent with these interim transfer functions that also satisfies the ex post participation constraints.

C Supplementary proofs

C.1 Proof of Lemma 1

(i) Fix $p \in [0,1]$. It suffices to show that $(\mu \tilde{F}_\mu)[0,p] = p$. For each $s \in \mathbf{R}$, let $F_\mu(s-)$ denote the left limit of F_μ at s. With this notation, we have

$$\tilde{F}_\mu(s, [0,p]) = \begin{cases} 1 & \text{if } F_\mu(s) \leq p, \\ \frac{p - F_\mu(s-)}{F_\mu(s) - F_\mu(s-)} & \text{if } F_\mu(s-) \leq p < F_\mu(s), \\ 0 & \text{if } F_\mu(s-) > p. \end{cases}$$

Set $F_\mu^+(p) = \sup\{t \in \mathbf{R} : F_\mu(t-) \leq p\}$.[30] By the left-continuity of the map $t \mapsto F_\mu(t-)$, we have $F_\mu(F_\mu^+(p)) \leq p$, with equality if μ is continuous at $F_\mu^+(p)$. If μ is continuous at $F_\mu^+(p)$, then

$$(\mu \tilde{F}_\mu)[0,p] = \mu(-\infty, F_\mu^+(p)] = F_\mu(F_\mu^+(p)) = p.$$

If μ is discontinuous at $F_\mu^+(p)$, then

$$(\mu \tilde{F}_\mu)[0,p] = F_\mu(F_\mu^+(p)-) + \mu(\{F_\mu^+(p)\}) \frac{p - F_\mu(F_\mu^+(p)-)}{F_\mu(F_\mu^+(p)) - F_\mu(F_\mu^+(p)-)},$$

and the right side simplifies to p.

(ii) Fix $s \in \mathbf{R}$. For each $p \in [0,1]$, recall that for all $s \in \operatorname{supp}\nu$, we have $Q_\nu(p) \leq s$ if and only if $p \leq F_\nu(s)$. Therefore,

$$\tilde{Q}_\nu(p, (-\infty, s]) = \begin{cases} 1 & \text{if } p \leq F_\nu(s), \\ 0 & \text{if } p > F_\nu(s). \end{cases}$$

We conclude that

$$(U_{[0,1]} \tilde{Q}_\nu)(-\infty, s] = U_{[0,1]}[0, F_\nu(s)] = F_\nu(s) = \nu(-\infty, s].$$

---

[30]This is the right-continuous inverse of F_μ and is more commonly defined as $\inf\{t \in \mathbf{R} : F_\mu(t) > p\}$.

(iii) This follows immediately from (i) and (ii).

C.2   Proof of Lemma 2

By Lemma 1 (iii), we have (ii) $\implies$ (iii). We prove that (i) $\implies$ (ii) $\implies$ (i).

(i) $\implies$ (ii). Suppose $\mu \succeq_{\mathrm{SD}} \nu$. Set $k = \tilde{F}_\mu \tilde{Q}_\nu$. Fix $s \in S$ and let $S_0 = (-\infty, s] \cap S$. Recall that the left-continuous quantile function satisfies the Galois inequality

$$Q_\nu(p) \le s \iff p \le F_\nu(s).$$

Thus,

$$k(s, S_0) = \int_0^1 \tilde{F}_\mu(s, \mathrm{d}p) \tilde{Q}_\nu(p, S_0) = \int_0^{F_\nu(s)} \tilde{F}_\mu(s, \mathrm{d}p).$$

By first-order stochastic dominance, $F_\nu(s) \ge F_\mu(s)$. Thus, the right side is at least $\tilde{F}_\mu(s, [0, F_\mu(s)])$, which equals 1.

(iii) $\implies$ (i). Let k be a downward transition. Fix s in S, and let $S_0 = S \cap (-\infty, s]$. We have

$$(\mu k)(S_0) = \int_{\mathbf{R}} \mu(\mathrm{d}t) k(t, S_0) \ge \int_{(-\infty, s]} \mu(\mathrm{d}t) k(t, S_0) = \mu(S_0).$$

C.3   Proof of Lemma 3

(i) Suppose m is monotone on S. If $\mu \ge_{\mathrm{SD}} \nu$, then by Lemma 2 (iii) there is a downward transition k on S such that $\mu k = \nu$. Fix $s_0 \in S$, and set $S_0 = (-\infty, s_0] \cap S$. Then

$$\begin{aligned}
(\nu m)(S_0) - (\mu m)(S_0) &= (\mu k m)(S_0) - (\mu m)(S_0) \\
&= \int_S \mu(\mathrm{d}s) \big[(km)(s, S_0) - m(s, S_0) \big] \\
&= \int_S \mu(\mathrm{d}s) \int_{-\infty}^s k(s, \mathrm{d}t) \big[m(t, S_0) - m(s, S_0) \big] \\
&\ge 0,
\end{aligned}$$

where we get the second equality because k is downward, and the inequality because m is monotone.

For the other direction, take $\mu = \delta_s$ and $\nu = \delta_t$ for $s > t$. Then

$$m_s = \delta_s m \ge_{\mathrm{SD}} \delta_t m = m_t.$$

(ii) Let m and m' be monotone transitions on S. Suppose μ and ν are measures on S satisfying $\mu \ge_{\mathrm{SD}} \nu$. Applying (i) twice, we have $\mu m \ge_{\mathrm{SD}} \nu m$ and hence

$$\mu(mm') = (\mu m)m' \ge_{\mathrm{SD}} (\nu m)m' = \mu(mm').$$

By (i), mm' is monotone.

C.4 Proof of Lemma 4

Consider a Markov transition $k\colon X\times\mathcal{Y}\to[0,1]$. Define $\bar{k}\colon X\times\overline{\mathcal{Y}}\to[0,1]$ by setting $\bar{k}_x$ equal to the extension of k_x to $\overline{\mathcal{Y}}$, for each x in X. For each B in $\mathcal{Y}$, we have $\bar{k}(x,B)=k(x,B)$, so the map $x\mapsto\bar{k}(x,B)$ is measurable and hence universally measurable. We need to check universal measurability for sets B in $\overline{\mathcal{Y}}$. Let μ be an arbitrary probability measure on $(X,\mathcal{X})$. It suffices to show that $\bar{k}(\cdot,B)$ is $\overline{\mathcal{X}}_\mu$-measurable for all B in $\overline{\mathcal{Y}}_{(\mu k)}$. If B is in $\overline{\mathcal{Y}}_{(\mu k)}$, then we can sandwich B between B_1 and B_2 satisfying

$$0=(\mu k)(B_2\setminus B_1)=\mu(k(\cdot,B_2))-\mu(k(\cdot,B_1)).$$

So the function $\bar{k}(\cdot,B)$ is sandwiched between the $\mathcal{X}$-measurable functions $k(\cdot,B_1)$ and $k(\cdot,B_2)$, which agree μ–almost surely. Hence $\bar{k}(\cdot,B)$ is $\overline{\mathcal{X}}_\mu$-measurable.

C.5 Proof of Lemma 5

For the inclusion, it suffices to show that every probability measure μ on $(X\times Y,\mathcal{X}\otimes\mathcal{Y})$, we have

$$\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}}\subset\overline{(\mathcal{X}\otimes\mathcal{Y})}_\mu.$$

Fix such a probability measure μ. Define measures μ_1 and μ_2 on $\mathcal{X}$ and $\mathcal{Y}$ by

$$\mu_1(A)=\mu(A\times Y)\quad\text{and}\quad\mu_2(B)=\mu(X\times B).$$

If A is in $\overline{\mathcal{X}}_{\mu_1}$, then there exist A_1 and A_2 in $\mathcal{X}$ sandwiching A such that

$$0=\mu_1(A_2\setminus A_1)=\mu((A_2\setminus A_1)\times Y)=\mu((A_2\times Y)\setminus(A_1\times Y)).$$

Similarly, if B in $\overline{\mathcal{Y}}_{\mu_2}$, then $X\times B$ is in $\overline{(\mathcal{X}\otimes\mathcal{Y})}_\mu$. Taking intersections we conclude that $A\times B$ is in $\overline{(\mathcal{X}\otimes\mathcal{Y})}_\mu$. Therefore,

$$\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}}\subset\overline{\mathcal{X}}_{\mu_1}\otimes\overline{\mathcal{Y}}_{\mu_2}\subset\overline{(\mathcal{X}\otimes\mathcal{Y})}_\mu.$$

Now we turn to the last equality. For each probability measure μ on $\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}}$, let μ_0 be the restriction of μ to $\mathcal{X}\otimes\mathcal{Y}$. Then

$$\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}}\subset\overline{(\mathcal{X}\otimes\mathcal{Y})}_{\mu_0}\subset\overline{(\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}})}_\mu.$$

Taking the intersection over all such μ gives $\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}}\subset\overline{\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}}}$.

Now we prove the reverse inclusion. Each probability measure ν on $\mathcal{X}\otimes\mathcal{Y}$ has a complete extension $\bar{\nu}$ to $\overline{(\mathcal{X}\otimes\mathcal{Y})}_\nu$. Let ν_0 be the restriction of $\bar{\nu}$ to $\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}}$. Then

$$\overline{(\mathcal{X}\otimes\mathcal{Y})}_\nu=\overline{(\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}})}_{\nu_0}\supset\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}}.$$

Taking the intersection over all such ν gives $\overline{\mathcal{X}\otimes\mathcal{Y}}\supset\overline{\overline{\mathcal{X}}\otimes\overline{\mathcal{Y}}}$.

References

ALIPRANTIS, C. D. AND K. C. BORDER (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, 3 ed.

AULETTA, V., P. PENNA, G. PERSIANO, AND C. VENTRE (2011): "Alternatives to Truthfulness are Hard to Recognize," *Autonomous Agents and Multi-Agent Systems*, 22, 200–216.

BALBUZANOV, I. (2019): "Lies and Consequences," *International Journal of Game Theory*, 1–38.

BEN-PORATH, E., E. DEKEL, AND B. L. LIPMAN (2014): "Optimal Allocation with Costly Verification," *American Economic Review*, 104, 3779–3813.

——— (2017): "Disclosure and choice," *The Review of Economic Studies*, 85, 1471–1501.

——— (2019): "Mechanisms With Evidence: Commitment and Robustness," *Econometrica*, 87, 529–566.

BLACKWELL, D. (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

BORDER, K. C. AND J. SOBEL (1987): "Samurai Accountant: A Theory of Auditing and Plunder," *Review of Economic Studies*, 54, 525–540.

BULL, J. AND J. WATSON (2004): "Evidence Disclosure and Verifiability," *Journal of Economic Theory*, 118, 1–31.

——— (2007): "Hard Evidence and Mechanism Design," *Games and Economic Behavior*, 58, 75–93.

CARAGIANNIS, I., E. ELKIND, M. SZEGEDY, AND L. YU (2012): "Mechanism Design: From Partial to Probabilistic Verification," in *Proceedings of the 13th ACM Conference on Electronic Commerce*, New York, NY, USA: ACM, EC '12, 266–283.

COHN, D. L. (2013): *Measure Theory*, Birkhäuser Advanced Texts Basler Lehrbücher, Birkhäuser Basel, 2 ed.

CROCKER, K. J. AND J. MORGAN (1998): "Is Honesty the Best Policy? Curtailing Insurance Fraud through Optimal Incentive Contracts," *Journal of Political Economy*, 106, 355–375.

DEB, R. AND C. STEWART (2018): "Optimal Adaptive Testing: Informativeness and Incentives," *Theoretical Economics*, 13, 1233–1274.

DEMARZO, P. M., I. KREMER, AND A. SKRZYPACZ (2019): "Test Design and Minimum Standards," *American Economic Review*, 109, 2173–2207.

DENECKERE, R. AND S. SEVERINOV (2008): "Mechanism Design with Partial State Verifiability," *Games and Economic Behavior*, 64, 487–513.

———— (2017): "Screening, Signalling and Costly Misrepresentation," Working paper.

Dziuda, W. and C. Salas (2018): "Communication with Detectable Deceit," *Available at SSRN 3234695.*

Erlanson, A. and A. Kleiner (2015): "Costly Verification in Collective Decisions," Working paper.

Ferraioli, D. and C. Ventre (2018): "Probabilistic Verification for Obviously Strategyproof Mechanisms," in *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, AAMAS '18, 1930–1932.

Fotakis, D. and E. Zampetakis (2015): "Truthfulness Flooded Domains and the Power of Verification for Mechanism Design," *ACM Transactions on Economics and Computation*, 3, 20:1–29.

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

Grossman, S. J. (1981): "The informational role of warranties and private disclosure about product quality," *The Journal of Law and Economics*, 24, 461–483.

Halac, M. and P. Yared (2017): "Commitment vs. Flexibility with Costly Verification," Working paper.

Hart, S., I. Kremer, and M. Perry (2017): "Evidence Games: Truth and Commitment," *American Economic Review*, 107, 690–713.

Kallenberg, O. (2017): *Random Measures, Theory and Applications*, vol. 77 of *Probability Theory and Stochastic Modelling*, Springer.

Kartik, N. (2009): "Strategic Communication with Lying Costs," *Review of Economic Studies*, 76, 1359–1395.

Kartik, N., M. Ottaviani, and F. Squintani (2007): "Credulity, Lies, and Costly Talk," *Journal of Economic Theory*, 134, 93–116.

Kephart, A. and V. Conitzer (2016): "The Revelation Principle for Mechanism Design with Reporting Costs," in *Proceedings of the 2016 ACM Conference on Economics and Computation*, 85–102.

Koessler, F. and E. Perez-Richet (2017): "Evidence Reading Mechanisms," Working paper.

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

Li, Y. (2017): "Mechanism Design with Costly Verification and Limited Punishments," Working paper.

LIPMAN, B. L. AND D. J. SEPPI (1995): "Robust Inference in Communication Games with Partial Provability," *Journal of Economic Theory*, 66, 370–405.

MAGGI, G. AND A. RODRIGUÉZ-CLARE (1995): "Costly Distortion of Information in Agency Problems," *RAND Journal of Economics*, 26, 675–689.

MERTENS, J.-F., S. SORIN, AND S. ZAMIR (2015): *Repeated Games*, vol. 55 of *Econometric Society Monographs*, Cambridge University Press.

MILGROM, P. AND I. SEGAL (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583–601.

MILGROM, P. R. (1981): "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 380–391.

MUSSA, M. AND S. ROSEN (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

MYERSON, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

MYLOVANOV, T. AND A. ZAPECHELNYUK (2017): "Optimal Allocation with Ex Post Verification and Limited Penalties," *American Economic Review*, 107, 2666–2694.

NISAN, N. AND A. RONEN (2001): "Algorithmic Mechanism Design," *Games and Economic Behavior*, 35, 166–196.

RILEY, J. AND R. ZECKHAUSER (1983): "Optimal Selling Strategies: When to Haggle, When to Hold Firm," *Quarterly Journal of Economics*, 98, 267–289.

ROCHET, J.-C. (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context," *Journal of Mathematical Economics*, 16, 191–200.

SINGH, N. AND D. WITTMAN (2001): "Implementation with Partial Verification," *Review of Economic Design*, 6, 63–84.

STRAUSZ, R. (2016): "Mechanism Design with Partially Verifiable Information," Cowles Foundation Discussion Paper No. 2040.

TOWNSEND, R. M. (1979): "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, 21, 265–293.

VOHRA, R. V. (2011): *Mechanism Design: A Linear Programming Approach*, Econometric Society Monographs, Cambridge University Press.

YU, L. (2011): "Mechanism Design with Partial Verification and Revelation Principle," *Autonomous Agents and Multi-Agent Systems*, 22, 217–223.