Evidence in Games and Mechanisms[1]

Elchanan Ben-Porath[2] Eddie Dekel[3] Barton L. Lipman[4]

Current Draft
August 2025

[1] We thank the National Science Foundation, grant SES–1919319 (Dekel and Lipman), and the US–Israel Binational Science Foundation for support for this research.

[2] Department of Economics and Center for Rationality, Hebrew University. Email: benporat@math.huji.ac.il.

[3] Economics Department, Northwestern University, and School of Economics, Tel Aviv University. Email: eddiedekel@gmail.com.

[4] Department of Economics, Boston University. Email: blipman@bu.edu.

Abstract

We survey theoretical work on the use of evidence, including work in game theory and in mechanism design.

1 Introduction

Standard models of strategic communication are primarily based on the key insight of Spence (1974) and Crawford–Sobel (1982). The idea of these models is to use variation in preferences across types to induce different choices by different types. In Spence, the agent's costs of actions depend on her type and so seeing these actions reveals information about her type. In Crawford and Sobel, the actions she wants the receiver to take depend on her type and so seeing her message (implicitly, her request for an action) reveals something about her preferences and hence her private information.

Similiarly, traditional work in mechanism design (e.g., Myerson (1981) and Maskin–Riley (1984)) introduces monetary transfers and differences across types in willingness to pay. Again, differences in preferences that are correlated with private information are exploited to obtain an outcome that depends on the private information.

By contrast, the focus of this survey is on the role of evidence — hard information that establishes facts regardless of the incentives of the presenter of the information.

Evidence plays an important role in many contexts. A lawyer presents evidence in court to try to persuade a judge to rule in her favor. Sellers may provide evidence to buyers about their product to persuade them to purchase. An entrepreneur can show evidence about her company to influence potential investors. A department head in an organization presents evidence regarding the abilities of his department to try to persuade the management to give them more resources. A potential employee may be asked for evidence of her ability by a potential employer. In many of these cases, what the presenter of evidence wants is independent of the facts they are giving evidence about. For example, the lawyer wants a ruling in favor of her client, in any situation. The seller wants the buyer to purchase, the entrepreneur wants funding, and the potential employee wants to be hired, regardless the value of this to the other party. In such settings, information cannot credibly be communicated without the use of evidence.

While the incentives of the informed party to reveal evidence will also be important to the equilibrium belief in response to the evidence, without evidence, incentive considerations alone would not allow information transmission in many of these settings.

Section 2 presents the basic model of evidence. We discuss applications in game theory in Section 3, beginning with the classical single–agent model and unraveling in Section 3.1 before turning to models without unraveling in Sections 3.2, 3.3, and 3.4. Section 3.5 discusses games with multiple agents. We turn to mechanism design in Section 4. We focus primarily on one–agent mechanisms, discussing the Revelation Principle in Section 4.1, some applications in Section 4.2, and the value of commitment in Section 4.3. Section 4.4 discusses some mechanism design problems with multiple agents. In Section 5, we turn to related models, discussing verification in Section 5.1 and evidence acquisition in

Section 5.2. Section 6 briefly concludes.

Because of space constraints, there are many outstanding papers we only briefly mention or even omit entirely. The discussion of many of our own papers is not because we believe these are the best papers in the literature but because these are the papers we know the best.

2 Modeling Evidence

Throughout, we use t (or, in the multi–agent case, t_i) to denote the *type* of the (an) agent and T (respectively T_i) the set of possible types. As usual, we assume the agent knows her type but this is private information. We assume a common prior μ over T. Types vary in aspects similar to those in the usual literature — e.g., different types may have different productivities or different tastes — but also differ in the evidence they can present.

There are two equivalent ways to model evidence which we use interchangably. In the first, we specify a function $\mathcal{M} : T \to \mathcal{L}$ where $\mathcal{L}$ is the set of all possible evidence messages. Here $\mathcal{M}(t)$ is the set of evidence messages type t is able to send.

In this formulation, sending message m is *proof* that the agent is able to send message m. That is, it proves that the agent's type is in the set of t with $m \in \mathcal{M}(t)$, i.e., proves $t \in \mathcal{M}^{-1}(m)$. If I show you a deed to a house with my name on it, that unambiguously proves that either I have owned a house or that I have acquired a forged document. There is no other way I could show you this document, so I prove that my type is in this set.

For example, suppose the agent's private information is whether she can play the piano. So $T = \{n, p\}$ where p is the piano–playing type and n can't play. Let $\mathcal{L} = \{c, r\}$ where c means playing a classical sonata and r means random banging on the keyboard. Then $\mathcal{M}(p) = \{c, r\}$, while $\mathcal{M}(n) = \{r\}$. That is, the type who can play the piano can play the sonata or simply bang on the keyboard, while the type who can't play can only do the latter.

Note that playing c proves that the agent's type is p. On the other hand, playing r proves nothing as either type could do this. As this illustrates, provability can be asymmetric in the sense that it might be possible to prove event E when it is true, but not possible to prove E^c when E is false.

Alternatively, we could focus on the set of events a type can prove. That is, we can let $\mathcal{E}(t)$ be the set of $E \subseteq T$ that type t can prove. Given a function $\mathcal{M}$, we can define this by

$$\mathcal{E}(t) \equiv \{E \subseteq T \mid E = \mathcal{M}^{-1}(m), \text{ for some } m \in \mathcal{M}(t\}. \tag{1}$$

The second model takes as the primitive a function $\mathcal{E}$. It is not difficult to show that there exists $\mathcal{M}$ generating $\mathcal{E}$ in the sense of equation (1) if and only if $\mathcal{E}$ satisfies the following two properties. First, evidence must be true. That is, $E \in \mathcal{E}(t)$ implies $t \in E$. Second, $\mathcal{E}$ is consistent in the sense that any type not ruled out by evidence E must be a type who has evidence E available. That is, if $E \in \cup_{t \in T} \mathcal{E}(t)$ (so E is something some type can prove) and $s \in E$, then $E \in \mathcal{E}(s)$.

Returning to the piano player example, it is not hard to see that we could equally well describe that example by saying that type n can only prove T, while type p can prove either $\{p\}$ or T. So $\mathcal{E}(n) = \{T\}$ and $\mathcal{E}(p) = \{\{p\}, T\}$.

With either approach, we assume the agent can only present *one* piece of evidence — that is, present one message or prove one event. At least for some purposes, this is without loss of generality. If the agent could present, say, K messages, we could redefine what "a message" is and replace messages with K–tuples of messages.

Much of the literature adds an assumption which makes the restriction to one piece of evidence irrelevant. Intuitively, this assumption implies that there are no costs of or time constraints on evidence presentation in the sense that it is as if the agent could present an unlimited number of messages. This condition was first given in Lipman–Seppi (1995) who referred to it as the *full reports condition.* It says that for every $t \in T$,

$$E_t^* \equiv \bigcap_{E \in \mathcal{E}(t)} E \in \mathcal{E}(t).$$

E_t^*, as defined above, is what the agent would prove if she were able to prove every $E \in \mathcal{E}(t)$. So the statement of the assumption is that this is itself an event the agent of type t can prove. We refer to the event E_t^* as *maximal evidence.* Bull and Watson (2007) gave an equivalent definition using the messages model which they referred to as *normality*, the name most commonly used in the literature. Their formulation is that for every t, there exists $m_t^* \in M(t)$ such that $m_t^* \in M(t')$ iff $M(t) \subseteq M(t')$. This condition implies that presenting m_t^* proves the event $\{t' \mid M(t) \subseteq M(t')\}$, exactly what presenting every message in $M(t)$ would prove.

Returning to the piano player example, it is easy to see that this satisfies normality. In the event version, we have $E_p^* = \{p\} \in \mathcal{E}(p)$ and $E_n^* = T \in \mathcal{E}(n)$.

For an analogous example that violates normality, suppose there are two kinds of music and four types of agent. Type n cannot play the piano, type c can play classical music, type b can play blues music, and type a can play anything. If there's only time to play one piece of music, then (focusing on the event version of the model)

$$\mathcal{E}(n) = \{T\}$$

$$\mathcal{E}(c) = \{\{c, a\}, T\}$$

$$\mathcal{E}(b) = \{\{b, a\}, T\}$$

$$\mathcal{E}(a) = \{\{c, a\}, \{b, a\}, T\}.$$

For each of the first three sets, $E_t^* \in \mathcal{E}(t)$. However, $E_a^* = \{a\} \notin \mathcal{E}(a)$, so normality is violated.

3 Games

In this section, we discuss game–theoretic models with evidence. Unless stated otherwise, "equilibrium" refers to perfect Bayesian equilibrium. We use the terms "sender" and "agent" interchangably to refer to the privately informed player who communicates information and "receiver" or "principal" to refer to the player who receives the communication.

3.1 Classical Single–Agent Model

In the classic model, there is a single sender who first learns her type $t \in T$, then sends an evidence message to the receiver. The receiver then chooses some action $a \in A$ which affects both of their utilities. We let $u(a, t)$ denote the utility of the sender and $v(a, t)$ the utility of the receiver.

The seminal papers on evidence are Grossman (1981) and Milgrom (1981). To present the key ideas most simply, we focus on a square–error loss utility function for the receiver. More specifically, we assume T is a finite subset of $\mathbf{R}_+$, $A = \mathbf{R}_+$, $u(a, t) = a$, and $v(a, t) = -(a - t)^2$. In other words, the receiver is trying to estimate the sender's type and so the receiver's optimal response is his conditional expectation or estimate of the type. The sender wants the receiver's estimate to be as large as possible.

For example, we could think of the receiver as an employer and the sender as an employee. As in Spence, the receiver's action is a wage and the receiver wants to set this equal to the sender/employee's productivity t. As another example which we will use later, the receiver could be "the market" wanting to price a firm's stock at its true value t. In this formulation, the sender is the manager of the firm and a is the stock price.

Grossman and Milgrom assume *complete provability* — that is, the sender has access to evidence that can prove any true fact. More formally, $\mathcal{E}(t) = \{E \subseteq T \mid t \in E\}$.

The result shown by Grossman and Milgrom is remarkable at first sight: In every equilibrium, the receiver's action equals the sender's type for every t. That is, the receiver always learns the sender's type.

While we show this result for a finite type space and pure strategy equilibria, it is easy to generalize. We write T as $\{t_1, \ldots, t_N\}$ where $t_1 < t_2 < \ldots < t_N$. Fix any equilibrium, say $(m^*(\cdot), a^*(\cdot), \mu^*(\cdot))$ where $m^* : T \to \mathcal{L}$ is the sender's strategy and must satisfy $m^*(t) \in \mathcal{M}(t)$, $a^* : \mathcal{L} \to A$ is the receiver's response, and $\mu^* : \mathcal{L} \to \Delta(T)$ is the receiver's belief. Note that the receiver's belief must respect the evidence presented in the sense that $\mu^*(\mathcal{M}^{-1}(m)) = 1$ for all $m \in \mathcal{L}$. Sequential rationality implies that $a^*(m) = \sum_t \mu^*(m)t$ — that is, it is the conditional expectation of t given the receiver's equilibrium belief.

Suppose that $a^*(m^*(t_N)) \neq t_N$. That is, the receiver's response to the message sent by type t_N in the equilibrium is not t_N. Sequential rationality implies it must be weakly below t_N, so $a^*(m^*(t_N)) < t_N$.

But then type t_N could strictly improve by proving her type is t_N. The belief in response to such a message must be degenerate on t_N so sequential rationality implies the response by the receiver must be $a = t_N$. Since t_N strictly gains from the deviation, we have a contradiction, implying $a^*(m^*(t_N)) = t_N$.

This implies that $m^*(t_N)$ must prove t_N. Otherwise, some other type could also send this message and get a response strictly above what she would get in equilibrium, a contradiction.

Hence the receiver's belief in response to $m^*(t_{N-1})$ must put zero probability on t_N and so his action in response must be weakly less than t_{N-1}. So suppose $a^*(m^*(t_{N-1})) < t_{N-1}$. Then, again, t_{N-1} would strictly gain by proving her type, a contradiction. Hence $m^*(t_{N-1})$ must prove $t \geq t_{N-1}$.

Clearly, the reasoning works down to t_1. Hence there is no equilibrium in which the response to some type's equilibrium message differs from her type. This argument is often referred to as *unraveling* since the top type must reveal, implying the second highest must, implying ...

This argument shows that there is no equilibrium in which $a^*(m^*(t)) \neq t$ for some t. That there is an equilibrium where $a^*(m^*(t)) = t$ for all t is easily shown using Milgrom's notion of *skeptical beliefs*. In fact, every equilibrium has skeptical beliefs.[1]

Formally, given any message $m \in \mathcal{L}$, define the skeptical belief for the receiver, say $\mu_S^*(m)$, to put probability 1 on the smallest t such that $m \in \mathcal{M}(t)$ and define a_S^* to equal this smallest t, as required by sequential rationality given this belief.

It is easy to see that the sender's best response to this strategy for the receiver is to rule out any type below her true type. It is impossible for her to provide evidence ruling out her true type. Ruling out types above her own is feasible but does not affect

[1] See Rappoport (2024) for a broader study of skeptical beliefs in equilibria of games with evidence.

her payoff. So the best she can do is to rule out all types below her true type. Given this strategy by the sender, the receiver's belief is correct in equilibrium. Thus we get an equilibrium where $a^*(m^*(t)) = t$ for all t. Furthermore, if the receiver had a different belief in response to some message, the worst type who could send that message, say t, would get a response $a > t$, contradicting our observation that $a^*(m^*(t)) = t$ in any equilibrium.

One gets the same result with various forms of less than complete provability. The simplest is what is referred to as a *disclosure game*. Here each type can either prove her type ("disclose") or prove nothing. That is, $\mathcal{E}(t) = \{\{t\}, T\}$ for all t. In this case, it's clear that the highest type will prefer disclosing her type to pooling with lower types. Hence the second–highest type cannot pool with the highest type and so will definitely want to disclose as well, etc.[2]

With complete provability, both the existence and uniqueness of this equilibrium outcome are easily generalized to a wide range of other preference structures. (See Seidmann and Winter (1997).) However, the *existence* of such an equilibrium outcome is much more general than its *uniqueness*. To see the point, consider the following payoff structure.

Suppose the receiver has two actions, $A = \{0, 1\}$ where 1 is "accepting" the receiver and 0 is "rejecting." The sender has a finite set of types, say, $\{t_1, \ldots, t_N\}$ where $t_1 < t_2 < \ldots < t_N$. The sender's payoff is again equal to a. The receiver's payoff is

$$v(a, t) = a(t - \bar{t})$$

for some $\bar{t} \in (t_1, t_N)$. In other words, the receiver (strictly) prefers to accept the sender (set $a = 1$) if the sender's type is (strictly) above $\bar{t}$ and prefers to reject otherwise.

Again, there are many natural economic examples fitting this setup. Suppose that the receiver is an employer who cannot set the wage $\bar{t}$ he must pay, only decide whether or not to hire the sender. The sender always wishes to be hired, while the receiver wants to hire only if the sender's productivity t is above the wage.

Using skeptical beliefs, we again get an equilibrium where the receiver effectively learns the sender's type and chooses her action accordingly. If the receiver always believes the worst thing consistent with whatever the sender proves, the sender will prove her type is above $\bar{t}$ if she can.

However, there can be other equilibria. For example, suppose that $\mathrm{E}(t) > \bar{t}$. Then it is an equilibrium for the sender to never prove anything and for the receiver to always accept.

Why do we not get unraveling? The highest type of the sender could deviate and prove

[2] More generally, one can show that this is the unique equilibrium outcome iff every type t can prove an event for which t is the smallest type.

something strictly better about herself. But this strictly better belief doesn't translate into a strictly better response by the receiver, so the highest type has no incentive to do so.

See Titova (2023), Zhang (2024), and Ali, Kleiner, and Zhang (2024) for more general characterizations of the set of equilibrium payoffs for the sender in these games, including when the sender can do as well as in the Kamenica–Gentzkow (2011) Bayesian persuasion model. See also Callander, Lambert, and Matouschek (2021) and Ali, Lewis, and Vasserman (2023) for economic applications.

3.2 Games with Incomplete Provability: Dye Evidence

Complete provability is a natural starting point, but does not seem realistic. In this section, we discuss plausible relaxations of complete provability which given interesting models which are useful in economic applications.

The most widely used such relaxation is often referred to as *Dye evidence* in honor of the first author to use the idea, Dye (1985).[3]

The Dye evidence model is a variation on the disclosure game discussed in Section 3.1. In the Dye model, some types can disclose but some types cannot. In this model, types are fundamentally (at least) two–dimensional because the type determines what the receiver wants to know but also determines whether the sender can disclose. For this reason, we generalize the model to assume the agent's type t determines a certain value, $v(t)$, which the receiver cares about and also determines the sender's evidence.

Specifically, Dye assumes that every type t either has no evidence or can prove her type. That is, for every t, either $\mathcal{E}(t) = \{T\}$ or $\mathcal{E}(t) = \{\{t\}, T\}$. Dye wrote this differently, saying that a given type t can disclose with some fixed probability q and not otherwise. In our formulation, this says that there are two types, t' and t'', with $v(t') = v(t'') = v$, where $\mathcal{E}(t') = \{T\}$ and $\mathcal{E}(t'') = \{\{t''\}, T\}$ and that conditional on $v(t) = v$, the probability that $t = t''$ is q. While the distinction between these formulations is not important for some purposes, it is useful for mechanism design to treat types as determining everything rather than have a second layer of randomness.

To illustrate this model, we return to the square–error loss formulation and augment it as follows. As before, $A = \mathbf{R}$ and $u(a, t) = a$. Now, though, we take the receiver's utility function to be $v(a, t) = -(a - v(t))^2$. So the receiver wants to set his action equal to $v(t)$, not t. For example, this is the natural formulation if we think of $v(t)$ as the value of a firm under a manager of type t and the receiver as "the market" setting the stock price a equal to the expected value of the firm.

[3] See also Jung and Kwon (1988) who corrected Dye's analysis.

Note that there *must* be a positive probability that the sender does not prove anything (other than the trivial event T) along the equilibrium path because there are types who have no other option. Let v^* denote the receiver's expectation of v when the sender presents no evidnece in equilibrium.

Given v^*, the sender's optimal strategy is immediate. If $\mathcal{E}(t) = \{T\}$, the sender cannot prove her type. Otherwise, if $v(t) < v^*$, proving her type would lead to a lower action than remaining silent, so the sender will provide no evidence. If $v(t) > v^*$, the sender will prove her type. Hence in equilibrium, we must have

$$v^* = \mathrm{E}\left[v(t) \mid t \text{ has no evidence or } v(t) \leq v^*\right]. \tag{2}$$

A remarkable fact about this model is that v^* exists and is unique, even without simplifying assumptions regarding the $v(\cdot)$ functions or the distribution of t.[4]

Note that v^* is independent of the behavior of indifferent types. If t is indifferent because $v(t) = v^*$, then the conditional expectation or "average" is the same whether we include this type or not as this value equals the average.

In short, we define v^* uniquely by equation (2) and (up to irrelevant indifference) this completely determines the sender's strategy and the receiver's.

This model is very widely used in applications in economics, finance, and accounting. We discuss some examples in the next section.

3.3 Applications of Dye Evidence

Shin (2003) uses the Dye evidence model to understand stock price responses to disclosure of information by the firm. The firm has N projects, each of which succeeds with probability $r \in (0, 1)$ and fails otherwise. If s projects succeed and $N - s$ fail, the value of the firm at date $D + 1$ is $h^s \ell^{N-s}$ where $0 < \ell < h$.

The sender in this model is the firm's manager. At each date $d = 1, \ldots, D$, for any given project, the manager has a probability q of receiving evidence which proves whether that project succeeds or fails. If the manager receives evidence, she can choose whether or not to disclose it. These evidence events are iid over time and independent of the success or failure of any projects.

The receiver is "the market" which observes the manager's disclosures (or lack thereof) and sets the price of the firm's stock at each date d equal to the expected value of the firm given observations up to that date. In other words, the market is a receiver with square–error loss utility.

[4] See, e.g., Lemma 2 of Guttman, Kremer,and Skrzypacz (2014).

The manager's utility is a function of the sequence of stock prices where her utility is strictly increasing in the stock price at any given date. For intutiion, think of the manager as maximizing the sum (or discounted sum) of the sequence of stock prices.

Clearly, one equilibrium of this game has the manager disclosing any success as soon as possible and never disclosing any failure. If this is the manager's strategy, then disclosing a success increases the market's expectation of the value of the firm at that and all subsequent dates, all else equal. Disclosing a failure would reduce the market's expectation. Hence it is optimal for the manager to follow this strategy.

The simplicity of this equilibrium makes it ideal for exploring the model's empirical implications. Shin contrasts the behavior of stock prices over time in this model of *strategic disclosure* versus a world of *exogenous disclosure*, where all evidence must be disclosed.

First, consider the effects of nondisclosure. With exogenous disclosure, the fact that nothing is shown means nothing was observed and hence has no effect on the stock price. With strategic disclosure, nondisclosure could mean nothing was observed but could also be because of bad news. Hence nondisclosure is bad news and reduces the stock price.

Second, consider how the effect of disclosure varies with time. With exogenous disclosure, information is disclosed as it arrives. Since the timing of the arrival of the information conveys no information regarding fundamentals, the price impact of disclosure is independent of time. With strategic disclosure, many periods of nondisclosure will lead the market to believe that it is likely the manager has learned that the project has failed. Consequently, a late disclosure of success increases the expected value of the firm by more than an early disclosure.

Finally, consider the uncertainty about (i.e., the variance of) the future stock price as a function of the current price. With exogenous disclosure, this uncertainty is not monotonic in the current price. Uncertainty will be a function of the number of projects for which the manager has disclosed the outcome, not what is disclosed. If a large number of projects have the outcome disclosed, there will be little uncertainty, whether these were all successes, giving a high current stock price, or all failures, giving a low current price. By contrast, uncertainty *is* monotonic in the current stock price under strategic disclosure. The only disclosures are successes, so more disclosure means both less uncertainty about future prices and a higher current stock price.

See Shin for more details on these comparisons, including a discussion of empirical evidence suggesting the data is more consistent with strategic disclosure than exogenous disclosure.

In Shin's model, the set of projects undertaken is exogenous. Ben-Porath, Dekel, and Lipman (2018) — henceforth BDL18 — explores the implications of disclosure on project selection and shows that the manager has incentives to choose inefficiently.

In BDL18, at period 0, the manager makes an unobserved choice of a *project* — a probability distribution over firm values — from a fixed set of options. In period 1, with probability $q \in (0, 1)$, the manager receives evidence proving the outcome of the project chosen. If so, she can disclose this to the market. The period 1 stock price is the market's expectation of the value of the firm given manager's disclosure or lack thereof. In period 2, the outcome of the project is observed by the market and we have a second stock price, equal to the true realization.

So, as in Shin, the market is the receiver and maximizes square–error loss. The manager is the sender and her utility is assumed to be a convex combination of the first and second period stock prices where α is the weight on the second or "long–run" stock price.

BDL18 show that strategic disclosure can lead to significant efficiency loss. Intuitively, the manager discloses good information and suppresses bad. Hence the manager has an incentive to take actions *ex ante* to influence this information revelation stage. These incentives are inefficient: the manager has an incentive to improve appearances even if this doesn't help (or even harms) reality.

Consider the following example. Suppose $\alpha = 0$ so the manager only cares about the first period stock price. Suppose there are two possible projects, F_1 and F_2. F_1 is a degenerate distribution giving a firm value of $x = 4$ with certainty. F_2 gives $x = 6$ with probability 1/2 and 0 otherwise. Obviously, F_1 gives a higher expected value and so is more efficient in this sense.

Is it an equilibrium for the manager to choose project F_1? If so, then in this equilibrium, in the strictly positive probability event that the manager discloses nothing at period 1, the market, knowing the manager's strategy is F_1, puts probability 1 on $x = 4$. Hence the first–period stock price is 4 if nothing is disclosed. Of course, it is also 4 if the manager discloses $x = 4$. Hence the manager's expected payoff in this hypothetical equilibrium is 4.

Suppose that the manager deviates to F_2. If the manager can't disclose anything or if the realization is 0 and the manager doesn't disclose this, the first–period stock price will be 4. If the realization is 6, the manager can disclose this and make the stock price equal to 6. Hence, the manager's expected payoff is a convex combination of 4 and 6 which is strictly larger than 4. So this is not an equilibrium as the manager would deviate.

The key observation is that the market cannot be fooled *in equilibrium*. Out of equilibrium, market can be fooled and this might be better for the manager, as in the example. This eliminates some equilibria, potentially (as in the example) making the manager and stockholders worse off.

To see the intuition, suppose we have an equilibrium where the manager chooses project F. The market must expect the distribution of profits to be F and so if the manager does not disclose, the market's belief will be given by the Dye value associated with F, say $\hat{x}_F$. Given this, the manager's expected payoff as a function of the realized value of the firm x is

$$\alpha x + (1-\alpha)[(1-q)\hat{x}_F + q \max\{\hat{x}_F, x\}].$$

Note that max is a convex function, so the manager is, in effect, risk–loving. Thus the manager prefers riskier projects, even if this requires accepting a lower mean. Similarly, note that this is strictly increasing in q. Thus the manager also prefers projects where she is more likely to have the option to disclose information, again, even if this requires accepting a lower mean.

BDL18 characterize worst–case outcomes. For example, they show that in any equilibrium with any set of possible projects, the equilibrium value of the firm can never be below 50% of the maximum possible expected value but can be arbitrarily close to 50%.

Aghamolla and An (2025) and Guttman, Kremer, Skrzypacz, and Wiedman (2025) develop further models along these lines.

DeMarzo, Kremer, and Skrzypacz (2019) study a formally similar model focused on test selection rather than production. An agent selects a test with the goal of convincing a receiver that her value is high. The main result is that the agent will select a test for which the belief in response to nondisclosure is minimal. They examine the implications for the quality of public information.

There are many other interesting game–theoretic applications of Dye evidence in the literature. For example, Acharya, DeMarzo, and Kremer (2011) use a dynamic Dye model to show that public events can trigger clustering of negative disclosure announcements by firms in response to bad market news. Guttman, Kremer, and Skrzypacz (2014) show the surprising result that otherwise equivalent disclosures have a more positive effect on market prices when they come later.

3.4 Costs of/Constraints on Evidence Presentation

A different class of models involve costs of or constraints on evidence presentation. One of the standard models in this literature, due to Verrecchia (1983), assumes that presenting evidence is costly to the sender. Describing Verrecchia in the language of this paper, he assumes the sender can exactly prove her type but this is costly. This implies that only types who gain enough from disclosure will provide evidence, again preventing unraveling.

While this model seems to have fewer applications than Dye in economics and finance,[5] it is very widely applied in accounting. We are less familiar with that literature and hence do not attempt to survey it. We note that Verrecchia's model seems especially influential in the empirical literature in accounting which explores the correlation of disclosure with other parameters to see if the empirical regularities conform to a full disclosure model or are more in line with the kind of partial disclosure induced by disclosure costs. For surveys in accounting on financial reporting that also discuss disclosure costs, see Beyer, Cohen, Lys, and Walther (2010) or Leuz and Wysocki (2016).

A related approach models such costs only indirectly, assuming that there is a limitation on how much evidence the sender can present. In other words, this approach focuses on evidence structures which are not normal.

Building on an example in Milgrom (1981), Fishman and Hagerty (1990) give a model of such constraints and illustrate the implications for communication. They consider a seller of a good with N attributes, each either h (high) or ℓ (low). The value of the good to a buyer is the number of attributes that are h. The seller has time to prove the value of any one attribute. Clearly, this evidence structure violates normality.

While there is an obvious equilibrium where the seller shows any one h chosen at random if she has one, there is a subtler and more informative equilibrium. Suppose the seller shows the *first* attribute that is h (if there is one). In this case, if the seller shows h for the kth attribute, the buyer infers that the first $k-1$ attributes are all ℓ. Hence the buyer's inference is worse for the seller the larger is k, so the seller will optimally show the lowest attribute with an h.

In Section 4.2, we discuss similarly motivated work in the area of mechanism design.

3.5 Multi–Agent Games

Additional issues arise in the multi–agent case. Suppose we have N agents, all of whom know the state of the world θ. Suppose they have potentially different abilities to provide evidence about θ and potentially different incentives to reveal information about it to the receiver.

It is intuitive that senders competing to persuade a receiver may reveal more information than any one of them would reveal without competition. This idea was first explored in the context of evidence by Milgrom and Roberts (1986) who showed that with complete provability but limited rationality by the receiver, the receiver can learn the state with conflicting interests among the senders.

[5] At least some applications of Dye evidence, such as the Shin model discussed in Section 3.3, seem unlikely to change much if redone using Verrecchia's approach.

Lipman and Seppi (1995) show separation with conflicting interests among the senders with much weaker evidence structures. Consider the following example, based on an example in their paper. The state of the world, d, is the level of damages lawyer 2's client has done to lawyer 1's client. Lawyer 1 wants the judge to conclude d is large, while lawyer 2 wants the judge to conclude it is small.

Only lawyer 2 has evidence. The structure of evidence is very restrictive. For every damage level d' not equal to the true level, there is a piece of evidence available which proves that damages are not equal to d' but proves nothing else. Lawyer 2 can only show one of these pieces of evidence. Clearly, this violates normality.

Consider the following three–stage game. First, lawyer 1 makes a claim d_1 about the true d. Next, lawyer 2 observes d_1, provides one piece of evidence, and makes her own claim d_2 about d. Finally, the judge rules on the value of d.

Suppose the judge believes lawyer 1's claim if lawyer 2 does not prove it to be false and believes lawyer 2's claim otherwise. In this case, it is clear that lawyer 1 will report truthfully. Lawyer 2 cannot refute the truth, so this will be the judge's inference. If she lied instead, lawyer 2 would always refute the lie if there is some other belief that would be better for lawyer 2 and hence worse for lawyer 1.

Note that the judge learns the true level of damages, even if he doesn't know the preferences of the lawyers or even the range of possible d's.

Hagenbach, Koessler, and Perez–Richet (2014) give a characterization of equilibria with separation with multiple agents and partial provability. Roughly speaking, the key condition on the evidence structure is that one can use Milgrom's skeptical beliefs to support a fully revealing equilibrium.

Onuchic and Ramos (2025) consider a very different multi–agent model, where agents are a team who *jointly* control disclosure decisions. For example, suppose there are two agents with independent types, t_1 and t_2. Suppose that the true $t = (t_1, t_2)$ can be disclosed or not. That is, the events that can be proved are $\mathcal{E}(t_1, t_2) = \{\{(t_1, t_2)\}, T_1 \times T_2\}$. Assume preferences are as in the square–error loss model, so the receiver chooses two actions, a_1 and a_2, his payoff is $-\sum_i (a_i - t_i)^2$, and agent i's utility is a_i. Suppose disclosure occurs only if both agents agree — either can block disclosure unilaterally and assume that the receiver sees only what, if anything, is disclosed.

One equilibrium of this game is a version of the Dye model where the availability of evidence is endogenous. If we let t_i^* be the receiver's expectation of t_i given that nothing is disclosed, it seems natural that we get disclosure iff $t_i > t_i^*$ for both i. Hence we have an equilibrium with

$$t_i^* = \mathrm{E}\left[t_i \mid t_1 \leq t_1^* \text{ or } \mathrm{t}_2 \leq \mathrm{t}_2^*\right], \ i = 1, 2.$$

Here i's ability to disclose is determined by j's incentive to do so.

4 Mechanism Design

We primarily focus on mechanism design with one agent, discussing applications with multiple agents in Section 4.4.

4.1 Revelation Principle

To consider optimal mechanisms with evidence, we first develop a useful form of the Revelation Principle identifying a relatively simple class of mechanism structures we can reduce attention to. Throughout, we maintain the usual assumption that the agent can only present one evidence message.

As above, T denotes the set of types of the agent and A the set of outcomes or actions for the principal. We let $u : A \times T \to \mathbf{R}$ denote the agent's utility function and $v : A \times T \to \mathcal{M}$ the principal's.

An outcome of a mechanism is a function $f : T \to \Delta(A)$. As usual, f is *implementable* if there exists a mechanism and an equilibrium of that mechanism whose outcome is f.

The most general result we discuss is the following. If f is implementable, then it is implementable in a multi–stage mechanism with the following structure. First, the agent makes a cheap talk report of her type. Next, the principal requests an evidence message, possibly at random, as a function of this report. Third, the agent sends an evidence message. Finally, the mechanism specifies an outcome, possibly random, as a function of the history. The mechanism is chosen so that it is optimal for the agent to report her type truthfully and send the requested evidence.

The argument for why this result is true is similar to a standard intuition for the Revelation Principle. Given any other game and equilibrium, we ask the agent her type and play her equilibrium strategy for her, asking her to present evidence when we hit the points at which this is required. If the original strategies were an equilibrium, she would have no incentive to deviate.

Note that if the mechanism makes a deterministic request for evidence as a function of the agent's reported type, then we could omit this step without changing anything. Hence this step is only important if the request is random. Why is such randomness needed?

As first observed by Glazer and Rubinstein (2004), when evidence is not normal, this randomization can make all of the agent's available evidence messages relevant.[6] To see

[6] See Carroll and Egorov (2019) for a more general analysis of when such randomization can induce

the idea, consider the following example. Suppose the agent has three types, t_1, t_2, and t_3. Suppose the evidence sets are $\mathcal{M}(t_1) = \{m_1\}$, $\mathcal{M}(t_2) = \{m_2\}$, and $\mathcal{M}(t_3) = \{m_1, m_2\}$. Note that this evidence structure violates normality.

Consider the outcome f where we give t_3 \$1 and nothing to t_1 or t_2. Take the mechanism to be as follows. If the agent reports that her type is t_1 or t_2, the mechanism requests the only message the reported type has and gives the agent 0 regardless of the evidence she provides. If the agent reports type t_3, the mechanism requests m_1 with probability 1/2 and m_2 otherwise. If the agent provides the requested evidence, she receives \$1. Otherwise, she is fined \$10.

It is easy to see that this mechanism implements f as it induces the agent to report truthfully. Only t_3 will report t_3 since only t_3 has both messages and hence can be sure to avoid the fine. By being willing to claim t_3, the agent, in effect, does show both messages as she signals clearly that she has both.

In many settings, large fines like the ones above, even off the equilibrium path, are unnatural and such mechanisms are unintuitive. If we assume normality, we obtain an alternative simplification in which this kind of mechanism is not useful.

With normal evidence, it is without loss of generality to focus on *simple, truth–telling, maximal evidence* mechanism. By "simple," we mean that the agent makes a report of her type and sends evidence without any need for the principal to act between the report and the sending of evidence. After this, the principal chooses an outcome. By "truth–telling," we mean that the mechanism induces the agent to report her type truthfully. By "maximal evidence," we mean that the mechanism also induces the agent to report maximal evidence.

The reason the maximal evidence is used is that this ensures the agent proves as much as she could ever prove, thus eliminating the largest possible number of incentive constraints.

It is not hard to use this to show that f is implementable iff the following two conditions hold. First, we have a restricted form of incentive compatibility:

$$\sum_{a \in A} f(a \mid t) u(a,t) \geq \sum_{a \in A} f(a \mid t') u(a,t), \qquad \forall t, t' \text{ with } m^*_{t'} \in \mathcal{M}(t). \tag{3}$$

In other words, if type t can send the maximal evidence of t', then f must induce t not to do so. Second, we must have a way of deterring "obvious deviations." If the agent reports type t but does not send the maximal evidence for t, we know the agent is deviating from the mechanism and need to punish this. More specifically, we require that

truth–telling.

for every provable event E which is not maximal evidence, there exists a "punishment" that is worse for every type who can prove E than what that type is supposed to get under f. That is, for every $m \in \cup_t \mathcal{M}(t)$, there exists $q_m \in \Delta(A)$ such that

$$\sum_{a \in A} f(a \mid t) u(a, t) \geq \sum_{a \in A} q_m(a) u(a, t), \ \forall t \text{ with } m \in \mathcal{M}(t).$$

See Green and Laffont (1986), Bull and Watson (2007), Deneckere and Severinov (2008), and Forges and Koessler (2005), and Schweighofer-Kodritsch and Strausz (2024) for various versions of the Revelation Principle for games with evidence. Ben-Porath, Dekel, and Lipman (forthcoming) provide Revelation Principles for more general evidence models.

4.2 Characterizations of Optimal Mechanisms

Glazer and Rubinstein (2004, 2006) study a binary action principal–agent problem with a general evidence structure. An agent wants a principal to choose action a (accept) regardless of her type while the principal wants to accept only some types and to reject (action r) otherwise. The focus in both papers is on the structure of optimal mechanisms with particular interest in situations like those discussed in Section 3.4 where costs of or constraints on evidence presentation lead to a failure of normality. In the 2004 paper, Glazer–Rubinstein examine the full protocol described at the beginning of this section. That is, the agent sends a cheap talk message, the principal requests evidence, possibly in a random way, the agent presents evidence, and finally the principal takes a (possibly random) action.[7] As mentioned in Section 4.1, this paper was the first to recognize the imporance of a random request for evidence by the principal. In the 2006 paper, they study a simpler protocol, which in some applications is more plausible, where the agent presents evidence and then the principal selects (possibly randomly) an action. Both papers give various interesting examples, obtain characterizations that are useful for computing optimal mechanisms, give conditions under which an optimal mechanism is deterministic, and show that there is no value for commitment. That is, there is an equilibrium of the game where the agent and principal use the protocol but without prior commitment by the principal which has the same outcome as the optimal mechanism. We discuss this issue further in Section 4.3.

Sher and Vohra (2015) study optimal selling mechanisms when the agent (the buyer) can present evidence about her value for the good. They assume normality of the evidence

[7] The 2004 paper is written as if the principal can "check" claims by the agent and hence seems more a model of verification (see Section 5.1). As observed in their 2006 paper, the model can be interpreted as an evidence model as we describe it here.

and develop graph–theoretic techniques to cleanly identify which incentive constraints remain relevant and their impacts.

Koessler and Skreta (2019) study a problem with a different structure than most of those discussed here. Their principal (a seller) also has private information and is the party that provides evidence.

4.3 Value of Commitment

In most of the literature on mechanism design without evidence, commitment to the mechanism is critical. For example, in the adverse selection principal–agent model in Mas-Collel, Whinston, and Green (1995), the "low type" takes inefficiently low effort ro reduce information rents to the "high type." Without commitment to the mechanism, the principal would want to renegotiate the contract, undermining the incentives.

As noted in Section 4.2, Glazer and Rubinstein (2004, 2006) show that commitment is not valuable in their mechanism design models with evidence and Sher (2011) and Hart, Kremer and Perry (2017) have generalized this result. As we show in this section, these results are not due to some "magical" property of evidence, but rather that with evidence, we are led to consider economic environments that are not of interest without evidence. For example, settings where the agent's utility is independent of her type are typically not interesting without evidence.

Formally, when we discuss the value of commitment, we consider a particular game between the agent and principal. Without commitment, we consider equilibria of this game as above. With commitment, we mean that the principal can choose her strategy in this game first and is committed to it. The principal chooses this strategy knowing the agent will best–respond to it. When there is an equilibrium without commitment giving the principal the same payoff as the maximum possible with commitment, we say that commitment has no value.

To illustrate, consider the following example. Suppose $T = \{1, \ldots, 1000\}$ and that the prior distribution over T satisfies $\mu(1000) = 2/1001$ and $\mu(t) = 1/1001$ for $t \neq 1000$. The evidence structure is a simple version of Dye: Type $t = 1000$ cannot prove anything, but any other type can disclose her type. That is, $\mathcal{E}(1000) = \{T\}$ and $\mathcal{E}(t) = \{\{t\}, T\}$ for $t \neq 1000$. Suppose the set of actions is $A = \mathbf{R}_+$ and the sender/agent's utility function is $u(a, t) = a$.

Consider the simplest game where the agent proves some event (possibly "proving" T) and then the principal chooses an action. Under commitment, the principal commits to an action as a function of what the agent proves. Without commitment, we have some equilibrium of this game.

We contrast the results under two different utility functions for the principal. First, suppose

$$v(a,t) = \begin{cases} 1, & \text{if } a = t \\ 0, & \text{otherwise.} \end{cases}$$

That is, the principal gets a payoff of 1 if he estimates the agent's true type and 0 otherwise.

In this case, in every equilibrium without commitment, the principal's response to observing no evidence (i.e., proof of T) is $a = 1000$. To see this, observe that type $t = 1000$ cannot provide evidence. Hence the principal's posterior belief on $t = 1000$ relative to any other t' must be at least the prior likelihood ratio. Since the prior has $t = 1000$ more likely than any other type, the posterior has this property too. So the principal's optimal action is $a = 1000$. Since this is the best action for the agent, no type presents evidence and the principal's expected payoff is $\mu(1000) = 2/1001$.

With commitment, the principal can obtain a strictly higher expected payoff. Suppose the principal commits to $a = 0$ if no evidence is presented and $a = t$ if the agent proves her type is t. In this case, every type $t \neq 1000$ will prove her type. So the principal will choose correctly with probability 999/1001.

Next, suppose that the principal's utility function is the square–error loss function used above, $v(a,t) = -(t-a)^2$. In this case, commitment has no value.

To see the intuition, consider the commitment mechanism above where the principal commits to $a = 0$ if the agent does not prove her type and $a = t$ if she proves she is type t. With this utility function, the principal could improve his payoff by instead committing to $a = 1$ if no evidence is provided. This gives the principal a strictly higher payoff when $t = 1000$ and does not interfere with the incentives of other types to report truthfully.

The principal can improve further. This mechanism pools types 1 and 1000 with the same action. With square–error loss, the best action when pooling these types is the conditional expectation given $t \in \{1, 1000\}$ and changes in this direction strictly increase the principal's payoff. Hence changing this common action to 2 strictly improves the principal's expected payoff conditional on $t \in \{1, 1000\}$ without interfering with the incentives of other types.

Similar reasoning implies that it would be better to pool types 1, 2, and 1000 at a higher action, such as 3. But then it would be better to pool 1, 2, 3, and 1000 at 4, etc. This continues until we reach a mechanism where types $1, \ldots, n$ and 1000 are pooled at the conditional expectation given this set of types and the expectation is between n and $n+1$. But at this point, we have the Dye equilibrium: the conditional expectation is the Dye cutoff previously defined. Hence commitment does not help the principal.

Ben-Porath, Dekel, and Lipman (2025) (henceforth BDL25) give a result which unifies the various results in the literature on when commitment has no value. The key hypothesis of the result is on endogenous variables, making it difficult to interpret. However, this condition is easily verified for all the earlier results in the literature, so it provides a unified way of understanding what is critical for these results.

To formalize this result, continue to let T denote the set of types, A the set of outcomes, $u(a,t)$ the agent's utility function, and $v(a,t)$ the principal's. Fix any *protocol.* By this, we mean a specification of the structure of the mechanism — i.e., stages of cheap talk, evidence presentation, etc., and ultimately the principal's choice of a. It would be natural to focus on the kind of mechanism the Revelation Principle says is without loss of utility for the principal, but this is not necessary.

Let B denote the set of pure strategies for the agent in this protocol and $\Delta(B)$ the set of mixed strategies with typical element β. For example, in the protocol for normal evidence where the agent reports a type and sends an evidence message, B would denote the set of functions $b : T \to T \times \mathcal{L}$ where we require $b(t) \in T \times \mathcal{M}(t)$.

Similarly, let G denote the set of pure strategies for the principal in this protocol and $\Delta(G)$ the set of mixed strategies with typical element γ. In the protocol for normal evidence, G is the set of functions from possible type reports and evidence messages to A.

Given mixed strategy profile (β,γ), let $U(\beta,\gamma)$ denote the expected utility for the agent and $V(\beta,\gamma)$ the expected utility for the principal, where these expectations are taken with respect to the randomness in the mixing, any randomness in the mechanism itself, and over the type of the agent.

Let $BR_u(\gamma)$ denote the set of best replies for the agent to γ. That is, the set of $\beta \in \Delta(B)$ that maximize $U(\beta,\gamma)$.

We take the payoff to the principal under commitment to be

$$V^* \equiv \max_{\gamma\in\Delta(G)} \left[\max_{\beta\in BR_u(\gamma)} V(\beta,\gamma) \right].$$

Implicitly, we let the principal choose the best reply for the agent. BDL25 compare this to the Nash equilibrium of the simultaneous move game with strategy sets B and G and payoff functions U and V.

BDL25 makes two assumptions. First, for simplicity, the set of pure strategies for the agent and for the principal are finite. This assumption can be replaced with appropriate continuity conditions.

The more substantive assumption is the following. BDL25 assumes that there is some mixed strategy for the principal, γ^*, with the property that for *every* $\beta \in BR_u(\gamma^*)$,

$$V(\beta, \gamma^*) = V^*.$$

That is, there is an optimal commitment strategy for the principal with the property that changes by the agent to alternative best replies do not affect the principal's payoff.

BDL25 shows that under these assumptions, there is a Nash equilibrium in the game giving the principal an expected payoff of V^*. Hence commitment has no value.

Because this result is about *Nash* equilibrium, it is natural to wonder about issues of sequential rationality. Without further assumptions on the protocol, we have no information about the structure of unreached information sets and so cannot say anything about sequential rationality. BDL25 show that one can extend the result to perfect Bayesian equilibrium with appropriate additional structure on the protocol.

To see how this result unifies results in the literature, first, suppose there are two possible outcomes and that no type is indifferent between these. In this case, the best response for the agent to any γ^* must give each t the highest possible probability of her preferred outcome. Hence any other best reply has the same probability distribution over outcomes for every type. Therefore, the principal's payoff is unchanged if the agent switches to a different best reply, so the BDL25 condition is satisfied. This includes all accept/reject problems and hence generalizes Glazer–Rubinstein (2004, 2006).

Next, consider any setting where there is an optimal mechanism which is deterministic. Assume that for every $t \in T$, the agent's utility is either strictly increasing or strictly decreasing in a (where this can vary across t). Just as above, this implies any best response for the agent gives each t the highest/lowest action she can generate. So as we vary the agent's best reply, the outcome for each type is the same, so the principal is indifferent across the agent's best replies. This includes the type–independent square–error loss settings discussed in Section 3 and generalizes Sher (2011) and Hart, Kremer, and Perry (2017).

4.4 Multi–Agent Mechanisms

There are a few papers considering mechanism design with evidence in settings with many agents. For example, there are papers on full implementation with evidence. Since the techniques of full implementation are quite different, we omit details here. See Kartik and Tercieux (2012), Ben-Porath and Lipman (2012), and Banerjee and Chen (2025).

Ben-Porath, Dekel, and Lipman (2019), henceforth BDL19, give a set of results applying to a range of multi–agent mechanism design problems. One of the problems covered, which we also discuss in other applications below, is the *simple allocation problem.* In this problem, the principal has one unit of a good to allocate to one of N agents. Each agent receives a payoff of 1 if she receives the good and 0 otherwise. The payoff to the principal of giving the good to agent i is $v_i(t_i)$ where t_i is agent i's type. No monetary transfers are possible. For example, the principal could be the dean of a College, the agents departments in the College, and the good a job slot the dean can allocate. Alternatively, the principal is a regional government choosing a city in the region for a new hospital, where the agents are the cities. In both cases, the agents want the good and have private information which affects the principal's payoff from the allocation. The result also covers a public goods problems where the principal chooses whether or not to provide a public good, charging each agent $1/N$ of the cost.

For brevity, we discuss only the simple allocation problem. There are N agents and T_i is the finite set of types for agent i. $A = \{0, 1, \ldots, N\}$ is the set of actions available to the principal where $a = 0$ means the principal keeps the good, while $a = i \neq 0$ means the principal gives the good to agent i. Agent i's utility is 1 if $a = i$ and 0 otherwise, regardless of her type. The principal's utility is $v_i(t_i)$ if $a = i$ and 0 if $a = 0$.

In the multi–agent setting, agents can have evidence about the profile of types or only about their own types. BDL19 assume the latter. So $\mathcal{E}_i(t_i)$ is the collection of subsets of T_i that type t_i can prove. They assume the evidence structure satisfies normality.

The main result in BDL19 gives several properties of optimal mechanisms. First, there is always a deterministic mechanism which is optimal. Second, there is always an optimal mechanism which satisfies a variation on dominant strategy incentive compatibility which BDL19 call robust incentive compatibility. So there is no cost to the principal of requiring this robustness.

Third, there is no value to commitment. That is, there is an equilibrium of the game without commitment giving the same outcome as an optimal mechanism. In addition, this equilibrium has two desirable properties. First, it can be constructed in a simple fashion. Specifically, this construction considers a family of auxiliary games, one for each agent, where the auxiliary game for agent i is just a game between agent i and the principal. Agent i's equilibrium strategy in her auxiliary game is the same as her strategy in the overall game without commitment and the optimality of her strategy in the auxiliary game implies its optimality in the game without commitment.

Second, this property implies the equilibrium of the game without commitment has a certain robustness property. The fact that agent i's strategy is constructed from the equilibrium of a game without any of the other agents suggests, correctly, that agent i's strategy in the game without commitment is optimal *regardless* of the strategies of the other agents.

BDL19 construct the auxiliary game between agent i and the principal as follows. First, i learns her type, t_i. Next, i sends a cheap talk message $s_i \in T_i$ and proves an event $E_i \in \mathcal{E}_i(t_i)$ which the principal observes. Finally, the principal chooses $\hat{v} \in \mathbf{R}$. The payoff to agent i is $\hat{v}$. The payoff to the principal is $-(v_i(t_i) - \hat{v})^2$. In other words, we have the square–error loss payoff structure discussed in Section 3.

To construct an equilibrium for the game without commitment, for each agent i, fix an equilibrium of the auxiliary game for i and take i's strategy in the game without commitment to be her equilibrium strategy in the auxiliary game. The principal's strategy in the game without commitment is her best reply to these strategies by the agents. BDL19 show that we can choose the equilibria of the auxiliary games so that these strategies form an equilibrium of the game without commitment with the same outcome as in the optimal mechanism.

Intuitively, in the auxiliary game, agent i wants to induce the principal to believe that $v_i(t_i)$ is large. In the game without commitment, the principal will allocate the good to that agent for whom his expectation of $v_i(t_i)$ is largest given the reports and evidence he sees. Hence in this game, again, agent i wants the principal to believe that $v_i(t_i)$ is large and so finds it optimal to use the strategy from the auxiliary game. The proof that there is an equilibrium of this form that gives the same outcome as the optimal mechanism is more involved.

We can illustrate this result using Dye evidence. Recall from Section 3.2 that the equilibrium outcome with Dye evidence is unique, so here there is no need to select equilibria from the auxiliary games — there is only one. Recall that the structure of the equilibrium of the auxiliary game for agent i is that there is a unique v_i^* defined by

$$v_i^* = \mathrm{E}[v_i(t_i) \mid \mathcal{E}_i(t_i) = \{T\} \text{ or } v_i(t_i) \leq v_i^*].$$

Types with no evidence, of course, cannot disclose. Types with evidence disclose iff $v_i(t_i) > v_i^*$.

This structure is easily used to show that there is a *favored–agent mechanism* which is optimal. For such a mechanism, there is an agent, say i^*, who is the favored agent and a threshold v^*. If no agent $i \neq i^*$ proves a value above v^*, then the favored agent receives the good regardless of what she proves. If some agent $i \neq i^*$ does prove a value above v^*, then the good goes to an agent who proves the highest value. In particular, in this case, no agent who doesn't prove something can receive the good.

To see that this is implied by the structure of the auxiliary game, let i^* be any agent i with the highest value of the Dye cutoff. Let the threshold be $v^* = v_{i^*}^*$, that is, the highest Dye cutoff.

Suppose no agent $i \neq i^*$ proves a value above v^*. Then either i proves a *smaller* value or else i proves nothing and the principal's expectation of v_i is $v_i^* \leq v_{i^*}^* = v^*$. On the other hand, i^* either proves some value above v^* or proves nothing and the principal's expectation of v_{i^*} is v^*. Hence i^* will be the agent the principal most wants to give the good to as the expectation of v_{i^*} will be larger, at least weakly, than the expectation of any v_i for $i \neq i^*$. In the case where some agent $i \neq i^*$ does prove some value above v^*, obviously, the agent who proves the highest value is the one the principal will most want to give the good to.

In short, the favored–agent mechanism is a description of the equilibrium under Dye evidence. The result of BDL19, then, implies that this is an optimal mechanism.

5 Other Directions

In this section, we discuss other models which incorporate evidence but which take some different approach than the models discussed above. In Section 5.1, we discuss models of costly verification, which can be thought of as changing which party has access to evidence. In Section 5.2, we discuss models of evidence acquisition.

5.1 Verification

In the models discussed so far, the agent has control of evidence and can choose whether to provide it to the principal. The principal may *incentivize* the agent to provide evidence but cannot take evidence from the agent without the agent's consent.

Verification can be thought of as reversing the property rights over evidence. Here the principal is able to get evidence about the agent without requiring the agent's agreement. Naturally, for the model to be interesting, this verification process must be costly for the principal. Otherwise, obviously, the principal will always take free evidence.

Townsend (1979) initiated the study of optimal mechanisms with verification. Unlike the work we discuss below, Townsend made critical use of monetary transfers. Other papers following this vein include Gale and Hellwig (1985), Border and Sobel (1987), and Mookherjee and Png (1989). Instead, we focus on models where monetary transfers are not possible, as in Glazer and Rubinstein (2004).[8]

We illustrate these ideas through a discussion of Ben-Porath, Dekel, and Lipman (2014), henceforth BDL14. This paper considered the simple allocation problem discussed in Section 4.4, but with costly verification rather than evidence.

It will be convenient to alter the notation slightly. Instead of writing the types as t_i and the value to the principal of giving the good to agent i as $v_i(t_i)$, we will write the

[8] See also Ball and Gao (2025) for a related game–theoretic model.

types directly as v_i's, the value of giving the good to i. In this section, there is no need to distinguish between the evidence available and the value, so we do not need the extra notation of t_i's.

To be more explicit, the type of agent i is denoted v_i, where this is continuously distributed on some interval $[\underline{v}, \bar{v}]$ with $0 < \underline{v} < \bar{v}$. Types are independent across agents, but the distributions may vary across agents. The principal has one unit of a good to allocate and the payoff to the principal of giving the good to agent i is v_i, i's type. The payoff to any agent i of receiving the good is 1 and the payoff to not receiving it is 0.

Agents know their types. The principal can learn the type of any agent at a cost $c > 0$ per agent. (BDL14 allowed costs to vary across agents, but we simplify for brevity.)

Agents do not have evidence in this model, so a mechanism simply has them reporting types. As usual, we can focus on incentive compatible mechanisms which induce agents to report honestly. Given a profile of reports, the principal decides which agents, if any, to verify — i.e., to pay the cost to learn their types. After this, the principal chooses which agent (if any) to give the good to.

Perhaps surprisingly, BDL14 show that the optimal mechanism in this problem is again a favored–agent mechanism. In this setting, they define such a mechanism as follows. There is a favored agent, i^*, and a threshold $v^* \in [\underline{v}, \bar{v}]$. If all agents other than the favored agent report types below the threshold, then the principal does not verify any agent and gives the good to the favored agent. If some agent other than the favored agent reports a type above the threshold, then the principal verifies the report of the agent with the highest reported type and if (as will happen in equilibrium) the principal learns that the agent reported truthfully, the principal gives her the good.

BDL14 characterize the optimal agent to favor and the optimal threshold. For any i, define v_i^* by

$$\mathrm{E}(v_i) = \mathrm{E}\max\{v_i, v_i^*\} - c.$$

The optimal favored agent is any agent with the highest value of v_i^* and the optimal threshold is her v_i^*. While the definition of this cutoff is different in the evidence case, otherwise the selection of the favored agent and the threshold is the same as in that case. Later, we explain this unexpected similarity across these different models.

Another similarity between this favored–agent mechanism and the one discussed in Section 4.4 is that the mechanism is robustly incentive compatible. A difference between the two favored–agent mechanisms we have introduced is that commitment is necessary for this one. Clearly, when the principal verifies an agent, he is spending a cost expecting he will find that the agent didn't lie. If he were not committed, he would deviate. Since the agents would anticipate this, incentive compatibility would break down.

BDL19 identify the reason for the resemblance between these favored–agent mechanisms, also explaining the similarity between the results of Erlanson and Kleiner (2020) on costly verification in a public goods problem and BDL19's solution to the analogous Dye evidence problem. Loosely, these models are related by a change of variables. Intuitively, in BDL14, the principal pays a cost for evidence, while in BDL19, he incentivizes agents to reveal evidence and thus pays a "shadow price" for the implied distortions. Interestingly, these two distinct forms of cost enter the structure in mathematically similar ways, creating this similarity of results.

There are a number of papers which consider versions of verification. Mylovanov and Zapechelnyuk (2017) explore the implications of costless ex post verification and limited penalties, while Li (2020) considers a different form of limited penalties with costly verification. Li and Libgober (2023) consider a dynamic verification problem with projects proposed over time. Patel and Urgun (2025) consider an allocation problem with costly verification similar to BDL14 but where the principal can request agents to burn money. Kattwinkel and Preusser (2025) combine an evidence model of the kind considered in Section 4, evidence acquisition of the kind discussed in the next section, and costly verification by the principal. Ball and Kattwinkel (forthcoming) develop a framework for studying probabilistic verification, showing that every implementable social choice function can be implemented via what they call most discerning tests.

5.2 Evidence Acquisition

In the standard evidence model, we exogenously specify what evidence each type has. A natural and economically significant extension is to study agent choices to acquire evidence.

Note that evidence acquisition is related to but distinct from information acquisition. If the agent doesn't know her type, then acquiring evidence about it can also help her learn her type. On the other hand, evidence is acquired, at least in part, to persuade a principal, not just for learning.

There are a number of interesting game–theoretic models of evidence acquisition in the literature. Che and Kartik (2009) consider how evidence acquisition incentives shape the preferences of a receiver regarding what sender he would like to get information from. They consider a model where the state of the world, t, is normally distributed with a certain variance. The receiver and potential senders differ regarding their prior mean for the state. The receiver has to choose an action $a \in \mathbf{R}$. The receiver and all potential senders have utility function $-(a-t)^2$.

The receiver chooses one of the potential senders who then exerts costly effort to obtain a signal about the true state. The probability the sender gets a signal is increasing in the effort expended. If the sender receives a signal, this is evidence which she can show to the receiver if she wishes. That is, we have a Dye evidence model where the probability the sender has evidence is a function of the effort she chooses.

One might expect the receiver to prefer a sender with the same prior mean on the state as her own. If the probability the sender has evidence were exogenous instead of endogenous, this would be optimal for the receiver. The sender would not hide evidence since the receiver would use the evidence to make the choice the sender would like him to make. On the other hand, if the sender has a different mean than the receiver, the sender will sometimes withhold evidence. Intuitively, if the sender's prior mean is much larger than the receiver's and the signal suggests a value below the receiver's prior mean, the sender's preferred action will fall but will remain above the receiver's. So the sender would not want the receiver to see the signal and choose an even lower action.

But when the probability the sender gets evidence is endogenous, the receiver may prefer a sender with a different prior than her own. Intuitively, if the sender's prior beliefs are different than the receiver's, then the sender will expect evidence to pull the receiver's beliefs closer to her own. If the difference in their beliefs is larger, the desire by the sender to influence the receiver's beliefs will be stronger and the sender will exert more effort to obtain evidence. Of course, as observed above, a sender with very different beliefs than the receiver will withhold evidence more often, so there is a tradeoff for the receiver.

Chade and Pram (2024) give an interesting application of evidence acquisition in a game–theoretic model. They consider college entrance exams. Suppose students have imperfect information about their abilities. A student can take a costly test which reveals noisy but more precise information about her ability and which can be shown as evidence to a university to which she applies. Assume the qualities of the universities are known and all students agree on the ranking of universities. Assume also that universities all want students with higher ability so that if ability were known, we would have positive assortative matching of students and universities.

Among other things, Chade and Pram ask what effect various disclosure requirements have on the students' utility. For example, suppose students can choose whether to take the test or not. Compare the utility of students in a world where they are free to disclose or hide their test results versus a world where anyone who takes the test must reveal the score. Perhaps surprisingly, students whose prior beliefs are that they are low ability prefer the world where scores *must* be revealed. The reason is that these students are skeptical about the likelihood they will perform well on the test and so will not take the test in either case. When revealing one's score is not required, students who took the test and did badly are pooled with students who did not take it. Since the test is more informative than one's priors, this group makes the inference in response to nondisclosure worse.

Other interesting papers in this area include DeMarzo, Kremer, and Skrzypacz (2019) and Shiskin (2022), both of which develop models of optimal test design.

Ben-Porath, Dekel, and Lipman (forthcoming) discusses mechanism design for evidence acquisition for the case where the agent knows her type ex ante but wishes to acquire evidence to persuade the principal. For example, the agent may take a test that generates a probability distribution over evidence messages. The paper characterizes the class of mechanisms the principal can use without loss of utility for this setting and a broad class of models with stochastic evidence. It also gives conditions on the evidence structure which allow for simplifications in the required structure and hence relatively tractable analysis.

Ben-Porath, Dekel, and Lipman (2024), henceforth BDL24, give a model of mechanism design with costly evidence acquisition in a setting where agents don't know their types ex ante, so that evidence acquisition goes hand–in–hand with information acquisition. The model is again the simple allocation problem discussed earlier. The principal has one unit of a good to allocate to one of N agents. Each agent receives a payoff of 1 if she gets the good, 0 otherwise (not including costs discussed below). The value to the principal of giving the good to agent i is v_i. The v_i's are unknown to the principal or any agent at the outset and are independently and continuously distributed with support $[0, 1]$. Agent i can pay a cost $c_i \in (0, 1)$ to both learn her v_i and to obtain evidence proving this value to the principal.

The tension in the model is that an agent must have a high enough chance to obtain the good to be willing to pay for evidence, while the principal would like information from as many agents as possible. Ideally, the principal would like to ask all agents to obtain and provide evidence. In the symmetric case, for example, this cannot be incentive compatible if $c > 1/N$ since the expected payoff to any agent from obtaining evidence if all others do so is $(1/N) - c$. So in this case, the agent would prefer to opt out of the mechanism.

The optimal mechanism in BDL24 is easiest to understand when the costs and distributions are the same across agents. In this case, the principal begins by choosing an agent at random (where each agent has probability $1/N$ of being chosen first). The chosen agent is asked to obtain and report evidence. If the agent proves a value above a certain threshold, v^*, the mechanism ends and this agent receives the good. Otherwise, the principal selects one of the remaining agents at random (now each with probability $1/(N-1)$) and asks this agent to obtain and provide evidence. Again, if this agent proves her value is above v^*, she receives the good and otherwise we continue this process. If all agents have values below v^*, the principal will end up asking all for evidence. In this case, the principal will give the good to the agent with the highest value.

An important aspect of the mechanism is that an agent is given no information when she is asked for evidence (other than knowledge of the mechanism itself, that is).[9] So the agent does not know if she is first to be asked, last, or somewhere in between. To see the idea, suppose the principal uses the mechanism above but does tell each agent how many others have already obtained evidence. In this case, the last agent to be asked knows that all the other agents have values below the threshold and hence she has a relatively high probability of succeeding in getting the good. The first agent to be asked has a much lower probability of receiving the good. If the first agent asked has a high enough probability of receiving the good that she is willing to pay the cost to obtain evidence, the other agents' incentive constraints must be slack. Hence it must be possible to improve the mechanism.

In the asymmetric case, instead of comparing the agent's value to a threshold or other agents' values, the mechanism is based on virtual values, equal to the value plus an agent–specific constant, λ_i, where this is the Lagrange multiplier on the incentive constraint for agent i. Second, the distribution over the order in which agents are asked for evidence is asymmetric. Third, the threshold may decrease over time.

The proof of optimality combines Weitzman (1979) and Border (1991). BDL24 show that if one treats the Lagrange multipliers for the optimization as if they were exogenous, the Lagrangian is equivalent to Weitzman's objective function. In this reinterpretation, agents correspond to the boxes in Weitzman, the cost of opening box i is $\lambda_i c_i$, and the prize in box i is $v_i + \lambda_i$. One can then apply Weitzman's characterization of optimal search procedures to characterize the solution given the Lagrange multipliers.

Also, in Weitzman, the randomizations over the order are irrelevant as he considers a decision problem, but they affect incentive compatibility for BDL24, so they must characterize these along with the multipliers. BDL24 show that one can characterize the set of feasible Lagrange multipliers and probabilities of asking each agent for evidence using methods based on those of Border (1991). The parameters of this dynamic mechanism, such as the randomizing probabilities, could depend on the values observed along the way, leading to a very complicated mechanism. The characterization of the multipliers and probabilities using Border shows that it is without loss of utility for the principal to restrict to mechanisms which do not use such information.

6 Conclusion

As we hope this survey demonstrates, evidence is relevant in a wide range of economic environments and can be fruitfully studied to understand the nature of the distortions

[9] The optimal mechanism in Gershkov and Szentes (2009) is similar in this respect. Their model has pure information acquisition, not evidence acquisition, and involves a public good rather than private, so the nature of the incentives and constraints differ.

it alleviates and creates. The literature has only started to ask these questions and we expect to see interesting applications in other directions.

References

[1] Acharya, V., P. DeMarzo, and I. Kremer, "Endogenous Information Flows and the Clustering of Announcements," *American Economic Review*, December 2011.

[2] Aghamolla, C., and B.-J. An, "Mandatory vs. Voluntary ESG Disclosure, Efficiency, and Real Effects," working paper, February 2025.

[3] Ali, S. N., A. Kleiner, and K. Zhang, "From Design to Disclosure," working paper, November 2024.

[4] Ali, S. N., G. Lewis, and S. Vasserman, "Voluntary Disclosure and Personalized Pricing," *Review of Economic Studies*, **90**, 2, March 2023.

[5] Ball, I., and X. Gao, "Checking Cheap Talk," working paper, May 2025.

[6] Ball, I., and D. Kattwinkel, "Probabilistic Verification in Mechanism Design," *Theoretical Economics*, forthcoming.

[7] Banerjee, S., and Y.-C. Chen, "Implementation with Uncertain Evidence," working paper, January 2025.

[8] Ben-Porath, E., E. Dekel, and B. Lipman, "Disclosure and Choice," *Review of Economic Studies*, **85**, July 2018, 1425–1470.

[9] Ben-Porath, E., E. Dekel, and B. Lipman, "Mechanisms with Evidence: Commitment and Robustness," *Econometrica*, **87**, March 2019, 529–566.

[10] Ben-Porath, E., E. Dekel, and B. Lipman, "Mechanism Design for Acquisition of/Stochastic Evidence," *Theoretical Economics*, forthcoming.

[11] Ben-Porath, E., E. Dekel, and B. Lipman, "Sequential Mechanisms for Evidence Acquisition," working paper, current draft April 2024.

[12] Ben-Porath, E., E. Dekel, and B. Lipman, "When Does Commitment Not Have Value?," working paper, August 2025.

[13] Ben-Porath, E., and B. Lipman, "Implementation and Partial Provability," *Journal of Economic Theory*, **147**, September 2012, 1689–1724.

[14] Beyer, A., D. Cohen, T. Lys, and B. Walther, "The Financial Reporting Environment: Review of the Recent Literature," *Journal of Accounting and Economics*, **50**, 2010, 296–343.

[15] Border, K., "Implementation of Reduced Form Auctions: A Geometric Approach," *Econometrica*, **59**, July 1991, 1175–1187.

[16] Border, K., and J. Sobel, "Samurai Accountant: A Theory of Auditing and Plunder," *Review of Economic Studies*, **54**, 4, 1987, 525–540.

[17] Bull, J., and J. Watson, "Hard Evidence and Mechanism Design," *Games and Economic Behavior*, **58**, January 2007, 75–93.

[18] Callander, S., N. Lambert, and N. Matouschek, "The Power of Referential Advice," *Journal of Political Economy*, **129**, November 2021, 3073–3140.

[19] Carroll, G., and G. Egorov, "Strategic Communication with Minimal Verification," *Econometrica*, **87**, November 2019, 1867–1892.

[20] Chade, H., and K. Pram, "Matching and Disclosure," working paper, February 2024.

[21] Che, Y.-K., and N. Kartik, "Opinions as Incentives," *Journal of Political Economy*, **117**, October 2009, 815–860.

[22] Crawford, V., and J. Sobel, "Strategic Information Transmission," *Econometrica*, **50**, November 1982, 1431–1451.

[23] DeMarzo, P., I. Kremer, and A. Skrzypacz, "Test Design and Disclosure," *American Economic Review*, **109**, June 2019, 2173–2207.

[24] Deneckere, R. and S. Severinov, "Mechanism Design with Partial State Verifiability," *Games and Economic Behavior*, **64**, November 2008, 487–513.

[25] Dye, R. A., "Disclosure of Nonproprietary Information," *Journal of Accounting Research*, **23**, 1985, 123–145.

[26] Erlanson, A., and A. Kleiner, "Costly Verification in Collective Decisions," *Theoretical Economics*, **15**, number 3, 2020, 923–954.

[27] Fishman, M., and K. Hagerty, "The Optimal Amount of Discretion to Allow in Disclosure," *Quarterly Journal of Economics*, **105**, May 1990, 427–444.

[28] Forges, F., and F. Koessler, "Communication Equilibria with Partially Verifiable Types," *Journal of Mathematical Economics*, **41**, 2005, 793–811.

[29] Gale, D., and M. Hellwig, "Incentive Compatible Debt Contracts: The One–Period Problem," *Review of Economic Studies*, **52**, 4, 1985, 647–663.

[30] Gershkov, A., and B. Szentes, “Optimal Voting Schemes with Costly Information Acquisition,” *Journal of Economic Theory*, **144**, January 2009, 36–68.

[31] Glazer, J., and A. Rubinstein, “On Optimal Rules of Persuasion,” *Econometrica*, **72**, November 2004, 1715–1736.

[32] Glazer, J., and A. Rubinstein, “A Study in the Pragmatics of Persuasion: A Game Theoretical Approach,” *Theoretical Economics*, **1**, December 2006, 395–410.

[33] Green, J., and J.-J. Laffont, “Partially Verifiable Information and Mechanism Design,” *Review of Economic Studies*, **53**, July 1986, 447–456.

[34] Grossman, S., “The Informational Role of Warranties and Private Disclosure about Product Quality,” *Journal of Law and Economics*, **24**, 1981, 461–483.

[35] Guttman, I., I. Kremer, and A. Skrzypacz, “Not Only What but also When: A Theory of Dynamic Voluntary Disclosure,” *American Economic Review*, **104**, August 2014, 2400–2420.

[36] Guttman, I., I. Kremer, A. Skrzypacz, and E. Wiedman, “Investment Decisions, Voluntary Disclosure, Myopia, and Bounded Inefficiency,” working paper, February 2025.

[37] Hart, S., I. Kremer, and M. Perry, “Evidence Games: Truth and Commitment,” *American Economic Review*, **107**, March 2017, 690–713.

[38] Hagenbach, J., F. Koessler, and E. Perez-Richet, Certifiable Pre–Play Communication: Full Disclosure,” *Econometrica*, **82**, May 2014, 1093–1131.

[39] Jung, W., and Y. Kwon, “Disclosure When the Market is Unsure of Information Endowment of Managers,” *Journal of Accounting Research*, **26**, 1988, 146–153.

[40] Kamenica, E., and M. Gentzkow, “Bayesian Persuasion,” *American Economic Review*, **101**, October 2011, 2590–2615.

[41] Kartik, N., and O. Tercieux, “Implementation with Evidence,” *Theoretical Economics*, **7**, May 2012, 323–355.

[42] Kattwinkel, D., and J. Preusser, “The Division of Surplus and the Burden of Proof,” working paper, March 2025.

[43] Koessler, F., and V. Skreta, “Selling with Evidence,” *Theoretical Economics*, **14**, number 2, 2019, 345–371.

[44] Leuz, C., and P. Wysocki, “The Economics of Disclosure and Financial Reporting Regulation: Evidence and Suggestions for Future Research,” *Journal of Accounting Research*, **54**, May 2016, 525–622.

[45] Li, Y., "Mechanism Design with Costly Verification and Limited Punishments," *Journal of Economic Theory*, **186**, January 2020, 1–54.

[46] Li, Z., and J. Libgober, "The Dynamics of Verification when Searching for Quality," working paper, March 2024.

[47] Lipman, B., and D. Seppi, "Robust Inference in Communication Games with Partial Provability," *Journal of Economic Theory*, **66**, August 1995, 370-405.

[48] Maskin, E., and J. Riley, "Monopoly with Incomplete Information," *Rand Journal of Economics*, **15**, Summer 1984, 171–196.

[49] Milgrom, P., "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, **12**, 1981, 380–391.

[50] Milgrom, P., and J. Roberts, "Relying on the Information of Interested Parties," *Rand Journal of Economics*, **17**, 1986, 18–32.

[51] Mookherjee, D., and I. Png, "Optimal Auditing, Insurance, and Redistribution," *Quarterly Journal of Economics*, **104**, 2, 1989, 399–415.

[52] Myerson, R., "Optimal Auction Design," *Mathematics of Operations Research*, **6**, February 1981, 58–73.

[53] Mylovanov, T., and A. Zapechelnyuk, "Optimal Allocation with Ex Post Verification and Limited Penalties," *American Economic Review*, **107**, September 2017, 2666–2694.

[54] Onuchic, P., and J. Ramos, "Disclosure by Groups," working paper, March 2025.

[55] Patel, R., and C. Urgun, "Costly Verification and Money Burning," working paper, March 2025.

[56] Rappoport, D., "Evidence and Skepticism in Verifiable Disclosure Games," working paper, January 2024.

[57] Schweighofer-Kodritsch, S., and R. Strausz, "Principled Mechanism Design with Evidence," working paper, May 2024.

[58] Seidmann, D., and E. Winter, "Strategic Information Transmission with Verifiable Messages," *Econometrica*, **65**, January 1997, 163–169.

[59] Sher, I., "Credibility and Determinism in a Game of Persuasion," *Games and Economic Behavior*, **71**, March 2011, 409–419.

[60] Sher, I., and R. Vohra, "Price Discrimination through Communication," *Theoretical Economics*, **10**, May 2015, 597–648.

[61] Shin, H. S., "Disclosures and Asset Returns," *Econometrica*, **71**, January 2003, 105–133.

[62] Shishkin, D., "Evidence Acquisition and Voluntary Disclosure," working paper, December 2020.

[63] Spence, A. M., *Market Signaling*, Cambridge: Harvard University Press, 1974.

[64] Titova, M., "Persuasion with Verifiable Information," Vanderbilt University working paper, November 2023.

[65] Townsend, R., "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, **21**, October 1979, 265–293.

[66] Verrecchia, R., "Discretionary Disclosure," *Journal of Accounting and Economics*, **5**, 1983, 179–194.

[67] Weitzman, M., "Optimal Search for the Best Alternative," *Econometrica*, **47**, May 1979, 641–654.

[68] Zhang, K., "Withholding Verifiable Information," working paper, April 2024.