Information Control in the Hold-up Problem*

Anh Nguyen† Teck Yong Tan‡

April 2019

Abstract

We study the use of information control to mitigate hold-up risks. We identify a distinction between asymmetric information that creates an *ex-ante investment incentive* and asymmetric information that causes *ex-post inefficiency*, which then allows ex-post inefficiency to be eliminated without compromising the ex-ante investment incentive. We characterize the properties of the optimal information structure and the payoffs and welfare achievable with information control in the presence of hold-up risks.

Keywords: Hold-up; Information control; Asymmetric information

JEL Classification: D42, D82, D83.

---

*We are grateful to the Editor and three anonymous referees for helpful comments and suggestions. We thank Charles Angelucci, Yeon-Koo Che, Navin Kartik, Kate Ho, Johannes Hörner, Mike Riordan, Tobias Salz, Bernard Salanié, Ran Spiegler and participants at various conferences for their useful feedback.

†Tepper School of Business, Carnegie Mellon University. Email: anhnguyen@cmu.edu.

‡Department of Economics, Nanyang Technological University. Email: teckyongtan@ntu.edu.sg.

1 Introduction

This paper revisits the classic hold-up problem: an economic agent has the opportunity to make a costly and sunk investment that increases the surplus of a potential partnership, but anticipating that his partner would alter the terms of the partnership and expropriate the gains from the investment later, the agent chooses not to invest in the first place. Such hold-up risks are ubiquitous in many situations, including relationship-specific investment in joint business ventures, acquisition of firm-specific skills by employees, provision of general training by firms, campaign contributions in political lobbying and quality investment by hospitals in the healthcare market.

Some papers (e.g., Gul, 2001; Lau, 2008) have suggested that hold-up risks could potentially be mitigated by hiding the investment from the partner. The resulting asymmetric information limits the partner's ability to extract the investment gains, thus improving *ex-ante efficiency* by partially restoring the investing agent's investment incentive. However, *ex-post inefficiency* can also arise when the parties negotiate the terms of the partnership under asymmetric information. This suggests that a tradeoff exists between creating an ex-ante investment incentive and minimizing ex-post inefficiency, and the optimal information control must balance the two effects.[1] Our goals in this paper are to show that this tradeoff is not necessary under more general forms of information control and to distinguish asymmetric information that generates an ex-ante investment incentive from asymmetric information that causes ex-post inefficiency.

As an example, consider a relationship between an upstream supplier who supplies inputs for the production process of a downstream manufacturer. The supplier could make manufacturer-specific investments that lower its cost to produce the inputs; however, since

---

[1] As Lau (2008) wrote in her abstract, "The optimal information structure balances the tradeoff between ex-ante efficiency (the "information rent" effect) and ex-post efficiency (the "bargaining disagreement" effect)."

the quality and type of investment are difficult to verify, the parties cannot write contracts that are contingent on the supplier's investment, which generates hold-up risks. Nevertheless, the two parties are often able to contract on the terms that govern the manufacturer's ability to inspect the supplier's facilities, such as the frequency and areas of inspection.[2] In turn, these terms affect the probability that the manufacturer can learn about the supplier's investment, and we show that when appropriately designed, these terms can improve the supplier's investment incentive without incurring any post-investment "haggling" cost.

We follow the related literature and frame the hold-up problem throughout this paper in a monopoly setting where a buyer (he) can make an uncontractible investment to increase his valuation for a good sold by a monopolist seller (she). We add to this setup a signal structure that generates signals to the seller regarding the buyer's eventual valuation. To see how information control can disentangle the effects of different types of asymmetric information, notice that the buyer's ex-ante investment incentive comes from the seller being (at least sometimes) unaware of the buyer's higher valuation and undercharging him—this concerns hiding information from the seller in the "investment state". On the other hand, ex-post inefficiency from a trade breakdown arises when the seller is unaware of the buyer's low valuation and still charges a high price—this concerns hiding information about the "non-investment state", which does not affect the buyer's investment incentive. In turn, we can exploit this distinction to eliminate ex-post inefficiency while maintaining the buyer's investment incentive.

Our analysis begins by showing that when the seller cannot directly observe the investment, the buyer's investment decision must be randomized in equilibrium. This is readily understood under binary investment: if the seller anticipates that the buyer always invests, she will charge a high price, which destroys the buyer's ex-ante investment incentive; on the

---

[2] See, for example, Smith (2007) for an exposition on including clauses on auditing the various aspects of production in contracts for outsourcing.

other hand, if she anticipates that the buyer never invests, she will charge a low price, which then induces the buyer to invest. A stochastic investment strategy implies that the first best is not achievable by information control, but we show that it can still always improve welfare relative to a complete hold-up which arises when the seller perfectly observes the investment.

To simplify the problem, we first derive a technical result that allows us to, without loss of generality, restrict our attention to "direct" signal structures making incentive-compatible price recommendations to the seller. While this is straightforward from the revelation principle on the seller's side, the novelty lies in showing that direct signal structures can preserve the buyer's investment incentive, which is based on the signal structure as a whole.

Our aim is to characterize the set of investment probabilities and social welfare that are sustainable in equilibrium and the signal structures that implement them. In general, the buyer's investment incentive comes from the seller being too pessimistic about the buyer's valuation and undercharging him, whereas an ex-post trade breakdown occurs when the seller is too optimistic and overcharges the buyer. Pessimism (respectively, optimism) arises when the signal structure muddles information about the buyer's true investment with signals that suggest *lower* (respectively, *higher*) investments. Given this separation, the optimal signal structure is characterized by just pinning down the required pessimism, while optimism (and hence any ex-post inefficiency) is eliminated.

With binary investment, we show that every implementable investment probability can be optimally implemented by the same signal structure, which is essentially unique. The characterization under multiple investment levels is more difficult, but we provide a characterization for the implementation of a large class of investment strategies that can still generate welfare gain relative to the hold-up outcome. In both cases, the optimal signal structure only muddles the true valuation with the next lower valuation. Therefore, under a small number of investment levels, the optimal signal structure is simple, which potentially allows it to be replicated by practical arrangements and thus also alleviates the usual concern in the information design literature on the commitment assumption to a signal structure.[3]

The rest of the paper proceeds as follows. The next section reviews the related literature and Section 3 provides the main analysis under a hold-up setting with binary investment. Section 4 generalizes the main result of no ex-post inefficiency to a setting with multiple investment levels and Section 5 concludes. All omitted proofs are found in Appendix A.

2 Related Literature

Our paper is primarily related to the literature on the use of asymmetric information to mitigate hold-up risks. Gibbons (1992) and Gul (2001) show that when the investment is completely unobservable to the seller, the ex-ante efficiency and ex-post inefficiency created from asymmetric information exactly cancel out; thus, the welfare is unchanged from the hold-up outcome. Lau (2008) shows that the two effects change at different rates when the probability of the seller observing the investment varies. Therefore, welfare can be improved if the seller observes the investment with an intermediate probability. In addition, both Gul (2001) and Lau (2008) emphasize how welfare is always improved if the seller can make repeated and frequent offers after rejection by the buyer.

Our paper contrasts with Gul (2001) and Lau (2008) in two main ways. First, we allow for more general forms of information control. Second, we show that conditional on the optimal information structure, allowing for repeated offers upon the buyer's rejection has no effect because the seller's offer will be accepted immediately. However, this result also complements the works of Gul and Lau because it shows why allowing for repeated offers in their setup always improves welfare: repeated offers improve welfare only after a trade breakdown (which occurs with positive probability in the two papers), and we show that such inefficiency can

---

[3] See Min (2017), Fréchette, Lizzeri, and Perego (2017), Guo and Shmaya (2018), Lipnowski, Ravid, and Shishkin (2018) and Nguyen and Tan (2019) for recent works on information design under partial commitment.

be eliminated without needing to sacrifice the buyer's ex-ante investment incentive.

Moreover, although Lau (2008) illustrates a way that endogenous asymmetric information can improve welfare under hold-up risks, her setup does not distinguish the type of asymmetric information generated, so it is not clear how the welfare gain actually arises. By contrast, we distinguish the type of asymmetric information that creates an ex-ante investment incentive from the type of asymmetric information that causes ex-post inefficiency. This helps us understand the channel through which endogenous asymmetric information mitigates the hold-up problem, which allows us to obtain a higher welfare than that under the signal structures considered in Lau (2008).

Other papers that also study asymmetric information in the hold-up problem include Riordan (1990), González (2004), Hermalin and Katz (2009), Hermalin (2013), Halac (2015), and Tan (2019). As in Gul (2001), these papers study a variety of related issues while restricting their attention to perfect observability versus perfect unobservability of the investment; however, they do not consider more general forms of information control as we do here.

Our paper is also related to the information design literature—see, for example, Rayo and Segal (2010), Ostrovsky and Schwarz (2010) and Kamenica and Gentzkow (2011). The difference is that the papers in this literature need not restrict the signal structure, whereas the signal structure in our setup must also satisfy an equilibrium condition because information control is embedded in a hold-up problem. Consequently, we cannot appeal to the "concavification" argument (Aumann and Maschler, 1995) commonly used in the Bayesian persuasion literature, and as will be discussed, the technique used to solve for the optimal signal structure and the logic behind it are also very different.

Away from pure information design, Bergemann, Brooks, and Morris (2015) study the welfare effects of the signal structure that generates signals to the seller regarding the buyer's valuation, but the buyer's valuation distribution there is exogenous. On the other hand, Condorelli and Szentes (2018) consider the effects of the buyer's valuation distribution, while Roesler and Szentes (2017) endogenize the buyer's learning process of his own valuation. However, in both of these papers, the seller receives no information about the buyer's valuation. Our paper has elements of both approaches: the buyer's valuation distribution is endogenous, and the seller receives partial information about the buyer's valuation.

3 Binary Investment

We consider information control in the hold-up problem with binary investment in this section. We first describe the baseline model and provide a few benchmarks before solving for the optimal signal structure.

Model

A buyer (he) has valuation $v = L$ for a good that a seller (she) can produce at a cost normalized to zero. Before interacting with the seller, the buyer can privately increase his valuation to $v = H$ at a cost c. Increasing the valuation is henceforth termed an investment. We assume that $H - L > c$ so that it is socially efficient to invest. However, due to incomplete contracts, the investment is not contractible.

After the investment decision is made but before trade occurs, the seller receives a signal s regarding v. Let the set of signals be S. For expositional clarity, we assume that S is a finite set, although this is without loss of generality. A signal structure is defined by $\{S, \pi\}$, where $\pi(s|v)$ denotes the conditional probability of $s \in S$ given valuation v. This signal structure is common knowledge to both players at the start of the game.

After observing the signal, the seller makes a take-it-or-leave-it price offer p to the buyer. If it is accepted, the seller's payoff is p, while the buyer's payoff is $v - p - \mathbb{I}c$, where $\mathbb{I}$ is an indicator function for investment;[4] if the offer is rejected, the seller's payoff is 0, and the buyer's payoff is $-\mathbb{I}c$.

Our equilibrium concept is sequential equilibrium (hereafter equilibrium). Heuristically, given a signal structure, the buyer optimally chooses to invest or not, taking into account the distribution of signals that his investment decision will generate and his conjecture about the seller's pricing strategy after observing each signal $s \in S$. Upon observing a signal s, the seller forms a posterior that depends on both her conjecture about the buyer's investment strategy and the distribution of the signals under the signal structure, and she then optimally sets a price based on her posterior. The buyer then accepts the offer if $p \leq v$ and rejects it otherwise. In equilibrium, each player's conjecture about the other player's strategy is correct.[5] It is readily noted that the seller will only set $p = L$ or $p = H$ in equilibrium.

Throughout, we let $q \in [0, 1]$ denote the probability that the buyer invests. Thus, we say that a signal structure *implements* q if the buyer investing with probability q can be sustained as an equilibrium under the signal structure. In addition, all payoffs will be expected payoffs; henceforth, we drop the "expected" quantifier for ease of exposition.

Proposition 1. *The buyer's payoff is always zero in equilibrium. Moreover,* $q = 1$ *cannot be implemented.*

Proof. Since the seller will never set p lower than L, the buyer's payoff is zero if he does not invest (i.e., $q = 0$). If $q \in (0, 1)$ in equilibrium, the buyer must be indifferent between investing and not investing, which means that his payoff is also 0. Lastly, q cannot be 1 in equilibrium. This is because if $q = 1$, the seller will correctly conjecture it in equilibrium and always set $p = H$, in which case the buyer will never invest in the first place. □

Proposition 1 thus implies that the seller's payoff in equilibrium is also the social welfare,

---

[4] $\mathbb{I} = 1$ if the buyer invests; $\mathbb{I} = 0$ if the buyer does not invest.

[5] In addition, sequential equilibrium requires that the seller's posterior about v is the limit of a sequence of beliefs formed by Bayes rule and a sequence of fully mixed buyer's investment strategies that converges to his equilibrium investment strategy.

so there is no need to distinguish between the two. Henceforth, the term "optimality" will refer to the optimality of the seller's payoff. We say that a signal structure π *achieves* a seller's payoff W if π can implement an investment strategy in which the seller's payoff in the equilibrium is W. We are interested in the set of the implementable buyer's investment strategies and achievable seller's payoffs. Since the first best investment strategy is $q = 1$, which cannot be implemented, the first best seller's payoff cannot be achieved.

Benchmarks

Fully Informative Signal Structure (Complete Hold-up). The hold-up problem arises when the signal structure is fully informative. When the seller perfectly knows v, she always sets $p = v$; hence, the buyer never invests in equilibrium. The social welfare is thus L, all of which is given to the seller.

Fully Uninformative Signal Structure. Next, consider a fully uninformative signal structure: $\pi(s|L) = \pi(s|H)\ \forall s \in S$. The following restates the result of Gibbons (1992) and Gul (2001) that the players' payoffs under no information are the same as under a complete hold-up:[6]

Proposition 2. *The equilibrium under the fully uninformative signal structure is unique: the buyer invests with probability* $\frac{L}{H}$*, and the seller sets* $p = L$ *with probability* $\frac{c}{H-L}$ *after every (uninformative) signal. The seller's equilibrium payoff is* L*. Therefore, the players' payoffs are the same as with a fully informative signal structure with a complete hold-up.*

Proof. From Proposition 1, $q \neq 1$. Similarly, $q \neq 0$; if $q = 0$, the seller will also correctly conjecture that in equilibrium and always set $p = L$, in which case the buyer will deviate to choosing $q = 1$ instead. Let α be the probability that the seller sets price $p = L$, which is

---

[6] See problem 2.23 in Gibbons (1992) and Proposition 1 in Gul (2001).

independent of the signal because the signal has no information. Since $q \in (0,1)$, the buyer must be indifferent between investing and not investing, which implies that $\alpha(H-L)-c=0 \iff \alpha=\frac{c}{H-L} \in (0,1)$. This implies that the seller also randomizes over H and L in equilibrium, which means that she must be indifferent between the two prices: $L=qH \iff q=\frac{L}{H}$. □

The seller's ignorance about the buyer's investment limits her ability to expropriate the investment gains, thus improving ex-ante efficiency by (partially) restoring the buyer's ex-ante investment incentive. As a result, the buyer invests with positive probability. However, asymmetric information at the trading stage creates ex-post inefficiency because trade breaks down when the buyer did not invest but the seller sets $p=H$. These two effects exactly cancel each other out in equilibrium.

Lau (2008): "Truth-or-Noise" Signal Structures. Next, consider a signal structure where $S=\{l,h,\phi\}$, $\pi(l|L)=\pi(h|H)=1-\mu$ and $\pi(\phi|L)=\pi(\phi|H)=\mu$. This implies that the seller perfectly observes the buyer's investment with probability $1-\mu$ and has no information with probability μ, as considered in Lau (2008). We call such an information structure a *truth-or-noise (TN)* signal structure, which is parametrized by the noise probability μ.

Proposition 3. *q is implementable by a TN signal structure if and only if $q \leq \frac{L}{H}$. For any $q \in (0,\frac{L}{H}]$, the optimal TN signal structure that implements q is uniquely $\mu=\frac{c}{H-L}$, and the seller's equilibrium payoff is $L+q(H-L-c)$.*[7]

Proof. In equilibrium, the seller must offer $p=H$ (respectively, $p=L$) after $s=h$ (respectively, $s=l$). If $q>\frac{L}{H}$, the seller will offer $p=H$ after $s=\phi$. In turn, the buyer will never invest, so $q>\frac{L}{H}$ is not implementable. If $q<\frac{L}{H}$, the seller will offer $p=L$ after $s=\phi$. The buyer will then randomize if and only if $\mu(H-L)-c=0 \iff \mu=\frac{c}{H-L}$, which

---

[7] $q=0$ is implementable by any $\mu \leq \frac{c}{H-L}$ and the seller's equilibrium payoff is always L. Therefore, the uniqueness property does not hold at $q=0$.

uniquely pins down the signal structure that implements $q < \frac{L}{H}$. Therefore, this is also the unique optimal signal structure that implements $q < \frac{L}{H}$, and it is readily verified that the seller's equilibrium payoff is $L + q\left(H - L - c\right)$. Next, notice that $\mu = \frac{c}{H-L}$ also implements $q = \frac{L}{H}$ with the seller always offering $p = L$ after $s = \phi$,[8] and the seller's equilibrium payoff is $L + \frac{L}{H}\left(H - L - c\right)$. This is the highest seller's payoff under $q = \frac{L}{H}$ because trade always occurs, so $\mu = \frac{c}{H-L}$ is an optimal signal structure for implementing $q = \frac{L}{H}$. We relegate the proof that $\mu = \frac{c}{H-L}$ is the unique optimal signal structure that implements $q = \frac{L}{H}$ to Appendix A. $\square$

As established in Lau (2008), the seller's equilibrium payoff can be higher than L using TN signal structures. Clearly, within this class of signal structures, the highest payoff is achieved when $q = \frac{L}{H}$ is optimally implemented. We show later that the TN signal structure is always suboptimal when a more general form of information control is available.

Optimal Signal Structure

We now fix a $q \in (0, 1)$ and consider the optimal signal structure that implements it.[9] The subsequent variables depend on q, but we omit the argument for ease of notation. Let β_s be the seller's posterior belief that $v = H$ after observing signal s under signal structure $\{S, \pi\}$. With q correctly conjectured by the seller in equilibrium,

$$\beta_s = \Pr\left(v = H|s\right) = \frac{\pi\left(s|H\right) q}{\pi\left(s|H\right) q + \pi\left(s|L\right)\left(1-q\right)}. \tag{1}$$

Conditional on s, the seller's payoff from setting $p = H$ is $\beta_s H$, and that from setting $p = L$ is L. Denote $x_s := \pi\left(s|H\right) q + \pi\left(s|L\right)\left(1-q\right)$ as the ex-ante probability that s is realized. We say that a signal structure $\{S, \pi\}$ is *direct* if $S = \{l, h\}$, x_l and x_h are both strictly positive,

---

[8] At $q = \frac{L}{H}$, the seller is indifferent between offering $p = L$ and $p = H$ after $s = \phi$.

[9] $q = 0$ can be optimally implemented by the fully informative signal structure. $q = 1$ is never implementable by Proposition 1.

and π generates a set of beliefs satisfying $\beta_l \leq \frac{L}{H}$ and $\beta_h > \frac{L}{H}$, which thus provides incentive-compatible pricing recommendations to the seller. Clearly, the "directness" property of a signal structure here is with respect to the fixed q, but we again drop this argument for ease of exposition.

Lemma 1. *Suppose that there exists a signal structure that implements q and the seller's payoff in the equilibrium is U. There exists a direct signal structure that also implements q, and in that equilibrium, the seller charges $p = H$ (respectively, $p = L$) when she observes $s = h$ (respectively, $s = l$), and her payoff is also U.*

On the seller's side, the intuition behind Lemma 1 is similar to the revelation principle. When the seller's posterior is higher (respectively, lower) than $\frac{L}{H}$, her best response is to charge $p = H$ (respectively, $p = L$); therefore, all signals that generate posteriors higher (respectively, lower) than $\frac{L}{H}$ can be grouped together. In turn, since the buyer's investment incentive is ultimately determined by the equilibrium price distribution, preserving the seller's pricing incentive would then also preserve the buyer's investment incentive.[10]

Owing to Lemma 1, we can henceforth restrict our attention to direct signal structures without loss of generality. The choice of a signal structure is equivalent to choosing a distribution of posteriors $\{x_s, \beta_s\}_{\sum_s x_s = 1}$ that satisfies the Bayes plausibility constraint:

$$\sum_{s \in \{l,h\}} x_s \beta_s = q. \tag{2}$$

Unlike a Bayesian persuasion problem, the signal structure here must also satisfy the following equilibrium condition:

---

[10] Notice that the convention here is for the seller to always charge $p = L$ when she is indifferent between the two prices (i.e., at a posterior of $\frac{L}{H}$). In the Online Appendix, we prove that in an optimal equilibrium, the seller never randomizes when she is indifferent; however, this result is not required for the subsequent analysis.

$$\pi(l|H)(H-L)-c=0. \tag{3}$$

The left-hand side of (3) is the buyer's payoff from investing. For him to be indifferent between investing and not investing (so that $q \in (0,1)$ in equilibrium), this payoff must be the same as his payoff from not investing, which is 0. Using (1), condition (3) is equivalent to

$$x_l\beta_l = q\left(\frac{c}{H-L}\right). \tag{4}$$

Therefore, a direct signal structure implements q if the resulting distribution of posteriors satisfies (2) and (4). The resulting seller's payoff is

$$x_lL + x_h\beta_hH \;=\; L + x_h(\beta_hH-L) \tag{5}$$

The following theorem gives the optimal signal structure:

Theorem 1. *A signal structure that implements q exists if and only if $q \leq \frac{L}{L+c}$. For any $q \in \left(0, \frac{L}{L+c}\right]$, the signal structure that maximizes the seller's payoff while implementing q is unique within the set of direct signal structures.*[11] *It consists of*

$$\begin{pmatrix} \pi(l|L) & \pi(h|L) \\ \pi(l|H) & \pi(h|H) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{c}{H-L} & 1-\frac{c}{H-L} \end{pmatrix} \tag{6}$$

The resulting posteriors are $\beta_h = 1$ and $\beta_l = \frac{1}{1+\frac{1-q}{q}\left(\frac{H-L}{c}\right)} \leq \frac{L}{H}$. Trade always takes place (i.e., zero ex-post inefficiency), and the seller's equilibrium payoff is $L + q(H-L-c)$.

We emphasize that the equilibrium existence condition in Theorem 1 takes into account all possible signal structures instead of only direct signal structures. Therefore, Theorem

---

[11] The optimal direct signal structure that implements $q=0$ is $\begin{pmatrix} \pi(l|L) & \pi(h|L) \\ \pi(l|H) & \pi(h|H) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \kappa & 1-\kappa \end{pmatrix}$ for any $\kappa \leq \frac{c}{H-L}$. Therefore, the uniqueness property does not hold at $q=0$.

1 provides an upper bound on the implementable investment probability, and the bound is strictly decreasing in the ratio $\frac{c}{L}$ but (perhaps surprisingly) independent of H. Moreover, the optimal signal structure is independent of q; thus, the direct signal structure that optimally implements q is the same for all implementable q. With regard to payoffs, the maximum seller's payoff under the optimal signal structure is higher than L. Therefore, partial information revelation in the form of the signal structure in Theorem 1 can always mitigate hold-up risks and improve upon the perfectly uninformative signal structure.

To explain the derivation of the optimal direct signal structure in (6), we first note that the buyer's ex-ante investment incentive comes from the probability of the seller not detecting his investment, which allows him to keep the investment gain. This probability is set at $\frac{c}{H-L}$ so that he is ex-ante indifferent between investing or not, which pins down $\pi(\cdot|H)$. As for $\pi(\cdot|L)$, since $\pi(h|L)$ is the conditional probability of a trade breakdown and $\pi(\cdot|L)$ does not affect the buyer's investment incentive, $\pi(h|L)$ is set to zero to eliminate all ex-post inefficiency, thus pinning down $\pi(\cdot|L)$.

To see why the implementable q is bounded below 1, notice that in fixing a q, a signal structure is a hypothesis test for the investment with the type I (false positive) error of $(1-q)\pi(h|L)$ and the type II (false negative) error of $q\pi(l|H)$. The type II error of the test is what creates an investment incentive for the buyer. However, the seller will follow the price recommendation from the test only if it is sufficiently accurate, which implies that it cannot make too much of either error. In turn, this puts an upper bound on q because as q increases, signal l needs to detect L more accurately so that the type II error is not too high; however, this destroys the buyer's ex-ante incentive to invest.

Although the signal structure in (6) is reminiscent of the optimal signal structure in the leading "prosecutor" example in Kamenica and Gentzkow (2011) (hereafter KG)—in that one state is always revealed (the "L" state here and the "guilty" state in KG)—the reasonings behind the two signal structures are very different. In particular, the seller's payoff is convex in her belief here,[12] and the optimal signal structure under Bayesian persuasion in KG when the receiver has such a payoff function is fully informative. However, a fully informative signal structure leads to the complete hold-up outcome here. More generally, in Bayesian persuasion, the optimal signal structure is derived from pooling "favorable" states with "unfavorable" states while maintaining the credibility of the signals. By contrast, the optimal signal structure here determines the conditional probabilities at each state *separately*.

Corollary 1. *The set of the seller's payoffs achievable in equilibrium is* $\left[L, L + \frac{L}{L+c}\left(H - L - c\right)\right]$. *The optimal signal structure can achieve a strictly higher seller's payoff than any TN signal structure.*

From Proposition 3 and Theorem 1, we see that the optimal signal structure can implement a higher investment probability than any TN signal structure and thus achieve a higher seller's payoff. To see how this gain arises, note that the buyer's investment incentive is driven by the seller having incentive to choose $p = L$ when she is uncertain about v. Under the TN signal structure, this happens when $s = \phi$, in which case the seller has absolutely no information, so the seller bases her inference about v solely from her conjecture on q, which is correct in equilibrium. In turn, the equilibrium q cannot be too high. On the other hand, under the optimal signal structure, the seller is uncertain about v when $s = l$, but unlike the situation under the TN signal structure, signal l carries information about v. This implies that the effect of the equilibrium q on the seller's inference about v is now smaller, so the downward constraint on q is also relaxed.

---

[12] At belief $\beta \leq \frac{L}{H}$, the seller's payoff is L; at belief $\beta \geq \frac{L}{H}$, her payoff is βH.

4 Multiple Investment Levels

In this section, we generalize the no ex-post inefficiency result in Theorem 1 to the hold-up setting with an arbitrary (but finite) number of investment levels.[13]

Model

Suppose now that the buyer has $m+1$ possible investment levels. Investment $i \in M := \{0, 1, \ldots, m\}$ results in a valuation of v^i and leads to a cost of c^i for the buyer. We assume that $v^0 < v^1 < \cdots < v^m$, $c^0 = 0$, and $c^i > 0$ for any $i \geq 1$. Therefore, $i = 0$ corresponds to no investment, and $m = 1$ is the case considered in the previous section. We also assume that $v^i - c^i > v^0$ for some $i \geq 1$ so that non-investment is never the socially efficient level. Beyond this, however, we do not make further assumptions on c^i or v^i at this point.

As before, an ex-ante determined signal structure generates signals about the buyer's valuation before the seller makes her price offer. The signal structure consists of $\{S, \pi\}$, where S is the signal space and $\pi\left(s|v^i\right)$ is the conditional probability of signal $s \in S$ after the buyer has chosen investment i. Let $\vec{q} = (q^0, q^1, \ldots, q^m)$ be a probability vector where q^i is the probability that the buyer chooses investment i. It is readily noted that in equilibrium, the seller will choose a price from only the set of possible ex-post valuations $\{v^0, v^1, \ldots, v^m\}$, and she will never offer a price v^i if she believes that $v = v^i$ with zero probability. The following statements are the analogs of Propositions 1 and 2 under multiple levels of investment:

Proposition 4.

1. *If $\vec{q}$ is implementable, then $q^0 > 0$. In any equilibrium, the buyer's payoff is zero.*
2. *Under the fully uninformative signal structure, the seller's equilibrium payoff is v^0. Therefore, the players' payoffs are the same as with a fully informative signal structure with a complete hold-up.*

---

[13] We consider the case with a continuum of investment levels in the Online Appendix.

Proposition 4.1 states that the buyer chooses not to invest with positive probability in equilibrium. This is because the seller will correctly conjecture the lowest investment that is played with positive probability by the buyer in equilibrium; hence, she will never charge a price below that resulting valuation. Therefore, the buyer receives a negative payoff unless the lowest investment played is of zero cost, which is possible only if it is the no-investment level. In turn, since the buyer must be indifferent toward any investment level played with positive probability and the no-investment level gives him a zero payoff, the buyer's payoff in equilibrium is zero. As in the single investment setting, the seller's payoff is the social welfare, so there is no need to distinguish between the two.

If the buyer's investment is perfectly observed by the seller, the hold-up problem arises: the only equilibrium outcome is that the buyer chooses not to invest and is charged price v^0. The social welfare with a complete hold-up is v^0, all of which is extracted by the seller. Proposition 4.2 generalizes Proposition 2: with multiple investment levels, the fully uninformative signal structure still does not improve welfare relative to the hold-up outcome.

Optimal Signal Structure

We consider the optimal signal structure next. As in Section 3, we proceed by fixing a $\vec{q}$ and then consider the optimal signal structure that implements it. It remains without loss of generality that we restrict our attention to direct signal structures, which we describe next. The basic intuition is similar to Lemma 1, but proving this claim requires additional notations that are not needed for the main analysis; therefore, we relegate the details to Appendix A (see Lemma 2).

Upon observing signal s, the seller's posterior is a probability vector $\vec{\beta}_s = \{\beta_s^0, \beta_s^1, \ldots, \beta_s^m\}$, where β_s^i is the probability that the seller assigns to $v = v^i$, and the updating formula is

$$\beta_s^i = \Pr\left(v = v^i | s\right) = \frac{\pi\left(s|v^i\right) q^i}{\sum_{j=0}^m \pi\left(s|v^j\right) q^j}. \tag{7}$$

Let $x_s = \sum_{i=0}^{m} q^i \pi\left(s|v^i\right)$ be the ex-ante probability of signal s being realized.

A direct signal structure consists of a signal space $S = M$ and a set of conditional probabilities π that result in posteriors satisfying the following condition:

$$\text{for any } s \in M\text{:} \quad \text{if } x_s > 0 \text{, then} \quad \sum_{j \geq s} \beta_s^j v^s \geq \sum_{j \geq i} \beta_s^j v^i \quad \forall i \neq s. \tag{8}$$

The term $\sum_{j \geq i} \beta_s^j$ in (8) is $\Pr\left[v \geq v^i|s\right]$, which is the seller's subjective probability upon observing signal s that the buyer will accept price v^i. Therefore, $\sum_{j \geq i} \beta_s^j v^i$ is her interim expected payoff from offering price v^i after signal s. Condition (8) thus implies that the direct signal structure provides signals that give incentive-compatible price recommendations to the seller—the seller sets $p = v^s$ upon receiving signal s.

Next, analogous to condition (4), a signal structure must satisfy an equilibrium condition. In particular, a direct signal structure $\{M, \pi\}$ implements $\vec{q}$ if

$$\text{for any } i \in M\text{:} \quad \text{if } q^i > 0 \text{, then} \quad \sum_{j \leq i} \pi\left(j|v^i\right)\left(v^i - v^j\right) - c^i = 0. \tag{9}$$

Condition (9) is the buyer's incentive-compatible condition—his payoff from playing any investment i in which $q^i > 0$ must be zero (Proposition 4.1).

Therefore, a direct signal structure $\{M, \pi\}$ implements $\vec{q}$ if they jointly satisfy (8) and (9). As $\sum_{j \leq i} \pi\left(j|v^i\right)$ is the conditional probability of trade occurring after the buyer has chosen investment i, the seller's ex-ante payoff, which is the ex-ante expected welfare, is

$$\sum_{i=0}^{m} q^i \left[\sum_{j \leq i} \pi\left(j|v^i\right)\right]\left(v^i - c^i\right). \tag{10}$$

Theorem 2.

1. *Suppose that $\vec{q}$ is implementable and the direct signal structure $\{M, \pi\}$ optimally implements $\vec{q}$. It holds that for all i such that $q^i > 0$, $\pi\left(j|v^i\right) = 0$ for all $j > i$, and the*

seller's equilibrium payoff is $\sum_{i=0}^{m} q^i \left(v^i - c^i\right)$.

2. *The set of achievable seller's payoffs is a closed interval* $[v^o, W^*]$ *with* $v^0 < W^*$. *Moreover, if a signal structure* π^* *can achieve* W^*, π^* *can also achieve any* $W \in [v^0, W^*]$.

Corollary 2. *By Theorem 2.1, under the optimal signal structure, the posterior satisfies* $\Pr\left(v \geq v^s | s\right) = 1$ *for any* $s \in M$ *such that* $x_s > 0$. *Therefore, trade always takes place.*

$$\begin{array}{cc} & \begin{array}{ccccc} s=0 & s=1 & s=2 & \dots & s=m \end{array} \\ \begin{array}{c} v^0 \\ v^1 \\ v^2 \\ \vdots \\ v^m \end{array} & \begin{pmatrix} \boldsymbol{q^0\pi_{00}} & \boxed{q^0\pi_{01}} & \boxed{q^0\pi_{02}} & \dots & \boxed{q^0\pi_{0m}} \\ \underline{q^1\pi_{10}} & \boldsymbol{q^1\pi_{11}} & \boxed{q^1\pi_{12}} & \dots & \boxed{q^1\pi_{1m}} \\ \underline{q^2\pi_{20}} & \underline{q^2\pi_{21}} & \boldsymbol{q^2\pi_{22}} & \dots & \boxed{q^2\pi_{2m}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \underline{q^m\pi_{m0}} & \underline{q^m\pi_{m1}} & \underline{q^m\pi_{m2}} & \dots & \boldsymbol{q^m\pi_{mm}} \end{pmatrix} \end{array}$$

Figure 1: Unconditional probabilities of the three types of outcomes. An ex-ante investment incentive relies on the "underlined" outcomes; ex-post inefficiency arises in the "boxed" outcomes.

Since $v^0 < W^*$, information control can always allow the seller to earn a higher payoff relative to the hold-up outcome. Theorem 2.1 provides a partial characterization of the optimal signal structure that achieves it. To understand it, consider the unconditional probabilities of the outcomes. Denoting $\pi\left(j|v^i\right)$ by π_{ij} (to simplify notation), Figure 1 shows a probability matrix where entry $q^i\pi_{ij}$ is the unconditional probability of the "outcome" in which the buyer had taken investment i and signal j is realized. The set of outcomes can be categorized into three types: the "boxed outcomes", the "underlined outcomes" and the "diagonal outcomes". The boxed outcomes are in the upper triangular section of the matrix, where the probabilities of their occurrences are boxed; in these outcomes, the seller is overly optimistic about the buyer's valuation and hence charges too high a price, which then results in a trade breakdown and thus ex-post inefficiency. On the other hand, the underlined outcomes are in the lower triangular section of the matrix, where the probabilities of their occurrences are underlined; in these outcomes, the seller is too pessimistic about the buyer's valuation and hence undercharges him, which then allows the buyer to keep some of the investment gain. Finally, the diagonal outcomes are the diagonals of the matrix; in these outcomes, the seller has the exact judgement of the buyer's valuation and hence charges him his true valuation, which results in the seller extracting all the investment gains.

Given $\vec{q}$, the seller strictly prefers only the diagonal outcomes to occur. However, the buyer's ex-ante investment incentive can come from only the underlined outcomes. In particular, along each row, the underlined entries in the matrix (and only these underlined entries) have to satisfy the buyer's incentive compatibility condition in (9). Therefore, at least some of the underlined outcomes must occur with positive probability. The other concern is the seller's incentive compatibility condition in (8). This requires that along each column, the seller believes that the outcome is sufficiently likely to be the diagonal outcome. Notice that moving all the probabilities of the boxed outcomes (where ex-post inefficiency occurs) along each row to their corresponding diagonal outcomes actually helps satisfy (8) and does not affect (9). This illustrates the separability between eliminating ex-post inefficiency and creating ex-ante investment incentive for the buyer. Therefore, under the optimal signal structure, the boxed outcomes (and hence ex-post inefficiency) never occur.

The characterization of the optimal signal structure is completed by choosing the occurrences of the underlined outcomes to generate the buyer's investment incentive to play $\vec{q}$. As these choices interact with the occurrences of the diagonal outcomes to simultaneously satisfy both conditions (8) and (9), this problem is significantly more difficult when $m > 1$, and there is no general property that one can exploit.

A2 Signal Structure

In this subsection, we provide a signal structure—termed the A2 signal structure—that is readily computed and could potentially be optimal. We first define the A2 signal structure in Definition 1 and then characterize its sets of implementable buyer's investment strategies and achievable seller's payoffs in Proposition 5.

Definition 1. An adjacent type-II error (A2) signal structure is a direct signal structure $\{M, \pi\}$ satisfying

- $\pi\left(0|v^0\right) = 1$, and $\pi\left(j|v^0\right) = 0 \ \forall j \geq 1$.
- For all $i \geq 1$, $\pi\left(i|v^i\right) = 1 - \frac{c^i}{v^i - v^{i-1}}$, $\pi\left(i-1|v^i\right) = \frac{c^i}{v^i - v^{i-1}}$, and $\pi\left(j|v^i\right) = 0 \ \ \forall j \neq i-1, i$.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{c^1}{v^1 - v^o} & 1 - \frac{c^1}{v^1 - v^0} & 0 & 0 & 0 \\ 0 & \frac{c^2}{v^2 - v^1} & 1 - \frac{c^2}{v^2 - v^1} & 0 & 0 \\ 0 & 0 & \frac{c^3}{v^3 - v^2} & 1 - \frac{c^3}{v^3 - v^2} & 0 \\ 0 & 0 & 0 & \frac{c^4}{v^4 - v^3} & 1 - \frac{c^4}{v^4 - v^3} \end{pmatrix}$$

Figure 2: The A2 signal structure for $m = 4$, where each entry is $\pi\left(s|v^i\right)$, the row indexes v^i and the column indexes signal s.

Figure 2 provides an illustration of the A2 signal structure for $m = 4$. It has the simple feature that at each investment level i, the seller's pessimism, which is what creates an incentive for the buyer to choose i, is created only via the possibility that it could be the next lower investment $i-1$, hence the name "adjacent type-II error" signal structure. Under this restriction, the buyer's incentive-compatibility condition in (9) for $i \geq 1$ reduces to

$$\pi\left(i-1|v^i\right)\left(v^i - v^{i-1}\right) - c^i = 0,$$

which completely pins down $\pi\left(\cdot|v^i\right)$ as described in Definition 1. Clearly, the A2 signal structure is well-defined if and only if $v^i - v^{i-1} \geq c^i$ for all $i \geq 1$. For ease of exposition, we assume that this holds for now, and we discuss how to modify the A2 signal structure when this assumption is violated later.

In terms of practical considerations, the A2 signal structure is also attractive. First, each signal is generated by only two different valuations, so the posterior at each signal can be easily computed by the seller. Moreover, the signal structure either provides the correct information about the valuation or makes the smallest possible downward estimation of it, which means that it never makes any "big" mistake.

Proposition 5. *Strategy $\vec{q}$ is implementable by the A2 signal structure if and only if*

$$q^1 \leq \frac{v^0}{c^1} q^0 \quad \text{and} \quad q^{i+1} \leq \frac{\left(v^i - v^{i-1} - c^i\right) v^i}{\left(v^i - v^{i-1}\right) c^{i+1}} q^i \;\; \forall i \in \{1, 2, \ldots, m-1\}. \tag{11}$$

The set of the seller's payoffs achievable by the A2 signal structure is a closed interval $\left[v^0, \bar{W}\right]$ with $\bar{W} > v^o$.

As $\bar{W} > v^0$, Proposition 5 implies that the A2 signal structure can always mitigate hold-up risks. The value of $\bar{W}$ and the $\vec{q}$ that is implemented to achieve it are characterized in the proof of Proposition 5. It is readily verified that when $m = 1$, the A2 signal structure is the optimal binary signal structure in Theorem 1, so $\bar{W} = v^o + \frac{v^o}{v^o + c^1}\left(v^1 - v^0 - c^1\right)$. While we are unable to say when the A2 signal structure is optimal in general, we can characterize the set of parameters for which the A2 signal structure is indeed the optimal signal structure when $m = 2$. We show this next.

When $m = 2$. The optimal signal structure for the case of $m = 2$ is represented as a matrix in (13) below. From Theorem 2, all the entries in the upper triangle are zero. Next, the buyer's incentive-compatibility condition in (9) for $i = 1$ pins down $\pi\left(0|v^1\right)$ and $\pi\left(1|v^1\right)$, and the same condition for $i = 2$ requires that

$$\pi\left(0|v^2\right)\left(v^2 - v^0\right) + \pi\left(1|v^2\right)\left(v^2 - v^1\right) - c^2 = 0. \tag{12}$$

Denoting $\pi\left(0|v^2\right)$ by γ, (12) implies that $\pi\left(1|v^2\right) = f\left(\gamma\right) = \frac{c^2 - \gamma\left(v^2 - v^0\right)}{v^2 - v^1}$; therefore, the signal structure is fully parameterized by γ, with $\gamma \leq \frac{c^2}{v^2 - v^0}$.

Proposition 6. *Let* $\xi = \left(v^2 - v^0\right)\left(1 - \frac{v^1\left(v^1 - v^0 - c^1\right)}{\left(v^2 - v^0\right)\left(v^0 + c^1\right) - v^0\left(v^1 - v^0 - c^1\right)}\right)$. *The optimal direct signal structure for* $m = 2$ *is uniquely:*

$$\begin{pmatrix} \pi\left(0|v^0\right) & \pi\left(1|v^0\right) & \pi\left(2|v^0\right) \\ \pi\left(0|v^1\right) & \pi\left(1|v^1\right) & \pi\left(2|v^1\right) \\ \pi\left(0|v^2\right) & \pi\left(1|v^2\right) & \pi\left(2|v^2\right) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{c^1}{v^1 - v^0} & 1 - \frac{c^1}{v^1 - v^0} & 0 \\ \gamma & f\left(\gamma\right) & 1 - \gamma - f\left(\gamma\right) \end{pmatrix} \tag{13}$$

where

- *if* $c^2 < \xi$, *then* $\gamma = \frac{c^1 c^2}{\left(v^1 - v^0\right)\left(v^2 - v^1 + c^1\right)} > 0$;
- *if* $c^2 \geq \xi$, *then* $\gamma = 0$ *(i.e., the A2 signal structure).*[14]

Proposition 6 implies that the A2 signal structure is optimal when $c^2 \geq \xi$, which is not a knife-edge case. The optimal signal structure in (13) is derived by solving a continuum of linear programming problems—the details are provided in the appendix. To give some intuition, we note that condition (11) in Proposition 5 implies that the A2 signal structure (i.e., $\gamma = 0$) can only implement strategies $\vec{q}$ in which the q^2 is bounded relative to the q^1. When we increase γ (i.e., deviate from the A2 signal structure), we can relax this bound and implement a higher q^2, but this is only possible by decreasing q^1 *and* increasing q^0. Therefore, implementing such a $\vec{q}$ is valuable only if the gain from investment level $i = 2$

---

[14] To be precise, the uniqueness property does not apply at the knife-edge case of $c^2 = \xi$. In this case, the optimal γ is anything in the set $\left[0, \frac{c^1 c^2}{\left(v^1 - v^0\right)\left(v^2 - v^1 + c^1\right)}\right]$.

(i.e., $v^2 - c^2$) is high. It can be verified that ξ is increasing in v^2; hence, a lower value of investment level $i = 2$ makes it more likely that $c^2 \geq \xi$, which is the condition for $\gamma = 0$ to be optimal.

On the Assumption that $v^i - v^{i-1} \geq c^i$. The A2 signal structure is well-defined if and only if $v^i - v^{i-1} \geq c^i \; \forall i \geq 1$. We describe how to modify the A2 signal structure in Definition 1 when some investment i violates this assumption. This is done by "removing" the investment by having the signal structure always fully revealing it—so the buyer will never choose this investment—and then adjusting the indexes of the other investments accordingly. For example, suppose that $m = 5$, but $v^4 - v^3 < c^4$. We can remove investment $i = 4$ by setting $\pi\left(4|v^4\right) = 1$ and $\pi\left(4|v^i\right) = 0 \; \forall i \neq 4$ and then adjust the indexes by keeping $i = 0, 1, 2, 3$ as they are and making $i = 5$ the new "4". The A2 signal structure is then constructed based on only the 5 remaining investment levels.[15] Alternatively, we could remove investment $i = 3$ instead in the same way. A removal does not guarantee that the condition $v^i - v^{i-1} \geq c^i$ is immediately satisfied for the "remaining" investments, so we might have to perform multiple removals. However, since there exists some i such that $v^i - c^i > v^0$, there exists at least one way to remove the investment(s) that assures that the assumption is satisfied. The set of implementable payoffs depends on which investment(s) was removed. However, since Proposition 5 characterizes the set of achievable seller's payoffs by the associated A2 signal structure for each set of valuations and costs satisfying the assumption, it is possible to compare the sets of achievable payoffs across the different choices of removals.

---

[15] i.e., $\pi\left(0|v^0\right) = 1$; for $i = 1, 2, 3$, $\pi\left(i|v^i\right) = 1 - \frac{c^i}{v^i - v^{i-1}}$, $\pi\left(i-1|v^i\right) = \frac{c^i}{v^i - v^{i-1}}$; $\pi\left(4|v^4\right) = 1$; $\pi\left(5|v^5\right) = 1 - \frac{c^5}{v^5 - v^3}$, $\pi\left(3|v^5\right) = \frac{c^5}{v^5 - v^3}$; and the other probabilities are 0.

5 Conclusion

The literature has noted that introducing information asymmetry regarding the buyer's investment can prevent the seller from abusing her bargaining power and hence alleviate the hold-up problem. Implicitly suggested in these earlier papers is a tradeoff between creating an ex-ante investment incentive and ex-post inefficiency due to asymmetric information. In this paper, we make the point that such a tradeoff is unnecessary because the type of asymmetric information that creates an ex-ante investment incentive is different from the type of asymmetric information that causes ex-post inefficiency. Consequently, by hiding and revealing the right information, ex-post inefficiency can be eliminated without compromising the ex-ante investment incentive. We show that information control can always improve welfare and does not require overly complex arrangements in the relationship to attain some welfare gain. In turn, we hope that our results can serve as a guide for future work on how to make better use of information control to mitigate the hold-up problem in various settings.

References

Aumann, Robert J and Michael Maschler. 1995. *Repeated Games with Incomplete Information.* MIT press.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2015. "The Limits of Price Discrimination." *American Economic Review* 105 (3):921–57.

Condorelli, Daniele and Balazs Szentes. 2018. "Information Design in the Hold-up Problem." .

Fréchette, Guillaume, Alessandro Lizzeri, and Jacopo Perego. 2017. "Rules and Commitment in Communication." .

Gibbons, Robert. 1992. *Game Theory for Applied Economists*. Princeton University Press.

González, Patrick. 2004. "Investment and Screening under Asymmetric Endogenous Information." *The RAND Journal of Economics* :502–519.

Gul, Faruk. 2001. "Unobservable Investment and the Hold-Up Problem." *Econometrica* 69 (2):343–376.

Guo, Yingni and Eran Shmaya. 2018. "Costly Miscalibration in Communication." .

Halac, Marina. 2015. "Investing in a relationship." *The RAND Journal of Economics* 46 (1):165–185.

Hermalin, Benjamin E. 2013. "Unobserved Investment, Endogenous Quality, and Trade." *The RAND Journal of Economics* 44 (1):33–55.

Hermalin, Benjamin E. and Michael L. Katz. 2009. "Information and the Hold-up Problem." *The RAND Journal of Economics* 40 (3):405–423.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review* 101 (6):2590–2615.

Lau, Stephanie. 2008. "Information and Bargaining in the Hold-up Problem." *The RAND Journal of Economics* 39 (1):266–282.

Lipnowski, Elliot, Doron Ravid, and Denis Shishkin. 2018. "Persuasion via Weak Institutions." .

Min, Daehong. 2017. "Bayesian Persuasion under Partial Commitment." .

Nguyen, Anh and Teck Yong Tan. 2019. "Bayesian Persuasion with Costly Messages." .

Ostrovsky, Michael and Michael Schwarz. 2010. "Information Disclosure and Unraveling in Matching Markets." *American Economic Journal: Microeconomics* 2 (2):34–63.

Rayo, Luis and Ilya Segal. 2010. “Optimal Information Disclosure.” *Journal of Political Economy* 118 (5):949–987.

Riordan, Michael. 1990. “What is Vertical Integration?” In *The Firm as a Nexus of Treaties*, edited by Oliver E. Williamson Masahiko Aoki, Bo Gustafsson. London: Sage, 94–111.

Roesler, Anne-Katrin and Balázs Szentes. 2017. “Buyer-Optimal Learning and Monopoly Pricing.” *American Economic Review* 107 (7):2072–80.

Smith, Gordon. 2007. “Outsourcing Problem: The Right-to-audit Clause.” *Journal of Corporate Accounting & Finance* 18 (5):19–22.

Tan, Teck Yong. 2019. “Knowledge as Property Rights under the Ratchet Effect of Innovation.” .

A Proof Appendix

Proof of Proposition 3

Proof. We show that any TN signal structure (besides $\mu = \frac{c}{H-L}$) that implements $q = \frac{L}{H}$ gives a lower equilibrium payoff for the seller. At $q = \frac{L}{H}$, the seller is indifferent between offering either prices. Let α be the probability that the seller offers L. The buyer's indifference condition then requires that $\mu\alpha(H-L) - c = 0 \iff \mu = \frac{c}{\alpha(H-L)}$. Therefore, the signal structure $\frac{c}{\alpha(H-L)}$ implements $q = \frac{L}{H}$ with the seller offering $p = L$ with probability α after $s = \phi$. Since the seller is indifferent between either price at $s = \phi$, her payoff after $s = \phi$ is L. Thus, her ex-ante expected payoff is $q(\mu L + (1-\mu)H) + (1-q)L = L + q\left(H - L - \frac{c}{\alpha}\right)$, which is decreasing in α. Therefore $\mu = \frac{c}{(H-L)}$ uniquely implements $q = \frac{L}{H}$ optimally . $\square$

Proof of Lemma 1

Proof. Follows from Lemma 2 with $m = 1$, $v^0 = L$, $v^1 = H$ and $c^1 = c$. $\square$

Proof of Theorem 1

Proof. We prove the "only if" direction of the existence result first. Suppose, for a contradiction, that $q > \frac{L}{L+c}$ but there exists a signal structure that implements q. From Lemma 1, there exists a direct signal structure that implements q. Let $\beta_l \leq \frac{L}{H}$ and $\beta_h > \frac{L}{H}$ be the resulting posteriors. From (4), $\beta_l = \frac{qc}{x_l(H-L)}$; from (2), $x_l = \frac{\beta_h - q}{\beta_h - \beta_l}$. Combining the two, we obtain

$$\beta_l = \frac{q\left(\frac{c}{H-L}\right)\beta_h}{\beta_h - q + q\left(\frac{c}{H-L}\right)} = \frac{q\left(\frac{c}{H-L}\right)}{1 - \frac{q}{\beta_h}\left(1 - \frac{c}{H-L}\right)} \tag{14}$$

Since $\frac{c}{H-L} < 1$, β_l is decreasing in β_h. $\beta_h \leq 1$ then implies that $\beta_l \geq \frac{q\left(\frac{c}{H-L}\right)}{1-q\left(1-\frac{c}{H-L}\right)}$. When $q > \frac{L}{L+c}$, $\beta_l > \frac{L}{H}$, which contradicts $\beta_l \leq \frac{L}{H}$. Next, for the "if" direction, it is readily verified that when $q \leq \frac{L}{L+c}$, the signal structure in the theorem results in posteriors $\beta_h = 1$ and $\beta_l \leq \frac{L}{H}$, and it satisfies the equilibrium condition (4).

For optimality, it suffices to consider the set of direct signal structures (Lemma 1). For any β_h, the corresponding β_l is in (14), and the seller's payoff, from (5), is $U(\beta_h) = L + x_h\beta_h H - x_h L$. From (2) and (4),

$$\begin{aligned} x_h\beta_h H &= (q - x_l\beta_l) H = \left[q - q\left(\tfrac{c}{H-L}\right)\right] H \\ x_h L &= \tfrac{q-\beta_l}{\beta_h-\beta_l} L = \left(\frac{q - \frac{q\left(\frac{c}{H-L}\right)\beta_h}{\beta_h - q + q\left(\frac{c}{H-L}\right)}}{\beta_l - \frac{q\left(\frac{c}{H-L}\right)\beta_h}{\beta_h - q + q\left(\frac{c}{H-L}\right)}}\right) L = \tfrac{q}{\beta_h}\left(1 - \tfrac{c}{H-L}\right) L \quad . \end{aligned}$$

Therefore, $U(\beta_h) = L + q\left(H - \frac{L}{\beta_h}\right)\left[1 - \left(\frac{c}{H-L}\right)\right]$. As $U(\beta_h)$ is strictly increasing, the optimal

β_h is 1, and $\beta_l = \frac{q\left(\frac{c}{H-L}\right)}{1-q+q\left(\frac{c}{H-L}\right)}$. The seller's payoff is $U\left(1\right) = L + q\left(H - L - c\right)$. The signal structure is then backed out via $\pi\left(s|H\right) = \frac{\beta_s x_s}{q}$ and $\pi\left(s|L\right) = \frac{\left(1-\beta_s\right)x_s}{1-q}$. $\square$

Proof of Proposition 4

Proof. The first statement follows from the discussion in the main text. For the second statement, under a fully uninformative signal structure, the seller plays the same pricing strategy after every (uninformative) signal. If the seller offers $p = v^0$, her payoff is v^0 because trade is guaranteed; therefore, it suffices to show that the seller's equilibrium pricing strategy must involve offering $p = v^0$ with strictly positive probability. Suppose, for a contradiction, that the lowest price offered by the seller with positive probability is some $v^{\underline{i}} > v^0$. This implies that the buyer's equilibrium strategy must involve $q^{\underline{i}} > 0$. As noted in the main text, if $q^i > 0$ in equilibrium, the buyer's payoff from playing i must be zero. However, as the price is never below $v^{\underline{i}}$, the buyer's payoff from choosing investment $\underline{i}$ is $-c^{\underline{i}}$, which is negative when $\underline{i} > 0$; contradiction. $\square$

Sufficiency of Direct Signal Structures

We formalize the claim in Section 4 that it is without loss of generality to restrict our attention to direct signal structures. We begin by setting up the problem from its primitives and prove this claim in Lemma 2 below.

Fix a $\vec{q} = (q^0, q^1, \ldots, q^m)$, which is correctly conjectured by the seller in equilibrium. Under an arbitrary signal structure $\{S, \pi\}$, the seller's posterior upon observing signal $s \in S$ is a probability vector $\vec{\beta}_s = (\beta_s^0, \beta_s^1, \ldots, \beta_s^m)$, where β_s^i is the probability that the seller assigns to $v = v^i$ and the updating formula is in (7). The ex-ante probability of signal s is thus $x_s = \sum_{i=0}^m q^i \pi\left(s|v^i\right)$. Let $\vec{\sigma} := \{\vec{\sigma}_s\}_{s\in S}$ and $\vec{\sigma}_s := (\sigma_s^0, \sigma_s^1, \ldots, \sigma_s^m)$ be a probability vector where σ_s^i is the probability that the seller offers price $p = v^i$ upon observing signal s.

The signal structure has to satisfy equilibrium conditions whereby both players' strategies are best responses against each other. $\vec{\sigma}$ is a best response price strategy for the seller if

$$\text{for all } s \in S, \text{ if } \sigma_s^i > 0, \text{ then } \sum_{k \geq i} \beta_s^k v^i \geq \sum_{k \geq j} \beta_s^k v^j \quad \forall j \neq i,$$

which can equivalently be written as

$$\text{for all } s \in S, \text{ if } \sigma_s^i > 0, \text{ then } \sum_{k \geq i} q^k \pi\left(s|v^k\right) v^i - \sum_{k \geq j} q^k \pi\left(s|v^k\right) v^j \geq 0 \quad \forall j \neq i. \tag{15}$$

For the buyer, for every investment played with strictly positive probability under $\vec{q}$, his expected payoff from it must be 0 (Proposition 4). This is equivalent to

$$\text{for all } i \in M, \text{ if } q^i > 0, \text{ then } \sum_{s \in S} \pi\left(s|v^i\right) \sum_{j \leq i} \sigma_s^j \left(v^i - v^j\right) = c^i \ . \tag{16}$$

Therefore, a signal structure implements $\vec{q}$ and $\vec{\sigma}$ if the signal structure satisfies (15) and (16). The seller's ex-ante expected payoff in equilibrium is

$$\sum_{s \in S} x_s \left[\max_{i \in M} \sum_{k \geq i} \beta_s^k v^i\right].$$

Lemma 2. *Suppose that the signal structure $\{S, \pi\}$ implements $\vec{q}$ and the seller's equilibrium payoff is U. There exists a direct signal structure $\{M, \hat{\pi}\}$ that also implements $\vec{q}$, and in that equilibrium, the seller's strategy $\vec{\hat{\sigma}}$ satisfies $\hat{\sigma}_i^i = 1$ $\forall i \in M$, and her payoff is also U.*

Proof. Let $\vec{\sigma}$ be the seller's equilibrium under $\{S, \pi\}$ when implementing $\vec{q}$. For each $\hat{s} \in M$ and $i \in M$, set

$$\hat{\pi}\left(\hat{s}|v^i\right) = \sum_{s \in S} \pi\left(s|v^i\right) \sigma_s^{\hat{s}}. \tag{17}$$

For any $i \in M$, $\sum_{\hat{s} \in M} \left[\sum_{s \in S} \pi\left(s|v^i\right) \sigma_s^{\hat{s}}\right] = \sum_{s \in S} \pi\left(s|v^i\right) = 1$, so $\hat{\pi}$ is a valid signal structure.

$\vec{\hat{\sigma}}$ is the seller's best response if $\forall i \neq j$,

$$\begin{aligned} & \textstyle\sum_{k\geq i} q^k \hat{\pi}\left(i|v^k\right) v^i \geq \sum_{k\geq j} q^k \hat{\pi}\left(i|v^k\right) v^j \\ \iff \quad & \textstyle\sum_{k\geq i} q^k \left[\sum_{s\in S} \pi\left(s|v^k\right) \sigma_s^i\right] v^i \geq \sum_{k\geq j} q^k \left[\sum_{s\in S} \pi\left(s|v^k\right) \sigma_s^i\right] v^j \\ \iff \quad & \textstyle\sum_{s\in S} \left[\sum_{k\geq i} q^k \pi\left(s|v^k\right) v^i - \sum_{k\geq j} q^k \pi\left(s|v^k\right) v^j\right] \sigma_s^i \geq 0 \end{aligned} \quad ,$$

where the last inequality holds from (15). Next, if playing $\vec{q}$ is a best response for the buyer under $\hat{\pi}$, then for all $i \in M$, if $q^i > 0$, $\hat{\pi}$ satisfies

$$\begin{aligned} & \textstyle\sum_{j\leq i} \hat{\pi}\left(j|v^i\right)\left(v^i - v^j\right) = c^i \\ \iff \quad & \textstyle\sum_{j\leq i} \left[\sum_{s\in S} \pi\left(s|v^i\right) \sigma_s^j\right]\left(v^i - v^j\right) = c^i \end{aligned} \quad ,$$

which holds from (16). Therefore, $\hat{\pi}$ also implements $\vec{q}$. To check that the seller's payoff is also U under $\hat{\pi}$, recall from Proposition 4 that the seller's payoff is the social welfare. Thus, it suffices to check that the probabilities of a trade breaking down at each valuation are the same across π and $\hat{\pi}$. Under π, conditional on v^i, the probability of no trade is $\sum_{s\in S} \pi\left(s|v^i\right) \sum_{j\geq i} \sigma_s^j$; under $\hat{\pi}$, the corresponding probability is $\sum_{j\geq i} \hat{\pi}\left(j|v^i\right) = \sum_{j\geq i} \left[\sum_{s\in S} \pi\left(s|v^i\right) \sigma_s^j\right] = \sum_{s\in S} \pi\left(s|v^i\right) \sum_{j\geq i} \sigma_s^j$. □

Lemma 2 thus establishes that it is without loss of generality to restrict our attention to direct signal structures as is done in the main text.

Proof of Theorem 2

Proof. Suppose that π implements $\vec{q}$. Therefore, π satisfies (8) and (9). Consider the following signal structure $\{M, \hat{\pi}\}$:

$$\text{for any } i, j \in M,\;\; \hat{\pi}\left(j|v^i\right) = \begin{cases} \pi\left(j|v^i\right) & \text{, if } j < i \\ \sum_{j' \geq i}^{m} \pi\left(j'|v^i\right) & \text{, if } j = i \\ 0 & \text{, if } j > i \end{cases}$$

Under $\hat{\pi}$, the buyer's payoff from choosing investment i is $\left[\sum_{j \leq i}^{m} \hat{\pi}\left(j|v^i\right)\left(v^i - v^j\right)\right] - c^i = \left[\sum_{j \leq i}^{m} \pi\left(j|v^i\right)\left(v^i - v^j\right)\right] - c^i$. Therefore, $\hat{\pi}$ satisfies (9). Next, we check that $\hat{\pi}$ satisfies (8), which can be equivalently written as

$$\text{for any } i \in M, \quad \sum_{k \geq i}^{m} q^k \pi\left(i|v^k\right) v^i \geq \sum_{k \geq j}^{m} q^k \pi\left(i|v^k\right) v^j \quad , \forall j \in M. \tag{18}$$

Fix $i \in M$. Consider $j < i$ first. Notice that this implies $\sum_{k \geq j}^{m} q^k \hat{\pi}\left(i|v^k\right) = \sum_{k \geq i}^{m} q^k \hat{\pi}\left(i|v^k\right)$, so $\sum_{k \geq i}^{m} q^k \hat{\pi}\left(i|v^k\right) v^i > \sum_{k \geq j}^{m} q^k \hat{\pi}\left(i|v^k\right) v^j$. Consider $j > i$ next. Note that $\hat{\pi}\left(i|v^k\right) = \pi\left(i|v^k\right)$ for any $i > k$, while $\hat{\pi}\left(i|v^i\right) \geq \pi\left(i|v^i\right)$. Therefore,

$$\begin{aligned} & \sum_{k \geq i}^{m} q^k \hat{\pi}\left(i|v^k\right) v^i - \sum_{k \geq j}^{m} q^k \hat{\pi}\left(i|v^k\right) v^j \\ = & \left[q^k \hat{\pi}\left(i|v^i\right) - q^k \pi\left(i|v^i\right)\right] v^i + \underbrace{\sum_{k \geq i}^{m} q^k \pi\left(i|v^k\right) v^i - \sum_{k \geq j}^{m} q^k \pi\left(i|v^k\right) v^j}_{\geq 0 \text{ from } (18)} \quad \geq \; 0, \end{aligned}$$

which hence satisfies (8).

This means that $\hat{\pi}$ also implements $\vec{q}$, and the ex-ante payoff of the seller is

$$\sum_{i=0}^{m} q^i \left[\sum_{j \leq i}^{m} \hat{\pi}\left(j|v^i\right)\right]\left(v^i - c^i\right) = \sum_{i=0}^{m} q^i \left(v^i - c^i\right) \geq \sum_{i=0}^{m} q^i \left[\sum_{j \leq i}^{m} \pi\left(j|v^i\right)\right]\left(v^i - c^i\right),$$

where the inequality is strict if there exists $j > i$ such that $q^i > 0$ but $\pi\left(j|v^i\right) > 0$ (i.e., $\sum_{j\leq i}^{m} \pi\left(j|v^i\right) < 1$).

The highest achievable seller's payoff is the solution to the program that chooses π and $\vec{q}$ to maximize (10) subject to constraints (8) and (9). The feasible set is clearly compact and the objective function is continuous. Therefore, the maximum exists. Let the maximum payoff be W^*. As the fully informative signal structure (and also the fully uninformative signal structure) can achieve a seller's payoff of v^0, we know that $W^* \geq v^0$. To show that $W^* > v^0$, consider a j such that $v^j - c^j > v^0$, which exists by assumption. It is readily verified (using the proof of Theorem 1) that the strategy $\vec{q}$ in which $q^0 = 1 - \frac{v^0}{v^0+c^j}$, $q^j = \frac{v^0}{v^0+c^j}$ and $q^i = 0\ \forall i \neq j$ is implementable by the signal structure $\pi\left(0|v^0\right) = 1$, $\pi\left(0|v^j\right) = \frac{c^j}{v^j - v^0}$, $\pi\left(j|v^j\right) = 1 - \frac{c^j}{v^j - v^0}$ and $\pi\left(i|v^i\right) = 1\ \forall i \neq j$. The resulting seller's payoff is $v^0 + q^j\left(v^j - v^0 - c^j\right) > v^0$. As this payoff is weakly lower than W^* by definition, $W^* > v^0$.

Next, to show that the set of achievable seller's payoff is a closed interval and every payoff in the interval is achievable by the same signal structure, let π^* be a signal structure that implements some $\vec{q}^*$ that achieves a seller's payoff of W^*. Let $\bar{i}$ be the highest investment level played with strictly positive probability under $\vec{q}^*$. Notice that if we move probability $\varepsilon > 0$ from $q^{\bar{i}*}$ to q^{0*}, constraint (8) is relaxed while constraint (9) is unaffected. Thus, the resulting $\vec{q}$ is still implemented by π^*. The seller's payoff achieved under the new $\vec{q}$ is $W^* - \varepsilon\left(v^{\bar{i}} - c^{\bar{i}} - v^0\right)$. Therefore, by varying ε, any payoff in the interval $\left[W^* - q^{\bar{i}}\left(v^{\bar{i}} - c^{\bar{i}} - v^0\right), W^*\right]$ is achievable. To achieve a payoff lower than $W^* - q^{\bar{i}}\left(v^{\bar{i}} - c^{\bar{i}} - v^0\right)$, we begin with the $\vec{q}$ that has shifted the entire $q^{\bar{i}*}$ to q^{0*} and induct the argument on the remaining highest investment. At $q^0 = 1$, the payoff is v^0; therefore, any payoff in $[v^0, W^*]$ is achievable by π^*. $\square$

Proof of Proposition 5

Proof. Under the A2 signal structure, for all i,

$$\sum_{j\leq i}\pi\left(j|v^i\right)\left(v^i-v^j\right)-c^i=\pi\left(i-1|v^i\right)\left(v^i-v^{i-1}\right)-c^i=0.$$

Therefore, constraint (9) is always satisfied.[16] Next, a strategy $\vec{q}$ and the A2 signal structure jointly satisfy constraint (8) if and only if for all $i\leq m-1$,

$$\left[q^i\pi\left(i|v^i\right)+q^{i+1}\pi\left(i|v^{i+1}\right)\right]v^i\geq q^{i+1}\pi\left(i|v^{i+1}\right)v^{i+1}. \tag{19}$$

For $i=0$, (19) is equivalent to $q^1\leq\frac{v^0}{c^1}q^0$; for $i\geq 1$, (19) is equivalent to $q^i\geq\frac{\left[v^i-v^{i-1}\right]c^{i+1}}{\left[v^i-v^{i-1}-c^i\right]v^i}q^{i+1}$, which jointly make condition (11).

Next, let $\mathcal{Q}$ be the set of the buyer's investment strategies satisfying condition (11), and let $\theta^i:=\frac{\left(v^i-v^{i-1}\right)c^{i+1}}{\left(v^i-v^{i-1}-c^i\right)v^i}$ with the convention that $v^{-1}=0$. Therefore, $\vec{q}\in\mathcal{Q}$ if and only if $q^i\geq\theta^iq^{i+1}\ \forall i\leq m-1$. We claim that the highest seller's payoff among all $\vec{q}\in Q$ is the $\vec{q}$ that satisfies $q^i=\theta^iq^{i+1}\ \forall i\leq m-1$. To see this, suppose for a contradiction that $q^i>\theta^iq^{i+1}$. Denote $Q=q^i+q^{i+1}$, and denote $q^{i+1}=\lambda Q$ and $q^i=(1-\lambda)Q$. Therefore, $\frac{1-\lambda}{\lambda}>\theta^i$. Consider $\hat{\lambda}>\lambda$ such that $\frac{1-\hat{\lambda}}{\hat{\lambda}}=\theta^i$, and decrease q^i to $\hat{q}^i=\left(1-\hat{\lambda}\right)Q$ and increase q^{i+1} to $\hat{q}^{i+1}=\hat{\lambda}Q$. It is readily seen that condition (11) for all other $j\neq i$ will still be satisfied after this change, so this new $\vec{\hat{q}}$ is still in $\mathcal{Q}$. As $v^{i+1}-v^i-c^{i+1}>0\ \forall i\leq m-1$, it implies that $v^{i+1}-c^{i+1}>v^i-c^i$, so the change increases the seller's payoff because the probability of taking the better investment $i+1$ increases at the expense of the inferior investment level i. Therefore, the highest seller's payoff among all $\vec{q}\in\mathcal{Q}$, which we denote by $\bar{W}$, is achieved under the $\vec{q}$ such that $q^i=\frac{1}{\theta^{i-1}}q^{i-1}\ \forall i\geq 1$. We can write the probability of taking each

---

[16] If $q^i=0$, then $\pi\left(\cdot|v^i\right)$ does not affect the equilibrium conditions (8) and (9).

investment level i as

$$q^i \; = \; \frac{1}{\theta^{i-1}} q^{i-1} \; = \; \ldots \; = \; \left(\frac{1}{\theta^{i-1}} \times \frac{1}{\theta^{i-2}} \times \cdots \times \frac{1}{\theta^0} \right) q^0 \; = \; \underbrace{\prod_{j=0}^{i-1} \frac{\left(v^j - v^{j-1} - c^j\right) v^j}{\left(v^j - v^{j-1}\right) c^{j+1}}}_{\alpha^i} q^0$$

Note that $q^0 = 1 - \sum_{i=1}^m q^i = 1 - \sum_{i=1}^m \alpha^i q^0$. Therefore, $q^0 = \frac{1}{1+\sum_{j=1}^m \alpha^j}$. Let $\alpha^0 = 1$ and we have $q^i = \frac{\alpha^i}{\sum_{j=0}^m \alpha^j}$ $\forall i$. The seller's payoff under this $\vec{q}$ is thus

$$\bar{W} = \sum_{i=0}^m q^i \left(v^i - c \right) = \sum_{i=0}^m \frac{\alpha^i}{\sum_{j=0}^m \alpha^j} \left(v^i - c^i \right).$$

Following the argument that establishes Theorem 2.2, any seller's payoff in $\left[v^0, \bar{W}\right]$ is achievable by the A2 signal structure. □

Proof of Proposition 6

Proof. Under signal structure γ, the highest achievable seller's payoff comes from solving the following linear program:

$$W(\gamma) \quad := \quad \max_{q^1 q^2, \in [0,1]} \; \left(1 - q^1 - q^2\right) v^0 + q^1 \left(v^1 - c^1\right) + q^2 \left(v^2 - c^2\right) \tag{20}$$

subject to

$$\left[\left(1 - q^1 - q^2\right) + q^1 \left(\frac{c^1}{v^1 - v^0} \right) + q^2 \gamma \right] v^0 \quad \geq \quad \left[q^1 \left(\frac{c^1}{v^1 - v^0} \right) + q^2 \gamma \right] v^1 \tag{21}$$

$$\left[\left(1 - q^1 - q^2\right) + q^1 \left(\frac{c^1}{v^1 - v^0} \right) + q^2 \gamma \right] v^0 \quad \geq \quad q^2 \gamma v^2 \tag{22}$$

$$\left[q^1 \left(\frac{v^1 - v^0 - c^1}{v^1 - v^0} \right) + q^2 f(\gamma) \right] v^1 \quad \geq \quad q^2 f(\gamma) v^2 \tag{23}$$

$$q^1 + q^2 < 1 \tag{24}$$

The objective is the sender's payoff, and the three constraints are the seller's pricing incentive constraints in (8): (21) (respectively, (22)) is the constraint that upon observing $s = 0$, the seller prefers charging $p = v^0$ to $p = v^1$ (respectively, $p = v^2$), and (23) is the constraint that upon observing $s = 1$, the seller prefers charging $p = v^1$ to $p = v^2$. The optimal signal structure is then derived from solving $\max_{\gamma\in\left[0,\frac{c^2}{v^2-v^0}\right]} W\left(\gamma\right)$.

We first note that (24) is subsumed by (21). To see this, we rewrite (21) as $\frac{\left(1-q^1-q^2\right)}{q^1\frac{c^1}{v^1-v^0}+q^2\gamma}+1\geq\frac{v^1}{v^0}$; if $q^1+q^2\geq 1$, then the left-hand side is less than 1, but $\frac{v^1}{v^0}$ is greater than 1. Therefore, we can ignore (24).

For each program γ, the optimal q^1 and q^2 must be strictly positive: first, it is readily verified that $q^2 = 0$ is never optimal under any γ. Next, by writing (23) as $q^1\left(\frac{v^1-v^0-c^1}{v^1-v^0}\right)\geq q^2 f\left(\gamma\right)\left(v^2-v^1\right)$, if $q^1 = 0$, then $q^2 = 0$; therefore, $q^1 = 0$ also cannot be optimal.

Next, note that constraint (21) subsumes (22) if $\left[q^1\left(\frac{c^1}{v^1-v^0}\right)+q^2\gamma\right]v^1\geq q^2\gamma v^2$, which is equivalent to $\frac{q^1}{q^2}\geq\frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}}$. Moreover, constraint (23) can be rewritten as $\frac{q^1}{q^2}\geq\frac{f(\gamma)\left(v^2-v^1\right)}{\frac{v^1-v^0-c^1}{v^1-v^0}}$. Therefore, if $\frac{f(\gamma)\left(v^2-v^1\right)}{\frac{v^1-v^0-c^1}{v^1-v^0}}\geq\frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}}$, which is equivalent to $\gamma\leq\frac{c^1c^2}{(v^1-v^0)(v^2-v^1+c^1)}=:\hat{\gamma}$, constraint (22) is always subsumed by (21) and (23). As this is a linear program, the solution must always lie on a vertex. This implies that when $\gamma\leq\hat{\gamma}$, the solution is (q^1, q^2) such that (21) and (23) bind. The following follows from this and some algebra:

Lemma 3. *When $\gamma\leq\hat{\gamma}$, the solution to program $W\left(\gamma\right)$ is*

$$q^1 = \left(\frac{v^0}{1+\left[c^1+v^0\right]h\left(\gamma\right)}\right)h\left(\gamma\right)$$
$$q^2 = \left(\frac{v^0}{\left[1+\left[c^1+v^0\right]h\left(\gamma\right)\right]}\right)\frac{1}{\left[\gamma\left(v^1-v^0\right)+v^0\right]}$$

where $h\left(\gamma\right)=\frac{\left[c^2-\gamma\left(v^2-v^0\right)\right]\left[v^1-v^0\right]}{\left[\gamma(v^1-v^0)+v^0\right]\left[v^1-v^0-c^1\right]v^1}$. The value is:

$$W\left(\gamma\right)=v^0+\left(\frac{v^0}{\left[1+\left[c^1+v^0\right]h\left(\gamma\right)\right]}\right)\left[h\left(\gamma\right)\left(v^1-v^0-c^1\right)+\frac{v^2-v^0-c^2}{\gamma\left(v^1-v^0\right)+v^0}\right],$$

and

$$\frac{dW(\gamma)}{d\gamma}=\frac{v^0v^1\left(v^1-v^0\right)\left(v^1-v^0-c^1\right)\left[v^0v^1\left(v^1+c^2-c^1\right)-v^2\left(\left(v^0+c^1\right)\left(v^0+c^2\right)+v^1\left(v^1-c^1\right)\right)+\left(v^0+c^1\right)\left(v^2\right)^2\right]}{\left[c^1(v^0v^1+\gamma(v^1-v^0)(v^2+v^1-v^0)-c^2(v^1-v^0))-(v^1-v^0)\left(v^0c^2+v^0v^1+\gamma\left((v^0)^2+(v^1)^2-v^0(v^1+v^2)\right)\right)\right]^2}.$$

Thus $\frac{dW(\gamma)}{d\gamma}\leq 0$ *if and only if:*

$$\begin{aligned} & v^0v^1\left(v^1+c^2-c^1\right)-v^2\left(\left(v^0+c^1\right)\left(v^0+c^2\right)+v^1\left(v^1-c^1\right)\right)+\left(v^0+c^1\right)\left(v^2\right)^2\leq 0 \\ \iff & c^2\leq\left(v^2-v^0\right)\left(1-\frac{v^1\left(v^1-v^0-c^1\right)}{(v^2-v^0)(v^0+c^1)-v^0(v^1-v^0-c^1)}\right)=\xi. \end{aligned}$$

Next, consider $\gamma>\hat{\gamma}$. Suppose, for a contradiction, that constraint (22) does not bind. This implies that constraints (21) and (23) must bind. The binding (23) implies that $\frac{q^1}{q^2}=\frac{f(\gamma)\left(v^2-v^1\right)}{\frac{v^1-v^0-c^1}{v^1-v^0}}<\frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}}$, where the inequality follows from $\gamma>\hat{\gamma}$. This implies that $\left[q^1\left(\frac{c^1}{v^1-v^0}\right)+q^2\gamma\right]v^1<q^2\gamma v^2$, so if constraint (21) binds, then constraint (22) must be violated—contradiction. Therefore, if $\gamma>\hat{\gamma}$, then constraint (22) is binding under the solution of program $W(\gamma)$. There are thus two cases to consider:

Case A: Constraints (21) and (22) bind. Let $\left(q_A^1(\gamma),q_A^2(\gamma)\right)$ be the solution to the system of equations where (21) and (22) hold with equality. Solving it, we have

$$\begin{aligned} q_A^1(\gamma)= & \frac{\gamma v^0\left(v^1-v^0\right)\left(v^2-v^1\right)}{\gamma v^2\left(c^1+v^0\right)\left(v^1-v^0\right)+v^0v^1\left(c^1-\gamma\left(v^1-v^0\right)\right)} \\ q_A^2(\gamma)= & \frac{v^0v^1c^1}{\gamma v^2\left(c^1+v^0\right)\left(v^1-v^0\right)+v^0v^1\left(c^1-\gamma\left(v^1-v^0\right)\right)} \end{aligned}$$

Under $\left(q_A^1(\gamma),q_A^2(\gamma)\right)$, the value is

$$\begin{aligned} W_A(\gamma):= & v^0+q_A^1(\gamma)\left(v^1-v^0-c^1\right)+q_A^2(\gamma)\left(v^2-v^0-c^2\right) \\ = & \frac{v^0v^1\left[\gamma\left(v^1-v^0\right)\left(v^2-v^1\right)+c^1\left(v^2-c^2+\gamma\left(v^1-v^0\right)\right)\right]}{\gamma v^2\left(v^0+c^1\right)\left(v^1-v^0\right)-v^0v^1\left(\gamma\left(v^1-v^0\right)-c^1\right)}. \end{aligned}$$

When constraints (21) and (22) bind, $\left[q^1\left(\frac{c^1}{v^1-v^0}\right)+q^2\gamma\right]v^1 = q^2\gamma v^2 \iff \frac{q^1}{q^2} = \frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}} > \frac{f(\gamma)\left(v^2-v^1\right)}{\frac{v^1-v^0-c^1}{v^1-v^0}}$, where the last inequality follows from $\gamma > \hat{\gamma}$. The last inequality implies constraint (23); therefore, $W_A(\gamma)$ is achievable, and

$$\begin{aligned}\frac{dW_A(\gamma)}{d\gamma} &= -\frac{c^1v^0v^1\left(v^1-v^0\right)\left[v^0\left(v^2-v^1\right)\left(v^2-v^1-c^2\right)+c^1\left(v^2\left(v^2-c^2\right)-v^0v^1\right)\right]}{\left[v^0\left(c^1+\gamma\left(v^0-v^1\right)\right)v^1-\gamma\left(c^1+v^0\right)\left(v^0-v^1\right)v^2\right]^2}\\ &< -\frac{c^1v^0v^1\left(v^1-v^0\right)\left[v^0\left(v^2-v^1\right)\left(v^2-v^1-c^2\right)+c^1v^2\left(v^2-v^0-c^2\right)\right]}{\left[v^0\left(c^1+\gamma\left(v^0-v^1\right)\right)v^1-\gamma\left(c^1+v^0\right)\left(v^0-v^1\right)v^2\right]^2} < 0\end{aligned}$$

Case B: (22) and (23) bind. Let $\left(q_B^1(\gamma), q_B^2(\gamma)\right)$ be the solution to the system of equations where (22) and (23) hold with equality. Solving it, we have

$$\begin{aligned}q_B^1(\gamma) =& \frac{v^0\left(v^1-v^0\right)\left(c^2+\gamma\left(v^2-v^0\right)\right)}{\left(v^1-v^0-c^1\right)\left[v^0c^2+v^0v^1+\gamma\left(v^1-v^0\right)\left(v^2-v^0\right)\right]}\\ q_B^2(\gamma) =& \frac{v^0v^1}{v^0c^2+v^0v^1+\gamma\left(v^1-v^0\right)\left(v^2-v^0\right)}\end{aligned}$$

Under $\left(q_B^1(\gamma), q_B^2(\gamma)\right)$, the value is

$$\begin{aligned}W_B(\gamma) :=& v^0+q_B^1(\gamma)\left(v^1-v^0-c^1\right)+q_B^2(\gamma)\left(v^2-v^0-c^2\right)\\ =& \frac{v^0v^1v^2}{c^2v^0+v^0v^1+\gamma\left(v^1-v^0\right)\left(v^2-v^0\right)}.\end{aligned}$$

The binding (23) implies that $\frac{q^1}{q^2} = \frac{f(\gamma)\left(v^2-v^1\right)}{\frac{v^1-v^0-c^1}{v^1-v^0}} < \frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}}$. The last inequality $\frac{q^1}{q^2} < \frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}}$ implies that $\left[q^1\left(\frac{c^1}{v^1-v^0}\right)+q^2\gamma\right]v^1 < q^2\gamma v^2$, which means that constraint (21) is satisfied. Therefore, $W_B(\gamma)$ is achievable. It is immediate that $W_B(\gamma)$ is strictly decreasing.

The following lemma thus follows:

Lemma 4. *When $\gamma > \hat{\gamma}$, the value function of program $W(\gamma)$ is $W(\gamma) = \max\left\{W_A(\gamma), W_B(\gamma)\right\}$. Since both W_A and W_B are strictly decreasing, W is also strictly decreasing.*

Proposition 6 then follows from Lemmas 3 and 4. □