How Effective Can Ex Post Destruction Alleviate the Hold-up Problem?

Huan Wang[*], Juyan Zhang[†], and Yi Zhang[‡]

Revised: July 2014[§]

Abstract

We first investigate whether or not ex post destruction can possibly alleviate the hold-up problem in a one-shot game between a supplier and a buyer. The answer is yes but only when the buyer believes that the supplier might be a Homo reciprocans agent with sufficiently strong propensity for reciprocity. Under incomplete information with informed supplier, investment is made feasible by the "mismatch" between the buyer's belief of stronger supplier reciprocal propensity and a de facto weaker one. Under incomplete information with uninformed supplier, the "mismatch" between the buyer's belief of weaker supplier reciprocal propensity and a stronger ex post one results in "destruction." Thereafter, we inquire how effective the alleviation can be. We find out that the degree of efficiency improvement is positively correlated with the intensity of potential ex post destruction for the case of uninformed supplier while non-monotonically correlated with the buyer's prior belief about the supplier's uninformedness.

JEL classification: C72, L14

Keywords: Hold-up, Homo Reciprocans, Destruction Intensity, Uninformedness

[*]CASS; hw276@nyu.edu.

[†]SUFE; zhangjuyan@gmail.com.

[‡]Singapore Management University; yizhang@smu.edu.sg.

1 Introduction

According to Che and Sákovics (2008), "hold-up arises when part of the return on an agent's relationship-specific investments is ex post expropriable by his trading partner." With an incomplete contract, which arises due to such causes as unforeseen contingencies and the enforcement problem, relationship-specific investments are typically distorted by the hold-up problem and therefore inefficient.

The main stream literature on hold-up assumes that human beings are Homo economicus agents who maximize their own ex post monetary payoffs. Therefore, the parties end up with non-trade payoffs (outside option) if ex post negotiation breaks down. However, the real world is actually filled with stories of "destruction" that makes both parties worse off than in non-trade. For example, Sobel (2005) points out that some recently fired employees engage in sabotage activities, costly for both the employers and the employees, such as ruining computer systems before they leave their offices.

In general, empirical studies of social behaviors indicate that human beings are mostly Homo reciprocans agents who are willing to engage in destructive behaviors at their own costs to punish the perceived perpetrators. There are some studies that compare games with and without added punishment schemes. For example, Fehr and Gächter (2000) report the empirical evidence that prospects of cooperation will be greatly improved if there exists an option to punish defectors, even though punishments are costly for the punishers in a one-shot context.[1]

Two main questions of our inquiry are: can the option of ex post destruction in general enhance ex ante efficiency? If yes, how effective can be the enhancement? We construct a simple hold-up game in the following context : (1) only one party (the supplier) invests ex ante; (2) the other party (the buyer) makes a take-it-or-leave-it offer during the ex post negotiation; (3) the supplier has the option of destruction if the ex post negotiation breaks down.[2]

Logically, under complete information, if the supplier is a Homo economicus agent who maximizes her own ex post payoff, she will never conduct costly destruction.[3]

---

[1]Dufwenberg and Kirchsteiger (2004) investigate a modified version of the game "So Long, Sucker," invented by John Nash, in which the subsequent players incur some monetary costs to punish the first movers. They show that there exists a sequential reciprocity equilibrium (SRE) if the marginal costs for the subsequent players are "swamped by the sweetness of revenge."

[2]There are some literature studying the ultimatum game preceded by a sunk cost investment of the responder (supplier). For instance, Ellingsen and Johannesson (2004) conduct experiments to study the mitigation of the hold-up through communication. In their experimental setting, if the ex post negotiation breaks down, both parties get zero, leaving the seller with a net loss. In contrast, with option of destruction in our setting, if the ex post negotiation breaks down, both parties could end up with a net loss, including the buyer.

[3]Ng's (2011) conclusion that the ex post ability to destroy the relationship-specific investment can solve the hold-up problem is based on the assumption of supplier indifference. Only when the

If, instead, the supplier is a Homo reciprocans agent who will reject an offer below her threshold and engage in some destructive behavior, then the buyer may have an incentive to propose an offer greater than or equal to the threshold to induce the supplier to invest ex ante.

However, is Homo reciprocans condition alone sufficient? The answer is no. We find out that the hold-up problem can be possibly solved only when the buyer believes that the supplier has a sufficiently strong propensity for reciprocity. Specifically, under complete information, only if the supplier reciprocal propensity is strong enough to cover the investment cost, will the supplier invest. In contrast, under incomplete information with informed supplier, investment is made feasible by the "mismatch" between the buyer's belief of stronger supplier reciprocal propensity and a de facto weaker one. Under incomplete information with uninformed supplier, the "mismatch" between the buyer's belief of weaker supplier reciprocal propensity and a stronger ex post one results in "destruction."

We find out further that the degree of efficiency improvement is positively correlated with the intensity of potential ex post destruction for the case of uninformed supplier while non-monotonically correlated with the buyer's prior belief about the supplier's uninformedness.

Our paper is related to the literature of behavioral game theory, which suggests that fairness (Fehr and Schmidt 1999) or reciprocity (Dufwenberg and Kirchsteiger 2004) may affect inefficiency of the hold-up problem. For example, Dufwenberg, Smith, and Van Essen (2013) argue that vengeance (negative reciprocity) may alleviate the inefficiency caused by the hold-up problem, provided that the investor holds the residual rights of control once ex post negotiation breaks down.[4] von Siemens (2009) investigates a model of private information in which the seller[5] knows her own type as either "selfish" or "fair-minded." Therefore the seller's investment plays a role of signalling that entails behavioral costs as the major incentive for investment.[6]

The rest of the paper is organized as follows. Section 2 introduces the basic settings. Section 3 discusses the case of complete information, in which both the supplier and the buyer are informed about the supplier's reciprocal propensity. Section 4 discusses the case of incomplete information with informed and uninformed supplier

---

supplier is indifferent between destructing and not destructing, will she destruct, which is in fact non-trade since it incurs no additional cost for the supplier.

[4]In Dufwenberg, Smith, and Van Essen (2013), the residual value for the investor (player 1) is normalized to zero, while the residual value for the opponent (player 2) is $(1-\alpha)v_2$, where $\alpha \in [0,1]$. If ex post negotiation breaks down, player 1 punishes player 2 at player 1's own cost by forgoing the "consolation payoff" t, which is less than the residual value.

[5]Equivalent to the supplier in our model.

[6]von Siemens (2009) assumes behavioral costs for the "fair-minded" seller if she accepts the unfair offer below some threshold. In contrast, we focus on the option of ex post destruction, in which the supplier rejects an offer below some threshold and engage in some destructive behavior to punish the buyer at the supplier's own cost.

respectively. The degree of efficiency improvement is discussed in section 5 and 6 for the intensity of potential ex post destruction and the prior belief about the supplier's uninformedness respectively. Section 7 concludes.

2   Basic Settings

Consider a trading relationship between a buyer and a supplier, who are both risk neutral. At date 0, the supplier decides whether or not to make an investment that costs her $c > 0$. Investment (I) enables her to split a pie of size m with the buyer in the subsequent period. Assume there is no discounting and it is socially efficient to invest, i.e., $c < m$. If the supplier does not make the investment (O), they both get nothing.

The timing of the model is as follows. If at date 0 the supplier invests, at date 1 the buyer makes a take-it-or-leave-it offer, $p \geq 0$. Then at the following date 2 the supplier has three options: to accept the offer (A), to reject the offer (R), or to engage in some destructive behavior (D). If the supplier accepts the offer, they will split the pie according to the offer and the game ends. Rejecting an offer ends the game and they will stay with their own non-trade payoffs, which are normalized to zero. If the supplier engages in some destructive behavior (D), the payoffs for the supplier and the buyer will be $(-X, -Y)$ on top of their non-trade payoffs, where $X > 0, Y > 0$.

The supplier could be a Homo economicus agent or a Homo reciprocans agent. Specifically, if the supplier is a Homo economicus agent, who maximizes her own ex post payoff, the game tree is illustrated in Figure 1.
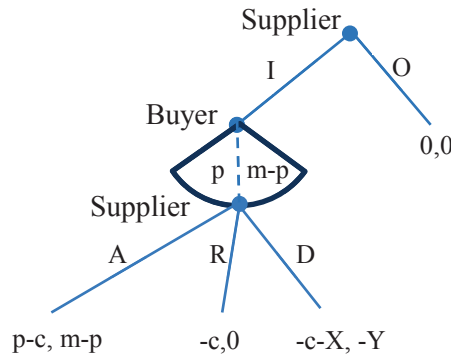

Figure 1: Game tree with a Homo economicus supplier

If the supplier is a Homo reciprocans agent, who will reject an offer below some threshold $T \in (0, m]$ and engage in some destructive behavior to punish the buyer at the supplier's own cost, the game tree is illustrated in Figure 2.
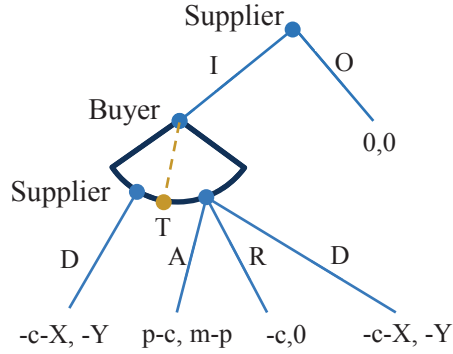
4

Figure 2: Game tree with a Homo reciprocans supplier

2.1 Information

There are two folds of information: first whether the supplier is a Homo economicus or a Homo reciprocans agent; second if she is a Homo reciprocans agent, how high her threshold for destruction T is. In the following sections, we discuss the complete information case and incomplete information case respectively. In the case of complete information, both the supplier and the buyer know whether the supplier is a Homo economicus or a Homo reciprocans agent, and know the supplier's threshold T if she is a Homo reciprocans agent.

In contrast, in the incomplete information case, the supplier and the buyer may not know whether the supplier is a Homo economicus or a Homo reciprocans agent let alone her threshold T. Suppose the prior belief follows some distribution. Following Fehr and Schmidt (1999), assume with probability s the supplier is Homo economicus,[7] while with probability $1 - s$ she is a Homo reciprocans to various degrees $T \in (k, m)$, which follows some continuous distribution F, with the density f, where $k \in (0, m)$. In addition, assume the density function f decreases in T to capture the idea that the supplier is less likely to be strong Homo reciprocans.[8]

For incomplete information, there are two sub-cases. 1. In the case of incomplete information with **informed supplier**, the supplier knows whether she is a Homo economicus or a Homo reciprocans agent and her threshold T if she is a Homo reciprocans agent, while the buyer does not, and the buyer knows that the supplier knows these information. 2. In the case of incomplete information with **uninformed supplier**, neither the supplier nor the buyer knows ex ante whether the supplier is a Homo economicus or a Homo reciprocans agent let alone her threshold T if she is

---

[7]It is equivalent to say that for the Homo economicus agent, $T = 0$.

[8]In reality, from the supplier's point of view, the "fair" ex post offer from the buyer has to at least cover the ex ante investment cost. An example of the distribution is the beta distribution: $Beta(1, \beta, c, m)$ with the support (c, m), where $\beta > 1$.

a Homo reciprocans agent. In this case, since both the supplier and the buyer are uninformed, their decisions will be based on their prior beliefs.

3   Complete Information

Now consider the case with complete information: both the supplier and the buyer know whether the supplier is a Homo economicus or a Homo reciprocans agent, and know her threshold T if she is a Homo reciprocans agent.

3.1   Homo Economicus Supplier

Suppose the supplier is a Homo economicus agent, and both the supplier and the buyer know this information. Back to figure 1, as $X > 0$, $-c > -c - X$. The Homo economicus supplier will never destruct. Anticipating this, the buyer will only offer the minimum amount to the supplier.[9] Backward to the beginning of the game, as long as the offer is less than the cost of the investment, there is no incentive for the Homo economicus supplier to invest ex ante. Therefore, if the supplier is a Homo economicus agent, who maximizes her own ex post payoff, the option of ex post destruction does not alleviate the inefficiency of the hold-up problem.

3.2   Homo Reciprocans Supplier

Suppose the supplier is a Homo reciprocans agent, and both the supplier and the buyer know that the supplier is a Homo reciprocans agent and her threshold T. Back to figure 2, the existence of the Homo reciprocans supplier, who will "destruct" if the offer is less than T, enables the supplier to threaten the buyer not to fully exploit her ex post. Therefore, the buyer has an incentive to propose an offer $p = T$ ex post.[10] Back to the beginning of the game, as long as, $T \geq c$, the supplier has the incentive to invest ex ante. In this case, the intensity of ex post destruction Y (the damage caused by the destruction to the buyer) is irrelevant because no offer is less than T. The underinvestment caused by the hold-up problem is only partially solved, as there is no investment ex ante if $T < c$.

---

[9]As pointed out by Fehr and Schmidt (1999), if there is a smallest money unit ϵ, then the buyer may offer $p = \epsilon$ to the supplier ex post, other than $p = 0$ in the continuous case.

[10]$p > T$ is not necessary, as $p = T$ is sufficient to induce the supplier to "accept" the offer.

4 Incomplete Information

4.1 Incomplete Information – Informed Supplier

In the case of incomplete information with informed supplier, the supplier knows whether she is a Homo economicus or a Homo reciprocans agent and knows her threshold T if she is a Homo reciprocans agent, while the buyer does not, and the buyer knows that the supplier knows these information. The following proposition shows that under incomplete information with informed supplier, only if the buyer's prior belief of the supplier's propensity for reciprocity is sufficiently strong, will the supplier have the incentive to invest ex ante.

Proposition 1 *In the case of incomplete information with informed supplier, there exists a p^*, such that*

$$\begin{cases} D(p^*) = 0 & if \quad D(k) \geq 0 \\ p^* = 0 & if \quad D(k) < 0 \end{cases}$$

where $D(p) = -s + (1-s)[-F(p) + (m-p)f(p)]$.

If $p^* \geq c$, we have a unique equilibrium, in which the supplier with $T \leq p^*$ invests at date 0 and the buyer offers $p = p^*$ at date 1; if $p^* < c$, we have a unique equilibrium, in which the supplier does not invest at date 0.

Proof. See Appendix. ∎

Intuitively, under incomplete information with informed supplier, the supplier will not invest ex ante if she is a Homo reciprocans agent with threshold greater than the anticipated ex post offer from the buyer at date 1. Therefore, even with the option of destruction ex post, there will be no destruction at all in equilibrium.

Further, to solve the underinvestment caused by hold-up, there are two conditions. First of all, the buyer's prior belief of the supplier's propensity for reciprocity has to be sufficiently strong, such that $p^* \geq c$. Second, the supplier has to be either a Homo economicus agent or a Homo reciprocans agent with $T \leq p^*$. That is to say, it is the "mismatch" between the buyer's belief of stronger supplier reciprocal propensity and a de facto weaker one that makes the investment feasible.

4.2 Incomplete Information – Uninformed Supplier

In the case of incomplete information with **uninformed supplier**, neither the supplier nor the buyer knows whether the supplier is a Homo economicus or a Homo reciprocans agent let alone her threshold T if she is a Homo reciprocans agent. In

this case, since both the supplier and the buyer are uninformed, their decisions will be based on their prior belief. The following proposition shows that under incomplete information with uninformed supplier, only if the prior belief of the supplier's propensity for reciprocity is sufficiently strong, will the supplier have the incentive to invest ex ante.

Proposition 2 *In the case of incomplete information with uninformed supplier, there exists a p^{**}, such that*

$$\begin{cases} p^{**} = m & if \quad \widetilde{D}(m) > 0 \\ \widetilde{D}(p^{**}) = 0 & if \quad \widetilde{D}(k) \geq 0 \geq \widetilde{D}(m) \\ p^{**} = 0 & if \quad \widetilde{D}(k) < 0 \end{cases}$$

where $\widetilde{D}(p) = -s + (1-s)[-F(p) + (m - p + Y)f(p)]$.

If $p^{**} \geq c$, we have a unique equilibrium, in which the supplier invests at date 0 and the buyer offers $p = p^{**}$ at date 1; if $p^{**} < c$, we have a unique equilibrium, in which the supplier does not invest at date 0.

Proof. See Appendix. ∎

Intuitively, as the supplier is uninformed ex ante, the buyer is willing to offer more to the supplier ex post, to avoid possible destruction by the supplier. Therefore, $p^{**} \geq p^*$ for any $Y > 0$. In this sense, the possibility of destruction may become an incentive to enhance efficiency in bilateral relationships with hold-up.

Further, to solve the underinvestment caused by hold-up, we only need one condition: the prior belief of the supplier's propensity for reciprocity has to be sufficiently strong, such that $p^{**} \geq c$. Compared with the informed supplier case in section 4.1, the condition is relaxed as $p^{**} \geq p^*$. Moreover, beyond the ex ante inefficiency, we could have ex post inefficiency. If the supplier's ex post realized threshold $T > p^{**}$, the supplier will reject the offer and engage in some destruction behavior. That is to say, it is the "mismatch" between the buyer's belief of weaker supplier reciprocal propensity and a stronger ex post propensity that results in "destruction."

5 How Effective Destruction Intensity Matters?

Now, we turn to inquire how effective the alleviation can be. In this section, we consider the intensity of potential ex post destruction. In the following section, we move on to consider the uncertainty of supplier informedness of reciprocal propensity.

5.1 Complete Information

In the case of complete information, both the supplier and the buyer are informed. Back to section 3, there is no investment if the supplier is a Homo economicus or a Homo reciprocans agent with $T < c$. Even with the option of destruction ex post, there is no destruction at all in equilibrium. Consequently, the intensity of ex post destruction Y (the damage caused by the destruction to the buyer) is irrelevant.

5.2 Incomplete Information – Informed Supplier

In the case of incomplete information with informed supplier, the supplier is informed, while the buyer is not, and the buyer knows that the supplier is informed. The supplier will not invest ex ante if she is a Homo reciprocans agent with threshold greater than the anticipated ex post offer from the buyer at date 1. Similar to the complete information case, even with the option of destruction ex post, there will be no destruction at all in equilibrium. Subsequently, p^* is independent of the intensity of ex post destruction Y. That is, the intensity of ex post destruction is irrelevant, which is described in the following corollary.

Corollary 1 p^* *is a constant function of* Y.

5.3 Incomplete Information – Uninformed Supplier

In the case of incomplete information with uninformed supplier, neither the supplier nor the buyer is informed ex ante. In this case, their decisions will be based on their prior belief. The following corollary shows that with uninformed supplier, the more intense the ex post destruction Y, the more the buyer is willing to offer to the supplier ex post, to avoid possible destruction by the supplier.

Corollary 2 p^{**} *is an increasing function of* Y. *In addition,* $\lim_{Y \to 0} p^{**} = p^*$.

Proof. See Appendix. ∎

In another word, the degree of efficiency improvement is positively correlated with the intensity of potential ex post destruction.

6 How Effective Supplier Informedness Uncertainty Matters?

We move on to consider the uncertainty of supplier informedness of reciprocal propensity. Suppose there is a prior belief that the probability for the supplier to be "unin-

formed" ex ante is $\pi \in (0, 1)$, while the the probability for her to be "informed" ex ante is $1 - \pi$.

The following proposition shows that under incomplete information with uncertainty of supplier informedness, only if the prior belief of the supplier's reciprocal propensity is sufficiently strong, will the supplier have the incentive to invest ex ante.

Proposition 3 *In the case of incomplete information with uncertainty of supplier informedness, there exists a $\overline{p}$, such that*

$$\begin{cases} \overline{p} = m & if \quad \overline{D}(m) > 0 \\ \overline{D}(\overline{p}) = 0 & if \quad \overline{D}(k) \geq 0 \geq \overline{D}(m) \\ \overline{p} = 0 & if \quad \overline{D}(k) < 0 \end{cases}$$

where $\overline{D}(p) = -s + (1 - s)[-F(p)] + \pi(1 - s)(m - p + Y)f(p)$.

if $\overline{p} \geq c$, we have a unique equilibrium, in which informed supplier with $T \leq \overline{p}$ and uninformed supplier invest at date 0 and the buyer offers $p = \overline{p}$ at date 1; if $\overline{p} < c$, we have a unique equilibrium, in which there is no investment at date 0.

Proof. See Appendix. ■

Similar to corollary 2, the following corollary shows that with uncertainty of supplier informedness, the more intense the ex post destruction Y, the more the buyer is willing to offer to the supplier ex post, to avoid possible destruction by the uninformed supplier.

Corollary 3 *$\overline{p}$ is an increasing function of Y.*

Proof. See Appendix. ■

Moreover, the following corollary shows that with uncertainty of supplier informedness, the larger π (the prior probability of "uninformed" suppliers), the more the buyer is willing to offer to the supplier ex post, to avoid possible destruction by the uninformed supplier.

Corollary 4 *$\overline{p}$ is an increasing function of π. $\lim_{\pi \to 1} \overline{p} = p^*$ and $\lim_{\pi \to 0} \overline{p} = 0 \leq p^*$.*

Proof. See Appendix. ■

Note, the probability for the supplier to be uninformed changes the structure of the game. The efficiency improvement caused by the probability of supplier uninformedness changing from 0 to 1 is non-monotonic. Initially, a minute probability of the supplier being uninformed will cause a sharp drop of the equilibrium offering
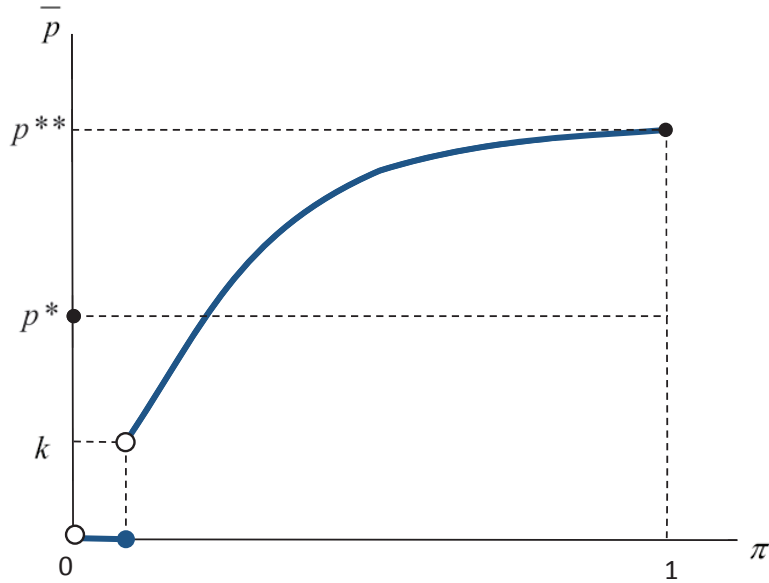
10

Figure 3: Equilibrium offering from the buyer

and therefore exacerbate the inefficiency of underinvestment, which can be alleviated thereafter only with the threat of ex post destruction by the uninformed supplier. These are illustrated in Figure 3.

Intuitively, according to the prior, the buyer believes that the supplier is uninformed with probability π and informed with probability $1 - \pi$. There is no way for the buyer to exploit the informed supplier as the supplier moves first. In contrast, it is possible for the buyer to exploit the uninformed supplier by cutting the offer, as long as the offer is greater than or equal to the ex ante investment cost. Due to this opportunistic behavior of the buyer, a slim chance for the supplier to be uninformed will cause a sharp drop of the equilibrium offering.

Therefore, the existence of the uninformed supplier is a double-edged sword. On the one hand, the option of ex post destruction enhances the likelihood of ex ante investment; on the other hand a slim chance of existence of uninformed Homo reciprocans supplier will cause a sharp drop of the equilibrium offering and subsequently exacerbate the inefficiency of underinvestment caused by the hold-up problem. To put it in another way, it is because of the existence of the uninformed Homo reciprocans supplier, that the option of ex post destruction becomes relevant to make the remedy.

Still, the underinvestment caused by the hold-up problem is only partially solved, as there is no investment ex ante if $\bar{p} < c$. Further, we could have the ex post inefficiency, beyond the ex ante inefficiency. For instance, consider the case that the buyer believes that the supplier is uninformed with probability π: the uninformed

supplier invests at date 0 and the buyer offers $\overline{p}$ at date 1. If the supplier's ex post realized threshold $T > \overline{p}$, the uninformed supplier will reject the offer and engage in some destructive behavior. This can be well justified by the existence of ex post destruction in the real world.

7   Concluding Remarks

We investigate whether or not and how effective ex post destruction can alleviate the hold-up problem in a one-shot game between a supplier and a buyer. Given the distinction between ex post destruction and non-trade, the option of ex post destruction cannot enhance ex ante efficiency if the supplier is a Homo economicus agent under complete information. However, if the buyer believes that the supplier might be a Homo reciprocans agent with sufficiently strong propensity for reciprocity, the problem can possibly be alleviated. Specifically, under complete information, only if the supplier reciprocal propensity is strong enough to cover the investment cost, will the supplier invest. In contrast, under incomplete information with informed supplier, investment is made feasible by the "mismatch" between the buyer's belief of stronger supplier reciprocal propensity and a de facto weaker one. Under incomplete information with uninformed supplier, the "mismatch" between the buyer's belief of weaker supplier reciprocal propensity and a stronger ex post one results in "destruction."

When it works, the degree of efficiency improvement is positively correlated with the intensity of potential ex post destruction under incomplete information with uninformed supplier. Further, in the case of incomplete information with uncertainty of supplier informedness, efficiency improvement caused by supplier uninformedness probability changing from 0 to 1 is non-monotonic. Initially, the minute probability of the supplier being uninformed will cause a sharp drop of the equilibrium offering and therefore exacerbate the inefficiency of underinvestment, which can be alleviated thereafter only with the threat of ex post destruction by the uninformed supplier. That is to say, it is because of the existence of the uninformed Homo reciprocans supplier, that the option of ex post destruction becomes relevant to make the remedy.

Appendix

Proof of Proposition 1

Consider the buyer's problem at date 1. Given the supplier invests at date 0, at date 1 if the buyer offers p, the supplier who is a Homo economicus agent or a Homo reciprocans agent with $T \leq p$ will accept the offer. As the supplier is informed, she will not invest at date 0 if she is a Homo reciprocans agent with $T > p$.

Suppose at date 0, a supplier who is a Homo economicus agent or a Homo recip-

rocans agent with $T \leq \widetilde{p}$ invests, otherwise not. Given this, the buyer's problem at date 1 is as follows.

$$\max_{p} \quad \frac{s(m-p) + (1-s)\int_0^p (m-p)f(T)\mathrm{m}T}{s + (1-s)F(\widetilde{p})}$$

Note, the buyer will never offer more than $\widetilde{p}$. Solving the optimization problem above, we have the first order condition (FOC) as follows.[11]

$$-s + (1-s)[-F(p) + (m-p)f(p)] = 0 \tag{1}$$

Denote the left hand side of FOC, $-s + (1-s)[-F(p) + (m-p)f(p)] = D(p)$. The cumulative distribution function (CDF) $F(p)$ is an increasing function, and the density $f(p)$ is a decreasing function as assumed. Clearly, $D(p)$ is decreasing in p. In addition, when $p = m$, $D(p) = -s - (1-s) = -1 < 0$.

If $D(k) \geq 0$, we can always find a $p \in [k, m)$ such that $D(p) = 0$. Further, differentiate $D(p)$ with respect to p and we have

$$D'(p) = (1-s)\left[-\frac{dF(p)}{dp} - f(p) + (m-p)\frac{df(p)}{dp}\right] = (1-s)\left[-2f(p) + (m-p)\frac{df(p)}{dp}\right] \leq 0$$

where $\frac{dF(p)}{dp} = f(p) \geq 0$ and $\frac{df(p)}{dp} \leq 0$ as assumed. Thus, the second order condition is satisfied.

If $D(k) < 0$, we end up with a corner solution, either 0 or k. If the buyer offers $p = 0$, the expected net benefit is $\frac{sy}{s+(1-s)F(\widetilde{p})}$, which is greater than the expected net benefit $\frac{s(m-k)}{s+(1-s)F(\widetilde{p})}$ if he offers $p = k$. Thus, if the prior probability that the supplier is Homo economicus is high, the buyer will offer 0.

To summarize, the optimal solution of the buyer's problem p^* is either solving the FOC above or 0, depending on the prior. Anticipating this, at date 0, the supplier who is a Homo economicus agent or a Homo reciprocans agent with $T \leq p^*$ invests, provided that $p^* \geq c$. Since the Homo reciprocans supplier will not invest ex ante if her threshold is greater than the anticipated ex post offer from the buyer at date 1, there is no destruction at all in equilibrium, even with the option of destruction ex post. Therefore, if $p^* \geq c$, we have a unique equilibrium, in which the supplier who is a Homo economicus agent or a Homo reciprocans agent with $T \leq p^*$ invests at date 0 and the buyer offers $p = p^*$ at date 1. If, instead, $p^* < c$, the anticipated ex post offer from the buyer at date 1 does not cover the cost of investment, the supplier will not invest at date 0. ■

Proof of Proposition 2

---

[11]Here, we ignore the positive constant terms.

Consider the buyer's problem at date 1. Given the supplier invests at date 0, at date 1 if the buyer offers p, the supplier who is a Homo economicus agent or a Homo reciprocans agent with $T \le p$ will accept the offer. The Homo reciprocans supplier with $T > p$ will reject the offer and engage in some destruction behavior.[12] The buyer's problem at date 1 is as follows.

$$\max_{p} \quad s(m-p) + (1-s)\left[\int_0^p (m-p)f(T)\mathrm{m}T + \int_p^m -Yf(T)\mathrm{m}T\right]$$

The first order condition is

$$-s + (1-s)[-F(p) + (m-p+Y)f(p)] = 0 \tag{2}$$

Similar to the proof of proposition 1, denote the left hand side of FOC, $-s + (1-s)[-F(p) + (m-p)f(p)] = \widetilde{D}(p)$. $F(p)$ is an increasing function, and the density $f(p)$ is a decreasing function as assumed. Clearly, $\widetilde{D}(p)$ is decreasing in p.

If $\widetilde{D}(k) \ge 0 \ge \widetilde{D}(m)$, we can always find a $p \in [k,m)$ such that $\widetilde{D}(p) = 0$. Further, differentiate $\widetilde{D}(p)$ with respect to p and we have

$$\widetilde{D}'(p) = (1-s)\left[-\frac{dF(p)}{dp} - f(p) + (m-p+Y)\frac{df(p)}{dp}\right] = (1-s)\left[-2f(p) + (m-p)\frac{df(p)}{dp}\right] \le 0$$

where $\frac{dF(p)}{dp} = f(p) \ge 0$ and $\frac{df(p)}{dp} \le 0$ as assumed. Thus, the second order condition is satisfied.

If $\widetilde{D}(k) < 0$, we end up with a corner solution, either 0 or k. If the buyer offers $p = 0$, the expected net benefit is $sy - (1-s)Y$, which is greater than the expected net benefit $s(m-k) - (1-s)Y$ if he offers $p = k$. Thus, if the prior probability that the supplier is Homo economicus is high, the buyer will offer 0. To the opposite, if $\widetilde{D}(m) > 0$, again we end up with a corner solution, m. There is no way for the buyer to offer more than m. Thus, if the prior belief of the supplier's propensity for reciprocity is sufficiently strong and the ex post destruction Y is intense, the buyer will offer m.

To summarize, the optimal solution of the buyer's problem p^* is either solving the FOC above, 0, or m, depending on the prior. Based on the prior belief, at date 1 the buyer offers p^{**}. Back to the beginning of the game, as long as $p^{**} \ge c$, the supplier has the incentive to invest ex ante. ∎

---

[12] Assume, once the buyer makes the take-it-or-leave-it offer, the buyer believes that the reciprocal propensity of the uninformed supplier will be realized and the uninformed supplier will make decision based on her true nature as Homo reciprocans, even though she is uninformed ex ante. This says that one will get to know one's true self in the crucial moment eventually.

Proof of Corollary 2

Consider the variation of the parameter Y. Take the derivative with respect to Y on both sides of FOC in proposition 2 (equation 2).

$$-\frac{dF(p)}{dp}\frac{dp}{dY} + \left(-\frac{dp}{dY} + 1\right)f(p) + (m - p + Y)\frac{df(p)}{dp}\frac{dp}{dY} = 0$$

where $\frac{dF(p)}{dp} = f(p)$. Rearrange and we have

$$\frac{dp}{dY} = \frac{f(p)}{2f(p) - (m - p + Y)\frac{df(p)}{dp}}$$

Since $f(p) \geq 0$ and $\frac{df(p)}{dp} \leq 0$ as assumed, we have $dp/dY \geq 0$. In addition, $\lim\limits_{Y \to 0} p^{**} = p^*$, as $\lim\limits_{Y \to 0} \widetilde{D}(p) = D(p)$. ∎

Proof of Proposition 3

Let's see the response of the buyer if the supplier invests at date 0. Given the supplier invests at date 0, at date 1 if the buyer offers p, the supplier who is a Homo economicus agent or a Homo reciprocans agent with $T \leq p$ will accept the offer. The Homo reciprocans supplier with $T > p$ will reject the offer and engage in some destructive behavior. The informed supplier will not invest at date 0 if $T > p$.

Suppose at date 0, an informed supplier who is a Homo economicus agent or a Homo reciprocans agent with $T \leq \widetilde{p}$ invests, otherwise not. Given this, if the buyer offers $p < \widetilde{p}$, the buyer's problem at date 1 is as follows.

$$\max_p \quad \frac{1}{\pi + (1-\pi)[s + (1-s)F(\widetilde{p})]}\left\{\pi\left[s(m - p) + (1-s)\left(\int_0^p (m - p)f(T)\mathrm{m}T + \int_p^m -Yf(T)\mathrm{m}T\right)\right] + \right.$$
$$\left. (1-\pi)\left[s(m - p) + (1-s)\left(\int_0^p (m - p)f(T)\mathrm{m}T + \int_p^{\widetilde{p}} -Yf(T)\mathrm{m}T\right)\right]\right\}$$

Solving the optimization problem above, we have the first order condition as follows, which is the same as equation 2 in proposition 2.[13]

$$-s + (1-s)[-F(p) + (m - p + Y)f(p)] = 0$$

From proposition 2, we have p^{**} solving the optimization problem above.

Instead, if the buyer offers $p \geq \widetilde{p}$, the buyer's problem at date 1 is as follows.

$$\max_p \quad \frac{\pi\left\{s(m - p) + (1-s)\left[\int_0^p (m - p)f(T)\mathrm{m}T + \int_p^m -Yf(T)\mathrm{m}T\right]\right\} + (1-\pi)(m - p)[s + (1-s)F(\widetilde{p})]}{\pi + (1-\pi)[s + (1-s)F(\widetilde{p})]}$$

[13]Here, we ignore the positive constant terms.

Solving the optimization problem above, we have the first order condition as follows.

$$\pi\left\{-s + (1-s)[-F(p) + (m - p + Y)f(p)]\right\} - (1-\pi)[s + (1-s)F(\widetilde{p})] = 0$$

Similar to the proof of proposition 1 and 2, there exists a $\widehat{p}$ solving the optimization problem above. Back to date 0, anticipating the buyer's offer $\widehat{p}$, the informed supplier who is a Homo economicus agent or a Homo reciprocans agent with $T \leq \widehat{p}$ will invest, provided that $\widehat{p} \geq c$. The informed Homo reciprocans supplier with $T > \widehat{p}$ will not invest.

We reach an equilibrium till to the point that there exists a $\overline{p}$ solving the following equation.

$$-s + (1-s)[-F(p)] + \pi(1-s)(m - p + Y)f(p) = 0 \qquad (3)$$

Similar to the proof of proposition 1 and 2, denote the left hand side of FOC, $-s + (1-s)[-F(p)] + \pi(1-s)(m-p+Y)f(p) = \overline{D}(p)$. $F(p)$ is an increasing function, and the density $f(p)$ is a decreasing function as assumed. Clearly, $\overline{D}(p)$ is decreasing in p.

If $\overline{D}(k) \geq 0 \geq \overline{D}(m)$, we can always find a $p \in [k, m)$ such that $\overline{D}(p) = 0$. Further, differentiate $\overline{D}(p)$ with respect to p and we have

$$\overline{D}'(p) = (1-s)\left[-\frac{dF(p)}{dp}\right] + \pi(1-s)\left[-f(p) + (m - p + Y)\frac{df(p)}{dp}\right]$$

$$= (1-s)\left[-(1+\pi)f(p) + \pi(m - p + Y)\frac{df(p)}{dp}\right] \leq 0$$

where $\frac{dF(p)}{dp} = f(p) \geq 0$ and $\frac{df(p)}{dp} \leq 0$ as assumed. Thus, the second order condition is satisfied.

If $\overline{D}(k) < 0$, we end up with a corner solution, either 0 or k. If the buyer offers $p = 0$, the expected net benefit is $\frac{\pi\{sy - (1-s)Y\} + (1-\pi)m[s + (1-s)F(\widetilde{p})]}{\pi + (1-\pi)[s + (1-s)F(\widetilde{p})]}$, which is greater than the expected net benefit $\frac{\pi\{s(m-k) - (1-s)Y\} + (1-\pi)(m-k)[s + (1-s)F(\widetilde{p})]}{\pi + (1-\pi)[s + (1-s)F(\widetilde{p})]}$ if he offers $p = k$. Thus, if the prior probability that the supplier is Homo economicus is large, the buyer will offer 0. To the opposite, if $\overline{D}(m) > 0$, again we end up with a corner solution, m. There is no way for the buyer to offer more than m. Thus, if the prior belief of the supplier's propensity for reciprocity is sufficiently strong and the ex post destruction Y is intense, the buyer will offer m.

To summarize, the optimal solution of the buyer's problem $\overline{p}$ is either solving equation 3, 0, or m, depending on the prior. Based on the prior belief, at date 1 the buyer offers $\overline{p}$. Anticipating these, at date 0, the informed supplier who is a Homo economicus agent or a Homo reciprocans agent with $T \leq \overline{p}$ and uninformed supplier invest, provided that $\overline{p} \geq c$. ∎

Proof of Corollary 3

Consider the variation of the parameters Y. Take the derivative with respect to Y on both sides of equation 3.

$$-(1-s)\frac{dF(p)}{dp}\frac{\partial p}{\partial Y} + \pi(1-s)\left[\left(-\frac{\partial p}{\partial Y}+1\right)f(p) + (m-p+Y)\frac{df(p)}{dp}\frac{\partial p}{\partial Y}\right] = 0$$

where $\frac{dF(p)}{dp} = f(p)$. Rearrange and we have

$$\frac{\partial p}{\partial Y} = \frac{\pi f(p)}{(1+\pi)f(p) - (m-p+Y)\frac{df(p)}{dp}}$$

Since $f(p) \geq 0$ and $\frac{df(p)}{dp} \leq 0$ as assumed, we have $\frac{\partial p}{\partial Y} \geq 0$. ∎

Proof of Corollary 4

Consider the variation of the parameters π. Take the derivative with respect to π on both sides of equation 3.

$$-(1-s)\frac{dF(p)}{dp}\frac{\partial p}{\partial \pi} + (1-s)(m-p+Y)f(p) + \pi(1-s)\left[-\frac{\partial p}{\partial \pi}f(p) + (m-p+Y)\frac{df(p)}{dp}\frac{\partial p}{\partial \pi}\right] = 0$$

where $\frac{dF(p)}{dp} = f(p)$. Rearrange and we have

$$\frac{\partial p}{\partial \pi} = \frac{(m-p+Y)f(p)}{2f(p) - \pi(m-p+Y)\frac{df(p)}{dp}}$$

Since $f(p) \geq 0$ and $\frac{df(p)}{dp} \leq 0$ as assumed, we have $\frac{\partial p}{\partial \pi} \geq 0$. In addition, $\lim\limits_{\pi \to 1} \overline{p} = p^{**}$, as $\lim\limits_{\pi \to 1} \overline{D}(p) = \widetilde{D}(p)$.

Consider the case $\pi \to 0$. Back to the proof of proposition 3, we have the first order condition as follows (equation 3).

$$-s + (1-s)[-F(p)] + \pi(1-s)(m-p+Y)f(p) = 0$$

As $\pi \to 0$, the left hand side goes to $-s + (1-s)[-F(p)] \leq -s + (1-s)[-F(p) + (m-p)f(p)] = D(p)$. It follows that $\lim\limits_{\pi \to 0} \overline{p} \leq p^*$. Further, $-s + (1-s)[-F(p)] < 0$, which implies $\lim\limits_{\pi \to 0} \overline{p} = 0$. ∎

References

[1] Che, Yeon-Koo, and József Sákovics. 2008. "Hold-up Problem." *The New Palgrave Dictionary of Economics*, 2nd edition.

[2] Dufwenberg, Martin, and Georg Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior*, 47: 268-298.

[3] Dufwenberg, Martin, Alec Smith, and Matt Van Essen. 2013. "Hold-up: With a Vengeance." *Economic Inquiry*, 51(1): 896-908.

[4] Ellingsen, Tore, and Magnus Johannesson. 2004. "Promises, Threats and Fairness." *Economic Journal*, 114: 397-420.

[5] Fehr, Ernst, and Klaus Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 114: 817-868.

[6] Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review*, 90(4): 980-994.

[7] Ng, Travis. 2011. "Destructing the Hold-up." *Economics Letters*, 111(3): 247-248.

[8] Sobel, Joel. 2005. "Interdependent Preferences and Reciprocity." *Journal of Economic Literature*, 43(2): 392-436.

[9] von Siemens, Ferdinand A. 2009. "Bargaining under Incomplete Information, Fairness, and the Hold-Up Problem." *Journal of Economic Behavior and Organization*, 71: 486-494.