A survey of the hold-up problem in the experimental economics literature

Yadi Yang

November 2, 2020

Abstract

This paper reviews experimental studies on the so-called “hold-up problem”. Common features in the experimental design and results are summarized. Most experimental studies show evidence of the hold-up problem, but to an extent less severe than what standard self-interest model predicts. Hold-up occurs at the individual level, but exhibits a less severe pattern than theoretically predicted at the aggregate level. A positive correlation is found between the investment stage decisions and subsequent bargaining behavior. Social preferences largely influence the results in hold-up games. Remedies that enhance the effect of social preferences can effectively alleviate the hold-up problem. These laboratory findings are also relevant to real-life hold-up situations, but differences in the specific institution and environment may require more variations in the experimental design.

JEL Codes: L14, C78
Key words: Hold-up problem, bargaining, social preferences
1 Introduction

When multiple parties make non-recoverable relationship-specific investments that generate a joint surplus to be divided through ex-post bargaining, underinvestment may occur. Since the final allocation is determined by the interplay of ex-post bargaining power of all participating parties, each agent is unlikely to fully appropriate the return from his investment. Therefore, agents refrain from investing at the efficient level for fear of being held-up by their counterparts. This underinvestment is referred to as the “hold-up problem” in the economic literature (Che and Sákovics, 2008). It is a common phenomenon in bilateral transactions that rely on incomplete contracts. Common examples of the hold-up problem can be found in procurement contracts where the manufacturer needs to make product-specific investment beforehand, employment contracts where the employee needs to invest in firm-specific skills, etc. In these cases, the specific investments are non-contractable, and cannot be appropriated by suppliers or employees if being held-up. As a result, manufacturers and employees may make insufficient investments in the investment stage.

Early theoretical analyses of the hold-up problem can be found in Williamson (1971) and Klein et al. (1978). (Klein, 1998) summarizes three main features of the hold-up problem as “specific investments”, “incomplete contracts”, and “renegotiation”. Grout (1984) develops a model for the hold-up problem in an intra-firm employment contract setting,\(^1\) while Tirole (1986) develops a similar model in inter-firm transactions.\(^2\) A simplified version of their two-stage model with one investor (seller) and one proposer (buyer) is summarized by Che and Sákovics (2008). A buyer and a seller decide upon a contract to trade a positive quantity \(q\) at price \(t\). Before the trade takes place, the seller makes a binary investment decision with fixed and non-recoverable costs to increase the joint surplus. The investment choice is observable but not verifiable, thus cannot be contracted upon. The price and quantity to be traded in the second stage are the only contractable variables in this case. The socially optimal solution that maximizes the overall payoff yields the seller to invest. However, Nash bargaining solution at the negotiation stage yields an equal split of the gross surplus between the seller and the buyer. When the seller decides whether to invest in the first stage, he anticipates that he bears the whole investment cost but receives only half of the investment return. There are conditions under which a rational seller refrains from investing in the first stage, despite it being socially optimal. A more general two-agent model is analysed by Grossman and Hart (1986). In the first stage, each agent independently makes an investment decision that contributes to a joint-surplus. Agent \(i\) invests \(I_i\) with per-unit cost \(c\). Investments \(I_1\) and \(I_2\) together generate
a joint surplus $R(I_1, I_2)$, with $\partial R(I_1, I_2)/\partial I_i \geq 0$, and $\partial^2 R(I_1, I_2)/\partial I_i^2 \leq 0$. In the second stage, the two agents negotiate over the division of the surplus. The (Nash) cooperative game solution gives a net payoff of $\frac{1}{2} R(I_1, I_2) - c I_i$ to agent $i$. Using backwards induction, anticipating the bargaining result, agent $i$ chooses the first stage investment level $I_i$ to maximize the net payoff, resulting in the selection of $\hat{I}_i$ that satisfies the first order condition $\frac{1}{2} \partial R/\partial \hat{I}_i = c$. However, the first-best investment levels that maximize the total payoff $R(I_1, I_2) - c I_1 - c I_2$ are given by $(I_1^*, I_2^*)$ that satisfies the first order condition $\partial R/\partial I_i^* = c$. With the assumptions about the first and second order derivatives of $R(\cdot)$, it can be shown that $\hat{I}_i \leq I_i^*$. In the absence of a contract on ex-ante investments, undesirable underinvestment occurs.

In recent years, conducting laboratory experiments has become a popular data collection method among economists. It allows researchers to implement the institutional environment that aligns best with theory, to insert strict control that reduces various confounding factors, and to create counterfactuals that establish causality. These features are usually difficult to obtain with field data. The experimental method has been frequently applied in examining people’s behaviors in strategic interactions. The typical setting of an experiment on the hold-up problem takes the form of a “nested bargaining game”, as phrased by Sonnemans et al. (2001). Two subjects are paired up to participate in a two-stage game. In the first stage, they make non-contractable investments with fixed and non-recoverable costs. In the second stage, they bargain over the surplus generated from previous investment decisions. The detailed set-up varies across different experiments. Investment may come from both subjects or only one subject, and could be either a continuous decision where the subject selects a level to invest or a binary decision where the subject selects whether to invest or not. The bargaining mechanism also differs across studies, with ultimatum game, dictator game, and alternating-offer Rubinstein (1982) bargaining game as the most common forms.

Being a bargaining game with a preceding investment stage, the hold-up problem refers to a very specific scenario. As a consequence, the number of experimental studies that specifically focus on “hold-up” games is limited. The rest of this paper reviews experimental papers on the hold-up problem. By comparing their experimental settings, the typical results, and manipulations they employ to solve the hold-up problem, it is found that social preferences play an important role in mitigating the hold-up problem, the effectiveness of which is found to largely depend on various aspects of the game structure.

To avoid inefficiency, economists have been looking into ways of diminishing the prevalence of hold-up and restore investment incentives. Miller (2011) provides a comprehensive summary of remedies for the hold-up problem. Conventional reme-
dies (classified as “formal controls” by Miller (2011)) can be categorized into two major types: organizational remedies and contractual remedies (Che and Sákovics, 2008). Laboratory experiment provides a convenient testbed to examine the treatment effect of a certain policy under strictly controlled conditions. Experiments testing the two types of strategic remedies are limited in numbers. Joint ownership and option contracts are found to effectively mitigate the hold-up problem. In addition to these conventional remedies, behavioral remedies (classified as “informal controls” by Miller (2011)) such as observable investment, costly punishment, communication, and the provision of social history are found to have a significant effect through the channel of social preferences.

The rest of the paper is organized as follows. The next section summarizes the experimental studies on the hold-up problem. Experiments are classified based on their experimental design. Section 3 reviews the different remedies for the hold-up problem, which includes conventional remedies as well as behavioral remedies. The main findings and results are summarized in Section 4. Section 5 discusses the generalisability of laboratory results and implications for further research.

2 Experiments about hold-up

Experimental studies on the hold-up problem differ in various aspects of their design. Nevertheless, they share common features in the basic set-up. A prototypical hold-up experiment consists of two stages: a production stage requiring sunk investments and a subsequent bargaining stage to divide the surplus earned from the joint-production. Following the theoretical model by Grossman and Hart (1986) and Tirole (1986), most experiments pair up two endowed subjects who each decide individually how much of their endowment to be invested in a joint production. In some experiments, the two subjects are explicitly framed as “buyers” and “sellers” (Hackett, 1993, 1994), representing the original incomplete contract problem. Some other experimenters frame the roles as “investor” and “trading partner” (Ellingsen and Johannesson, 2004a,b). In most other studies, a neutral framing that labels different roles by different letters is employed. Early experiments allow both parties to make the investment decisions. In more recent experiments, a more common setting is to allow only one investor. Subjects in all studies are undergraduate students with business and economics majors. A comparison of the general features of experimental design is shown in Table 1.
<table>
<thead>
<tr>
<th>Study</th>
<th>Periods</th>
<th>Matching</th>
<th>Framing</th>
<th>Fixed role</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hackett (1993)</td>
<td>10</td>
<td>Stranger</td>
<td>Buyer-Seller</td>
<td>No</td>
<td>Relative investment incentives</td>
</tr>
<tr>
<td>Hackett (1994)</td>
<td>10</td>
<td>Stranger</td>
<td>Buyer-Seller</td>
<td>No</td>
<td>Discount rate, investment observability</td>
</tr>
<tr>
<td>Berg et al. (1995)</td>
<td>1</td>
<td>One-shot</td>
<td>Neutral</td>
<td>-</td>
<td>Social history</td>
</tr>
<tr>
<td>Gantner et al. (2001)</td>
<td>2</td>
<td>Random</td>
<td>Neutral</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Sonnemans et al. (2001)</td>
<td>18</td>
<td>Block structure</td>
<td>Neutral</td>
<td>Within blocks</td>
<td>Level of outside option</td>
</tr>
<tr>
<td>Oosterbeek et al. (2003)</td>
<td>10</td>
<td>Block structure</td>
<td>Neutral</td>
<td>Within blocks</td>
<td>Level of outside option</td>
</tr>
<tr>
<td>Ellingsen and Johannesson (2004a)</td>
<td>1</td>
<td>One-shot</td>
<td>Investor-Partner</td>
<td>-</td>
<td>Communication (Promise and threat)</td>
</tr>
<tr>
<td>Ellingsen and Johannesson (2004b)</td>
<td>1</td>
<td>One-shot</td>
<td>Investor-Partner</td>
<td>-</td>
<td>Communication (Promise and threat)</td>
</tr>
<tr>
<td>Sloof et al. (2004)</td>
<td>20</td>
<td>Block structure</td>
<td>Neutral</td>
<td>Within blocks</td>
<td>Level of outside option</td>
</tr>
<tr>
<td>Charness and Dufwenberg (2006)</td>
<td>1</td>
<td>One-shot</td>
<td>Neutral</td>
<td>-</td>
<td>Communication (Promise)</td>
</tr>
<tr>
<td>Sloof et al. (2007)</td>
<td>36</td>
<td>Block structure</td>
<td>Neutral</td>
<td>Within blocks</td>
<td>Investment costs and observability</td>
</tr>
<tr>
<td>Fehr et al. (2008)</td>
<td>10</td>
<td>Random</td>
<td>Neutral</td>
<td>Yes</td>
<td>Ownership structure</td>
</tr>
<tr>
<td>Vanberg (2008)</td>
<td>1</td>
<td>One-shot</td>
<td>Neutral</td>
<td>-</td>
<td>Communication (Promise)</td>
</tr>
<tr>
<td>Charness and Dufwenberg (2010)</td>
<td>1</td>
<td>One-shot</td>
<td>Neutral</td>
<td>-</td>
<td>Communication (Promise)</td>
</tr>
<tr>
<td>Charness et al. (2011)</td>
<td>36</td>
<td>Random</td>
<td>Neutral</td>
<td>No</td>
<td>History information</td>
</tr>
<tr>
<td>Hoppe and Schnitz (2011)</td>
<td>1</td>
<td>One-shot</td>
<td>Neutral</td>
<td>-</td>
<td>Contractual changes</td>
</tr>
<tr>
<td>Huck et al. (2012)</td>
<td>30</td>
<td>Random/Endogenous</td>
<td>Neutral</td>
<td>Yes</td>
<td>History, competition</td>
</tr>
<tr>
<td>Dufwenberg et al. (2013)</td>
<td>5</td>
<td>Random</td>
<td>Neutral</td>
<td>Yes</td>
<td>Level of outside option</td>
</tr>
<tr>
<td>Morita and Servátka (2013)</td>
<td>1</td>
<td>One-shot</td>
<td>Neutral</td>
<td>-</td>
<td>Group identity</td>
</tr>
<tr>
<td>Eisenkopf and Nüschi (2016)</td>
<td>10</td>
<td>Stranger</td>
<td>Neutral</td>
<td>Yes</td>
<td>Third party</td>
</tr>
<tr>
<td>Ismayilov and Potters (2016)</td>
<td>1</td>
<td>One-shot</td>
<td>Neutral</td>
<td>-</td>
<td>Communication (Promise)</td>
</tr>
<tr>
<td>Eisenkopf and Nüschi (2017)</td>
<td>10</td>
<td>Stranger</td>
<td>Neutral</td>
<td>Yes</td>
<td>Third party</td>
</tr>
<tr>
<td>Davis and Leider (2018)</td>
<td>10</td>
<td>Random</td>
<td>Supplier-Retailer</td>
<td>Yes</td>
<td>Contractual changes</td>
</tr>
<tr>
<td>Morita and Servátka (2018)</td>
<td>1</td>
<td>One-shot</td>
<td>Neutral</td>
<td>-</td>
<td>Outside option</td>
</tr>
<tr>
<td>Haruvy et al. (2019)</td>
<td>100</td>
<td>Random</td>
<td>Neutral</td>
<td>Yes</td>
<td>Uncertainty, reputation</td>
</tr>
<tr>
<td>Zheng et al. (2020)</td>
<td>1</td>
<td>One-shot</td>
<td>Neutral</td>
<td>-</td>
<td>Reciprocity, veto power</td>
</tr>
</tbody>
</table>
2.1 Comparison of experimental settings

Despite variation in details, an experiment on the hold-up problem essentially consists of an investment stage and a subsequent bargaining stage. For each of the two stages, experiments differ in whether both subjects or only the non-investor is allowed to take actions and whether the decisions are dichotomous or continuous. In the bargaining stage, the specific bargaining mechanism employed in each study also differs. A summary of the hold-up experiments reviewed in this paper by their setup is shown in Table 2.

The earliest experiment on the hold-up problem dates back to Hackett (1993, 1994), the design of which closely resembles the model by Grossman and Hart (1986). Hackett conducts ten periods of a two-sided nested bargaining game, using stranger matching, so that each subject will not encounter another subject that he had interacted with twice, to avoid reputation effects. The roles of buyer or seller are randomly assigned to subjects at the beginning of every period. Subjects simultaneously select a value that induces an unrecoverable quadratic cost. The values chosen by both players generate a joint surplus with production cost, the value of which is known to both subjects. Both the surplus and the production cost can be either high or low. The value selected by the buyer corresponds to the probability of a high surplus, and the value selected by the seller corresponds to the probability of low production cost. The second-stage bargaining is a modified version of the Rubinstein (1982) procedure. In his implementation, Hackett fixed the maximum number of allowed bargaining rounds without informing the subjects, following the setting of Binmore et al. (1991), so that the subjects only know that there is a positive probability of a forced breakdown after each bargaining round. Before the bargaining stage, subjects each can decide whether to enter the bargaining stage or not. One or both subjects opting out terminates the game, and allocates all the surplus to the buyer and all the costs to the seller. Bargaining only proceeds when both parties agree to bargain. Subjects alternate in making offers, with the buyer proposing first in all odd-numbered periods and the seller proposing first in all even-numbered periods. In each bargaining round, the responder chooses whether to accept or to reject and come up with a counter-offer in the next bargaining round. The disagreement payoff is the same as the opting out results. After the allocation of the surplus is finalized, each subject has an independent option to veto the results. Selection of veto by at least one subject eradicates all surplus and production cost, leaving both subjects zero payoffs.

In retrospect, there are many unique features in the design of Hackett (1993, 1994). Both the production function and the sunk cost function are non-linear, while most later studies use either a linear function or a binary choice for these
Table 2: Experimental design of the hold-up problem (II)

<table>
<thead>
<tr>
<th>Investment Stage</th>
<th>Both invest</th>
<th>Only one invests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous</td>
<td>Continuous</td>
<td>Binary</td>
</tr>
<tr>
<td>Both propose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-shot bargaining</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nash demand</td>
<td>Gantner et al. (2001)</td>
<td>Ellingsen and Johannesson (2004a)</td>
</tr>
<tr>
<td>Rubinstein bargaining</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic with feedback</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| One subject proposes |             |                  |
| Dictator game        |             |                  |
| Binary Offer         |             |                  |
|                      | Eisenkopf and Nüesch (2016, 2017) |                  |
| Ultimatum game       |             |                  |
| Binary Offer         |             |                  |
two decisions. Roles are reassigned at the beginning of every period, indicating that it is possible for the same subject to play both roles, while most later studies have fixed roles for the subjects for the whole session. Different from one-round simple bargaining mechanisms as employed in many later experiments, Rubinstein alternating offer procedure is implemented in Hackett (1993, 1994), allowing him to gather more information on the bargaining behavior and examine the effect of discount rate on the results by manipulating the maximum allowed bargaining rounds. In his “high discount rate” treatment, the maximum number of allowed bargaining round for some periods is fixed to be one, which is equivalent to a one-round ultimatum game as in many other studies.

Apart from the above three aspects, the most distinctive feature of Hackett (1993, 1994) is the inclusion of a veto option. The hold-up behavior on the buyer’s side could take two different forms: a disadvantageous allocation in the bargaining or choosing to opt-out before the bargaining stage starts. The veto option also increases the bargaining power of the seller.

A later experiment by Gantner et al. (2001) has a more standard experimental setting. Similar to Hackett (1993, 1994), two randomly matched subjects bargain over a joint surplus generated from their preceding investments. The joint production function is a linear combination of investments from both subjects. Subjects paired together differ in marginal productivity. This manipulation alters the bargaining balance between the two partners and makes room for different equity standards. The bargaining stage takes two forms: an ultimatum game and a Nash demand game. Each subject participates in both games with different partners. The two possible orders of combination are both included to exclude order effect. In the bargaining stage, subjects simultaneously select a demand and a lowest acceptance bound. The roles of a proposer and a responder are assigned to each subject after the values are chosen. In the ultimatum game setting, the demand of the proposer is matched with the lowest acceptance level of the responder. An agreement is reached if and only if the sum of the two values is less than the total surplus, resulting in an allocation that gives the demanded amount to the proposer and the residual to the responder. In the Nash demand game, the demand of both subjects is matched first. If the sum does not exceed the total gain, an agreement is reached and subjects receive their respective demand, with the efficiency loss of any unclaimed amount. Otherwise, their least acceptance bounds are matched and the surplus is distributed by the same method in the case of an agreement. Under both schemes, the disagreement payoff is zero for both subjects.

In the two experiments discussed above, both subjects invest and have the chance to propose an allocation, therefore, they both have an incentive to hold-up the partner and the possibility to become a potential victim to hold-up. In a num-
ber of more recent experiments, only one party makes the investment decision, while the other party proposes an allocation. Under this design, the incentive to hold up his counterpart only falls on the non-investor. This design fits the model by Hart (1995). Under this design, it is easier to distinguish the motives and behaviors of the two parties. A number of experiments with such unilateral investment stage still adopt bargaining mechanisms which assign symmetric bargaining power to both parties, e.g., the one-shot Nash demand game (Ellingsen and Johannesson, 2004a) and the multi-round alternating offer game (Sonnemans et al., 2001; Oosterbeek et al., 2003; Sloof et al., 2004; Davis and Leider, 2018). The bargaining procedure of Sonnemans et al. (2001), Oosterbeek et al. (2003), and Sloof et al. (2004) is similar to the Rubinstein procedure as applied by Hackett (1993, 1994), except that here, the maximum number of allowed rounds is fixed to ten and is known to all subjects. In each bargaining round, the responder decides whether to accept the offer and end the bargaining, or to reject with a counteroffer in the following round (if applicable) at the cost of receiving the disagreement payoff. In some treatments of Sonnemans et al. (2001) and Sloof et al. (2004), subjects also have the option to opt-out and end the bargaining stage in each round. The outside option of Sonnemans et al. (2001) leaves a positive amount to the investor and zero to the non-investor, while the contrary is the case for Sloof et al. (2004). Sonnemans et al. (2001) and Sloof et al. (2004) employ a special “block” structure in their experiments. They divide their multiple periods into blocks: Sonnemans et al. (2001) had two blocks of nine periods and Sloof et al. (2007) had four blocks of five periods. This “block” structure allows the experimenters to adopt perfect stranger matching within each block to avoid reputation effect and gives them room for testing within-subject treatments. Oosterbeek et al. (2003) adopt perfect stranger matching in all ten of their experimental periods. However, subjects participated as both roles: they were assigned one role for the first five periods and the other role for the remaining five periods. Davis and Leider (2018) adopt a unique dynamic bargaining mechanism, where both parties are allowed to make offers and provide limited feedback with few restrictions on the order and the number of offers can be made by either player.

The ultimatum game, where one player proposes an allocation and the other player decides whether to accept or to reject, can be regarded as an extreme case of the Rubinstein alternating offer bargain with only one round of offer. It is common in most recent hold-up experiments (Ellingsen and Johannesson, 2004b; Sloof et al., 2007; Hoppe and Schmitz, 2011; Morita and Servátka, 2013, 2018; Haruvy et al., 2019; Zheng et al., 2020) to adopt a unilateral investment with a binary investment choice followed by an ultimatum bargaining stage where the non-investor proposes an allocation and the investor decides whether to accept or
to reject. Rejection leads to zero payoffs to both parties. Dufwenberg et al. (2013) use binary choices in the bargaining stage as well: instead of having to come up with his own allocation scheme, the non-investor chooses between an equal split and an option to “hold up” the investor by exploiting his payoff to almost zero. The investor then chooses between two allocations which are equivalent to accepting an unfair offer and punishing the unfair non-investor at his own cost in a standard ultimatum game.

Some experimental studies reduce the bargaining power of the investor by adopting a dictator game in the bargaining stage, so that the investor no longer has the possibility to reject an offer made by the non-investor. Such hold-up games with a dictator bargaining stage following the investment stage can be regarded as the trust game. The typical trust game was designed and conducted by Berg et al. (1995). The investor first decides a proportion of his show-up fee to be transferred to the non-investor, the amount of which will then be tripled by the experimenter. The non-investor then decides how much of the tripled amount to be sent back to the investor. Since the “investment stage” is explicitly framed as sending money to the partner, there is a stronger focus on willingness to trust and reciprocity in trust games. In addition, the incentive to hold up is stronger in this case, since the non-investor can propose any allocation without the fear of being rejected. Despite these differences, the trust game still constructs a hold-up situation. This experiment has been replicated by a large number of researchers as summarized in the meta study of Johnson and Mislin (2011). These subsequent studies closely follow the basic design of the Berg et al. (1995) experiment, with variations in minor setting details such as rate of return, location, subjects’ demographic heterogeneity, etc. It is beyond the scope of this paper to provide an exhaustive overview of all trust game experiments. A number of typical studies that explore different mechanisms to increase the level of investment and transfer (Berg et al., 1995; Sloof et al., 2007; Charness and Dufwenberg, 2006, 2010; Charness et al., 2011; Huck et al., 2012; Eisenkopf and Nüesch, 2016; Ismayilov and Potters, 2016; Eisenkopf and Nüesch, 2017) are selected to compare different potential remedies for the hold-up problem.

2.2 Stylized findings

The hold-up experiments discussed above adopt different experimental settings. Some experimenters also introduce variations to test for a solution. Therefore, it is difficult to quantitatively compare the results. However, a qualitative comparison of the simple hold-up game in the baseline treatment of these experiments yield some general results that hold across different studies.
2.2.1 Evidence of hold-up

The surveyed studies all find evidence showing that the hold-up problem does exist. The most straightforward representation of hold-up lies in proposals made in the bargaining stage. In experiments where only one subject is allowed to propose an allocation, cases of exploitation are found in a number of studies. Dufwenberg et al. (2013) find that more than half of non-investors choose to exploit the investors by choosing the allocation that leaves only a minimum amount to the investor. Ellingsen and Johannesson (2004b), Berg et al. (1995), and Sloof et al. (2007) also find quite a few cases where the non-investor extracts all payoffs from the investor. Fehr and List (2004) find that on average non-investors only offer a small proportion of the total surplus to the investors. In experiments with multi-round bargaining, Oosterbeek et al. (2003) find that a majority of the first proposals by the non-investor leaves a less than the equal-split of the surplus to the investor, which indicates that non-investors take the chance to hold up the sunk investment of the investors. In experiments where both parties can propose an allocation, evidence of hold-up is also found on both sides. Ellingsen and Johannesson (2004a) find a number of cases where the claims from both parties add up to more than the total surplus to be divided. Gantner et al. (2001) also find a few allocations that fit the game-theoretic predictions.

For fear of being held-up by their partners, underinvestment on the investors’ side occurs accordingly. In the studies reviewed, different degrees of underinvestment is found on the investors’ side. Hackett (1993, 1994) finds that a substantial amount of subjects choose investment amounts below the efficient level. Gantner et al. (2001) also discover a few occurrences of low investment or even zero investment. Sonnemans et al. (2001) and Sloof et al. (2004) find that the average investment levels in all treatments are below the socially efficient level. In Sonnemans et al. (2001), the majority of individual investments fall below the efficient level. In experiments with binary investment choices, a substantial amount of investors refrain from investing in the first stage of the game (Dufwenberg et al., 2013; Oosterbeek et al., 2003; Ellingsen and Johannesson, 2004a,b; Sloof et al., 2007). In most trust games following Berg et al. (1995), quite a number of investors invest below the efficient level.

2.2.2 Discrepancy with standard self-interest predictions

In spite of the individual occurrences found in a number of experiments, on the aggregate level, the hold-up problem is found to be less of a concern than in theoretical models under standard self-interest assumptions. Underinvestment, although present, is found less severe than what standard self-interest theory predicts. Hack-
ett (1993, 1994) finds that the average investment levels in the investing stage by both parties lie between the self-interest predictions and the socially optimal level, with a substantial portion of cases above the self-interest predictions. Gantner et al. (2001) find that the efficient investment level is selected in most cases. In experiments with binary investment choices (Oosterbeek et al., 2003; Ellingsen and Johannesson, 2004a,b; Sloof et al., 2007; Morita and Servátka, 2018; Zheng et al., 2020), underinvestment occurs less frequently than the self-interest predictions. Berg et al. (1995) and Fehr and List (2004) also find that a large proportion of investors trust their partners with the majority of their endowment.

In the bargaining stage, evidence also shows that exploitation of the investors is less severe than the standard self-interest prediction. In experiments where only one party is allowed to make an offer in the bargaining stage, namely the experiments using an ultimatum game (Gantner et al., 2001; Ellingsen and Johannesson, 2004b; Hoppe and Schmitz, 2011; Morita and Servátka, 2013, 2018; Haruvy et al., 2019; Zheng et al., 2020) or a dictator game (Berg et al., 1995; Fehr and List, 2004; Charness and Dufwenberg, 2006; Sloof et al., 2007; Charness and Dufwenberg, 2010; Charness et al., 2011; Huck et al., 2012; Eisenkopf and Nüesch, 2016; Ismayilov and Potters, 2016; Eisenkopf and Nüesch, 2017) as the bargaining mechanism, the proposer offers a positive amount to the partner under most circumstances. In the experiment by Ellingsen and Johannesson (2004b), the most common offer by the proposer is one that allocates equal net profit to both parties, chosen by almost 50 percent of the proposers. Consistent with the common robust results in pure ultimatum game or pure dictator game (Fehr and Schmidt, 2006), the majority of bargaining outcomes in these experiments exhibit a deviation from the standard self-interest prediction of full exploitation. In addition, similar results are also found in experiments with repeated alternate offers. Hackett (1993) finds an equal split of the total surplus to be the most frequent among all successfully negotiated final allocations, with other allocations clustered around it. Somemans et al. (2001) also find the finally agreed allocations to be different from the game-theoretic predictions with standard self-interest assumptions. Instead of an equal split of the total surplus, they find the average value of the final allocations closer to the “split-the-difference” result, which is defined as “both players receiving their no-trade pay-offs plus 50% of the remaining surplus”.

A third discrepancy is found in the link between investment behavior and bargaining results. Various experimental results exhibit a close positive correlation between the investment level and the allocation results (Hackett, 1993, 1994; Gantner et al., 2001; Oosterbeek et al., 2003; Ellingsen and Johannesson, 2004a). Hackett (1993) finds that the difference in the sunk investment cost has a significant and positive effect on the bargaining outcomes. The bargaining outcome suggests
the existence of an equity rule where the party investing the greater share receives the larger proportion of the surplus. Hackett (1994) discovers strong support for subjects adjusting their investment decisions after anticipating a linkage between observable investment levels and resulting allocations. In experiments with a Nash demand bargaining stage, the Nash cooperative bargaining solution predicts the unique equilibrium outcome of an equal split of the surplus, regardless of what happens in the investment stage. However, Gantner et al. (2001) find the share of surplus that subjects claim in the bargaining stage to be positively correlated with their share of input in the investment stage. The results of Ellingsen and Johannesson (2004a) also indicate that subjects take into consideration the sunk cost of the investors during the bargaining stage.

The discrepancy with standard self-interest predictions in many of the experiments above indicates that subjects’ decisions in hold-up games are influenced by more than just pure strategic concerns. Various theories taking into account social preferences of the subjects can offer a better explanation for the observations. Researchers fit their experimental results with predictions from standard models and social preference models. They typically find that various selected social preference models better explain the results (Berg et al., 1995; Gantner et al., 2001; Fehr and List, 2004; Sloof et al., 2004; Dufwenberg et al., 2013; Morita and Servátka, 2018). Some of the patterns discovered in the bargaining stage of hold-up experiments are similar to findings in respective simple bargaining games. In studies of simple ultimatum games, there is a common trend that responders reject strictly positive offers, and offers made by the proposers are clustered around the equal split and skewed to the left (Fehr and Schmidt, 1999), which is against standard theoretical predictions based on material payoffs. Similar results are found in hold-up experiments with an ultimatum bargaining stage. Ellingsen and Johannesson (2004b) and Sloof et al. (2007) both find that subjects reject unfair positive material offers. On the other hand, some findings are particular to hold-up experiments due to the distinct settings. For instance, Ellingsen and Johannesson (2004b) show that instead of an equal split of total surplus, the most common offer is an equal split of the net surplus, indicating that the proposer takes into account not only the final payoff but also the sunk investment cost of the investor.

Though predictions from models of social preferences fit many experimental findings, the specific mechanism through which social preferences work still remains to be identified. Various models are selected to fit the data from different experiments. Dufwenberg et al. (2013) use their intention-based reciprocity model (Dufwenberg and Kirchsteiger, 2004) to explain investors selecting an allocation to punish the non-investor at their own costs. Gantner et al. (2001), Ellingsen and Johannesson (2004b), and Sloof et al. (2004) show that different variations from
the inequality aversion model (Fehr and Schmidt, 1999) can explain the behaviors observed in their experiments. Fehr et al. (2008) find that their results largely depend on the individual heterogeneity in the degree of fair-mindedness among subjects. They claim that the behaviors they observed are results of the interaction between self-interested and fair subjects. Zheng et al. (2020) decompose the hold-up game and provide supporting evidence for the effect of reciprocity. These experimental results provide further support for the advantage of various social preference models over self-interest material payoff models in explaining people’s behavior in hold-up situations.

In addition to individual heterogeneity in the level of fair-mindedness, differences also arise in equity standards that individuals apply to their decisions. Individuals may have different understandings of what constitutes a fair allocation. Ellingsen and Johannesson (2004a,b) find that despite an equal split of the net surplus being the most common allocation, some subjects are found to demand an equal split of the total surplus, which leads to disagreement in a number of cases. By introducing asymmetric marginal contribution rate, Gantner et al. (2001) find three different equity standards among subjects whose decisions could be regarded as “equitable”. In addition to whether it is the net surplus or the total surplus to be divided, subjects also disagree on whether their contribution should be in line with their productivity, whether the surplus should be divided in proportion to their contribution, etc. Differences in equity standards create additional obstacles to reach an agreement in the bargaining stage, thus undermining the effect of social preferences on mitigating the hold-up problem. Furthermore, it also creates some moral wiggle room which a self-interested subject can exploit by choosing the equity standard to his own advantage.

2.2.3 Summary

Results from various hold-up experiments provide evidence for the existence of the hold-up problem, as well as a general pattern indicating that the problem is less severe than standard theoretical predictions with pure monetary payoff concerns. The few cases of complete hold-up occur only on the individual level in most experimental studies, while on the aggregate level, the investment level lies in between the self-interest strategic prediction and the socially optimal level. Most subjects take into account the link between investment decisions of both subjects and the allocation results. Models of social preferences explain the findings better than models of self-interest. Social preferences open a new channel that mitigates the hold-up problem. However, the detailed mechanism still remains to be understood. Individual heterogeneity in social preferences and equity standards may undermine
the effect of social preferences.

3 Remedies for the hold-up problem

Miller (2011) classifies remedies for the hold-up problem into two categories: formal remedies and informal remedies, depending on whether it requires formal changes in the institutional environment or not. Che and Sákovics (2008) further classifies the formal remedies into two categories: organizational remedies, such as vertical integration as proposed by Klein et al. (1978), and contractual remedies, such as contracting on the allocation before the investment decision is made. A small number of experimental studies examine the effectiveness of various formal remedies (Fehr et al., 2008; Hoppe and Schmitz, 2011; Eisenkopf and Nüesch, 2016, 2017; Davis and Leider, 2018). On the other hand, a growing number of experiments provide evidence for how various behavioral methods, such as changing the level of the outside option (Sonnemans et al., 2001; Oosterbeek et al., 2003; Sloof et al., 2004; Morita and Servátka, 2018), the observability of investment decisions (Hackett, 1993; Sloof et al., 2007), veto power and punishment possibilities (Hackett, 1993, 1994; Dufwenberg et al., 2013; Fehr and List, 2004; Zheng et al., 2020), communication (Ellingsen and Johannesson, 2004a,b; Charness and Dufwenberg, 2006, 2010; Vanberg, 2008; Ismayilov and Potters, 2016), information of past behaviors (Berg et al., 1995; Charness et al., 2011; Huck et al., 2012; Haruvy et al., 2019; Zheng et al., 2020), group identity (Morita and Servátka, 2013), etc., play an important role in mitigating the hold-up problem.

3.1 Formal remedies

3.1.1 Organizational remedies

As argued by Hart (1995) and Grossman and Hart (1986), changes in the organizational structure can help mitigate the hold-up problem. They argue that joint-ownership gives most room for underinvestment incentives to take effect, since both parties have the chance to hold up their partners in post-production renegotiations. On the contrary, single-party ownership solves this problem since the incentive is not eroded by allocating the ownership to only one party.

Fehr et al. (2008) conduct several experiments to examine the effect of different ownership schemes on subjects’ investment behaviors. Their basic experimental setting adds an additional ownership-setting stage prior to the two-party joint-investment hold-up game. The ownership scheme determines how the joint surplus from investment is allocated. Joint ownership is associated with an equal split of the surplus, while with single ownership, the owner receives all the surplus at
the cost of a fixed fee which is paid to his partner in order to obtain ownership. Subjects start with an initial ownership scheme, where one subject has the option to change the ownership scheme by selling his share of the surplus in exchange for a fixed fee. Two different ownership scenarios are tested in the experiment: 1) Subjects start with single ownership and the owner decides whether to give half of the ownership to his partner or to retain single ownership and pay a fixed fee to his partner. 2) Subjects start with joint ownership and one subject decides whether to sell his share for a fixed price or to retain joint ownership. In the investment stage, subjects sequentially choose the investment level, with the subject who does not have the option to switch ownership schemes moves first. The final surplus generated by the investments is then divided according to their agreed ownership scheme in the first stage.

Fehr et al. (2008) find that joint ownership is most frequently selected by the subjects, even though they start from different initial ownership schemes. In the joint initial ownership treatment, the majority of the subjects choose to maintain the initial ownership scheme. Among those who offer to purchase the ownership of their partners, the offer is rejected in 30% of the cases. In the single initial ownership treatment, the majority of initial owners offer joint ownership to their partners. Under joint ownership, the investment behaviors are the same no matter what initial ownership scheme they start from. There is a strong positive relationship between the investment level of the first mover and the second mover, indicating a reciprocal pattern. The efficient investment level is chosen by most of the first movers, and the second mover responds by choosing the efficient level as well, while only a small proportion of second movers exploit the first mover by choosing the minimum investment level. Fehr et al. (2008) find that in spite of different initial conditions, joint-ownership prevails as the most selected ownership structure and efficient investment is achieved in most cases under joint ownership, which is contrary to the predictions by Hart (1995) and Grossman and Hart (1986). Nevertheless, they provide evidence that changing the ownership structure indeed affects investment incentives and thus provides a solution to the hold-up problem. However, it is far from sufficient to make a decisive claim of joint ownership being the most efficient ownership structure. Fehr et al. (2008)’s findings are only valid when the number of partners is small, or if the free-rider problem is not too severe, and if there is no other way of contracting on the relationship-specific investment. The effectiveness of joint ownership can be partially attributed to different fairness concerns among subjects, but the effect is only limited to cases with less strong free-rider incentives.

Eisenkopf and Nüesch (2016, 2017) conduct experiments to test whether third-party ownership can mitigate the hold-up problem. In a Berg et al. (1995) trust
game, they introduce a third party after the investment stage to decide the allocation of the investment gains instead of the non-investor. They compare various selection mechanisms of the third party, varying the degree of independence of the third party: 1) The computer randomly selects a third party who receives a fixed payment (Eisenkopf and Nüssch, 2016, 2017). 2) The non-investor chooses a third party whose payment increases in the number of times being selected among all unidentified third parties (Eisenkopf and Nüssch, 2017). 3) The non-investor chooses a third party who receives a fixed payment based on non-binding messages sent by third parties prior to the trust game (Eisenkopf and Nüssch, 2016). 4) The non-investor chooses a third party whose payment increases in the number of times being selected, based on non-binding messages sent by the third parties prior to the trust game (Eisenkopf and Nüssch, 2016). 5) The non-investor chooses a third party whose payment increases in the number of times being selected with a fixed identification number to induce a one-sided reputation (Eisenkopf and Nüssch, 2017). Among all selection mechanisms, the selection of the third party is exogenous to the investor, while the non-investor has various degrees of information about the third party.

Compared with random selection, the selection mechanisms in the latter four treatments may induce various degrees of potential bias towards the non-investor and thus render the third party less independent. This in turn affects the investor’s trust of the third party. This is reflected in the proportion of back transfers and the level of investment. Randomly selected third parties are found to transfer back the highest amount for a given level of investment and induce the highest level of investment from the investors in both studies (Eisenkopf and Nüssch, 2016, 2017). When the non-investor selects the third party without identifying information, the proportion of back transfers is slightly lower than that with random selection, while the level of investment is similar (Eisenkopf and Nüssch, 2017). When the third party can communicate with the non-investor via a non-binding pre-play message but receives a fixed payment, the proportion of back transfers is only marginally higher than that in the baseline two-party hold-up treatment, while the investment level is no improvement from the baseline. However, if the payment of the third party depends on whether the third party is selected, the proportion of back transfers is similar to that in the two-party baseline, while the level of investment is lower than the two-party baseline (Eisenkopf and Nüssch, 2016). When the third-party can build a one-sided reputation with the non-investor, both the proportion of back transfers and the investment level are lower than that in the two-party baseline (Eisenkopf and Nüssch, 2017). These results indicate that transferring the allocation right to a third party with the appropriate level of independence can mitigate the hold-up problem, although the effect only takes
place after some positive experience. On the other hand, competition among the third parties for the chance of being selected and one-sided reputation with the non-investor offsets this effect and exacerbates the hold-up problem. Eisenkopf and Nüesch (2017) further investigate the selection mechanism of the third party by conducting an additional experiment where the non-investor is given the option to endogenously choose whether to delegate the allocation right to a third party. The results do not differ when the delegation choice is made endogenously and exogenously.

3.1.2 Contractual remedies

The source of the hold-up problem lies in the fact that investments are non-contractable in the first place, and thus both parties may have insufficient incentives to invest the efficient level. A natural solution would be to contract on the allocation of the surplus prior to the investment decision. Whether contracts can effectively mitigate the hold-up problem is widely debated in the economic literature. Maskin and Moore (1999) initiate the argument that contracts can solve the hold-up problem. Nöldeke and Schmidt (1995) discuss the possibility of using an option contract to solve the hold-up problem. Rogerson (1992) also discusses the contractual solution to the hold-up problem. According to Hart (1995), contractual solution is effective only in two limited cases: either when the widget type can be described in advance, or when the investment can be verified.

Hoppe and Schmitz (2011) add an additional contracting stage before a typical one-shot hold-up experiment with a single-party investing stage and an ultimatum bargaining stage to compare the effectiveness of different contractual arrangements. The results from the baseline treatment fit the general stylized findings of hold-up experiments: around 40% of investors already choose the high investment. Hoppe and Schmitz (2011) examines three different contractual treatments: a fixed-price contract where the non-investor pays a fixed price to the investor, an option contract where the non-investor has the option to pay the same fixed price but can decide whether to exercise the contract or not, and an option contract with renegotiation where the non-investor can make another offer should he choose not to exercise the option contract. According to theoretical predictions, only the option contract can ensure sufficient investment incentives. The findings of Hoppe and Schmitz (2011) are consistent with the predictions. The fixed-price contract does not induce higher investments than the baseline no-contract treatment. The option contract significantly increases the investment incentive. In the option contract treatment, all contracts were accepted and around 90% of investors choose the high investment. Allowing renegotiation undermines the effect of an option contract.
In the treatment of renegotiable option contract, the frequency of high investment is lower than that in non-renegotiable option contract treatment. However, the investment level is still higher than that in both fixed-price contract treatment and the baseline treatment. Hoppe and Schmitz (2011)’s findings indicate that an option contract significantly improves investment incentives and can effectively mitigate the hold-up problem. Allowing for renegotiation undermines the effect, but not as severely as theory predicts.

Davis and Leider (2018) also provide supporting evidence for the effectiveness of an option contract. They conduct a similar experiment to compare a wholesale price contract, a quantity premium contract, an option contract, and a service-level agreement in a hold-up game with random demand and a sophisticated dynamic bargaining procedure. An option contract and a service-level agreement are found to be the most efficient in increasing the investment level. They also find an indication of “superficial fairness” in their unique bargaining procedure consisting of an unstructured offer process and a structured communication process. The negotiated price often falls in the middle of the contracting space.

3.1.3 Summary

The number of experimental studies examining the effectiveness of formal remedies is limited. Contrary to standard theoretical predictions, joint ownership is found to be the most efficient ownership structure that significantly mitigates the hold-up problem. An independent third party can help increase both the investment level and transfers to the investor. An option contract on the post-investment allocation most effectively solves the hold-up problem. Renegotiation undermines the effectiveness of that contractual agreement but does not cancel it altogether.

3.2 Behavioral remedies

3.2.1 Outside options

Che and Sákovics (2008) remark that the effects of organizational remedies may depend on the bargaining solution. Changes in the outside options of a hold-up game can have a large effect on the incentives for the investor and the non-investor, and thus alter the bargaining positions. The Outside Option Principle (Binmore et al., 1989) indicates that a binding outside option has a stronger effect on the incentive of the investor, and can even induce the efficient level of investment. When one party is made residual claimant of the surplus, he then has the incentive to invest. On the other hand, the hold-up problem occurs when the non-investor’s outside option is non-binding. In a series of experiments, Sonnemans et al. (2001),
Sloof et al. (2004), and Oosterbeek et al. (2003) examine different levels of the outside option. Sonnemans et al. (2001) find that the relationship between the level of the outside option and the investment level depends on whether opting out is possible. When opting out is allowed, investment levels decrease as the value of the outside option increases; while when opting out is not available, the average investment level increases with the outside option. Sloof et al. (2004) and Oosterbeek et al. (2003) find little effect of different levels of the outside option. The overall investment level is constant over different values of the non-investors’ outside option. Contrary to the Outside Option Theory, more opportunism and lower investments are found when the non-investors’ outside option is high and binding, while the hold-up problem is less severe when the non-investor’s outside option is low and non-binding.

Morita and Servátka (2018) complements previous studies on the outside option by examining the effect of investing in outside options as a form of ex-post opportunistic behavior (Klein et al., 1978). They find that when the investor invests in outside options, the size of the outside option decreases the non-investor’s offers; when the investor does not invest in outside options, the non-investor’s offers increase with the size of the outside option. When the outside option is high, investing in outside options is regarded as opportunistic by the non-investor and can thus crowd out the non-investor’s other-regarding preferences.

### 3.2.2 Observability of investment

Standard theory with self-interest assumptions predicts that the hold-up problem can be alleviated by making specific investment unobservable, since private information can create an informational rent that boosts investment incentives. Sloof et al. (2007) conduct an experiment altering the observability of the first-stage investment to examine this. They also include changes in the cost of investment to test for the interplay of the two effects. They show that the effect of investment observability depends on the level of the sunk investment cost. Making investment unobservable can mitigate the hold-up problem when the cost of investment is high, while observability does not play a significant role when the cost is relatively low. When investment costs are high or intermediate, unobservable investment leads to higher investment levels. When investment costs are low, information condition does not significantly alter investment levels. The average investment levels are close to the predictions of standard economic theory with self-interest assumptions. When investment is observable, after observing the investment choice not selected, non-investors always demand full exploitation. After observing investment, non-investors leave room to cover the investment cost of the investor.
The results suggest that private information may partially crowd out the positive investment incentive effect of fairness and reciprocity motivations. Making investments unobservable also makes it difficult to determine whether the investor is being fair or not. As a result, unobservable investments interfere with the social preference effects, and undermines the effect of social preferences. Unobservable investments boost investments only when the costs of investment are relatively high, and thus there is insufficient scope for social preferences. It indicates that making investments observable might be a better remedy for the hold-up problem under the effect of social preferences.

In one treatment of Hackett (1993), he explicitly tests whether making observable investment is indeed an effective solution. Both players are informed of the investment decision by himself as well as by the partner. He finds that final bargaining allocation tends to be in line with ex-ante investment. Especially in the treatment where information about investment is made observable to the players. Hackett (1994) provides more supporting evidence that subjects in the treatment when investment is observable invest ten to eight percentage points more than subjects in treatment when investment is unobservable. With observable investments, subjects observe whether their counterparts have made a sufficiently “fair” investment, and reciprocate with respective “fair” allocations. Hackett (1993, 1994) also provide evidence that the treatment effect is larger when the investment cost is high, which is consistent with the findings of Sloof et al. (2007). The above experimental evidence indicates that the effectiveness of making investments common knowledge as a remedy for the hold-up problem largely depends on whether social preference is strong enough. When the investment cost is high, it is more likely for social preference to take effect. Therefore, making investments observable reinforces the effect of social preferences and can effectively increase investment incentives. On the other hand, when the investment cost is low, there is limited room for social preferences. In this case, making the investments unobservable can better alleviate the hold-up problem.

### 3.2.3 Veto power and punishment

The possibility for the responding party to reject a disadvantageous proposal can work as a tool of punishment to prevent the proposing party from exploiting with a low offer. The possibility to veto greatly changes the bargaining structure. The trust game can be regarded as a hold-up game with a veto-free bargaining stage. Zheng et al. (2020) isolate the hold-up game into a trust game which captures the effect of reciprocity and an ultimatum game which captures the effect of veto power. They separate the subjects and let them play one of these two games prior
to playing the complete hold-up game with an ultimatum bargaining stage. The investors are given information of their paired non-investors’ choices in the first game before they make the investment decision in the hold-up game. They find strong evidence for the effect of reciprocity. On the contrary, they did not find sufficient supporting evidence for the effect of veto power.

Hackett (1993, 1994) adds an additional veto option after the bargaining process, where subjects can choose whether to veto the bargaining results and receive zero payoffs instead. In his experiment, both players have the incentive to hold up. The incentive for the buyer is stronger since the no trade payoff of the bargain allocates all the gain from investment to the buyer and all the cost to the seller. Therefore, the buyer has an incentive to reject the proposals of the seller and receive the no-trade payoff. Veto on the sellers’ side gives credible threat when no agreement is reached or when the buyers choose not to bargain, but also acts as a non-binding threat when an agreement is reached.

Experiments on social preferences show that people are willing to punish others by sacrificing their own payoffs (Fehr and Schmidt, 2006). In one treatment of Dufwenberg et al. (2013), after the non-investor not choosing the equal allocation, the investor is given an option to reduce the payoff of the non-investor to a large extent at the cost of losing all of his own payoffs. The possibility of costly punishment largely increases the number of investors who choose to invest in the first stage, as compared to that in a control treatment without punishment opportunity. Compared with a 100% hold-up rate in the no-punishment control treatment, punishment significantly reduces the hold-up rate to 50%. Costly punishment effectively stops a proportion of hold-up behaviors and as a result restores investment incentives. Similarly, Fehr and List (2004) allow the investor to impose a fixed fine on the non-investor if the payback amount is lower than a “desired back-transfer” amount as announced by the investor. They also find that non-investors transfer back a higher amount of money to the investors when the punishment option is available. Moreover, within the punishment treatment, the amount sent back is higher when the punishment option is not enacted. Their results suggest that the existence of a costly punishment option can prevent hold-up behaviors, even though it is a non-binding threat and the investor chooses not to exercise it.

3.2.4 Communication

Ellingsen and Johannesson (2004a,b) conduct a series of experiments to test the effectiveness of cheap talk communication in mitigating the hold-up problem. They find that non-binding messages are indeed effective in mitigating the hold-up problem. Ellingsen and Johannesson (2004b) use an ultimatum game in the bargain-
ing stage, where communication works as a non-binding promise or threat, while Ellingsen and Johannesson (2004a) use the Nash demand game as the bargaining stage, where communication works as a coordination device. In both studies, they allow either the investor or the non-investor to send a free-form message. The investor sends a message to the non-investor simultaneously as he makes the investment decision, which can be viewed as a threat. The non-investor sends a message before the investor makes the investment decision, which can be viewed as a promise. Ellingsen and Johannesson (2004b) find that both communication treatments increase the offers made by the non-investors to the investors, though the difference is insignificant, partly due to the small number of observations. They also find that the treatment with non-investor communication leads to the highest offers to the investors than the treatment with investor communication. Comparing the consistency between the message being sent and the actions taken by the party who sends the message, promises are shown to be more credible than threats, since none of the promises were violated, while less than half of the threats that explicitly states that any offers less than the denoted amount will be rejected are executed. Applying the inequity aversion model by Fehr and Schmidt (1999), they show that inequity aversion makes promises by the non-investors more credible and threats by the investors less credible. Ellingsen and Johannesson (2004a) find similar results. Communication increase the investment rate and the investor’s profit. The rate of bargaining breakdowns is much lower when communication is allowed. Comparing the two studies, the increase in the rate of investment caused by communication is found to be higher in the Nash demand game (Ellingsen and Johannesson, 2004a) than the increase in Ultimatum game (Ellingsen and Johannesson, 2004b). Communication works better as a device to improve coordination. However, Ellingsen and Johannesson (2004a) also show that communication cannot fully eliminate the hold-up problem caused by coordination failure. There are still cases of investors refraining from investing in the investor-communication treatment. Together, these two experiments emphasise the role of communication in mitigating the hold-up problem.

A growing strand of experimental literature examines the mechanism through which communication, especially promises, increases the level of back transfer and investment in trust games.4 In a modified trust game with risk following the setup of Charness and Dufwenberg (2006), the non-investor is allowed to send a pre-play free-form message to the investor. Ellingsen and Johannesson (2004b) argue that people have a preference for promise-keeping, which fosters commitment to fulfill promises. Charness and Dufwenberg (2006) suggest an alternative mechanism of guilt aversion. People fulfill promises to avoid the guilt of deviating from others’ payoff expectations. Studies aiming to distinguish between
the two different mechanisms provide conflicting results. Vanberg (2008) finds supporting evidence for the commitment-based preference for promise-keeping, instead of the expectation-based guilt aversion. Charness and Dufwenberg (2010), however, find limited evidence for both mechanisms. In a treatment where the message of the non-investor is not delivered, Ismayilov and Potters (2016) find non-investors who make a promise are more likely to practice trustworthy behaviors than those who do not make a promise, which is in line with the commitment-based “internal consistency” explanation. However, Ismayilov and Potters (2016) fail to establish causality between promises and trustworthiness, and suggest a self-selection effect of communication such that more trustworthy non-investors are also more likely to send promises. In a trust game without communication, Ellingsen et al. (2010) fail to establish correlation between subjects’ elicited beliefs and more trusting/trustworthy behaviors, and Kawagoe and Narita (2014) find similar results in a trust game with pre-play communication, providing evidence against the expectation-based guilt aversion hypothesis. On the other hand, by exogenously varying the probability of the trustor keeping the promise, Ederer and Stremitzer (2017) provide supporting evidence for the expectation-based guilt aversion argument. In a survey paper, Cartwright (2019) compares the experimental evidence from various studies on the expectation-based guilt aversion with those on the commitment-based argument. He argues that the difference in the second-order beliefs elicitation approaches adopted in different studies renders it difficult to disentangle the two models. In a recent study, Di Bartolomeo et al. (2020) exogenously vary both promises and beliefs to compare the two approaches. Their results provide little support for the expectation-based argument.

3.2.5 Past behaviors

Reputation of past behaviors in the game can affect subjects’ expectation of what decisions the partner will make, and thus is believed to have an important effect in various strategic interactions (Schelling, 1960). Charness et al. (2011) investigate the effect of two types of reputation schemes of the non-investor: information of past behaviors in the same role (non-investor) or in the opposite role (investor). Even though only reputation in the same role is found to effectively induce the non-investor to select the trustworthy allocation, both reputation schemes have a similar positive effect on investment decisions of the investor. The reputation effect of the non-investor acting as an investor in the past indicates the role of “indirect reciprocity”. Huck et al. (2012) examine the interplay of competition and different degree of reputation on mitigating the hold-up problem. They find limited effect of reputation: reputation of identifiable non-investors can partially
increase investment decisions compared with a no-reputation baseline, the is still below the first-best; offering information of all past decisions of all non-investors does not lead to more investment decisions than limited information. However, after introducing competition, only the reputation of identifiable non-investors is sufficient to foster efficient first-best investment. Haruvy et al. (2019) finding supporting evidence for reputation of the non-investor leads to higher investment in a hold-up game with uncertainty and information asymmetry. Zheng et al. (2020) provide the investor with information of the non-investor’ decision in a prior trust game or ultimatum game. They find strong supporting evidence for information of non-investor’s previous trust game decisions leading to lower underinvestment, indicating the effect of reciprocity.

Economic models have shown that people get disutility from deviation from the behaviors of most other people (Bénaïbou and Tirole, 2006). Berg et al. (1995) show that a social norm of trust and cooperation can help initiate more trusting and cooperative behaviors from individuals, and thus mitigates the hold-up problem. In one treatment, subjects are given a summary of behaviors from subjects that have participated in previous sessions of the experiment without the social treatment. For each possible amount to be sent, the report details the frequency of every amount, as well as the average payback amount and net return. They find that after the social history treatment is introduced, there is a slight increase in the frequency of equal amount sent, as well as an increase in the frequency of higher amounts to be sent back by the non-investor. Although isolated, their findings suggest that social history tends to reinforce the effects of trust and reciprocity in mitigating the hold-up problem.

3.2.6 Group Identity

Morita and Servátka (2013) argue that group identity reinforces each member’s altruism towards other group members, and thus can lead to higher back transfers and higher investments. In their experiment, subjects are randomly divided into two teams, with team uniforms and a trivia question task to strengthen group identity. Each subject plays the hold-up game with either a member of his own team or of the other team. They find a higher investment rate and average back transfers in the same-team treatment than in the different-team treatment.

3.3 Summary

Experimental evidence has shown that the hold-up problem is less serious than standard theoretical predictions with self-regarding preferences of only monetary payoffs, mostly due to the fact that individuals take into consideration the pay-
off of their counterparts and will reciprocate in response to their counterparts’ behaviors. Methods that enhance this channel can effectively alleviate the hold-up problem. Making investments observable, giving subjects an option to veto “unfair” behaviors, allowing pre-play communication before any decision is made, providing information of past behaviors, and fostering a sense of group identity all prove to significantly improve investment incentives. Individuals’ behaviors in hold-up situations are also influenced by social history and social norms.

4 Conclusion

Being a nested bargaining game, experiments on the hold-up problem are more complicated and less commonly conducted than other simple bargaining games. This paper reviews the experimental literature on hold-up. Experiments on the hold-up problem take various forms, resulting in different representations of the problem. Though they differ in various aspects, these experiments share the same essential game structure as well as some common patterns in their results. Individual occurrences of hold-up and underinvestment have been discovered, whereas on the aggregate level the problem is less severe than what standard self-interest theory predicts. The overall behavior lies in between the social optimum and the self-interest strategic equilibrium. Subjects’ behaviors are believed to be strongly influenced by other-regarding preferences and reciprocity. A positive correlation between investment decision and bargaining results is discovered in many experiments. Individuals are found to take into consideration the investment level of their counterparts and incorporate it in their bargaining behaviors. On average, the experimental results on the hold-up problem are strongly influenced by social preferences and thus exhibits an overall pattern that is less bleak than predicted by theory.

Different variations are introduced in the experiments to examine how different types of remedies can mitigate the hold-up problem. Experimental results provide some evidence for contractual solutions and organizational solutions. Joint ownership is shown to be the most efficient ownership structure, which is contrary to theoretical results. Ownership of an independent third party is found to effectively mitigate the hold-up problem. An option contract is found to significantly boost investment incentives both when renegotiation is forbidden and allowed. However, due to the limited number of studies, further research needs to be done to study the effectiveness of these conventional remedies as well as their interplay with social preferences.

On the other hand, a number of experiments focus on behavioral remedies. The hold-up game is essentially a bargaining game embedded in an investment game.
Results from other simple bargaining games can also be applied to the bargaining stage of the hold-up game. Methods that are shown to instigate or to reinforce a “fair” allocation in other experiments on ultimatum games, dictator games, trust games, etc. can also be applied to the hold-up game. Changes in experimental settings such as making investment decision observable, the potential threat of veto power, cheap talk, reputation, and group identity are found to effectively mitigate the hold-up problem. These remedies work through various aspects of social preferences. However, further study is needed to better understand how social preferences can affect individuals’ behaviors in the hold-up problem. For example, conflicting evidence has been found regarding the mechanisms of how promises lead to higher back transfers in trust games. Various models of social preferences have been fitted with experimental data, but there is not conclusive evidence about the mechanism through which social preferences actually alters individuals’ behavior. In addition, individual heterogeneity in the degree of social preference and the judgment of fair allocations makes it more complicated to examine the social preference channel. The interplay of these individual heterogeneities creates a moral wiggle room which less prosocial individuals can exploit. Further research on the effectiveness of social preferences in hold-up situations can focus more on these heterogeneities.

5 Discussion

Laboratory experiments have many obvious advantages in collecting data for research on the hold-up problem. They enable researchers to easily implement the institutional environment that is aligned with theory, and to insert strict controls that reduce confounding factors. Though experiments on the hold-up problem are limited in numbers, they provide important insights about individuals’ behaviors in a hold-up situation. Most importantly, laboratory experiments make it easier to examine the treatment effect of a remedy for the hold-up problem. Various experimental results have provided important information to better understand the hold-up problem.

This being said, a few aspects of laboratory experiments raise concerns about the external validity of their results. The generalisability of applying laboratory results to the field is one of the most common criticisms of laboratory experiments. Laboratory experiments are mostly conducted with student subjects in a controlled environment, whose behaviors may differ from other economic players in real-world transactions from naturally occurring environments (Levitt and List, 2007). As a consequence, there have been questions about whether the results of laboratory experiments also pervade to similar situations in the real world. Among the hold-
up experiments reviewed, Fehr and List (2004) address this issue by conducting the same trust game with both student subjects and CEOs from the coffee mill sector in Costa Rica. By comparing the results from the two different subject pools, they find that CEO subjects exert more trust and exhibit more trustworthiness than student subjects. On average, CEO investors invest a larger share than students, and CEO non-investors offer back a larger share than students for any given investment level. This suggests that the hold-up problem could be less serious among CEOs. Though the difference could be partly due to the fact that the stakes used in the experiment may be considered too small to the CEOs but not to the students, the result provides an implication that there may be a discrepancy between lab results and real-world observations for hold-up games. Therefore, it is worth taking extra caution when attempting to draw implications from laboratory experiment results.

In addition to different subjects, laboratory experiments differ from real life situations in various ways. The hold-up problem can take on many different forms under different settings. The complications in institution and environment may not be fully abstracted in the current simple laboratory experiment. One example where the hold-up problem has been alleged to be present can be found in the standardization process of Standard Setting Organizations (SSOs). When the candidate technology to be incorporated in the standard is patent protected, the hold-up problem takes the form of patent hold-up, where patent owners fail to impose royalty fees according to the FRAND terms and charge high royalty fees; or it can also take the form of so-called “patent ambush”, in which patent owners withhold information about the patent in the standard setting process. Some results from the hold-up experiments in the laboratory could offer implications to the functioning of SSOs. The fair, reasonable, and non-discriminatory terms (FRAND) as is often imposed by most SSOs on members is an example. Members participating in the standard-setting process typically commit to charging license fees for their patents according to FRAND terms. The enforcement of FRAND terms reflects fairness concerns. Punishment of violating the FRAND terms can also work as a non-executed threat for the patent-holding parties to enforce FRAND terms. However, due to the heterogeneity of equity standards (Gantner et al., 2001), there could be different interpretations of the FRAND standard, which leads to difficulty in implementation. Different parties may have different interpretations of the FRAND pricing, which may still lead to some degree of patent hold-up. On the other hand, apart from different backgrounds of subjects, there are many more differences between actual SSOs and general laboratory hold-up experiments. First, in all laboratory experiments discussed above, each subject makes a decision individually in his or her own interest, while in the
standard-setting process, or many other real-life firm decisions, choices are decided collectively by a group within a firm. In the standard-setting example, the decision right is often delegated to a certain group of officers. There is already a large amount of literature studying the difference between individual and group decision making in non-strategic individual decisions such as risk and ambiguity attitudes. They seem to suggest that individuals act as if less risk-averse and more ambiguity aversive when they are part of a group than when they act individually (Brunette et al., 2015). Another common finding is that the degree of difference depends on the group decision rule. Studies comparing individual and group decision making on strategic social interaction games are quite limited in numbers. A recent study (Ambrus et al., 2009) also suggests that in a gift exchange game individuals act differently depending on the group decision rule. Whether social preferences prevail under group decision is essential to mitigating the hold-up problem. To study the individual-group decision difference in the hold-up game can offer implications for solving real-life problems such as patent hold-up. Similarly, the literature on delegation (Hamman et al., 2010; Bartling and Fischbacher, 2011) shows that due to responsibility shifting, “unfair” and “immoral” decisions are chosen more frequently under delegation. These studies focus on bargaining behavior in a dictator game. It is possible that delegation can also affect investment behavior. As the behavioral remedies for the hold-up problem work through the channel of social preferences, studying how delegation affects behaviors in both stages of a hold-up game can have an essential impact on finding a solution for the hold-up problem.

Another important difference lies in the decision process. In all experiments reviewed in this paper, subjects make decisions facing a computer screen or a paper questionnaire, without direct face-to-face interaction with their partners. Anonymity is strictly enforced in the laboratory. In addition, most experiments use a stranger setting, so that subjects interact with the same partner only once. A typical laboratory experiment usually takes no more than two hours, which limits the time span of the experiment. In the example of real business interactions, most interactions are conducted face-to-face through frequent meetings in as long as several months or even a few years. With face-to-face repeated interaction, it is possible for reputation or “self-image” effect to kick in, which may in turn strengthen the social preference channel. Therefore, whether hold-up in real-life interactions is similar to that in laboratory experiments remains an open issue.

The laboratory experiment is a simplified prototype of the real problem. This simplification is a double-edged sword that makes it possible for all the advantages of laboratory studies to be achieved, but on the other hand, also ignores some complications that could potentially induce different results. Current laboratory experiments provide important empirical evidence on subjects’ behaviors
and possible remedies of the hold-up problem with the basic structure. It is promising to continue adopting experimental methods for further understanding of the underlying mechanism behind the experimental evidence. Future experimental studies can gradually expand the scope by including variations that more closely resembles the real situation. For example, group decision and delegation could be introduced to the experimental process. In addition, a competitive environment could be introduced. Incorporating these variations will provide new evidence on the hold-up problem and its remedies under more “realistic” environments.

Notes

1Malcomson (1997) provides an overview of the hold-up problem in the labor market.
2See Schmitz (2001) for a survey of the hold-up problem using the incomplete contracts approach.
3Miller (2011) adopts the terminology from the management literature and uses terms such as “formal controls” and “informal controls”.
4The level of back transfer is often used as a proxy for trustworthiness while the level of investment for trust in this literature.
5Cartwright (2019) refers to the expectation-based guilt aversion as “belief-based model of guilt aversion”, and the commitment-based model as “reference-based model of guilt aversion”.

References


