The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers

www.cesifo.org/wp

Finding the Cost of Control

Judd B. Kessler
Stephen Leider

CESIFO WORKING PAPER NO. 4188
CATEGORY 13: BEHAVIOURAL ECONOMICS
APRIL 2013

Presented at CESifo Area Conference on Behavioural Economics, November 2012

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 4188*

Finding the Cost of Control

Abstract

A large and growing literature has demonstrated that explicit incentives, such as enforceable contracts, can lead agents to withhold effort. We investigate when this behavioral result arises. In an extensive laboratory experiment, we find that imposing control through an enforceable contract is only detrimental to principals in a special case when: (1) there is a preexisting norm that agents provide high effort; (2) control is imposed unilaterally and has an asymmetric effect on the agent; (3) control is weak (i.e. it cannot induce significant effort); and (4) the agent does not use control when acting as a principal.

JEL-Code: C900, J300, L200.

Keywords: experiment, principal-agent problem, hidden cost of control.

Judd B. Kessler
The Wharton School
University of Pennsylvania
3620 Locust Walk
1454 Steinberg Hall-Dietrich Hall
USA – Philadelphia, PA 19104
judd.kessler@wharton.upenn.edu

Stephen Leider
The Ross School of Business
University of Michigan
701 Tappan Street R4486
USA – Ann Arbor, MI 48109
leider@umich.edu

This Draft: January 21, 2013
The authors thank Rachel Croson, Florian Englmaier, Elena Katok, Muriel Niederle, Lise Vesterlund and seminar participants at Penn State, MIT, Wharton, and CESifo for helpful comments. The authors would also like to thank the staff at the Wharton Behavioral Lab.

I. Introduction

In many principal-agent relationships, a principal benefits when her agent takes costly actions and suffers when her agent shirks. Given this conflict, principals often consider the use of control (e.g. monitoring and contractual restrictions) or other forms of incentives (e.g. pay for performance and relational contracting) to induce costly action from agents. Standard economic theory suggests that these tools can better align the interests of agents with those of the principal and generate better outcomes.

Recent literature has demonstrated, however, that these control and incentive strategies may come at a cost. Control strategies that restrict an agent's actions may demonstrate distrust and may lead agents to respond with lower effort (e.g. Frey 1993, Barkema 1995, Falk and Kosfeld 2006).[1] This result has been referred to as the "hidden cost of control" and arises when a principal receives less effort from an agent when she controls the agent by taking away his most opportunistic actions. Similarly, extrinsic incentives put in place by a principal to motivate an agent might undermine an agent's intrinsic motivation and lead to lower effort (see e.g. Titmuss 1970, Frey 1994, Gneezy and Rustichini 2000a) or might undermine a norm and make misbehavior more transactional (Gneezy and Rustichini 2000b).[2]

While this research highlights the potential costs associated with control and other incentives, such strategies are common in principal-agent settings and becoming even more prevalent. Between 1960 and 1995 average supervisor-employee ratios in the non-farm economy increased for many developed countries (Vernon 2003).[3] Given the prevalence of control in principal-agent settings, the recent findings of its perverse consequences pose a bit of a puzzle and suggest a need to better understand when costs of control will arise.

More generally, demonstrations that control and other extrinsic incentives *can* undermine intrinsic motivation are an important first step in developing models of the principal-agent relationship that incorporate behavioral phenomena. The next step in developing

---

[1] A recent survey by Bowles and Polania-Reyes (2012) identifies four major mechanisms for a crowding out effect of incentives: (1) incentives providing "bad news" about the principal, (2) framing effects that lead to "moral disengagement", (3) aversion to a loss of autonomy, and (4) influence on the formation and updating of preferences.

[2] See also a rich literature in Psychology, which has shown that extrinsic incentives can undermine intrinsic motivations (see Lepper and Greene 1978; Deci 1975; Deci 1971; Kruglansky, Freedman, and Zeevi 1971), a notion which has been more recently studied in the economics literature (see for example Frey, Oberholzer and Eichenberger 1996 and Frey and Oberholzer-Gee 1997). A related literature argues that punishment of agents through fines may be less effective than incenting good behavior through bonuses (Fehr, Klein and Schmidt 2007).

[3] The use of incentives is also prevalent. In the United States, 37% of individuals have some form of pay-for-performance incentives (Lemieux et al. 2009).

these models is to understand when (i.e. in which environments) these perverse effects of control and other incentives will be severe enough that the standard model will fail to describe behavior.[4] Identifying the boundaries of the cost of control will help models better describe behavior and provide guidance for principals and firms.[5]

In this paper, we investigate the impact of control in a principal-agent relationship with a laboratory experiment. In the experiment, an agent has the opportunity to take a costly action that benefits a principal, who may control agent by eliminating the most opportunistic actions from the agent's action space.

Our main experiment has four treatments that vary the symmetry of the principal-agent relationship and the extent to which the control placed on the agent is imposed unilaterally. We vary the symmetry of the relationship by randomly assigning the roles of principal and agent in each round of the experiment and varying whether control is imposed before or after the identity of the agent is revealed. We vary whether control is imposed unilaterally by allowing one player to impose control or by requiring both players to choose control for it to be imposed (i.e. bilaterally).

In addition, we embed control as a decision amidst a broader set of contracting options, which allows us to identify the effect of control in settings where a prosocial norm for behavior is present and where one is not. Namely, we give the two players the opportunity to mutually agree to a non-binding high effort level in advance of the revelation of roles and the decision to control, and we interpret agreeing to this high effort level as establishing a prosocial norm for behavior (see Kessler and Leider 2012).

Since we observe subjects playing the game as both a principal and as an agent, we can also investigate individual differences in how subjects respond to control as an agent and when subjects use control as a principal.

Our experiment includes a set of contracting environments in which there is a significant and robust cost of control — where principals receive less effort on average when they impose control than when they do not.[6] By varying the contracting rules we are able to

---

[4] It is worth noting here that a number of recent experimental papers have investigated the impact of control on agent effort in designs very similar to Falk and Kosfeld (2006) and have in general failed to find a result from Falk and Kosfeld (2006) that principals received less effort on average when they imposed control (Hagemann 2007, Schnedler and Vadovic 2011, Ploner et al. 2011). We describe these papers in Section II.

[5] We see this exercise of putting boundaries on behavioral phenomena as a generally useful activity that pushes the field toward richer theories that incorporate these phenomena.

[6] We say that a "cost of control" arises when a principal receives less effort from controlling an agent than from giving the agent a larger action space. We say that we have observed a "behavioral response" or a "hidden cost of control" when subjects respond negatively to the imposition of control by providing less effort when control is imposed than when it is not. Falk and Kosfeld (2006) use the term "hidden cost of control" to title their paper.

eliminate the cost of control so that principals do as well, or better, by imposing control. In some of these treatments, we still observe a behavioral response in which agents respond with less effort when controlled than when not controlled, however the behavioral response is not large enough to overcome the benefit of control, and so the principal is no worse off when imposing control. By varying the contracting rules further, we able to turn off this behavioral response so that control works as predicted by standard theory and the principal is made better off by imposing control.

As the summary in the preceding paragraph suggests, a cost of control arises only in certain environments. We find that principals are only harmed by imposing control when all of the following conditions are met: (1) there is a prosocial norm motivating the agent; (2) control is imposed unilaterally and has an asymmetric effect on the agent; (3) control is weak, in that it cannot induce significant effort from the agent; and (4) the agent is does not use control himself when acting as a principal. In all other cases, we find that the principal is no worse off from imposing control. We also find that the principal is better off imposing control when either: (1) there is no prosocial norm motivating the agent or (2) control is agreed upon mutually between the principal and the agent.

These results provide guidance for when firms and principals should worry about imposing control and when they should not hesitate to use the contractual tools available to them. The results speak broadly to the literature on incomplete contracts (see Hart 1995 and Tirole 1999 for surveys) as the costs of control have been used as an explanation for why contracts might be left deliberately incomplete.[7]

The paper proceeds as follows. Section II highlights related literature about the principal-agent relationship. Section III describes the experimental design. Section IV presents the main experimental results. Section V describes extensions to our experiment and their associated results. Section VI discusses the implications of our results for economic theory and firm behavior and concludes.

[7] The logic of this argument is that if principals receive less effort from agents when they impose control, observed contracts will be left incomplete to avoid this outcome. Kessler and Leider (2012) make a related argument, observing that prosocial norms for a relationship can be established with the *unenforceable* clauses in contracts or with the conversations that take place during the contracting process. Kessler and Leider (2012) finds that once prosocial norms are established within a relationship, enforceable clauses rarely increase output; so if there is any cost to adding enforceable clauses to contracts doing so might not be worth it once prosocial norms are established.

II. Related Literature

Principal-agent relationships play an important role in the labor market and in markets with supply chains. Firms use control, monitoring, and incentives to manage the agency problem, and research has addressed the role of contracts in implementing these strategies. A striking fact is that many contracts are much simpler and less complete than standard theory would predict. Traditional explanations of this contractual incompleteness appeal to transaction costs (e.g. Coase 1937, Williamson 1975, 1985) or bounded rationality (e.g. Simon 1981) to argue that more complete contracts are impractical. Another line of research has suggested that leaving contracts incomplete may be suboptimal but necessary given that agents are asked to multitask (Holmstrom and Milgrom 1991). Additionally, some authors have provided theoretical justifications for why incomplete contracts may be optimal, such as complete contracts signaling negative information about the contract proposer (Allen and Gale 1992, Spier 1992), complete contracts leading the agent to infer that a less pro-social norm prevails (Sliwka 2007), or that incompleteness creates strategic ambiguity that helps enforce implicit agreements (Bernheim and Whinston 1998).

There is also an extensive experimental literature demonstrating that control, monitoring, and incentives can demoralize agents.[8] Falk and Kosfeld (2006) suggest that contractual incompleteness in control mechanisms could also arise to signal trust. The paper demonstrates that imposing control on agents — by eliminating their most opportunistic actions and forcing them to provide at least a minimum compulsory effort — can lead to worse outcomes for the principal. Falk and Kosfeld (2006) present the results of experiments in which an agent chooses a costly effort level, and the payoff to the principal is twice the cost paid by the agent. Before the agent chooses an effort level, the principal decides whether to impose control on her agent by limiting his action space; different treatments allow for control of different strengths (i.e. more extensive restrictions of the agent's action space). The principal receives higher effort from the agent, and thus higher profits, when she does not control the agent. This difference is statistically significant when control is weak and is only directionally negative when control is relatively strong. Their results are robust to whether the action of the agent is chosen by strategy method or direct choice and whether the principal has the opportunity to engage in gift exchange with the agent before the effort choice.

A number of recent papers have attempted and failed to replicate the result in Falk and Kosfeld (2006) that principals are made worse off by imposing control. The papers have

[8] In addition to the control mechanisms described in detail below, extrinsic incentives have been shown to crowd out intrinsic motivation (see Deci et al. 1999 and Gneezy, Meier and Rey-Beil 2011 for surveys).

consequently argued that costs of control are unlikely to be the cause of contractual incompleteness. Hagemann (2007) finds a non-significant negative effect of adding control in an attempted replication of Falk and Kosfeld (2006) using the strategy method.[9] The paper finds that uncontrolled agents respond with higher effort when the possibility of control is worded as the principal being able to "force" the agent to transfer points than when it is worded as the principal being able to "constrain" the agent or when control is described neutrally. This finding leads the author to argue that the original Falk and Kosfeld (2006) result arises from an experimenter demand effect. Similarly, Schnedler and Vadovic (2011) finds a behavioral response to control, but fail to replicate the cost of control result, instead finding that average effort is directionally higher when control is imposed. In a set of attempts to replicate Falk and Kosfeld (2006), Ploner et al. (2012) finds both directionally negative and directionally positive effects for the principal of imposing control, depending on the subject pool. Principals receive lower effort when they impose control on agents in only one of three incentivized experiments, and only in the condition where control is relatively weak.

While these papers fail to replicate the cost of control results from Falk and Kosfeld (2006), they generally do replicate the behavioral response (or "hidden cost of control") in which a number of agents contribute less when they are controlled than when they are not controlled. For example, these papers find that many subjects provide the minimum effort allowed when control is imposed but provide effort above that minimum when control is not imposed.

Research in other settings has observed the expected beneficial effect of control mechanisms without an offsetting behavioral response to imposing control. For example, Kessler and Leider (2012) have subjects play two-person games, including public good games, in which effort is personally costly but collectively beneficial. The authors find that adding an enforceable minimum (i.e. control) to a pre-game contract either increased effort or had no effect on effort in three of four games subjects played. The enforceable minimum decreased effort in only one game. For some of the games, adding the enforceable minimum did not even generate a behavioral response.[10]

The games in Kessler and Leider (2012) differ from a principal-agent setting on an important dimension: in Kessler and Leider (2012), both players make effort choices in a

[9] However, with only 30 agents in each treatment, Hagermann's experiment may be underpowered to identify a treatment effect in the baseline case. Hagermann (2007) finds a difference in average effort of 5.3, which is very similar to the difference of 5.5 (23 without control and 17.5 with control) in the equivalent treatment in Falk and Koslfeld (2006), which has 72 agents and identifies the effect as significant.

[10] As a consequence of identifying when costs of control arise, the experiment in this paper will help reconcile why some attempts to replicate Falk and Kosfeld (2006) have failed and why Kessler and Leider (2012) find no behavioral response associated with imposing control.

symmetric game. Consequently, when a minimum restriction is imposed, either bilaterally or unilaterally, it is imposed on both agents simultaneously. We hypothesize that this difference has the potential to create divergent results in the response to the imposition control in the principal-agent setting of Falk and Kosfeld (2006) and in the partnership setting of Kessler and Leider (2012). The experimental design in this paper, which is presented in the next section, starts with a principal-agent relationship where the principal can impose control on an agent and adds symmetry to the relationship. Symmetry is added both on who is controlled as well as on how control is imposed (i.e. unilaterally or bilaterally).

Most of the previous studies that find a cost of control usually observe high effort in the absence of incentives, suggesting that strong norms govern behavior in these settings. When norms are weak, however, we may expect that the benefits of control will outweigh any negative behavioral response associated with imposing control. The previous studies also suggest that the level of effort in the absence of control (and thus the "default" norm) can vary widely between subject pools. To increase the likelihood that our subjects will perceive a strong norm governing transfers, we use the pre-play agreement mechanism from Kessler and Leider (2012). In that paper, when players could make a non-binding agreement to play the first best action, a norm was established, and effort and profits were higher than when no agreement was available.[11] In related work, Dufwenberg et al. (2011) provide a theoretical model that identifies what agreements should form as binding contracts or as non-binding informal agreements, and test their model with a lost wallet game. They find that binding contracts are predominantly 50-50 splits, while non-binding informal agreements lead to higher payoffs for the second mover, which one can think of as the agent. We expect that pre-play agreements will more consistently lead to a strong norm for transfers, and therefore increase the likelihood of observing a cost of control.

III. Experimental Design

In the experiment, subjects in a laboratory sat at individual computer terminals and played an anonymous principal-agent transfer game a total of 20 times. Subjects were randomly matched with another subject in the lab in each round of the game.

Each round of the game, the agent (called "Player A" in the instructions) started with 120 experimental units (EUs) worth $0.05 each. The agent could transfer these units to the

[11] Other studies have found benefits of unilateral promises in holdup games (Ellingsen and Johannesson 2004), trust games (Charness and Dufwenberg 2006), and dictator games (Vanberg 2008).

principal (called "Player B" in the instructions) and any units transferred to the principal were doubled. Consequently, the payoffs for the principal agent game were:

$$\text{Agent ("Player A"): } \pi_A = 120 - x$$

$$\text{Principal ("Player B"): } \pi_P = 2x$$

where x represents the number of units transferred by the agent to the principal.

If control (called "a restriction" in the instructions) was not imposed, agents could choose to transfer any amount x from the range 0 to 120. If control was imposed, agents were restricted to transfer at least 4 EUs, so agents could transfer any amount x from 4 to 120.

Before subjects were assigned to the role of principal or agent for the round, and before they knew whether control would be imposed, they had the opportunity to make a non-binding agreement to transfer 40 EUs (i.e. x=40) if they ended up being the agent. We choose 40 EUs since that is the payoff-equalizing transfer, which leads both the principal and the agent to receive 80 experimental units.

Each of the players independently decided whether or not to suggest: "An agreement that says 'We agree that if we are Player A, we will transfer 40 EUs to Player B.'" If both players suggested the agreement, then the agreement was made. If one or both of the players did not suggest the agreement, then no agreement was made. After both players had decided whether or not to suggest the agreement, the players were told what the other had chosen and whether they had made an agreement. Allowing subjects to make an agreement to transfer 40 EUs provides an opportunity to investigate settings where a pro-social norm of behavior has been established. We expect, based on the results of Kessler and Leider (2012), that allowing subjects to make an informal agreement will more consistently lead to a strong pro-social norm than relying on the default background norm present in a subject pool.

All rounds began with subjects choosing whether or not to suggest an agreement, after which the instructions differed by treatment. The experiment has four treatments, which differ in whether control affects one or both agents and whether control was imposed unilaterally or bilaterally.

Figure 1 displays the treatments as a function of whether the control was imposed on one player (i.e. "single player") or on both players (i.e. "both players) as well as whether one player could impose control (i.e. "unilaterally") or whether both players needed to agree to control for it to be imposed (i.e. "bilaterally").

Note that in our experiment we make a distinction between whether control is imposed before or after the subjects are assigned to be the principal and the agent. When control is imposed before roles are assigned, we say the agent is "unknown," and when control is

imposed after, we say the agent is "known." As will be explained in detail below, this distinction is necessary in our design since in moving from a setting where control affects a single player to a setting where it affects both players, we must delay revelation of who is the agent and who is the principal.

Figure 1: Experimental Treatments

		Symmetry of Control		
		Single Player		**Both Players**
		Known Agent	*Unknown Agent*	
Control Imposed	**Unilaterally**	Baseline Treatment	Unknown Agent Treatment	Mutual Minimum Treatment
	Bilaterally			Consent Treatment

In the *Baseline Treatment*, the roles of principal and agent are assigned immediately after the players are told whether they have made an agreement. After the principal and the agent are assigned their roles, the principal is given the option of whether to impose control (called "a restriction on Player A's transfer" in the instructions). The principal decides between: "No restriction" and "A restriction that Player A must transfer at least 4 EUs." After the principal chooses, the choice is revealed to the agent. The agent then decides how many experimental units to transfer, and the transfer is restricted to be at least 4 EUs when control is imposed. Notice that for the *Baseline Treatment*, the minimum is imposed on a single player (i.e. "single player"), after the identity of the agent is known (i.e. "known agent"), and one of the players imposes control unilaterally on the other (i.e. "unilaterally").

We add symmetry to control in some treatments by having control affect both players. To achieve this, we did not assign the roles of principal and agent until after control had been imposed. By allowing control to be imposed before the role of agent was assigned, we had the opportunity to run treatments where control is imposed more symmetrically.

In the *Mutual Minimum Treatment*, before we assign the roles of principal and agent, we randomly give one of the players the option to impose control on *whichever player* becomes the agent. In this way, control is imposed symmetrically in that it affects both players (since either one could end up being the agent in that round). Once the player

decided whether to impose control, we assigned the roles of principal and agent. If the player had decided to impose control, whichever of the two players was randomly selected to be the agent was restricted to transfer between 4 and 120 EUs. If control was not imposed, the agent could choose any transfer between 0 and 120 EUs. Notice that for the *Mutual Minimum Treatment*, the minimum is imposed on both players (i.e. "both players"), while the identity of the agent is still unknown (i.e. "unknown agent"), and one of the players imposes control unilaterally on the other (i.e. "unilaterally").

Moving from the baseline treatment to the mutual minimum treatment is associated with two changes. We have allowed control to be imposed symmetrically but we have also imposed control before we have assigned the roles of principal and agent. To tease apart which of these two differences leads to differential outcomes between the baseline treatment and mutual minimum treatment, we also ran the *Unknown Agent Treatment*.

In the *Unknown Agent Treatment*, control is not imposed symmetrically (i.e. control is imposed on a "single player") but control is imposed before the role of agent as been assigned (i.e. "unknown agent"). In the *Unknown Agent Treatment*, before we assigned the roles of principal and agent, we randomly gave one of the players the option to impose control on *the other player* if that other player became the agent. Once the player decided whether to impose control, we assigned the roles of principal and agent. If the player who decided about control became the agent, he was always able to choose a transfer between 0 and 120 EUs. If the other player became the agent, the action space available to that agent depended on the choice of whether to impose control. Notice that for the *Unknown Agent Treatment*, the minimum is imposed on a single player (i.e. "single player"), while the identity of the agent is still unknown (i.e. "unknown agent"), and one of the players imposes control unilaterally (i.e. "unilaterally").

In all three of the treatments above, one subject has the opportunity to impose control unilaterally on either the other subject or on whichever subject becomes the agent. To investigate how control differentially impacts behavior when agreed to bilaterally, we ran the *Consent Treatment*.

In the *Consent Treatment*, before we assign the roles of principal and agent, we allow both players to suggest whether or not control should be imposed on *whichever player* becomes the agent. The decision to impose control is made in the same way as the agreement to transfer 40 EUs. Namely, each player can suggest the restriction or no restriction, and only if both players suggest the restriction is control imposed. After each has made a decision, the players are told who suggested the restriction and whether the restriction was imposed. We then assigned the roles of principal and agent. If both players suggested the restriction, the agent was restricted to transfer between 4 and 120 EUs. If at least one of the players had not suggested the restriction then there was no restriction, and the agent could choose to transfer any amount between 0 and 120 EUs.

Notice that for the *Consent Treatment*, the minimum is imposed on both players (i.e. "both players"), while the identity of the agent is still unknown (i.e. "unknown agent"), and both players impose control bilaterally (i.e. "bilaterally").

The cell in Figure 1 that is not associated with a treatment would require both subjects to agree bilaterally to impose control on one subject. For control to be imposed in this setting, a subject would need to choose to control himself, knowing that he alone would be affected. We do not consider this setting to be realistic or particularly relevant to our endeavor since the concept of control does not really apply.

Subjects in our experiment always played 10 rounds in the Baseline Treatment and 10 rounds in one of the three other treatments. Whether they played the Baseline Treatment first or second was randomly assigned by session.

We randomly paired subjects in each round and, as mentioned above, we randomly assigned the roles of principal and agent in each round. Consequently, the design allows us to observe the same subject playing as both a principal and an agent. In addition to identifying how agent behavior responds to the symmetry of control and whether control is imposed unilaterally or bilaterally, the experiment can also investigate how propensity to use control as a principal affects how subjects respond to control as an agent.

One important point about the experiment that is worth emphasizing here is that control may have two countervailing forces on behavior. Control may eliminate low transfers due to a minimum transfer being imposed — transfers that might have been in the range 0 to 3 EUs must be at least 4 EUs when control is imposed. In addition, control may lead to a behavioral response in which subjects are less likely to make large transfers. The net effect of these two forces will determine whether we observe a cost of control whereby the principal receives less effort from the agent when control is imposed. As we analyze the results starting in the next section, we want to document both potential effects. In the next section we look for a behavioral response as well as identify the net effect of both forces, which together determine whether there is a cost of control.

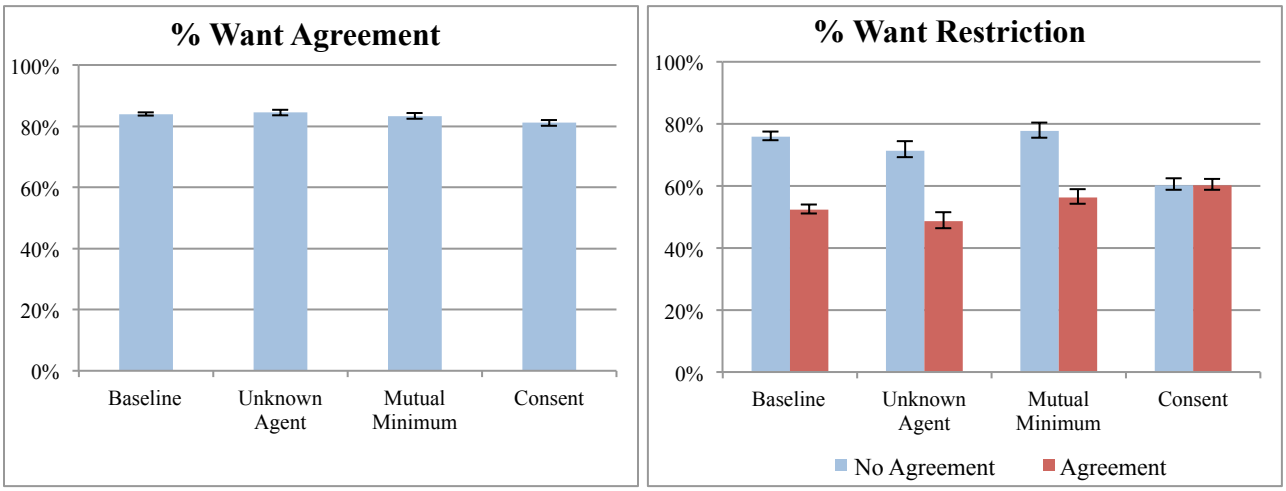
IV. Results

A total of 464 subjects participated in 25 sessions in the Wharton Behavioral Lab at the University of Pennsylvania. All subjects participated in the Baseline treatment and one other treatment. Of the 464 subjects, 148 subjects participated in the Unknown Agent treatment, 158 subjects in the Mutual Minimum treatment, and 158 subjects in the Consent treatment. Sessions lasted approximately one hour. Average subject pay was $17.28, including a $10 show-up fee.

IV.1 Agreement and Restriction Choices.

We begin by examining subjects' preferences for having an agreement or a restriction. Figure 2 displays the frequency with which subjects suggested the agreement in each treatment, as well as the frequency with which subjects imposed the restriction (or asked for the restriction in the Consent treatment) with or without an agreement. Subjects were strongly in favor of having an agreement across all four treatments, with very little difference between treatments — between 80 and 85% of subjects suggested the agreement in each treatment. This led subjects to form an agreement in 65 to 75% of periods.

Figure 2: Percent of Subjects who want an Agreement/Restriction



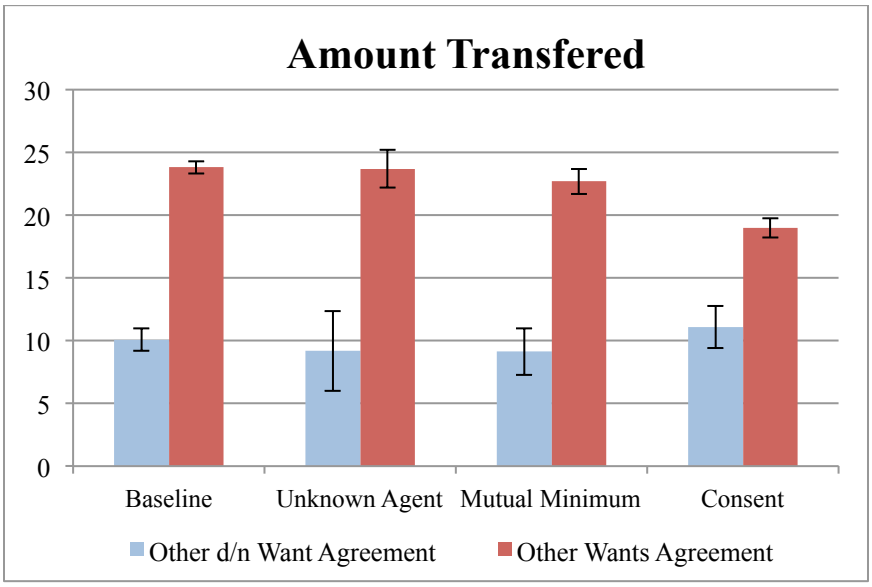
Desire to impose a restriction varies based on whether the subjects had previously made an agreement. In the Baseline, Unknown Agent, and Mutual Minimum treatments subjects impose a restriction approximately 50% of the time with an agreement, but nearly 75% of the time without an agreement. This difference is consistent with subjects anticipating lower transfers when there is no agreement, and therefore having an increased desire to rule out extremely low transfers. In the Consent treatment, by contrast, subjects request the restriction in 60% of periods both with and without an agreement.

IV.2 Effect of Agreement on Transfers

Next, we look at how making an agreement affects the amount transferred by the agent. Based on results in Kessler and Leider (2012), we expect that agreements will lead to higher transfers by the agent, and in particular an increase in the number of agents

transferring the agreed-upon amount of 40 EUs. Figure 3 shows the impact of the other subject wanting the agreement on the average amount transferred by agents.[12] Note that we split the data based on the other subject's desire for the agreement (rather than having the agreement) so that we do not introduce selection based on the agent's own preferences for the agreement, which may be correlated with that subject's choice of transfer.[13]

Figure 3: Effect of an Agreement on Average Transfers



Across all four treatments, we find that the other subject asking for the agreement leads to substantial increases in the average amount transferred. To test for statistical differences, we use non-parametric permutation tests on choices, aggregated first at the subject level and then at the session level.[14] The differences between transfers with an agreement and without an agreement are statistically significant for all treatments using subject-level comparisons ($p < 0.01$ for all) and for session-level comparisons ($p < 0.01$ for Baseline

[12] For this analysis, and subsequent analyses, we exclude observations in the Unknown Agent and Mutual Minimum treatments where the player who was randomly selected to decide whether or not there should be a restriction was also randomly selected to be the agent.
[13] The averages displayed by the "Other Wants Agreement" bars therefore represent all the observations where an agreement is formed (making up approximately 80% of the observations described by the bars) as well as the observations where the agent did not ask for the agreement and so no agreement was made (making up approximately 20% of the observations).
[14] We use unpaired permutation tests for subject-level data (because some subjects only have observations with an agreement or without an agreement) and paired permutation tests for session-level data.

and Mutual Minimum, $p = 0.08$ for Unknown Agent, and $p = 0.02$ for Consent). Additionally, we find that having an agreement substantially increases the number of subjects who choose the payoff-equalizing transfer of 40 EUs. Without an agreement only 2 to 12% of subjects transfer 40 EUs, while with an agreement between 31 and 47% of subjects transfer 40 EUs ($p < 0.01$ for all treatments at both the subject and session level). Hence, we find strong evidence in favor of the positive effect of agreements in our principal agent game, in line with the results of Kessler and Leider (2012) for symmetric games. We summarize this as Result 1.

Result 1: Offering to make an agreement significantly increases transfers in all treatments.

IV.3 Effect of Control without an Agreement

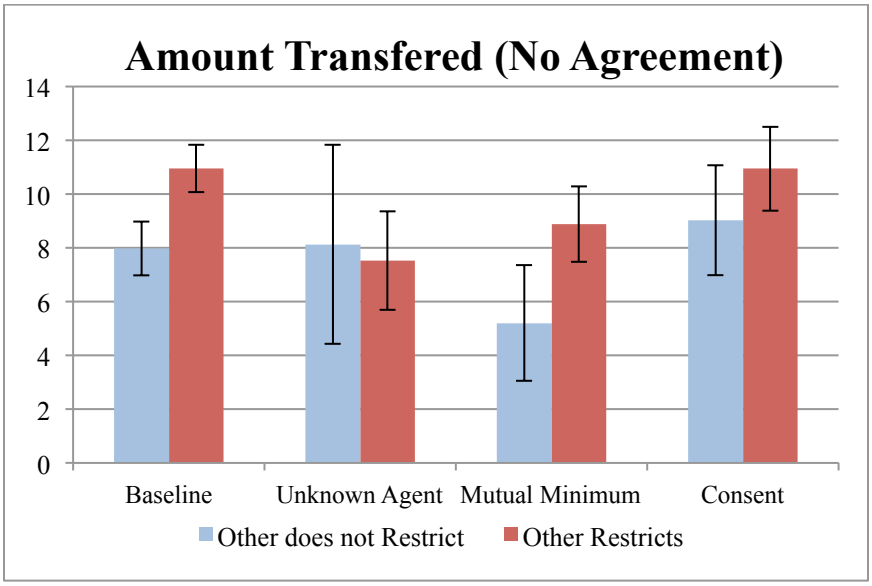
We now look at whether imposing control by restricting the transfer of the agent leads to a decrease in the amount the agent transfers. We begin by analyzing pairs who do not have an agreement.

As with the analysis of agreements, we analyze the effect on transfers associated with the other subject suggesting the restriction. (For the three unilateral treatments, the other subject suggesting for the restriction means the restriction is imposed. In the Consent treatment, however, agents who do not ask for the restriction do not have the restriction in place even when the other subject suggests it.) Analyzing the effect of the other subject suggesting the restriction avoids the potential selection into control based on an agent's own preference for control in the consent treatment.[15] Figure 4 shows the average amount transferred in each treatment.

When there is no agreement, the average transfer increases with the restriction in three of the four treatments, and remains essentially the same in the Unknown Agent treatment. Under our non-parametric tests, the increase is only significant in the Baseline treatment where we have the most data (subject-level: $p = 0.07$, session-level: $p < 0.01$; $p > 0.20$ for all other treatments). This result suggests that when there are only weak norms affecting behavior (due to the absence of an agreement) imposing a restriction is at worst neutral and at best beneficial to the principal, as in the Baseline treatment.

[15] In the other three treatments the other player wanting the restriction is equivalent to having the restriction. We exclude behavior of agents who were also the player who decided whether or not there should be a restriction in the Unknown Agent and Mutual Minimum treatments.

Figure 4: Amount Transferred with and without a Restriction (No Agreement)

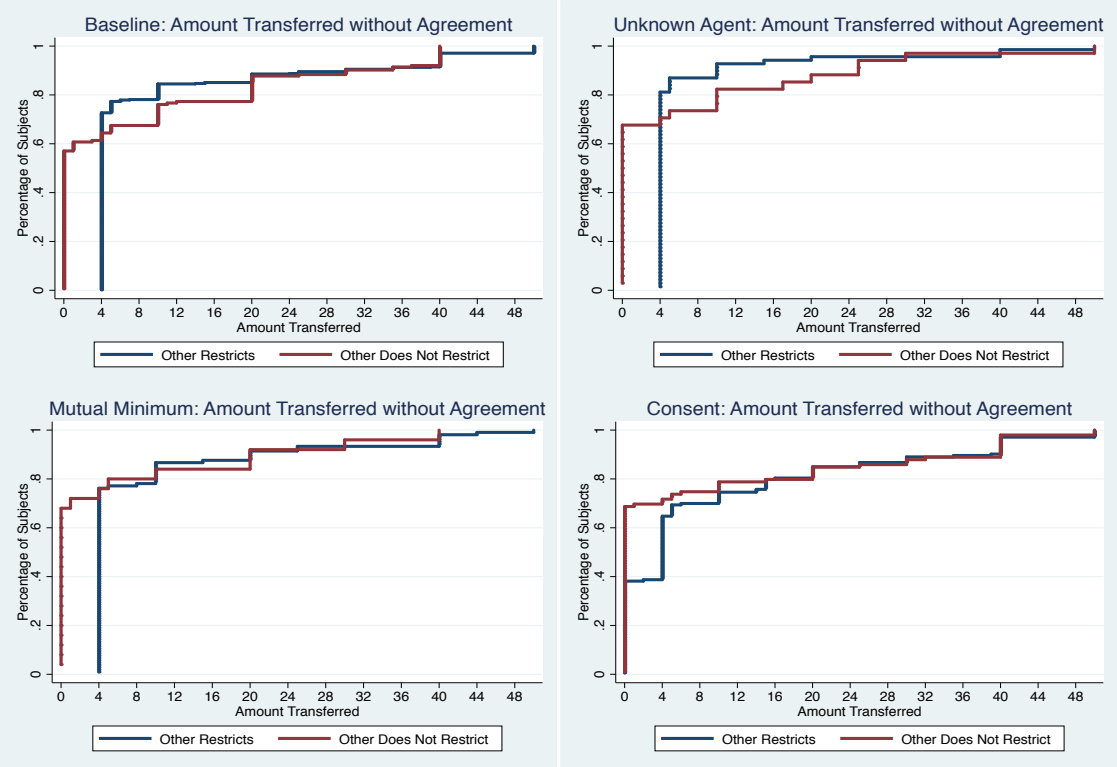


However, looking at the average transfer may mask two opposing effects. As was mentioned above, control may eliminate low transfers due to the enforcement of the minimum transfer and may also lead to a behavioral response in which subjects are less likely to make large transfers. We therefore want to examine whether there is a behavioral response to imposing control. In particular, we want to look at the fraction of subjects who transfer 4 units or less. If the restriction only affects those subjects who otherwise would have transferred less than the minimum, then the fraction of subjects transferring at or below the minimum of 4 should be the same. Alternatively, if subjects who would otherwise transfer more than the minimum react negatively to the imposition of control by transferring only the imposed minimum, than this fraction transferring 4 units or less should increase when control is imposed. Figure 5 plots the cumulative distribution of transfers by each treatment, based on whether or not the other subject wanted the restriction.

In all treatments, the vast majority of subjects transfer only a small amount: between 64 and 75% of subjects transfer 4 or less without the restriction and between 65 and 81% transfer 4 or less with the restriction. In the Baseline and the Unknown Agent treatments there is a slight directional increase in the percent of transfers that are at or below the minimum in responses to control (8 and 11 percentage points, respectively). In the Baseline treatment, the effect is marginally significant (subject-level: $p = 0.07$; session-level: $p = 0.40$), while in the Unknown Agent treatment the effect is not significant (subject-level: $p = 0.15$; session-level: $p = 0.16$). Recall that in the Baseline treatment we found a significant increase in agent effort in response to the minimum, so to the extent

that there is a behavioral response to the imposition of control, it is swamped by the effect of increasing transfers of less than 4 EUs up to at least 4 EUs.

Figure 5: Distribution of Transfers with and without a Restriction (No Agreement)



We also conduct a regression analysis of the individual agent transfer decisions, which is reported in Table 1. All specifications include subject fixed effects and cluster the standard errors by subject. Columns (1) and (2) use the amount transferred as the dependent variable, identifying any overall effect of the restriction on agent effort, while Columns (3) and (4) use an indicator variable for a transfer less than or equal to 4 as the dependent variable, to capture any behavioral response. Columns (2) and (4) use data from only the first treatment in a session.

Our regression results largely confirm our non-parametric analysis and find generally positive effects of imposing a restriction when there is no agreement. For the amount transferred, imposing the restriction significantly increases amount transferred in the Baseline treatment, and has directionally positive effects in all other treatments. Additionally, we find no increase in the number of subjects transferring very small amounts under the restriction in any treatment, with the Consent treatment having a

significant *decrease* in the frequency of transfers of 4 or less when the other subject asks for the restriction. When there is no agreement between the principal and agent, imposing a restriction is not costly for the principal; instead, it is weakly beneficial for the principal.

Table 1: Transfers without an Agreement

	Transfer		Transfer <= 4	
VARIABLES	(1)	(2)	(3)	(4)
Unknown Agent	1.454		-0.0111	
	(2.594)		(0.0903)	
Mutual Minimum	-0.476		-0.0990	
	(2.283)		(0.0792)	
Consent	0.372		0.123*	
	(2.445)		(0.0689)	
Other Restricted in Baseline	3.527***	3.356*	-0.00452	-0.0476
	(1.291)	(1.818)	(0.0414)	(0.0617)
Other Restricted in Unknown Agent	0.961	4.923	-0.0245	-0.154
	(2.033)	(6.683)	(0.0782)	(0.124)
Other Restricted in Mutual Minimum	1.992	2.800	0.0821	0.100
	(2.494)	(4.708)	(0.0863)	(0.174)
Other Restricted in Consent	3.218	7.087*	-0.130**	-0.192**
	(2.480)	(3.748)	(0.0640)	(0.0933)
First Treatment	4.444***		-0.158***	
	(1.061)		(0.0286)	
Constant	5.230***	8.752***	0.787***	0.676***
	(1.252)	(1.095)	(0.0391)	(0.0317)
Observations	1184	575	1184	575
Number of Subjects	401	298	401	298
R-squared	0.031	0.022	0.062	0.025

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors clustered at the subject level reported in parentheses. The sample is restricted to observations where there was no agreement, and for the Unknown Agent and Mutual Minimum treatments only observations where the principal had the opportunity to restrict the agent are included. In columns (2) and (4) the sample is further restricted to only the first treatment of a session. All specifications include subject fixed effects. The dependent variable in columns (1) and (2) is the transfer of the agent, in columns (3) and (4) it is an dummy variable that equals one of the transfer was less than or equal to 4.

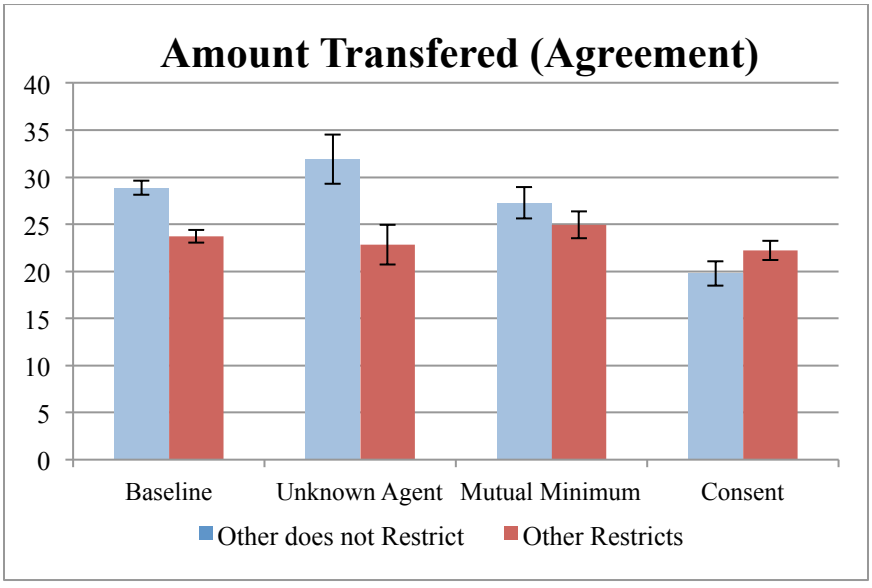
Result 2: When there is no agreement, there is no cost — and there is sometimes a benefit — of imposing control.

IV.4 Effect of Control with an Agreement

We did not find a cost of control when there was no agreement, but without an agreement most transfers were quite low in the absence of a restriction. We showed previously that transfers are much higher when subjects form an agreement. Consequently, when control is imposed after an agreement has been made, we may expect to find a smaller impact of control on raising low transfers, a larger behavioral response, and an accompanying cost of control.

Figure 6 displays the average amount transferred by agents when they have an agreement, and again we split the data based on whether the other subject *suggested* the restriction. For the three unilateral treatments, the other subject suggesting the restriction means the restriction is imposed. In the Consent treatment, agents who do not ask for the restriction do not have the restriction in place even when the principal suggests it.

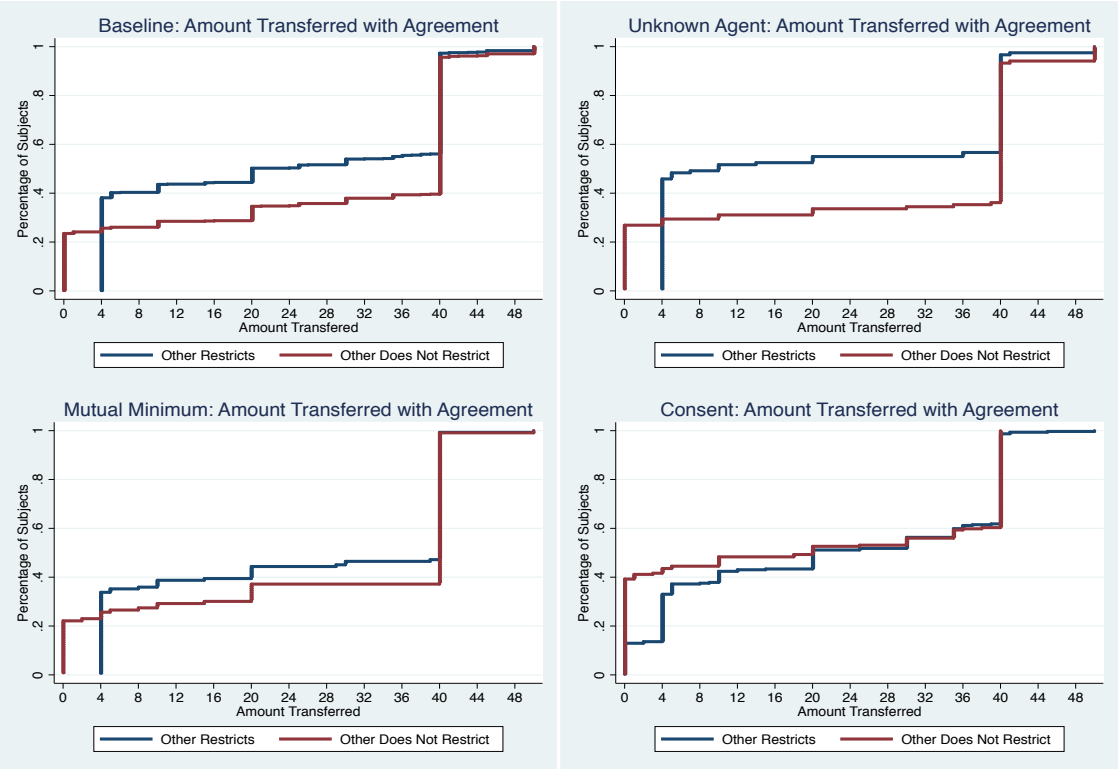
Figure 6: Amount Transferred with and without a Restriction (Agreement)



Our non-parametric tests find evidence of a cost of control in the Baseline and Unknown Agent treatments, (i.e. both the asymmetric treatments where control is imposed on a single player). In the Baseline treatment, the average transfer is 28.9 without a restriction, but transfers decrease to only 23.7 with an agreement (subject-level: $p = 0.10$, session-

level: p = 0.06). Similarly, in the Unknown Agent treatment the average transfer is 31.9 without a restriction, and 22.8 with a restriction (subject-level: p = 0.04, session-level: p = 0.07). However, this hidden cost is eliminated in the Mutual Minimum treatment, where the decrease in transfers of 2.3 units is not significant under either test (p > 0.20 for both). Furthermore, in the Consent treatment, average transfers *increase* from 19.8 without a restriction to 22.2 with a restriction (subject-level: p = 0.05, session-level: p = 0.08).

Figure 7: Distribution of Transfers with and without a Restriction (Agreement)



We also find a reversal in the behavioral response associated with control between treatments, both in the frequency of very small transfers of 4 units or less, and the frequency of transferring the agreed-upon 40 units. Figure 7 presents the cumulative distribution of transfers in each treatment when subjects have an agreement. In the Baseline treatment, the frequency of transfers of 4 or less increases from 26% without the restriction to 38% with the restriction (subject-level: p = 0.03, session-level: p = 0.03), and a decrease in the frequency of transferring 40 from 56% without the restriction to 41% with the restriction (subject-level: p < 0.01, session-level: p = 0.17). Similarly, in the Unknown Agent treatment, transfers of 4 units or less increase from 29% to 46%

(subject-level: $p = 0.03$, session-level: $p = 0.02$) and transfers of 40 decrease from 57% to 40% (subject-level: $p = 0.06$, session-level: $p = 0.05$). In both treatments, imposing control shifts the whole distribution to the left. In the Mutual Minimum treatment, the differences are much smaller and are not statistically significant: the frequency of small transfers increases from 26% to 34% and the frequency of transferring 40 units decreases from 62% to 52% ($p \geq 0.20$ for all tests). By contrast, in the Consent treatment, asking for control shifts the distribution to the *right* for all transfers below 20. The frequency of transfers of 4 or less *decreases* when the other subject asks for the restriction, from 44% to 33% (subject-level: $p = 0.06$, session-level: $p = 0.04$), while the frequency of transferring 40 remains approximately the same, moving from 40% to 37% ($p > 0.20$ for both tests).

We find essentially the same pattern with a regression analysis, presented in Table 2. Subject fixed effects and standard errors clustered by subject are used in all specifications. Columns (1) to (3) specify the amount transferred as the dependent variable, while Columns (4) to (6) use an indicator for transferring 4 or less, and Columns (7) to (9) use an indicator for transferring exactly 40 units.

In addition to using the full data set (columns (1), (4) and (7)) and just the first treatment of a session (columns (2), (5) and (8)), the third specification for each dependent variable (columns (3), (6) and (9)) includes only subjects who asked for the agreement in at least 8 out of 10 periods in both treatments (in both the baseline treatment and in whatever other treatment they played). This restriction avoids the possibility that differential selection into the agreement between treatments might generate our results.

Table 2: Transfers with an Agreement

	Transfer			Transfer <= 4			Transfer = 40		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Unknown Agent	2.724		3.001	-0.0425		-0.0532	0.0662		0.0649
	(1.975)		(2.039)	(0.0360)		(0.0376)	(0.0432)		(0.0447)
Mutual Minimum	0.160		-0.957	-0.00953		0.0158	0.0610		0.0378
	(1.813)		(1.854)	(0.0463)		(0.0487)	(0.0510)		(0.0538)
Consent	-8.639***		-8.892***	0.187***		0.169***	-0.132***		-0.156***
	(1.470)		(1.834)	(0.0391)		(0.0478)	(0.0399)		(0.0503)
Other Restricted in	-2.901***	-2.245*	-3.152**	0.0641***	0.0761***	0.0721***	-0.0719***	-0.0772**	-0.0791***
Baseline	(1.074)	(1.144)	(1.232)	(0.0213)	(0.0248)	(0.0237)	(0.0249)	(0.0352)	(0.0277)
Other Restricted in	-4.763*	-2.979	-4.628*	0.132**	0.0257	0.139**	-0.165***	-0.0514	-0.180***
Unknown Agent	(2.511)	(5.144)	(2.476)	(0.0552)	(0.0982)	(0.0599)	(0.0580)	(0.112)	(0.0610)
Other Restricted in	-2.334	1.541	-2.321	0.0474	-0.0349	0.0467	-0.0773	0.0174	-0.0863
Mutual Minimum	(1.930)	(2.689)	(2.116)	(0.0487)	(0.0688)	(0.0522)	(0.0561)	(0.0794)	(0.0620)
Other Restricted in	3.874***	2.008	4.330**	-0.125***	-0.0493	-0.145***	0.0310	0.00980	0.0499
Consent	(1.481)	(1.997)	(1.865)	(0.0451)	(0.0509)	(0.0553)	(0.0432)	(0.0699)	(0.0552)
First Treatment	5.779***		5.353***	-0.164***		-0.148***	0.188***		0.182***
	(0.796)		(0.888)	(0.0195)		(0.0211)	(0.0202)		(0.0224)
Constant	24.66***	29.14***	24.99***	0.371***	0.231***	0.361***	0.422***	0.607***	0.436***
	(0.818)	(0.455)	(0.895)	(0.0173)	(0.00996)	(0.0181)	(0.0186)	(0.0129)	(0.0201)
Observations	2653	1333	2056	2653	1333	2056	2653	1333	2056
Number of Subjects	443	410	306	443	410	306	443	410	306
R-squared	0.067	0.007	0.067	0.081	0.014	0.075	0.085	0.008	0.087

*** p < 0.01, ** p < 0.05, * p < 0.10. Standard errors clustered at the subject level reported in parentheses. The sample is restricted to observations where there was an agreement, and for the Unknown Agent and Mutual Minimum treatments only observations where the principal had the opportunity to restrict the agent are included. In columns (2), (5) and (8) the sample is further restricted to only the first treatment of a session. In columns (3), (6) and (9) only subjects who requested the agreement in at least 80% of periods for both treatments are included. All specifications include subject fixed effects. The dependent variable in columns (1) to (3) is the transfer of the agent, in columns (4) to (6) it is an dummy variable that equals one if the transfer was less than or equal to 4, in columns (7) to (9) it is a dummy variable that equals one if the transfer was equal to 40.

There is a significant cost to imposing control in both the Baseline and Unknown Agent treatments: agents transfer less on average to a principal who has imposed control. This cost of control is associated with a large behavioral response as agents are more likely to transfer 4 units or less and are less likely to transfer 40 units when control is imposed than when it is not. The coefficient for the restriction in the Mutual Minimum treatment is not significant in any specification, nor does it maintain a consistent sign. For the Consent treatment, we find that the restriction *increases* the average transfer and *decreases* the frequency of transferring 4 units or less. Our results are statistically weaker when we look only at the first treatment of the session (columns (2), (5) and (8)), however the results for the Baseline treatment stay at least marginally significant ($p=0.05$), and results for the Unknown Agent and Consent treatments maintain their sign. Restricting the data to subjects who demand the agreement with high frequency in both treatments (columns (3), (6) and (9)) does not change our results, suggesting the difference in the impact of the restriction between treatments is not driven by a selection effect.[1] Overall, we find that imposing control is detrimental to the Principal in the Baseline and Unknown Agent treatments, has no effect in the Mutual Minimum treatment, and is beneficial in the Consent treatment.

Result 3: When there is an agreement, the cost to the Principal of imposing control depends on the treatment. Control is costly in the Baseline and Unknown Agent treatments. This cost is eliminated in the Mutual Minimum treatment and is reversed in the Consent treatment.

[1] We also run a specification in the Consent treatment where we separately control for the agent wanting the restriction, the principal wanting the restriction and both wanting the restriction (full regression results are available from the authors on request). We find that there is a negative effect of the agent asking for the restriction if the other subject did not ($\beta = -5.57$, $p = 0.026$) — suggesting that the agent may be punishing the principal, possibly interpreting the principal's failure to ask for a restriction as a signal that the principal was intending to make a low transfer if the principal had instead ended up as the agent. We find no significant effect of only the principal requesting the restriction ($\beta = -2.61$, $p = 0.182$), but a strong positive effect if the principal joined the agent in requesting the restriction ($\beta = 10.81$, $p < 0.001$). This strengthens our result that control is beneficial, as there is both a positive effect of controlling and a negative effect of failing to control an agent who wants the restriction. Additionally, this "punishment" effect by agents helps explain the low average transfer shown in the "other does not restrict" bar in Figure 6. If neither subject requests the restriction the average transfer is 26.97, which is comparable to the other treatments.

IV.5 Who Responds Negatively To Control

Because we observe all subjects playing the role of the Principal in the Baseline treatment, we can use a subject's frequency of imposing control when a Principal in the Baseline treatment as a measure of their attitude towards control. This attitude towards control may affect how subjects respond to having control imposed upon them. For example, subjects who see control as a signal of distrust may be reluctant to impose control others and may react more negatively to being controlled. Conversely, subjects who see control as a reasonable precaution may prefer to restrict others and may not respond negatively to being controlled.

In the Baseline treatment, the median subject imposed control in 2/3 of periods as a Principal. To identify whether there is a different response for subjects with high and low usage of control, we estimate separate coefficients for the restriction in each treatment for subjects above and below the median usage. The results are reported in Table 3.

We find results that are quite reasonable across the treatments. In the Baseline treatment we find a "hidden cost of control" only among agents who used control infrequently as Principals. For this group, being restricted as an Agent led to an estimated decrease of 4.5 units, a 10 percentage point increase in the likelihood of making a transfer of 4 units or less, and a 12 percentage point decrease in the likelihood of transferring 40 units. By contrast, subjects in the Baseline treatment who used control frequently as a Principal had essentially zero response to the restriction as an Agent. In the Unknown Agent treatment we find a negative but insignificant effect of control on transfers for both groups of subjects, however subjects who infrequently restricted as Principals had a significant increase in the frequency of transfers of 4 or less and a significant decrease in the likelihood of transferring 40 when restricted as Agents. In the Mutual Minimum treatment we find somewhat insignificant results for all subjects, although subjects who used control frequently have directionally more positive reactions to being controlled. In the Consent treatment, the positive effect of the restriction was only observed among subjects who used the restriction frequently — for these subjects transfers increased by an estimated 6.8 units and the frequency of transfers of 4 or less decreased by 18 percentage points. Subjects who used the restriction infrequently have essentially a zero response to the restriction in the Consent treatment. Overall, the pattern of results suggests that there is important heterogeneity in how subjects perceived the restriction, with usage of the restriction as a Principal being correlated with more positive reactions to the restriction as an Agent.

Table 3: Effect of Subject Behavior as Principal in Baseline Treatment

Panel A: Amount Transferred

VARIABLES	Baseline (1)	Unknown Agent (2)	Mutual Minimum (3)	Consent (4)
Other Restricted & Used Restriction < 2/3 in Baseline	-4.498*** (1.550)	-2.785 (6.252)	-2.984 (3.561)	-0.167 (2.015)
Other Restricted & Used Restriction >= 2/3 in Baseline	-0.216 (1.696)	-3.406 (4.936)	-2.296 (2.660)	6.831*** (2.272)
Constant	27.25*** (0.616)	28.93*** (1.948)	27.41*** (1.190)	19.05*** (0.920)
Observations	1641	239	255	518
Number of Subjects	429	123	127	140
R-squared	0.012	0.008	0.013	0.029

Panel B: Transfer less than or equal to 4

VARIABLES	(5)	(6)	(7)	(8)
Other Restricted & Used Restriction < 2/3 in Baseline	0.102*** (0.0293)	0.196** (0.0916)	0.0676 (0.0932)	-0.0225 (0.0676)
Other Restricted & Used Restriction >= 2/3 in Baseline	-0.00238 (0.0326)	0.0401 (0.114)	0.0670 (0.0751)	-0.183*** (0.0661)
Constant	0.300*** (0.0118)	0.324*** (0.0387)	0.265*** (0.0325)	0.438*** (0.0283)
Observations	1641	239	255	518
Number of Subjects	429	123	127	140
R-squared	0.014	0.047	0.010	0.031

Panel C: Transfer equal to 40

VARIABLES	(9)	(10)	(11)	(12)
Other Restricted & Used Restriction < 2/3 in Baseline	-0.115*** (0.0395)	-0.301*** (0.101)	-0.138 (0.111)	-0.0227 (0.0645)
Other Restricted & Used Restriction >= 2/3 in Baseline	-0.0105 (0.0343)	-0 (0.118)	-0.106 (0.0782)	0.0769 (0.0661)
Constant	0.512*** (0.0135)	0.547*** (0.0406)	0.631*** (0.0359)	0.361*** (0.0277)
Observations	1641	239	255	518
Number of Subjects	429	123	127	140
R-squared	0.014	0.079	0.031	0.006

*** p < 0.01, ** p < 0.05, * p < 0.10. Standard errors clustered at the subject level reported in parentheses. The sample is restricted to observations where there was an agreement, and for the Unknown Agent and Mutual Minimum treatments only observations where the principal had the opportunity to restrict the agent are included. All specifications include subject fixed effects. The dependent variable in panel A is the transfer of the agent, in panel B it is an dummy variable that equals one if the transfer was less than or equal to 4, in panel C it is a dummy variable that equals one if the transfer was equal to 40.

Result 4: Subjects who imposed control more often as Principal in the Baseline treatment had a more positive reaction to being controlled as an Agent in the Baseline and Consent treatments. The cost of control results in the Baseline treatment is observed only in subjects who rarely impose control.

One concern with interpreting the results in the baseline condition is that subjects switch between playing as a principal and as an agent over the course of the 10 rounds in the Baseline condition. We are tempted to interpret these results as supportive of a story in which subjects who inclined to impose control as a Principal respond less negatively (or more positively) to control as an Agent. This interpretation would allow for a prescriptive suggestion that Principals can control Agents who themselves use control in settings where they are a Principal (e.g. a CEO could feel comfortable controlling middle managers who are observed to control their agents). However, an alternative explanation of Result 4 is that subjects who respond negatively to control eventually learn to avoid using it. To show that the former interpretation is still valid, we conduct a similar analysis but divide subjects by whether they chose to impose control the first time they were a Principal in the treatment being analyzed rather than whether they used control more than 2/3 of the time in the Baseline treatment. We then look only at behavior as an Agent in all subsequent rounds of that treatment. We replicate the results above and so can assert that subjects who are observed to use control as a Principal respond more favorably toward control when they are subsequently an Agent.

V. Additional Experiments

V.1 Control when there is no opportunity for an Agreement

In the Baseline treatment, we observe a cost of control when there is an agreement between the principal and the agent and a benefit of control when there is no agreement. This result suggests that control is detrimental when there is a strong pro-social norm governing behavior as a result of the agreement, and control is beneficial when there is a weak norm governing behavior due to failure to make an agreement.

Falk and Kosfeld (2006) find that control is detrimental in the "middle case" where no agreement was possible, and therefore only the default norm (or background norm) governs behavior. To investigate whether we would find this result in our data, we ran additional sessions with a *No Agreement Allowed Treatment* where no agreement

opportunity was presented.[2] The No Agreement Allowed treatment is the same as the Baseline treatment, except that subjects were not given the opportunity to make an agreement. We conducted an additional 5 sessions, with 94 subjects, in which we ran the No Agreement Allowed treatment followed by the Baseline treatment. We had subjects always play the Baseline treatment second so that subjects would not have been previously exposed to the agreement when playing in the No Agreement Allowed treatment.

Figure 8: Average Transfer with and without Restriction when no opportunity for an Agreement

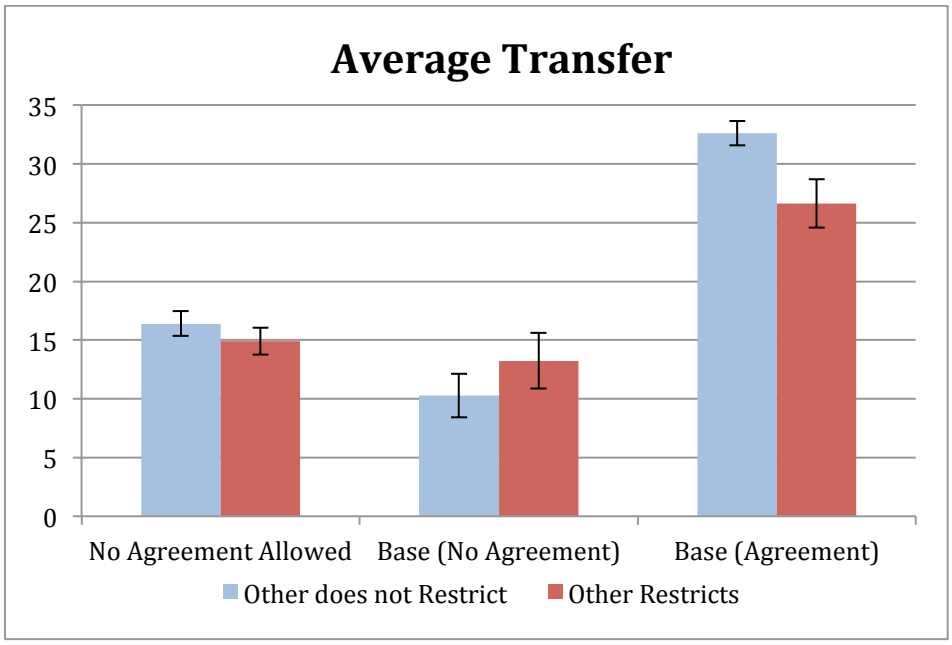


Figure 8 reports the average transfer with and without a restriction in the No Agreement Allowed treatment and compares it to behavior in the Baseline treatment for these new sessions.[3] In the No Agreement Allowed treatment, transfers decrease slightly from 16.4 when control is not imposed to 14.9 when control is imposed, and the difference is not significant ($p > 0.20$ for both subject-level and session-level tests). Similarly, while the fraction of subjects transferring 4 or less increases from 30% to 36% in response to

[2] Note that this design looks more like the Falk and Kosfeld (2006) experimental set-up; however, in our experiment the roles of Principal and Agent are randomly assigned across a number of rounds.

[3] We obtain essentially the same results if we instead compare the No Agreement Allowed treatment to the Baseline treatment when played first in a session in the original 25 sessions discussed in the preceding section.

control, the difference is not significant (subject-level: $p = 0.18$, session-level: $p > 0.20$). These small and insignificant differences contrast with the results in Falk and Kosfeld (2006), which finds that imposing a minimum transfer of 5 leads to a decrease in transfers from 25.1 to 12.2, and an increase in the fraction of subjects transferring 5 or less from approximately 20% to approximately 50%.

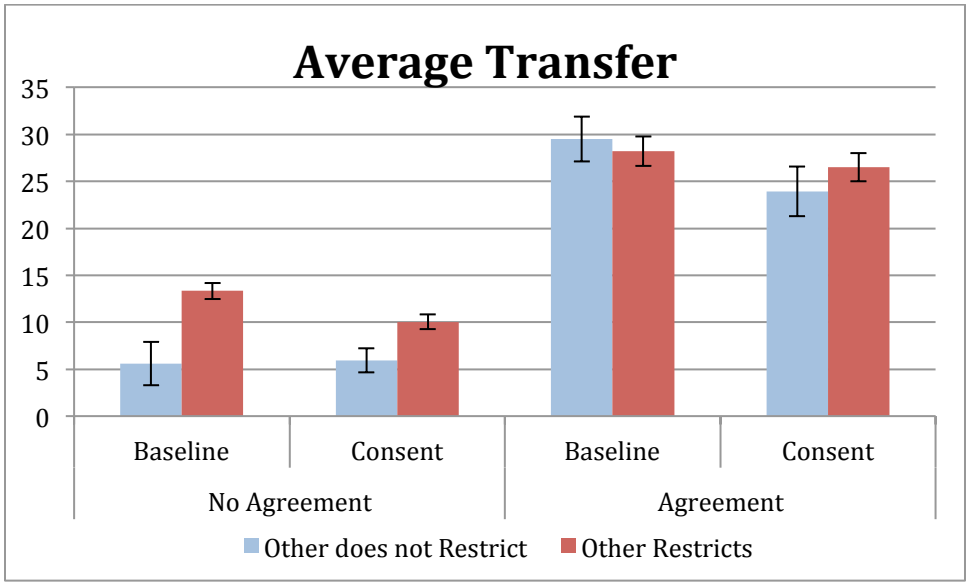
In the absence of control, however, transfers are much higher in Falk and Kosfeld's data than in ours, suggesting that there may be a difference in the background norm for their subject pool as compared to our subject pool. Taken together, we may expect a cost of control whenever there is a strong norm, either by default as in Falk and Kosfeld (2006) data, or due to a specific agreement as in our data.

V.2 Restrictions with a Higher Minimum Transfer

In our main experiment, we find that control is harmful to the principal in the Baseline treatment when there is an agreement. When control is imposed, there is a large behavioral response where subjects are much less likely to transfer 40 units and much more likely to transfer the minimum of 4 units. We observe a cost of control since this behavioral response is large relative to the benefit from raising transfers that would have been 0 to 3 units to the minimum of 4 units. A natural question is whether this cost of control persists if the principal has a somewhat more powerful controlling ability (e.g. if he has a better monitoring technology). To test the impact of more effective control, we ran 5 additional sessions with 94 subjects of the Baseline treatment and Consent treatments in which control required a minimum transfer of 10 units rather than 4 units. Figure 9 shows the average transfer in each treatment.

As in our main experiment, we find that when there is no agreement, imposing the powerful restriction (or asking for a restriction in the Consent treatment) leads to higher average transfers from the other subject. In the Baseline treatment, the average transfer increases from 5.6 to 13.3 (subject-level: $p < 0.01$, session-level: $p = 0.13$), while in the Consent treatment the average transfer increases from 5.6 to 10.3 (subject-level: $p < 0.01$, session-level: $p = 0.13$).

Figure 9: Average Transfer with and without High Restriction



When there is an agreement, however, we no longer find a cost of control in the Baseline treatment. When control forces a minimum transfer of 10, the average transfer decreases slightly from 29.5 without control to 28.2 with control, but the difference is not significant ($p > 0.20$ for both non-parametric tests). There is still evidence of a behavioral response: the fraction of transfers of 10 or less increases from 29% to 40% in the presence of control (subject-level: $p = 0.09$, session-level: $p = 0.38$), while the fraction of transfers equal to 40 decreases from 57% to 46% (subject-level: $p = 0.08$, session-level: $p = 0.25$). In this case, however, the benefit of the increase due to the binding minimum outweighs the decrease in larger transfers.

In the Consent treatment, we find that the restriction is somewhat beneficial for the Principal, increasing average transfers from 25.0 to 26.0, however this difference is not statistically significant ($p > 0.20$ for both tests). Overall these results suggest that a cost of control should only be a primary concern when the Principal's ability to monitor and control the Agent is relatively limited.

VI. Conclusion

In this paper, we investigate the conditions under which a Principal experiences a cost of control in that they are made worse off by imposing control on an Agent. The paper identifies settings where there is a significant cost of control, settings where agents exhibit a negative behavioral response associated with being controlled but in which that

response is offset by the increase in transfers up to the minimum imposed by control, and settings in which control operates as expected by standard theory, generating no behavioral response and leading to higher transfers on average.

In our experiment, subjects play a simple principal-agent game and have the opportunity to make a non-binding agreement before the roles of Principal and Agent are assigned. In the Baseline treatment, Principals can unilaterally impose a minimum transfer on the Agent, while additional treatments add symmetry to the contracting relationship by making the minimum binding on whichever subject is the agent, and by requiring both parties to agree to the minimum for it to be imposed (i.e. control is imposed bilaterally). We also conduct two additional treatments that (1) remove the agreement stage and (2) increase the minimum transfer associated with control.

Principals in our experiment face a cost of control only when four conditions are simultaneously met: (1) there is a prosocial norm motivating the agent; (2) control is imposed unilaterally and has an asymmetric effect on the agent; (3) control is weak, in that it cannot induce significant effort from the agent; and (4) the agent is does not use control himself when acting as a principal. In all other cases, we find that the principal is no worse off from imposing control.

A number of factors therefore mitigate the risk of a cost of control. First, no cost of control is observed when control has symmetric impact (i.e. it affects both players rather than just one). Second, no cost of control is observed if the minimum established by control is high enough. Third, no cost of control is observed if the agent being controlled has previously chosen to impose control on others. Furthermore, the cost is reversed so that the principal receives a benefit of control when the parties do not reach an agreement — and therefore the norms governing behavior are weak — or when the players decide on control bilaterally and control has symmetric effects on both parties.

These results are summarized in Figure 10, which highlights the types of settings where principals should be concerned about a cost of control.

Our results suggest that principals and firms should be most concerned about a cost of control when they have established a strong norm with the agent (e.g. via an informal agreement or corporate culture), when their monitoring and control technology is weak, and when their relationship with the agent is highly asymmetric (e.g. in an employment context or a supply chain setting with a dominant party). Control may be less problematic when both parties are on a more even footing (e.g. a joint venture). Firms may be able to diminish the cost if they can also credibly restrict their own bad actions or if they can allow agents to consent to the control.

Figure 10: Effect of Imposing Control

<table>
<tr><td></td><td colspan="3">Control is imposed asymmetrically</td><td>Control is imposed symmetrically</td></tr>
<tr><td rowspan="3">Control Imposed Unilaterally</td><td></td><td>No Prosocial Norm</td><td colspan="2">Prosocial Norm</td><td rowspan="3">None

• Supply chain contracts with a dominant party where the dominant party also makes contractual promises (e.g. strategic supplier relationships)</td></tr>
<tr><td rowspan="2">Control is weak</td><td rowspan="2">Positive

• Labor contracts in settings with usually low effort

• Supply chain contracts with a dominant party but with low visibility into upstream suppliers</td><td>Agent controls as a principal</td><td>Agent does not control as a principal</td></tr>
<tr><td>None

• Labor contracts with middle managers who use control</td><td>Negative

• Labor contracts in settings with usually high effort or good corporate culture</td></tr>
<tr><td>Control is strong</td><td>Positive

• Labor contracts where monitoring is effective in settings with usually low effort</td><td colspan="2">None

• Labor contracts where monitoring is effective in settings with unusually high effort or good corporate culture</td></tr>
<tr><td>Control Imposed with Bilateral Consent</td><td colspan="3">N/A</td><td>Positive

• Joint ventures

• Supply chain contracts where there is no dominant party</td></tr>
</table>

Note: This figure breaks down the results of the main experiment and additional experiments into three main categories differentiating between settings where we see a robust cost of control (labeled “negative”), no strong evidence of a cost of control (labeled “none”), and a benefit of imposing control (labeled “positive”)

Our results also reconcile results from a number of papers that have attempted and failed to replicate the results in Falk and Kosfeld (2006). One of the main results described above is that there is only a cost to control or a behavioral response to control in settings where there is a pro-social norm generating high effort when control is not imposed. Falk and Kosfeld (2006) observe subjects giving very high effort in their data when control is not imposed, hence there is a large scope for a behavioral response and little benefit in terms of raising low actions up to the minimum. In the papers that fail to replicate the cost of control, like Schnedler and Vadovic (2011), there are many low actions in the absence of control so there is a relatively large benefit to control which more than offsets the behavioral response they observe. Ploner et al. (2012) find both directionally negative and directionally positive effects of control depending on the subject pool, and this change in subject pool is also likely affecting the default norm for agent behavior. In addition, Schnedler and Vadovic (2011) find that when control is more "legitimate" (e.g. when a principal is preventing the agent from taking from the principal's endowment) control is less costly; having a good reason for control may make it seem less distrustful and may work in a similar was as allowing control to be imposed mutually. Furthermore, Kessler and Leider (2012) fail to find a behavioral response of imposing control, but their game is quite symmetric on the dimensions described here, namely control affects both parties symmetrically and in most of their treatments it is imposed bilaterally.

This paper used a laboratory experiment to identify boundaries of a behavioral phenomenon. This study shows when principals should be concerned about using the contractual tools available to them and when they can rely on the standard models to describe agent behavior. Future research should help map these experimental findings to other environments of interest. In doing so, future research will help principals interpret the boundaries of this phenomenon in their organizations. For example, we have introduced a number of relevant questions for a principal. How symmetric does the impact of control need to be for it to mitigate the cost of control? Does the contracting process have to be fully bilateral for control to have a benefit or can similar gains be made when an agent provides more limited "buy-in" to the imposition of control? Is the default norm high enough in a particular setting for a principal to be concerned about using control? We have provided a framework that suggests which questions the principal should be asking, and future research can help answer those questions even more concretely.

VII. References

Allen, Franklin and Douglas Gale, (1992). "Measurement Distortions and Missing Contingencies in Optimal Contracts." *Economic Theory*. 2(1): 1-26.

Barkema, Harry G., (1995). "Do Top Managers Work Harder when They Are Monitored?" *Kyklos*, 48(1): 19–42.

Bernheim, B. Douglas and Michael Whinston, (1998). "Incomplete Contracts and Strategic Ambiguity." *American Economic Review*. 88 (4), 902-932.

Bowles, Samuel and Sandra Polania-Reyes (2012). "Economic Incentives and Social Preferences: Substitutes and Complements." *Journal of Economic Literature*. 50(2), 368-425.

Charness, Gary and Martin Dufwenberg, (2006). "Promises and Partnerships." *Econometrica*, 74 (6): 1579-1601.

Coase, Ronald (1937). "The Nature of the Firm." *Economica*. 4, 386-405.

Deci, Edward, (1971) "Effects of externally mediated rewards on intrinsic motivation," *Journal of Personality and Social Psychology*, 18, 105-115.

Deci, Edward, (1975) *Intrinsic Motivation*. Plenum Press, New York and London.

Deci, Edward, Richard Koestner, and Richard Ryan, (1999). "A Meta-analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation." *Psychological Bulletin*, 125 (6), 627-668.

Dufwenberg, Martin, Simon Gächter, and Heike Hennig-Schmidt, (2011). "The framing of games and the psychology of play." *Games and Economic Behavior*, 73 (2), 459–478.

Ellingsen, Tore and Magnus Johannesson. (2004). "Promises, Threats and Fairness." *The Economic Journal*, 114(495), 397-420.

Falk, Armin and Michael Kosfeld, (2006). "The Hidden Cost of Control." *American Economic Review*, 96(5), 1611-1630.

Fehr, Ernst and Simon Gächter, (2002). "Do Incentive Contracts Undermine Voluntary Cooperation." *IEW Working Paper* No. 34.

Fehr, Ernst and List, John A., (2004). "The Hidden Costs and Rewards of Incentives." *Journal of the European Economic Association*, 2(5), 743-771.

Fehr, Ernst and Bettina Rockenbach, (2003). "Detrimental Effects of Sanctions on Human Altruism." *Nature*, 422 (6928), 137-140.

Fehr, Ernst, Alexander Klein, and Klaus M. Schmidt, (2007). “Fairness and contract design.” *Econometrica* 75:121–54.

Fischbacher, Urs, (2007). “z-Tree: Zurich Toolbox for Ready-made Economic experiments.” *Experimental Economics*, 10(2), 171-178.

Frey, Bruno S. (1993). “Does Monitoring Increase Work Effort? The Rivalry with Trust and Loyalty.” *Economic Inquiry*, 31(4): 663–70.

Frey, Bruno S. (1994) "How Intrinsic Motivation is Crowded in and Out," *Rationality and Society*, 6(3), 334-352.

Frey, Bruno, S., Felix Oberholzer-Gee and Reiner Eichenberger, (1996) "The old lady visits your backyard: a tale of morals and markets," *Journal of Political Economy*, 104(6), 1297-1313.

Frey Bruno S. and Felix Oberholzer-Gee, (1997) "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out," *American Economic Review*, 87(4), 746-755.

Gneezy, Uri and Aldo Rusticini, (2000a). “Pay Enough or Don’t Pay at All.” *Quarterly Journal of Economics*, 115(2), 791-810.

Gneezy, Uri and Aldo Rusticini, (2000b). “A Fine is a Price.” *Journal of Legal Studies*, 29(1), 1-18.

Gneezy, Uri, Stephan Meier and Pedro Rey-Biel, (2011). “When and Why Incentives (Don't) Work to Modify Behavior.” *The Journal of Economic Perspectives*, 25(4), 191-209.

Hagemann, Petra, (2007). “What’s in a Frame? Comment on: The Hidden Costs of Control,” *Unpublished manuscript*, University of Cologne.

Hart, Oliver (1995). *Firms, Contracts and Financial Structure*. Oxford University Press, Oxford.

Holmstrom, Bengt and Paul Milgrom, (1991). “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design,” *Journal of Law, Economics, & Organization*, Vol. 7, 24-52.

Kruglansky, Arie, Irith Freedman, and Gabriella Zeevi, (1971). "The effects of extrinsic incentives on some qualitative aspects of task performance," *Journal of Personality*, 39, 606-617.

Kruglansky, Arie, Sarah Alon, and Tirtzah Lewis, (1972) "Retrospective misattribution and task enjoyment," *Journal of Experimental Social Psychology*, 8, 493-501.

Lepper, Mark R., David Greene, and Richard E. Nisbett, (1973) "Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis," *Journal of Personality and Social Psychology*, 28, 129-137.

Lepper, Mark R. and Greene, David, (1978), *The Hidden Costs of Reward: New Perspectives in the Psychology of Human Motivation*. Lawrence Elbaum Associates, Publishers; John Wiley and Sons.

Lemieux, W. Thomas, Bentley MacLeod, and Daniel Parent, (2009), "Performance Pay and Wage Inequality" *The Quarterly Journal of Economics*, 124(1), 1-49.

Ploner, Matteo, Katrin Schmelz, and Anthony Ziegelmeyer, (2012). "Hidden Costs of Control: Four Repetitions and an Extension," *Experimental Economics,* 15(2), 323-340.

Simon, Herbert (1981). *The Sciences of the Artificial*. MIT Press, Cambridge, MA

Sliwka, Dirk, (2007). "Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes." *American Economic Review*, 97 (3), 999-1012.

Spier, Kathryn E. "Incomplete Contracts in a Model with Adverse Selection and Exogenous Costs of Enforcement." *RAND Journal of Economics*, 1992, 23, 432-443.

Tirole, Jean, (1999). "Incomplete Contracts: Where do we stand?" *Econometrica*, 67 (4), 741-781.

Titmuss, Richard M., (1970) *The Gift Relationship*. Allen and Unwin, London.

Vanberg, Cristoph, (2008). "Why do people keep their promises? An experimental test of two explanations." *Econometrica*, 76(6), 1467-1480.

Vernon, Guy (2003). "Comparative work organization, managerial hierarchies and occupational classification", *Employee Relations*, 25(4), 389 – 404.

Williamson, Oliver (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press, New York.

Williamson, Oliver (1985). *The Economic Institutions of Capitalism*. Free Press: New York.