Finding the Hidden Cost of Control*

By Judd B. Kessler† and Stephen Leider‡

This Draft: September 6, 2012

ABSTRACT

A large and growing literature has demonstrated that explicit incentives imposed by principals — such as enforceable contracts, fines, and monitoring regimes — can have detrimental effects on agent behavior. This literature, however, focuses on what *can* happen and provides little guidance about when detrimental effects will arise and when they should be incorporated into economic models. We investigate the hidden cost of control, a result in which adding enforceability to a principal-agent contract decreases agent effort. We show that principals are only harmed by imposing control when: (1) they have previously established a prosocial norm with their agent; (2) they impose control unilaterally on the agent; (3) the control is weak, in that it cannot induce significant effort from the agent; and (4) the agent is of a type that does not use control himself when acting as a principal.

* The authors thank Rachel Croson, Elena Katok, Lise Vesterlund and seminar participants at Penn State, MIT and Wharton for helpful comments. The authors would also like to thank the staff at the Wharton Behavioral Lab.

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

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

I. Introduction

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

Recent literature has demonstrated, however, that these incentive and control strategies may come at a cost.[1,2] 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).[3] 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).[4] This last result has been referred to as the "hidden cost of control" and arises when a principal receives less effort from an agent, and therefore earns less profit, when she controls the agent by taking away his most opportunistic actions.[5]

Despite this extensive research on the potential costs associated with extrinsic incentives and control, both are very common in principal-agent settings. In the United States, 37% of individuals have some form of pay-for-performance incentives (Lemieux et al. 2007).

---

[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 Eichenberger1996 and Frey and Oberholzer-Gee 1997).

[3] 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).

[4] In addition, 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 principals receiving less effort on average when they impose control (Hagemann 2007, Schnedler and Vadovic 2011, Ploner et al. 2011). We describe these papers in detail in Section II.

[5] Note that we define the "hidden cost of control" as arising when the principal earns less profit from controlling an agent than from giving the agent free rein. Falk and Kosfeld (2006) use the term to title their paper, and we infer this meaning as being in the spirit of their paper. Other authors have used the term to refer to the behavioral response that arises when some agents lower effort in response to being controlled; we simply refer to that behavioral response as a "behavioral response".

Control strategies are also prevalent and have been becoming more common. Average supervisor-employee ratios increased in the non-farm economy for many developed countries from 1960 to 1995 (Vernon 2003). Given the prevalence of extrinsic incentives and control strategies in the world, the recent findings on their perverse consequences poses a bit of a puzzle and suggests a need to better understanding when these "hidden costs" will arise.

More generally, demonstrations that extrinsic incentives and control *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 these models is to understand when — that is, in which environments — these perverse effects of incentives and control will be severe enough that the standard model will fail to predict behavior. Identifying the boundaries of the cost of control will help models better predict behavior and provide guidance for principals and firms.[6]

In this paper, we investigate the impact of using control in a principal-agent relationship by way of a laboratory experiment in which an agent has the opportunity to take a costly action that benefits a principal. In the experiment, the action space of the agent may be restricted to eliminate the most opportunistic actions of the agent. 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 (i.e. unilaterally) 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. 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). By allowing for a richer contracting environment, we are able to identify the effect of control on agents who have established a prosocial norm and those who have not.

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. By varying the contracting rules we are able to turn off the hidden cost of control so that the principal does as well or better by imposing control. In some of these treatments we still observe a behavioral response in which

[6] 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.

agents respond with less effort when controlled than when not controlled. By varying the contracting rules further, we able to turn off this behavioral response such that control works as predicted by standard theory.

As the title of the paper and the summary in the preceding paragraph suggest, the hidden cost of control arises only in certain environments. We find that principals are only harmed by imposing control when the following conditions are met: (1) they have previously established a prosocial norm with their agent; (2) they impose control unilaterally on the agent; (3) the control is weak, in that it cannot induce significant effort from the agent; and (4) the agent is of a type that does not use control himself when acting as a principal. In all other cases, we find that the principal is no worse off from using control. We also find that the principal is better off imposing control when either: (1) she has been unable to establish a prosocial norm with 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 hidden cost of control has 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.

II. Related Literature

Principal-agent relationships play an important role the labor market and in markets with supply chains. Firms use a variety of incentive and monitoring strategies to manage the agency problem. Consequently, significant attention has been placed on the effects of contracts in these settings. A striking fact is that many contracts are much simpler (and

[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) makes 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. The paper 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.

more incomplete) than standard theory would predict. Traditional explanations 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. Incomplete contracts may also be suboptimal but necessary due to multitasking problems (Holmstrom and Milgrom 1992). Additionally, some authors have given theoretical justifications for why incomplete contracts may in fact 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 incentives or monitoring and control can lower performance, for example by crowding out intrinsic motivation (see Deci et al. 1999 for a survey). We discuss several prominent examples here. Gneezy and Rustichini (2000a) asked subjects to perform one of two tasks — taking an IQ test or raising money for charity — with either zero, small, or large performance-based monetary incentives. The authors find that providing small incentives leads to worse performance than providing zero incentives. They argue that introducing incentives in an incomplete contracting setting changes the information and the implicit reference point, undermining the perception that reasonable performance is "due" in response to the flat compensation. Gneezy and Rustichini (2000b) investigate a field setting where a group of day care centers introduced a monetary fine for parents who arrived late to pick up their children. The authors find that while late arrivals were infrequent before the fine, the increased substantially when the fine was introduced, and continued even after the fine was removed. They argue that introducing the fine added new information to the incomplete contracting setting, in this case defining the upper bound on the penalties for late arrival.

Several principal-agent experiments have similarly found that incentives can have detrimental effects on performance. Fehr and Gächter (2002) study a buyer-seller game where the buyer can use either a flat price contract, which requires trusting the seller, or an incentive contract with a penalty for low quality. The authors find that incentive contract leads to lower average quality and the almost total elimination of voluntary cooperation. They argue that models of fairness and reciprocity can explain their results, although framing also plays a role (equivalent incentives described as a bonus lead to much smaller reductions in voluntary cooperation). Fehr and Rockebach (2003) find similar results from a trust game setting. In the base "trust" treatment, the first mover can request a specific back-transfer, but cannot impose any consequences for low returns. In the "trust with punishment" treatment, the first mover can also choose to specify a fine for an insufficient back transfer. The authors find that return transfers are lower when the fine is imposed than in the trust condition, but higher when the fine is available but not used. Fehr and List (2004) also find similar results from experiments using students and CEOs. CEOs return less when the punishment is imposed than when it was not available, while both CEOs and students increase their return transfers when the punishment was available but not used. The authors argue that refraining from punishment is seen as highly trusting, which induces reciprocal trustworthiness.

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. The paper presents the results of a set 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 chooses whether to impose control on her agent by limiting his action space; different treatments allow for control of different strengths. Results demonstrate that the principal does better, in terms of 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 main result in Falk and Kosfeld (2006). For example, Hagemann (2007) finds a non-significant negative effect of adding control in an attempted replication of Falk and Kosfeld (2006) using the strategy method.[8] The paper finds that when agents are not controlled, their effort is higher 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 is a function of experimenter demand effect. Similarly, Schnedler and Vadovic (2011) find a behavioral response to control, but fail to replicate the hidden cost of control. The authors find that average effort is directionally higher when control is used than when it is not. In a set of attempts to replicate Falk and Kosfeld (2006), Ploner et al. (2012) finds both directionally negative and directionally positive effects of 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

[8] However, with only 30 agents in each treatment, Hagermann’s experiment may be underpowered to identify a treatment effect in the baseline case. She 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.

papers fail to replicate the hidden cost of control, they generally do replicate the behavioral response 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 offer effort above the minimum when they are not controlled and choose the minimum effort allowed when they are controlled.

Research in other settings, however, has observed the expected beneficial effect of control mechanisms without an offsetting behavioral response to imposing control. For example, Kessler and Leider (2012) had subjects play two-person games, including public good games, in which effort was personally costly but collectively beneficial. The authors found that adding an enforceable minimum (i.e. control) to a pre-game contract had no effect on, or increased, effort in three of four games and decreased effort in only one. For some of the games, adding the enforceable minimum did not generate a behavioral response of lowing actions above the minimum. The results persist when the minimum is imposed bilaterally (i.e. when both agents agree to it) and when one agent imposes it unilaterally.

The games in Kessler and Leider (2012) differ from a principal-agent setting on an important dimension: in Kessler and Leider (2012), both players are making effort choices a 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 in Kessler and Leider (2012). Our experimental design, 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 on the impact of control as well as whether control is imposed unilaterally or bilaterally.

Most of the previous studies that find a crowding out effect for incentives or control find that effort and quality are generally quite high in the absence of incentives, suggesting that strong norms govern behavior in these settings. On the other hand, when social norms are weak, we may expect that the benefits of control will outweigh any behavioral response to control.

To increase the likelihood that our subjects will perceive a strong norm governing transfers, we will use the pre-play agreement mechanism from Kessler and Leider (2012). In that paper, we found that players when players could make a non-binding agreement to play the first best action, a norm for performance was established and effort and profits were higher than when no agreement was available. Other studies have found similar benefits of unilateral promises in holdup games (Ellingsen and Johannesson 2004), trust games (Charness and Dufwenberg 2006), and dictator games (Vanberg 2008). Dufwenberg et al. (2011) provide a theoretical model predicting what agreements should form when they are binding contracts or non-binding informal agreements, and test their model with a lost wallet game. They find that binding contracts are predominantly 50-50 splits, while informal agreements lead to higher payoffs for the second mover, which one can think of as the agent.

III. Experimental Design

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

In each round of the principal-agent game, the agent (called "Player A" in the instructions) started with 120 experimental units (EUs) that were worth \$0.05 each. The agent could transfer these units to the principal (called "Player B" in the instructions) and any units transferred were doubled for the principal. Consequently, the payoffs for the principal agent game were as follows:

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

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

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 make a transfer x of any number between 0 and 120. If control was imposed, it restricted transfers to be at least 4 EUs, so agents could make a transfer of any number between 4 and 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 units (i.e. x=40) if they ended up being the agent. 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 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 each other had chosen and whether an agreement was made.

All rounds began with subjects choosing whether or not to have an agreement, after which the instructions differed by treatment. There were four treatments in the experiment, which differed in the symmetry of control and whether control was imposed unilaterally or bilaterally.

Figure 1 displays the four treatments as a function of whether the control was imposed on one player (i.e. "single minimum") or on both players (i.e. "mutual minimum") 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").

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

In order to add symmetry to control, for some of the treatments we did not reveal which of the two players was the agent (and which of the two was the principal) until after control had been imposed. By allowing control to be imposed before the agent was revealed, we had the opportunity to run treatments with a variety of control rules.

In the *Unknown Agent Treatment*, before we assigned the roles of principal and agent, we randomly selected one of the players, and that player had the opportunity to impose control on *the other player* if that 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. 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, while the identity of the agent is still *unknown*, and one of the players imposes control *unilaterally*.

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

Finally, in the *Consent Treatment*, before we assigned the roles of principal and agent, we allowed both players to suggest whether or not control should be imposed on *whichever player* became the agent. The decision to impose control was made in the same way as the agreement to transfer 40 EUs. Namely, each player could suggest the restriction or no restriction, and only if both players suggested the restriction would control be imposed. After the decisions, the players were told whether each other had suggested control. Then the roles of principal and agent were assigned. If the players had both suggested control the agent was restricted to transfer between 4 and 120. If at least one of the players had not suggested control then there was no restriction, and the agent could choose any transfer between 0 and 120. Notice that for the *Consent Treatment*, the minimum is imposed on *both* players, while the identity of the agent is still *unknown*, and both players imposes control *bilaterally*.

Figure 1: Experimental Treatments

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

Subjects 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 assigned subjects to a new partner in each round and, as mentioned above, we randomly assigned the roles of principal and agent in each round. The design therefore allows us to observe the same subject playing as both a principal and an agent, so we can investigate how propensity to use control as a principal affects the likelihood of reacting negatively to control as an agent.

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, 148 subjects also 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.

Figure 2 here

We begin by examining subjects' preferences for having an agreement or a restriction. Figure 2 displays the frequency at which subjects suggested the agreement in each treatment, as well as the frequency at 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. This led subjects to form an agreement in 65 to 75% of periods.

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. Overall, the differences between the treatments are relatively small, however the difference between the cases with and without an agreement suggests that we should examine the effect of restrictions separately between the agreement and no agreement case.

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. Figure 3 shows the impact of the agreement on the average amount transferred by agents divided by whether the other subject suggested the agreement.[9] Note that we split the data based on the other subject's choice to hold fixed the agent's own preferences for an agreement (which could be correlated with the overall propensity to transfer)[10].

Figure 3 here

Across all four treatments we find that having an agreement leads to substantial increases in the average amount transferred. To test statistically for differences, we use non-parametric permutation tests on choices aggregated both at the subject level and the session level.[11] 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 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. Without an agreement only 2 to 12% of subjects transfer 40, while with an agreement between 31 and 47% of subjects transfer 40 ($p < 0.01$ 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: Making an agreement significantly increases transfers in all treatments.**

IV.3 Effect of Control without an Agreement

We now look at whether imposing control by implementing a restriction on the transfer of the agent leads to a decrease in the amount transferred by the agent. We begin by looking at pairs who do not have an agreement.

We first look at the average amount transferred, which demonstrates whether there is a "hidden cost of control", that is whether the principal is helped or hurt by suggesting the restriction. As with the agreement, we split the data based on whether the other person

---

[9] 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..

[10] 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 had rejected the agreement (making up approximately 20% of the observations).

[11] We use unpaired permutation tests for subject-level data (because some subjects only have observations with an agreement or only without an agreement) and paired permutation tests for session-level data.

suggested the restriction. Since partners were randomly assigned in each period, whether the other person asked for the restriction is exogenous. Notice that in the Control treatment, some of the agents may not have had the restriction in place if they did not ask for the restriction themselves. The benefit of this strategy, however, is the agent having asked for the restriction is not selected based on the agent's own preference for the restriction.[12] Figure 4 shows the average amount transferred in each treatment.

Figure 4 here

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. Only the increase in the Baseline treatment is significant under our non-parametric tests (subject-level: $p = 0.07$, session-level: $p < 0.01$; $p > 0.20$ for all other treatments). This 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 is beneficial to the principal in the Baseline treatment.

However, looking at the average transfer could mask two opposing effects if the reduction in very low transfers due to the minimum transfer was offset by a decrease in large transfers. Hence, we also want to examine whether there was any behavioral response to 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. However, if subjects who would otherwise transfer more than the minimum react negatively to the restriction by transferring only the minimum amount, than this fraction should increase. Figure 5 plots the cumulative distribution of transfers by each treatment, based on whether or not the other subject wanted the restriction.

Figure 5 here

In all treatments, the vast majority of subjects transfer only a small amount: between 64 and 75% 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 insignificant (subject-level: $p = 0.15$;

[12] In the other three treatments the other player wanting the restriction is equivalent to having the restriction. As noted above, 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.

session-level: p = 0.16). Recall that in the Baseline treatment we found a significant increase in the minimum, so to the extent that there is a behavioral response to control using the minimum, it is swamped by the effect of increasing transfers of less than 4 up to transfers of 4.

We also conduct a regression analysis of the individual 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 "hidden cost" of the restriction, 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. To avoid any results being driven by any contamination between the two treatments in a session, Columns (2) and (4) use data from only the first treatment in a session.

Table 1 here

Our regression results largely confirm our non-parametric analysis in finding generally positive effects of imposing a restriction. For the amount transferred, we find that 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 very small transfers. Hence, we conclude that when there is no agreement between the Principal and Agent, imposing a restriction is not costly for the Principal.

> **Result 2: When there is no agreement, there is no cost — and sometimes a benefit — of imposing a minimum transfer.**

IV.4 Effect of Control with an Agreement

While we did not find evidence of a hidden cost when there is no agreement, most transfers were quite low even without a restriction. However, we demonstrated previously that average transfers are much higher when subjects formed an agreement. Thus, we should expect a greater likelihood of finding a behavioral response, and a "hidden cost", of the restriction under an agreement. Figure 6 displays the average amount transferred by Agents when they have an agreement with the Principal split by whether the Principal imposed control on the Agent (or suggested the restriction in the baseline treatment).

Figure 6 here

We find evidence for a “hidden cost of control” in both the Baseline and Unknown Agent treatments. 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 contract 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 here

We also find a reversal in the behavioral effect of the restriction between treatments, both in the frequency of very small transfers of 4 or less, and the frequency of transferring the agreed-upon 40 units. Figure 7 presents the cumulative distribution of transfers in each treatment. 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 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 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 is essentially unchanged (40% vs 37%, $p > 0.20$ for both tests).

Table 2 here

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) use 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.

In addition to using the full data set and just the first treatment of a session, the third specification for each dependent variable includes only subjects who asked for the agreement in at least 8 out of 10 periods in both treatments. This restriction avoids the possibility that differential selection into the agreement between treatments might generate our results. By all three measures of behavior, there is a significant cost to imposing control with the restriction in both the Baseline and Unknown Agent treatments: agents transfer less on average, are more likely to transfer 4 or less, and are less likely to transfer 40. 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 or less. Our results are statistically weaker when we look only at the first treatment, however the results for the Baseline treatment stay at least marginally significant, 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 does not change our results, suggesting the difference in the impact of the restriction between treatments is not driven by a selection effect. 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.**

Because we observe all subjects playing the role of the Principal in the Baseline treatment, we can use a subject's frequency of imposing the restriction as a measure of their attitude towards control, which may affect how they react to having control imposed on them. For example, subjects who see control as a signal of distrust may be reluctant to restrict others, and may react more strongly to being restricted. Conversely, subjects who see control as a reasonable precaution may prefer to restrict others and may not be affected by others controlling them. 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.

Table 3 here

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 in transfers of 4.5 units, a 10% increase in the likelihood of making a transfer of 4 or less, and a 12% decrease in the likelihood of transferring 40 units. By contrast, subjects in the

Baseline treatment who frequently used the restriction 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 average transfers increased by an estimated 6.8 units and the frequency of transfers of 4 or less decreased by 18%. 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.

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 "hidden cost of control" in the Baseline treatment is observed in subjects who control less often.

V. Additional Experiments

V.1 Falk Kosfeld Replication

In our main experiment we observe a "hidden cost of control" in the Baseline treatment only when there is an agreement between the Agent and the Principal. This suggests that control is detrimental when there is a strong prosocial norm governing behavior (due to the agreement), and control is beneficial when there is a weak norm governing behavior (due to a failure to make an agreement). Falk and Kosfeld (2006) find that control is detrimental in the intermediate 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 setting, we ran additional sessions with a *Replication Treatment* that more closely matches Falk and Kosfeld's design. The Replication treatment is the same as the Baseline treatment, except that subjects were not given the opportunity to make an agreement. We conduct an additional 5 sessions, with 94 subjects, which included this treatment. In each session we ran the Replication treatment followed by the Baseline treatment, so that subjects would not have been exposed to the agreement when playing in the Replication treatment.

Figure 8 here

Figure 8 reports the average transfer with and without a restriction in each treatment. Again, we find in the Baseline treatment that control increases transfers when there is no agreement, and decreases transfers when there is an agreement. In contrast, the Replication 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 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's paper, 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%. However, in the absence of control 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 compared to ours. Consequently, we may expect there to be a hidden cost of control whenever there is a strong norm, either naturally (as in Falk and Kosfeld's 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 — that is, the cost from the decrease in transfers above the minimum overwhelms the benefit of the increase in transfers where the minimum is binding. A natural question is whether this hidden cost of control persists if the Principal has a more powerful controlling ability (e.g. if he has a better monitoring technology). To test the impact of more effective control, we ran 6 additional sessions, with 114 subjects, of the Baseline and Consent treatments in which the restriction required a minimum transfer of 10 units (rather than 4 units). Figure 9 shows the average transfer in each treatment.

Figure 9 here

As in our main experiment, we find that when there is no agreement imposing a restriction leads to higher average transfers. In the Baseline treatment, the average transfer increases from 5.1 to 13.5 (subject-level: $p < 0.01$, session-level: $p = 0.06$), while in the Consent treatment the average transfer increases from 6.0 to 10.0 (subject-level: $p < 0.01$, session-level: $p = 0.13$). However, when the minimum transfer is 10 we do not find a cost of control in the Baseline treatment under an agreement. In this case the average transfer decreases slightly from 29.3 to 28.8, but the difference is not significant ($p > 0.20$ for both). There is a small behavioral response — the fraction of transfers of 10 or less increases from 29% to 35% in the presence of control (subject-level: $p = 0.11$, session-level: $p = 0.44$), while the fraction of transfers equal to 40 decreases from 60% to 50% (subject-level: $p = 0.04$, session-level: $p = 0.16$). 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 23.9 to 26.5, however this difference is not significant ($p > 0.20$ for both tests). Overall these results suggest that the hidden cost of control is of primary concern when the Principal's ability to monitor and control the Agent is relatively limited.

VI. Conclusion

In this paper we investigate the circumstances under which a principal experiences a "hidden cost" from controlling an agent. In our experiment, subjects play a simple principal-agent game and have the opportunity to make a non-binding agreement before roles are determined. In our Baseline treatment, principals can unilaterally impose a minimum transfer on the agent, while additional treatments make the minimum mutually binding on whichever subject is the agent, and require both parties to agree to the minimum. We also conduct two additional treatments that either remove the agreement stage or increase the minimum transfer.

Principals in our experiment face a hidden cost of control in the Baseline and Unknown Agent treatments when the principal and agent had made an ex ante agreement. However, several factors reduce or eliminate this hidden cost of control. First, principals benefit from control when the parties do not reach an agreement, and therefore the norms governing behavior are weak. Second, the cost is eliminated when the restriction applies mutually (to both players rather than just one). Furthermore, the cost is reversed, so that the principal benefits from control, when both players consent to control being imposed. Third, the cost is eliminated, even in the Baseline treatment with an agreement, if the minimum established by control is high enough. Finally, principals avoid the cost of control if they are paired with agents who themselves were willing to impose control on others.

Our results suggest that principals and firms should be most concerned about a hidden 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). Costs of 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 hidden cost if they can also credibly restrict their own bad actions or if they can allow agents to consent to the control.

VII. References

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.

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.

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.

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.

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.

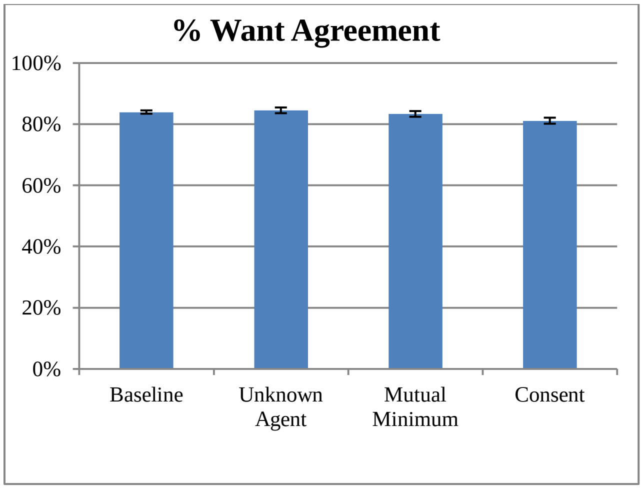
Vanberg, Cristoph, (2008). “Why do people keep their promises? An experimental test of two explanations.” *Econometrica*, 76(6), 1467-1480.

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.

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



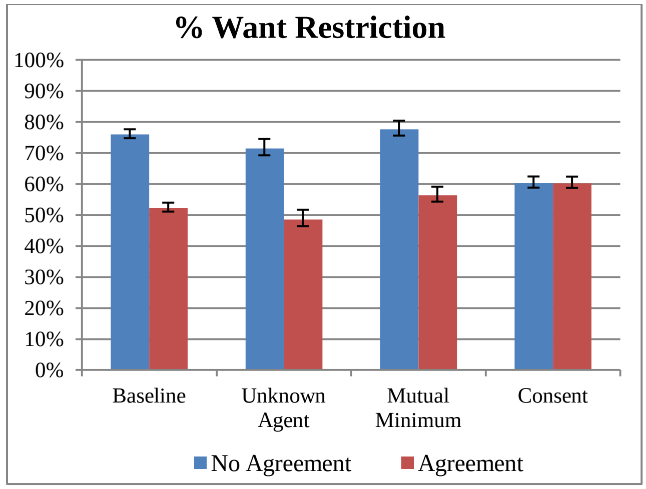


Figure 3: Effect of an Agreement on Average Transfers

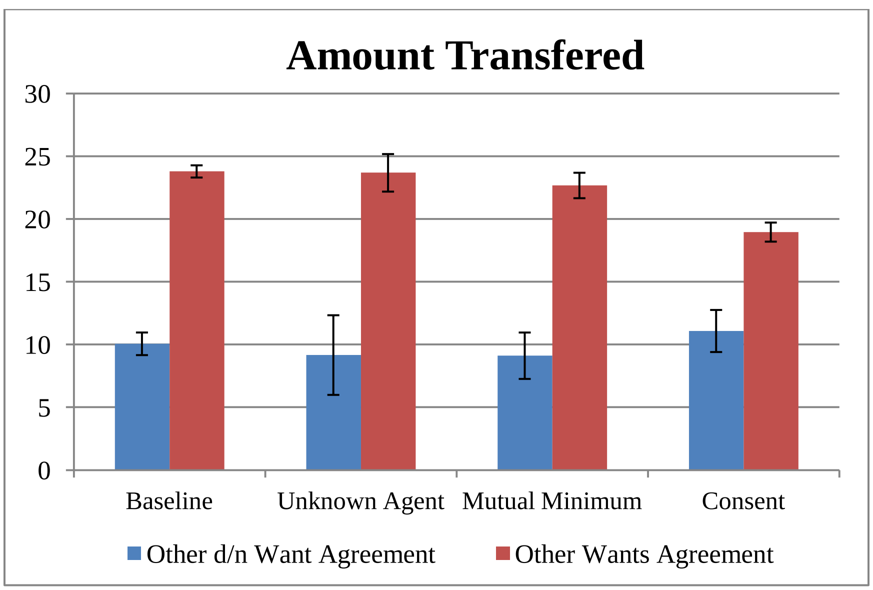


Figure 4: Amount Transferred with and without a Restriction (No Agreement)

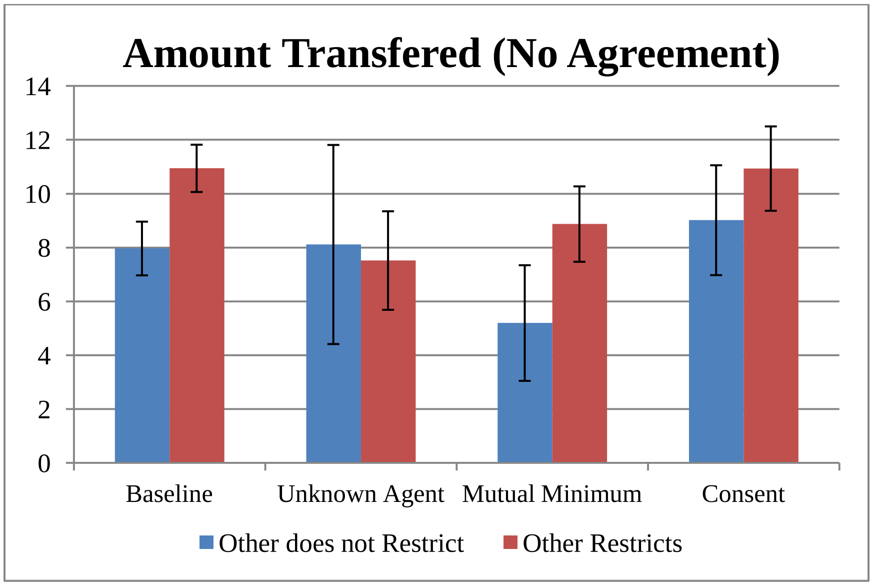


Figure 5: Distribution of Transfers with and without a Restriction (No Agreement)

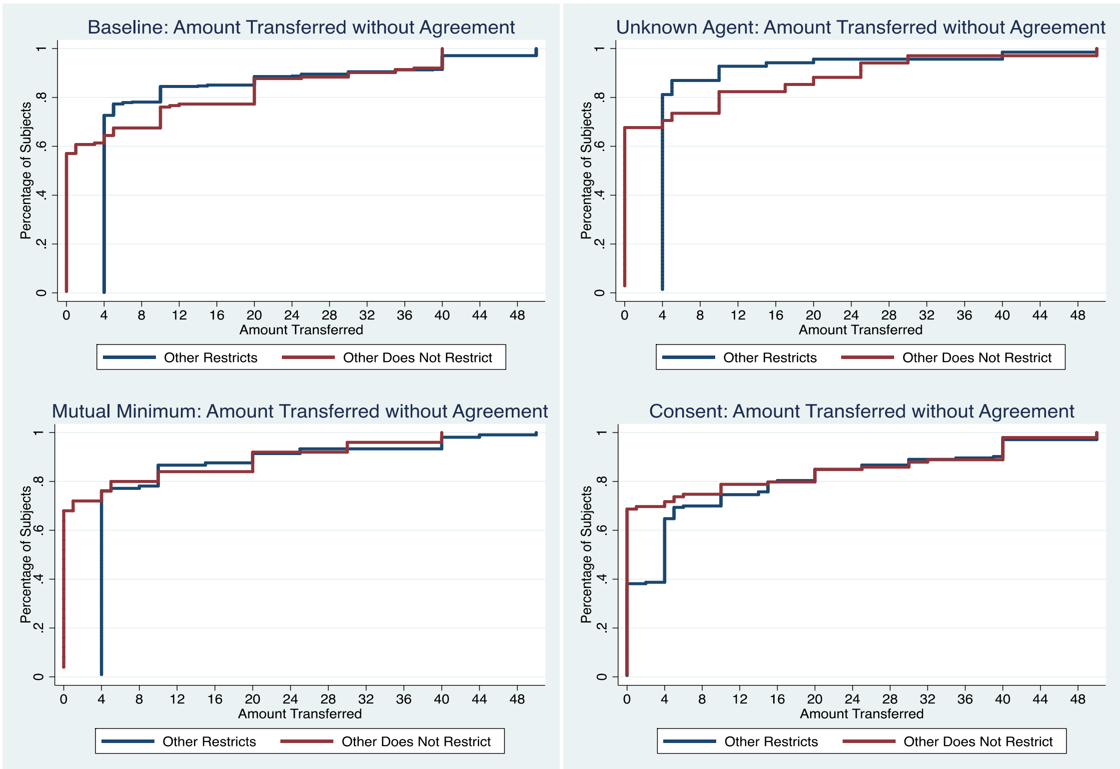


Figure 6: Amount Transferred with and without a Restriction (Agreement)

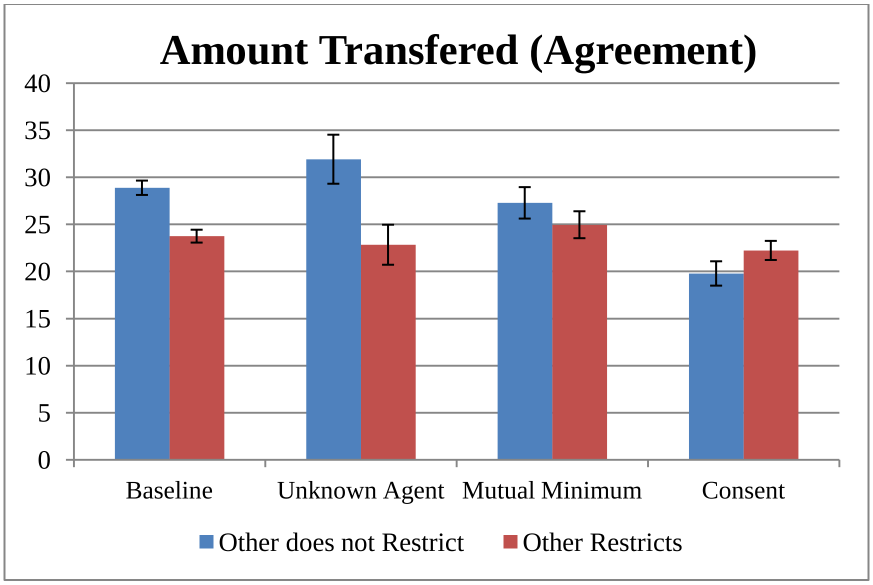


Figure 7: Distribution of Transfers with and without a Restriction (Agreement)

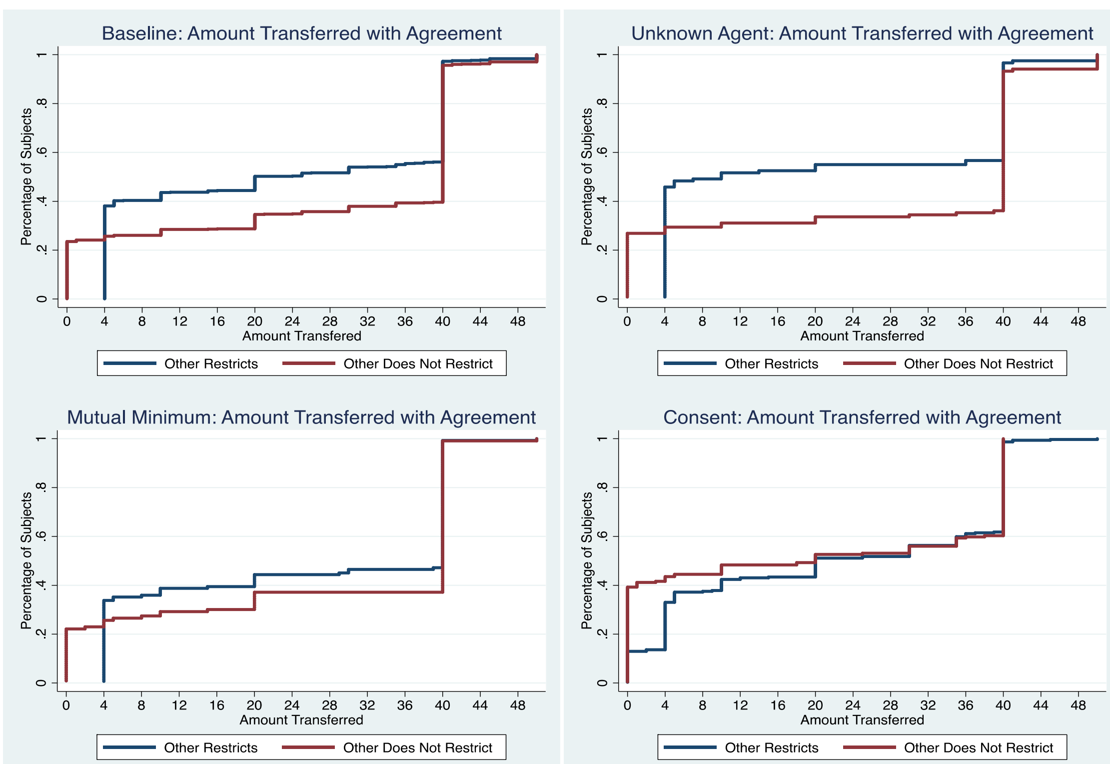


Figure 8: Average Transfer with and without Restriction (Falk-Kosfeld Replication)

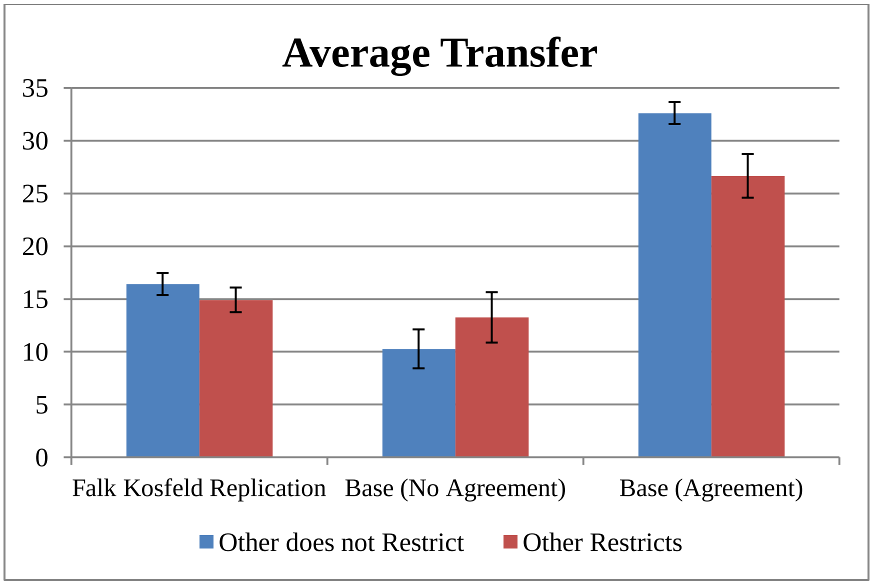


Figure 9: Average Transfer with and without High Restriction

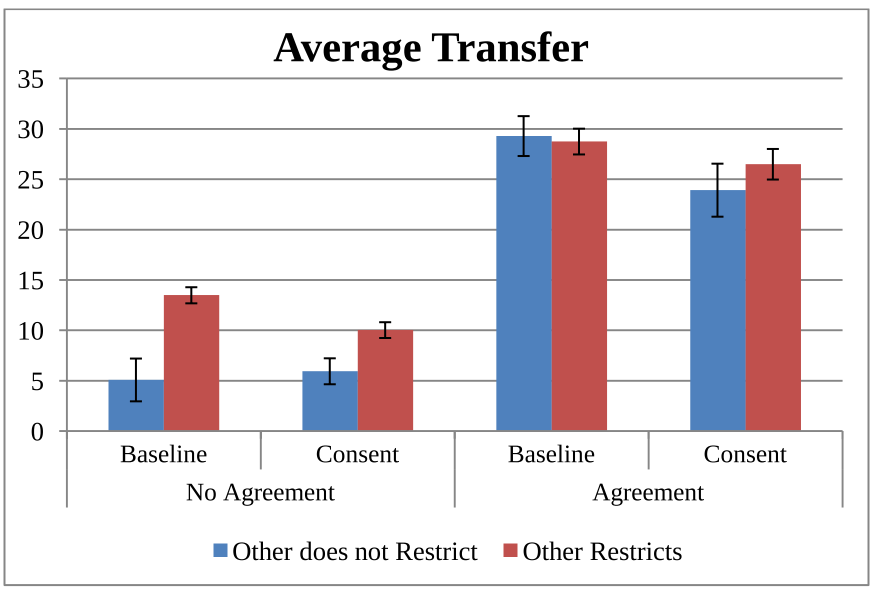


Table 1: Transfers without an Agreement

	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 wa the restricter 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.

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 Baseline	-2.901***	-2.245*	-3.152**	0.0641***	0.0761***	0.0721***	-0.0719***	-0.0772**	-0.0791***
	(1.074)	(1.144)	(1.232)	(0.0213)	(0.0248)	(0.0237)	(0.0249)	(0.0352)	(0.0277)
Other Restricted in Unknown Agent	-4.763*	-2.979	-4.628*	0.132**	0.0257	0.139**	-0.165***	-0.0514	-0.180***
	(2.511)	(5.144)	(2.476)	(0.0552)	(0.0982)	(0.0599)	(0.0580)	(0.112)	(0.0610)
Other Restricted in Mutual Minimum	-2.334	1.541	-2.321	0.0474	-0.0349	0.0467	-0.0773	0.0174	-0.0863
	(1.930)	(2.689)	(2.116)	(0.0487)	(0.0688)	(0.0522)	(0.0561)	(0.0794)	(0.0620)
Other Restricted in Consent	3.874***	2.008	4.330**	-0.125***	-0.0493	-0.145***	0.0310	0.00980	0.0499
	(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 was the restrictor 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.

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***	-2.785	-2.984	-0.167
	(1.550)	(6.252)	(3.561)	(2.015)
Other Restricted & Used Restriction >= 2/3 in Baseline	-0.216	-3.406	-2.296	6.831***
	(1.696)	(4.936)	(2.660)	(2.272)
Constant	27.25***	28.93***	27.41***	19.05***
	(0.616)	(1.948)	(1.190)	(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***	0.196**	0.0676	-0.0225
	(0.0293)	(0.0916)	(0.0932)	(0.0676)
Other Restricted & Used Restriction >= 2/3 in Baseline	-0.00238	0.0401	0.0670	-0.183***
	(0.0326)	(0.114)	(0.0751)	(0.0661)
Constant	0.300***	0.324***	0.265***	0.438***
	(0.0118)	(0.0387)	(0.0325)	(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***	-0.301***	-0.138	-0.0227
	(0.0395)	(0.101)	(0.111)	(0.0645)
Other Restricted & Used Restriction >= 2/3 in Baseline	-0.0105	-0	-0.106	0.0769
	(0.0343)	(0.118)	(0.0782)	(0.0661)
Constant	0.512***	0.547***	0.631***	0.361***
	(0.0135)	(0.0406)	(0.0359)	(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 was the restrictor 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.