Journal of Law, Economics, and Organization Advance Access published June 20, 2016

Procedural Fairness and the Cost of Control

Judd B. Kessler*
University of Pennsylvania

Stephen Leider
University of Michigan

A large and growing literature has demonstrated that imposing control on agents has the potential to backfire, leading agents to withhold effort. Consistent with principles of procedural fairness, we find that the way in which control is imposed—in particular whether control is imposed symmetrically on both principals and agents and whether both parties have a say in whether control is imposed—affects how agents respond to control. In our setting, control leads agents to withhold effort only when procedural fairness concerns are ignored and control is imposed unilaterally with an asymmetric effect on the agent. (*JEL* C7, C9, L2, M5)

1. Introduction

In many principal–agent, supply chain, and partnership relationships, parties can take opportunistic actions that harm their counterparts. Agents can shirk, principals can refuse to pay earned discretionary bonuses; suppliers can fail to deliver high quality, buyers can fail to pay on time; and partners can withhold effort or resources from each other. To prevent these opportunistic actions, parties often implement control mechanisms such as establishing contractual restrictions that require perfunctory performance. Standard economic theory suggests that these contractual tools can more effectively align the interests of parties and avoid inefficient outcomes, and control is regularly used in principal–agent settings.[1]

In many cases, these restrictions are instituted without much disruption. Firms successfully implement codes of conduct for employees that set standards for attendance and task completion. Suppliers meet qualification and certification procedures and undergo auditing of their processes.

*The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA. Email: judd.kessler@wharton.upenn.edu.

The authors thank Rachel Croson, Florian Englmaier, Elena Katok, Muriel Niederle, Lise Vesterlund and seminar participants at Penn State, MIT, Wharton, CESifo, and Columbia for helpful comments. The authors would also like to thank the staff at the Wharton Behavioral Laboratory.

1. Between 1960 and 1995 average supervisor–employee ratios in the non-farm economy increased for many developed countries (Vernon 2003).

The Journal of Law, Economics, and Organization, Vol. 0, No. 0
doi:10.1093/jleo/eww009

Partnership agreements establish the norms and expectations of how each partner should contribute to the partnership.[2] Recent literature has demonstrated, however, that control strategies may come at a cost. In principal–agent settings, control that restricts an agent's action has been shown to backfire, leading to *more* opportunistic behavior rather than *less*. The intuition behind these results is that imposing control in a contract demonstrates distrust and may lead controlled parties to withhold effort (e.g., Frey 1993; Barkema 1995; Falk and Kosfeld 2006).[3]

We propose that the fairness of the process by which control is imposed affects whether or not control will backfire. This hypothesis is based on an extensive literature on "procedural justice" and "procedural fairness," which has established that the fairness of the process by which an outcome is determined has a significant impact on the reception of the outcome and the ongoing relationship between the parties, with fairer processes leading to greater trust and better performance (see Thibaut and Walker 1975, 1978; Lind and Tyler 1988 for seminal works). For example, Konovsky and Cropanzano (1991) find that employees who perceived a drug-testing regime to be imposed in a more procedurally fair way had higher job satisfaction, lower turnover intentions, and better job performance.

In particular, following the principles of procedural fairness identified by Leventhal (1980), we hypothesize that control may backfire in principal–agent settings because these settings are highly asymmetric (violating Leventhal's "consistency" rule) and generally do not give agents a voice in whether or how control is imposed (violating Leventhal's "representativeness" rule). We would expect control to be more effective when control is applied symmetrically and when affected parties have a voice in imposing control. This hypothesis speaks to patterns of how control is implemented in practice. In employment relationships, pay and work rules are often set for job categories rather than for individuals, and unions and joint consultative committees frequently help determine pay and work practices. Even without these formal structures, managers often seek input from

2. For an example, see the Accenture LTD partnership agreement cited in Kessler and Leider (2012).

3. The feeling of distrust might be amplified when the relationship is more personal as in Frey (1993) and Dickinson and Villeval (2008). These findings are related to work showing that 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). See also a rich literature in Psychology, which has shown that extrinsic incentives can undermine intrinsic motivations (see Lepper and Greene 1978; Deci 1971, 1975; Kruglansky et al. 1972), a notion which has been more recently studied in the economics literature (see e.g., Frey et al. 1996; Frey and Oberholzer-Gee 1997). A 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.

employees.[4] Similarly, while a firm may require that its suppliers only purchase from companies on its Approved Vendor List, it typically would also require that of its own buying agents. In addition to explaining how control is imposed in practice, understanding the relationship between procedural fairness and control can help guide how control can be implemented optimally.[5]

To explore the implications of procedural fairness on the efficacy of control, we conduct a series of laboratory experiments.[6] In our experiment, pairs of subjects interact anonymously in a one-shot game. In the game, one party (the agent) has the opportunity to take a costly action that benefits the other (the prinicpal). If control is in place, the most opportunistic actions are eliminated from the action space so that the agent is required to provide perfunctory performance.[7] Our design allows us to vary whether control is imposed asymmetrically on symmetrically and whether one party imposes control unilaterally or whether control is imposed by an agreement that gives voice to both parties.

In order to vary how control is imposed, our design randomly assigns subjects to play in the role of principal or agent. Control that affects the agent is sometimes imposed before and sometimes imposed after subjects learn who is which role. To observe behavior when control is imposed unilaterally and asymmetrically, in our baseline treatment we reveal the identities of the principal and agent first and then allow the principal to impose control on the agent.[8] To observe behavior in a more symmetric environment where control is applied consistently across individuals, we run a treatment in which one subject is randomly selected to impose control before the identities of the principal and agent are revealed, so that control affects whichever of the two subjects is later revealed to be the

4. In the 2011 UK Workplace Employment Relations Study, 52% of employees rated management as "very good" at seeking input in decisions, and 46% rated it as very good at responding to suggestions (van Wanrooy et al. 2013). In particular, 95% of managers consulted employees if laying off two or more workers, and 40% or more consulted on decisions such as changing work techniques, changing work organization, introducing performance pay, or addressing health and safety concerns.

5. For example, in a labor relationship, a requirement to arrive to work by 9 a.m. or face punishment could apply only to front-line employees or to managers as well. The rule could have been imposed by the managers or negotiated with a worker union. Similarly, in a supply chain relationship, enforceable deadlines might only control the delivery of goods or also apply to accounts receivables. These contractual requirements could be part of a non-negotiable contract to transact with the supplier or be negotiated between the parties.

6. Schnedler and Vadovic (2011) consider a similar notion of "legitimacy of control," however they focus on the issue of who control is targeting—for example, varying whether there is a chance control will be imposed on a computer player rather than a real agent—while we address the process by which control is imposed.

7. In our setting, "higher" actions are more beneficial to the other party. Controlling the agent is therefore imposing a minimum action. This was described to subjects as "restricting" the other player.

8. This treatment allows us to embed a design quite similar to the experiment run in Falk and Kosfeld (2006).

agent. Finally, to model an environment in which both parties have a voice in imposing control, we run a treatment in which both players must agree to have control for it to be imposed on whichever subject is later revealed to be the agent.

Our results are consistent with theories of procedural fairness. We find that control only backfires when it is imposed unilaterally and asymmetrically on the agent. Control becomes somewhat more effective when we allow it to be imposed symmetrically, and it is most effective when both parties have a voice in whether control is imposed.

The article proceeds as follows. Section 2 highlights related literature. Section 3 describes the experimental design. Section 4 motivates our main hypotheses. Section 5 presents our experimental results. Section 6 describes implications of our results for firm behavior, speaks to existing economic models for the cost of control, helps to reconcile previous experimental results on the cost of control, and concludes.

2. Related Literature

Leventhal (1980) identifies six principles of procedural justice, two of which highlight the importance of symmetry in the imposition of control. The "consistency rule" argues that procedures to determine payments and outcomes should be consistent across persons, and the "bias suppression rule" states that in a process the decision-makers should seek to be impartial, rather than advancing their own personal self-interest. In addition, "voice"—the ability of those affected by a decision process to participate and state their interests and desires—has been identified as an important contributor to perceptions of procedural justice, highlighting the value of giving parties that might be affected by control a role in determining whether it is imposed (Leventhal 1980 calls this "representativeness," see also Bies and Shapiro 1988 and Lind et al. 1990).

The literature has also shown that perceptions of procedural justice (or injustice) affect behavior and performance. Alexander and Ruderman (1987) show that procedural justice, specifically giving workers voice, increases trust in management. Lind et al. (1993) find that procedurally fair dispute resolution mechanisms (i.e., those with impartial authorities and participant voice) were more successful in leading participants to accept the outcomes of the arbitration. De Cremer and van Knippenberg (2002, 2003) show that voice for participants increased cooperation in social dilemmas, particularly when decision-makers do not privately benefit from their power. Finally, a meta-analysis by Cohen-Charash and Spector (2001) finds that across both laboratory and field studies, procedural justice was positively associated with job performance (including both effort and outcome metrics) and negatively associated with counterproductive work behavior (such as improper work, theft, and damaging equipment).

An economics literature on procedural fairness, surveyed by Frey et al. (2004), finds similar results. Random allocation procedures such as lotteries have received particular attention, with lotteries being perceived of as quite fair, particularly if they are "symmetric" in the sense of choosing outcomes favorable to each party with equal probability (Kahneman et al. 1986; Bolton et al. 2005). Research on how procedural fairness affects performance mirrors the psychology literature. Frohlich and Oppenheimer (1990) find that workers who had the ability to vote on a tax system were more productive over time than those who had a tax system imposed. Greenberg (1990) finds that fair processes led to less negative reactions to pay decreases, while Benz and Stutzer (2003) finds that voice in pay determination led to increased job satisfaction. Benz and Frey (2004) compare self-employed workers to workers in firm hierarchies—holding fixed pay, hours, and other factors that might influence enjoyment of a job—and finds that the more self-directed employees had higher job satisfaction.

A parallel literature has focused on the efficacy of control and when it can backfire (i.e., leading controlled agents to provide less effort rather than more). Falk and Kosfeld (2006) demonstrate that imposing control on agents—by eliminating their most opportunistic actions and forcing them to provide at least minimum perfunctory effort—can lead to worse outcomes for the principal. Falk and Kosfeld (2006) find both a "hidden cost of control," in which some agents who would have provided high effort when unconstrained display a behavioral response and provide less effort when controlled, and an "average cost of control," in which this behavioral response is so large that it swamps the beneficial effect of control of raising low actions and makes principals worse off from imposing control.[9] Research in other settings, however, has observed the beneficial effect of control mechanisms without the offsetting behavioral response (Kessler and Leider 2012).[10]

The results from these literatures suggest that more procedurally fair impositions of control will lead control to backfire less often and be more effective. The experimental design in this article, which is presented in detail in the next section, investigates this hypothesis by analyzing whether the procedure by which control is imposed impacts the efficacy of control.

9. Falk and Kosfeld (2006) use the term "hidden cost of control" to title their paper. These findings are similar in spirit to research showing that extrinsic incentives can crowd out intrinsic motivation (see Deci et al. 1999 and Gneezy et al. 2011 for surveys).

10. This research on control aims to explore why contracts are regularly left incomplete. Previous explanations of contractual incompleteness appeal to transaction costs (e.g., Coase 1937; Williamson 1975, 1985) or bounded rationality (e.g., Simon 1981). Other lines of research have suggested that complete contracts may signal negative information about the contract proposer (Allen and Gale 1992; Spier 1992), complete contracts may lead the agent to infer that a less prosocial norm prevails (Sliwka 2007), incompleteness creates strategic ambiguity that helps enforce implicit agreements (Bernheim and Whinston 1998), or that leaving contracts incomplete may be suboptimal but necessary given that agents are asked to multitask (Holmstrom and Milgrom 1991).

3. Experimental Design

In the experiment, subjects played an anonymous, one-shot transfer game a total of 20 times. Subjects were randomly matched with another subject in the laboratory in each round of the game.[11]

In 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 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:

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

$$\text{Principal ("Player B")}: \ \pi_{\mathrm{P}} = 2x,$$

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

If control (called "a restriction on Player A's transfer" in the instructions) was not imposed, agents could choose to transfer any amount x from 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.

The experiment has three main treatments and varies the procedural fairness of control by changing whether control asymmetrically affects one subject or symmetrically affects both subjects and whether control is imposed unilaterally or bilaterally. Table 1 displays the three treatments.

In the *Baseline Treatment*, the roles of principal and agent were assigned before control was imposed. After the principal and the agent were assigned their roles, the principal was given the option to impose control (the principal decided between: "No restriction" and "A restriction that Player A must transfer at least 4 EUs"). This choice was revealed to the agent who decided how many EUs to transfer, with the transfer restricted to be at least 4 EUs when control was imposed. Notice that for the *Baseline Treatment*, the minimum is imposed asymmetrically and unilaterally.

We add symmetry to control in the other two treatments by having control affect both subjects; we implement this by giving subjects the opportunity to impose control before they learn who is the agent.[12] In the

11. The experiment was run on z-Tree (Fischbacher 2007).

12. This mechanism is similar to Rawls' (1971) "veil of ignorance," where individuals must establish the rules of a society before knowing their roles. The "veil of ignorance" has been used experimentally to examine issues such as taxation and redistribution (Frohlich and Oppenheimer 1990; Sutter and Weck-Hannemann 2003; Krawczyk 2010; Schildberg-Hörisch 2010). Here, we are not interested in analyzing the impact of the veil of ignorance *per se*; instead, we use it as a design tool to create initial symmetry between the two subjects, which allows us to explore changes in procedural fairness. In addition, while this particular form of role uncertainty—in which two parties make agreements without knowing who will be the principal and who will be the agent—may be rare in practice, it sometimes does arise (e.g., when cofounders begin a company without knowing who will be CEO). Finally, the structure of our game in which control is determined ex ante and roles are assigned ex post can also be

Table 1. Experimental Treatments

		Symmetry of control	
		Asymmetric	Symmetric
Control imposed	Unilaterally	Baseline treatments	Consistency treatment
	Bilaterally		Voice treatment

Table 1 shows the main experimental treatments in the experiment.

Consistency Treatment, we randomly give one of the subjects the option to impose control on *whichever subject* became the agent. After the subject decided whether to impose control, we assigned the roles of principal and agent. If the subject had imposed control, whichever of the two subjects 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 *Consistency Treatment*, the minimum is imposed symmetrically but is also imposed unilaterally by one subject.

In both treatments discussed so far, one subject has the opportunity to impose control unilaterally. In our third treatment, both subjects have voice in the process that imposes control in that they must bilaterally agree on control for it to be imposed. In the *Voice Treatment*, before we assigned the roles of principal and agent, we allowed both subjects to suggest whether or not control should be imposed on *whichever subject* became the agent. In particular, each subject could either suggest that the restriction be in place or not suggest it. Only if both subjects suggested the restriction be in place was control imposed. After each subject made a decision, the subjects were told who suggested the restriction and whether the restriction was imposed. We then assigned the roles of principal and agent. If both subjects suggested the restriction, the agent was restricted to transfer between 4 and 120 EUs. If at least one of the subjects did not suggest the restriction, then the agent could choose to transfer any amount between 0 and 120 EUs. Notice that for the *Voice Treatment*, the minimum is imposed symmetrically on both subjects and is imposed bilaterally, giving voice to both subjects.

thought to model very common scenarios in which one of two transacting parties can benefit or hurt the other, but it is not apparent who will have that opportunity at the time the parties contract. For example, either an upstream or a downstream firm might get new information about future demand or technology developments in the industry and might have an incentive to distort that information. Alternatively, either a supplier or a buyer may face a supply chain disruption (e.g., a shortfall of demand, unexpected production problems, or a commodity price change) that the other party may have the opportunity to mitigate (e.g., with contract renegotiation).

The cell in Table 1 that is not associated with a treatment would require both subjects to agree bilaterally to impose control on one particular 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 particularly relevant to our endeavor and so we did not run a treatment associated with this cell.

It is worth noting that two things change as we move from the *Baseline Treatment* to the *Consistency Treatment*. Control is being imposed symmetrically on both subjects but we have also imposed control before we assign the roles of principal and agent. To ensure that any difference between treatments was driven by the symmetry of control—as predicted by procedural fairness—and not the difference in timing, we also ran the *Unknown Agent Treatment* in which control was imposed asymmetrically on one subject but was imposed before the role of agent has been assigned. That is, before we assigned the roles of principal and agent, we randomly gave one of the subjects the option to impose control on *the other subject* if that other subject became the agent. If the subject who decided about control became the agent, he was unconstrained. In Appendix A, we show that results from the *Unknown Agent Treatment* are never significantly different from results in the *Baseline Treatment*. We collapse these treatments together in our main analysis and call them jointly the *Baseline Treatments*; we also show the results of the two treatments separately in Appendix B.

Subjects in our experiment always played 10 rounds in the *Baseline Treatment* and 10 rounds in one of the other three treatments. Whether they played the *Baseline Treatment* first or second was randomly assigned by session. Control can have two main effects on agent behavior. Control may raise transfers that would have been in the range 0–3 EUs to be at least 4 EUs when control is imposed. Additionally, control might lead some subjects who would have made a large transfer to transfer less, what we call a "behavioral response" or a "hidden cost of control". The net effect of these two forces in a given treatment will determine whether we observe an "average cost of control" (i.e., whether principals receive less effort on average from agents when control is imposed). As we analyze the results starting in the next section, we will look both for a behavioral response as well as identify the net effect of both forces on average transfers.

Since the experiment aimed to investigate the potential for control to backfire due to this behavioral response, we made two additional experimental design choices. First, we implemented a pre-stage to the game (i.e., before the control decision) with the intent of raising the average action when the agent was not controlled, so that we would have a better chance of observing a behavioral response. 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.[13] Results from our earlier work (Kessler and Leider 2012), as well as the work of others, suggest that allowing subjects to make such a non-binding agreement will lead to higher actions in the population and so would make behavioral responses to control easier to observe and measure.[14] We choose to make the agreement amount 40 EUs since that is the payoff-equalizing transfer, leading both the principal and the agent to receive 80 EUs.[15] Second, we chose a minimum to be 10% of the value of payoff-equalizing transfer, which set the minimum at 4 EUs in the hope of being able to identify a behavioral response and an average cost of control in the baseline case.[16]

Our design also allows us to speak to an additional question about how agents respond to the imposition of control. Since we observed subjects play this one-shot game a number of times, and since we randomly assigned the roles of principal and agent in each round, we observe the same subject playing as both a principal and an agent. Consequently, in addition to identifying how agent behavior responds to the symmetry of control and whether control is imposed unilaterally or bilaterally, the experiment addresses whether individuals' use control as a principal is correlated with how they respond to control as an agent. We address this question toward the end of the results section.

4. Behavioral Hypotheses

Control imposed in the *Baseline Treatments* exhibits little procedural fairness. The principal imposes control only on the agent (violating Leventhal's consistency rule); the principal directly and uniquely benefits from controlling the agent (contrary to the bias suppression rule); and the

13. Agreements were made in the same way control was implemented in the Voice Treatment. In particular, each of the subjects 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 subjects suggested the agreement, then the agreement was made. If one or both of the subjects did not suggest the agreement, then no agreement was made. After both subjects had decided whether or not to suggest the agreement, they were told what the other had chosen and whether they had made an agreement.

14. 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). 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. 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.

15. In Appendix C, we introduce additional experimental treatments, including a treatment in which no agreements are allowed. Results demonstrate that giving subjects the opportunity to make agreements raised average actions. Consequently, we were able to observe the behavioral response and an average cost of control more clearly due to this innovation.

16. Our minimum of 4 EUs is slightly below the benchmark minimums of 5, 10, and 20 EUs in Falk and Kosfeld (2006). We investigate the effect of raising the minimum in Appendix D.

agent has no voice in the decision process. We therefore hypothesize that controlled agents will feel untrusted and choose low effort. This would manifest as a behavioral response and potentially an average cost of control in which the principal receives lower effort on average when she imposes control.

Control in the *Consistency Treatment* displays more procedural fairness. Since both subjects face the same constraints when making an effort choice, the consistency rule is now satisfied. Additionally, the subject who can choose to impose control cannot privately benefit from this decision-making right, since the other subject would receive the same benefits of control if placed in the role of principal. Hence, we should expect agents to feel less distrusted by control. Effort should therefore be higher, and the behavioral costs of control should be lower, in *Consistency* when compared with the *Baseline Treatments*.

The *Voice Treatment* further increases procedural fairness, since both subjects now have a voice in the decision to impose control, which is imposed symmetrically as in the *Consistency Treatment*. Therefore, we should expect control to be most effective in the *Voice Treatment*.

Additionally, if there is heterogeneity between subjects in their concern about procedural fairness or how they view the procedural fairness of imposing control (e.g., whether unilateral, asymmetric control is an unjust sign of distrust or a pragmatic, reasonable precaution), this should affect behavior as a principal and as an agent. A subject's willingness to impose control as a principal in the *Baseline Treatment* may indicate that she does not care about procedural fairness or do not see control as unfair. This might lead that subject to have a smaller decrease in effort (i.e., a smaller behavioral response) when being controlled as an agent.

5. Results

A total of 464 student subjects participated in 25 sessions in the Wharton Behavioral Laboratory at the University of Pennsylvania. As noted in Section 3, all subjects participated in the *Baseline Treatment* and one other treatment. Of the 464 subjects, 158 subjects also participated in the *Consistency Treatment*, 158 subjects in the *Voice Treatment*, and the remaining 148 subjects in the *Unknown Agent Treatment*. Sessions lasted approximately 1 h. Average subject pay was $17.28, including a $10 show-up fee. Before we delve into analysis, we make two simplifications that allow us to more clearly communicate our data.

First, in our main results we combine data from the *Baseline Treatment* and the *Unknown Agent Treatment*. As described in Section 3, in both treatments control is imposed asymmetrically and unilaterally and the latter treatment was specifically designed to ensure the difference between *Baseline Treatment* and *Consistency* was not due to the timing of when principal and agent roles were revealed. Results demonstrate that there is no effect of the timing of role revelation and results from the *Baseline Treatment* and the *Unknown Agent Treatment* are nearly identical (see Appendix A for a comparison of summary statistics and Appendix B to see the main regression tables with these two treatments presented separately). Throughout this section we call the combined data the *Baseline Treatments*.

Second, as we expected, allowing subjects to make a non-binding agreement that whomever ends up as an agent will transfer 40 EUs raised the average transfer and made it easier to observe an average cost of control (see complete analysis in Appendix C). Most pairs decide to have the agreement, which significantly increases average actions in the absence of control and thus make it much more likely for us to observe a behavioral response when control is imposed. To show the effect of control on pairs that had and did not have an agreement, the graphs in this section condition on whether a pair had an agreement. However, it is worth emphasizing that whether the subjects in the pair choose to have an agreement is endogenous. Consequently, we present all our regression specifications twice, once pooling across *all pairs* in each treatment (i.e., combining together both pairs that had an agreement and those that did not) and again analyzing only the pairs that endogenously choose to have an agreement where we are more likely to see a cost of control.

5.1 Effect of Control across Treatments

In this subsection, we analyze how control affects transfers and show that, consistent with our hypotheses derived from theories of procedural fairness, the effect of control varies by whether it is imposed asymmetrically or symmetrically and unilaterally or bilaterally.

For control to be implemented in the *Baseline Treatments* and *Consistency Treatment*, *one subject* needs to have imposed control. In the *Voice Treatment*, on the other hand, *both subjects* need to want control for it to be imposed. To avoid selection issues arising from the fact that in the *Voice Treatment* control is only imposed on subjects who suggest it themselves, throughout this section we analyze agents' behavior as a function of whether the *other subject* wants control. Notice that whether the other subject wants control is exogenous to the agent who makes the transfer. This makes the *Voice Treatment* comparable with the other treatments.[17]

Figure 1 shows the average amount transferred in each treatment as a function of whether the other subject wanted to control the agent by having the minimum in place. Panel A shows the effect of the other subject wanting the minimum when the two subjects made an agreement to

17. In the analysis, we exclude agents who were also the player who decided whether or not there should be control in the *Unknown Agent* and *Consistency* treatments so that if control is active in those treatments it is always because the principal imposed control on the agent.

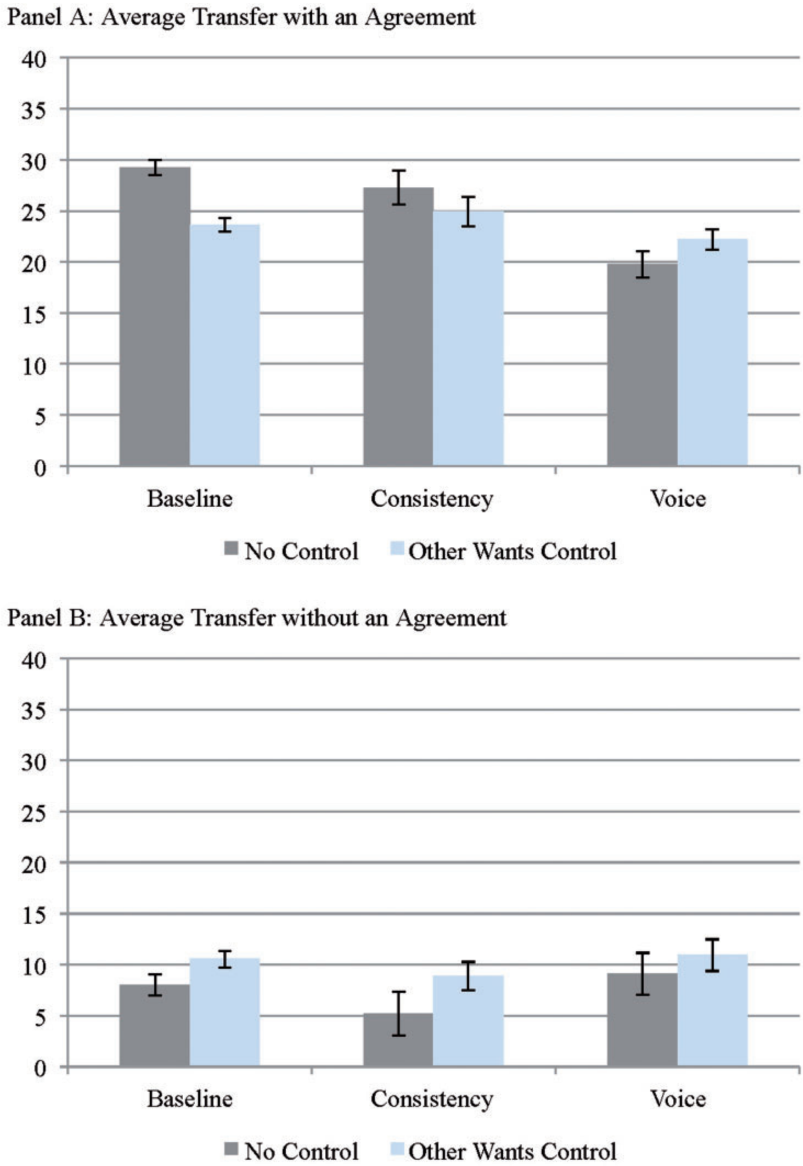


Figure 1. Effect of Control on Transfers by Treatment.
Figure 1 shows the average transfer as a function of whether the other player asks for control (right bar in each treatment) or does not ask for control (left bar in each treatment). Panel A shows transfers when subjects have an agreement to play 40 EUs. Panel B shows transfers when no agreement is in effect. Standard error bars are shown around each mean.

transfer 40 EUs. In this comparison we see differences between the treatments in line with our hypotheses. In the *Baseline Treatments*, where control is procedurally unfair by being asymmetric and unilateral, the average transfer when the other subject wants control is significantly lower than when the other does not want control, reflecting an average cost of control (29.28 when other does not want control, 23.62 when other wants control; subject level non-parametric permutation test: $p = 0.03$; session level permutation test: $p = 0.01$). This difference is smaller in the *Consistency Treatment* (27.27 when other does not want control, 24.95 when other wants control; subject-level: $p = 0.58$; session-level: $p = 0.30$) and flips sign in the *Voice Treatment* when asking for control leads to an increase in average transfer (19.78 when other does not want control, 22.22 when other wants control; subject-level: $p = 0.02$; session-level: $p = 0.08$). Adding more procedural fairness to the control process by making it more consistent between the two subjects, and giving the agent a voice in the control process, changes control from being a net negative to a net positive for the principal.

Panel B shows average transfers when subjects did not make an agreement to transfer 40 EUs. When no agreement is in place, asking for control always directionally increases transfers, with the increase being statistically significant in the *Baseline Treatments* ($p = 0.02$ for both) and marginally significant in the *Voice Treatment* (subject-level: $p = 0.09$; session-level: $p = 0.73$). This suggests that procedural fairness matters less when cooperation has already broken down in the sense that the parties had already failed to reach an agreement to take the equitable action.

We see a similar picture when we turn to the CDFs of transfers in Figure 2. In each graph the dashed line is the CDF when the other player asks for control and the solid line is the CDF when the other player does not ask for control. The left column displays CDFs when subjects have an agreement. Looking in the top row, we see that in the *Baseline Treatment* there is a behavioral response when control is imposed in that many fewer agents transfer 40 and many more make a transfer at the minimum of 4. This effect of asking for control is mitigated in the *Consistency Treatment*; while we still see directionally fewer subjects transfer 40 and directionally more transfer at the minimum when control is imposed, the CDFs appear closer together than in the *Baseline Treatments*. Finally, in the *Voice Treatment*, asking for control increases transfers by shifting the distribution up: it raises transfers from 0 up to 4 and there is no accompanying behavioral response.

We can use non-parametric tests to see if the distributions are significantly different. When subjects have an agreement in the *Baseline Treatments*, imposing control shifts the whole distribution above the minimum to the left (Kolmogorov–Smirnov test: $p < 0.01$ for both). In the *Consistency Treatment* there is no difference in the overall distribution of transfers (KS test: $p > 0.20$). By contrast, in the *Voice Treatment* there is a

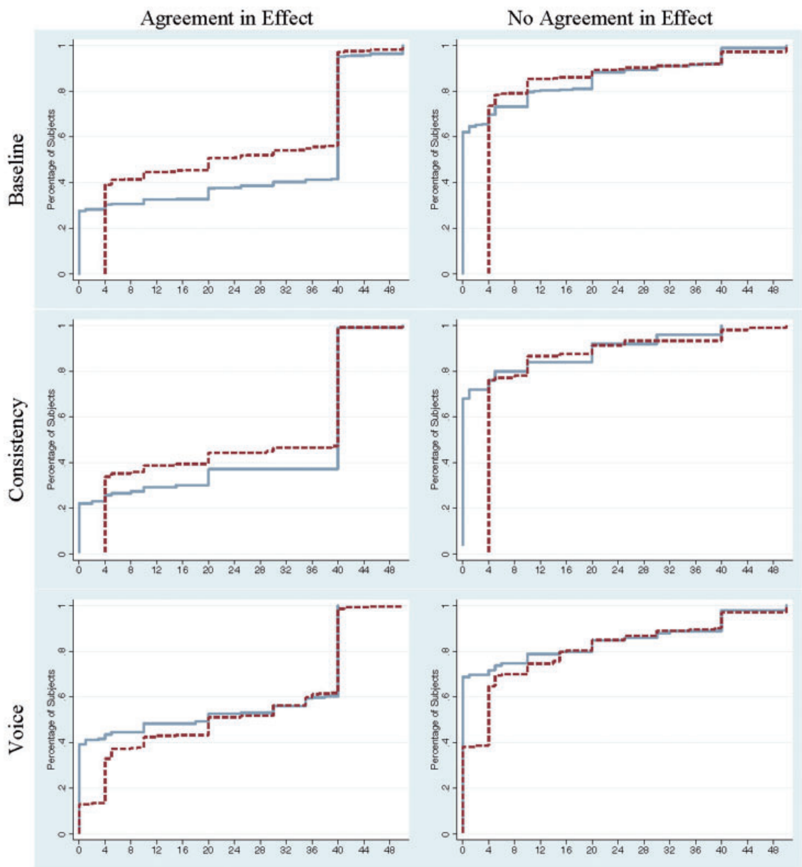


Figure 2. Distribution of Transfers by Treatment.
Figure 2 shows the CDF of transfers (with transfers censored at 50). The horizontal axis reports transfer size and the vertical axis reports the percentage of subjects. The solid lines are the CDFs when the other subject does not ask for control, the dashed lines are the CDFs when the other subject does ask for control. In the *Voice Treatment* the other subject asking for control only leads to control if the agent also asks for control and so some transfers can be below 4 even when the other subject asks for control in that treatment. In the *Baseline Treatments* and *Consistency Treatment* control is always imposed when the other subject asks for it.

nearly marginally significant shift of the overall distribution of transfers to the right (KS test: $p = 0.126$).

Meanwhile, when there is no agreement in place we see little-to-no difference between the CDFs for agents transferring to principals who want control and do not want control in any of three treatments. The only change is that having control increases transfers from 0 to 4 when control is imposed.

The results from Figures 1 and 2 are reflected in regression specifications shown in Table 2 and Table 3 (all specifications include subject fixed effects and cluster standard errors by session). Table 2 reports regressions including all the data from each treatment, pooling pairs that made the agreement and pairs that did not. The first two columns show the effect on average transfer for all rounds (column 1) and for only the first treatment played in a session (column 2). The coefficient on *Other Controlled in Baseline* is negative and significant demonstrating an average cost of control in the procedurally unfair *Baseline Treatments*. *Other Controlled in Consistency* is directionally positive and not significantly different from 0 and *Other Controlled in Voice* is positive and statistically significant, demonstrating that control increases transfers on average in the procedurally fair *Voice Treatment*.

Focusing on differences in average transfers can mask a behavioral response to control because control generates two opposing effects (as can be seen in the CDFs in Figure 2). Control increases transfers that would have been below the minimum transfer allowed up to the minimum. In addition, control may also induce a behavioral response in which subjects lower their transfers in response to being controlled. In Tables 2 and 3 we investigate two ways of identifying a behavioral response. The first is to look at the fraction of subjects who transfer 4 EUs 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 stay the same when control is imposed. 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. We analyze the probability the subjects take an action of 4 or less in columns (3) and (4) of Table 2. The second way of identifying a behavioral response is to look at the share of subjects who transfer 40 EUs or more. A transfer of 40 leads to equal earnings for both subjects and is the amount subjects promise to transfer when an agreement is in place. Subjects who display a behavioral response may be inclined to decrease their transfer to be below 40. We analyze the probability subjects take an action of 40 or more in columns (5) and (6) of Table 2.

Analyzing these additional regression specifications, we see that the cost of control identified in the *Baseline Treatments* is associated with a large behavioral response. Agents are more likely to transfer 4 units or less and are less likely to transfer 40 units or more when control is imposed. Meanwhile, the coefficient for asking for control in the *Consistency Treatment* is close to zero in most specifications and leads to a marginally significant increase in the likelihood of transferring 40 or more in the first half of the study. For the *Voice Treatment*, asking for control *decreases* the frequency of transferring 4 units or less.

Table 2. Transfers by Treatment (all pairs)

	Transfer		Transfer $\leq$ 4		Transfer $\geq$ 40	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Other controlled in baseline	−2.237***	−2.056***	0.0727***	0.0827***	−0.0947***	−0.0969***
	(0.728)	(0.694)	(0.0168)	(0.0199)	(0.0211)	(0.0241)
Other controlled in consistency	0.248	2.744**	0.00327	−0.0325	−0.0339	0.0564*
	(1.462)	(1.276)	(0.0421)	(0.0537)	(0.0411)	(0.0280)
Other controlled in voice	4.141***	6.509***	−0.128***	−0.162***	0.0370	0.0761*
	(1.045)	(2.094)	(0.0260)	(0.0386)	(0.0290)	(0.0427)
Had agreement	14.57***	16.72***	−0.360***	−0.378***	0.397***	0.471***
	(0.955)	(1.447)	(0.0246)	(0.0344)	(0.0251)	(0.0367)
Consistency	−1.221		0.0128		−0.0120	
	(1.917)		(0.0550)		(0.0483)	
Voice	−6.547***		0.185***		−0.126***	
	(1.652)		(0.0298)		(0.0357)	
First treatment	5.631***		−0.169***		0.145***	
	(0.815)		(0.0204)		(0.0187)	
Constant	9.492***	11.63***	0.731***	0.616***	0.0877***	0.149***
	(1.010)	(1.067)	(0.0277)	(0.0265)	(0.0244)	(0.0274)
Observations	3837	1908	3837	1908	3837	1908
Number of subjects	464	458	464	458	464	458
R-squared	0.187	0.197	0.218	0.203	0.247	0.266

Standard errors clustered at the session level reported in parentheses. All specifications include subject fixed effects. The sample is restricted for the Unknown Agent and Consistency treatments to include only observations where the principal had the opportunity to control the agent. In columns (2), (4), and (6) the sample is further restricted to only the first treatment of a session. For these specifications, treatment dummies are dropped as they are collinear with the 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 if the transfer was less than or equal to 4, in columns (5) and (6) it is a dummy variable that equals one if the transfer was greater than or equal to 40.

***$p < 0.01$,

**$p < 0.05$,

*$p < 0.10$.

Table 3 replicates the results in Table 2 but only includes subjects who end up as agents in pairs who made an agreement to transfer 40. This allows us to look at the behavior of agents who we think are likely to make a large transfer in the absence of control. It is worth noting that this means we are looking at subjects who endogenously chose to ask for the agreement (about 80% of subjects in each round in all treatments) and whose randomly chosen partner also asked for the agreement in that round.

We find essentially the same overall pattern of results from the agreement subsample as in the full sample. The *Baseline Treatment* has a strong cost of control across all measures: average transfers are significantly lower with control, the frequency of transferring 4 or less increases by 7 percentage points, and the frequency of transferring 40 or more decreases by 9 percentage points. Control has the opposite effect in the *Voice Treatment*: average transfers increase and the frequency of transferring 4 or less decreases. The results for the *Consistency* treatment are much more mixed, with both significance and direction of effect varying across specifications.

Table 3. Transfers by Treatment (pairs with an Agreement)

	Transfer			Transfer $\leq$ 4			Transfer $\geq$ 40		
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Other controlled in baseline	−3.142***	−2.306**	−3.343**	0.0732***	0.0719***	0.0814***	−0.110***	−0.0910**	−0.111***
	(0.994)	(1.019)	(1.255)	(0.0220)	(0.0245)	(0.0270)	(0.0258)	(0.0347)	(0.0315)
Other controlled in consistency	−2.328*	1.541**	−2.318	0.0472	−0.0349	0.0466	−0.0940**	0.0174	−0.0900
	(1.299)	(0.689)	(1.577)	(0.0379)	(0.0715)	(0.0356)	(0.0435)	(0.0221)	(0.0563)
Other controlled in voice	3.868***	2.008	4.322***	−0.125***	−0.0493	−0.144***	0.0401	0.0290	0.0628
	(0.983)	(2.503)	(1.304)	(0.0331)	(0.0513)	(0.0307)	(0.0338)	(0.0909)	(0.0423)
Consistency	0.0190		−1.068	−0.00423		0.0211	0.0324		0.00538
	(1.843)		(1.889)	(0.0441)		(0.0441)	(0.0525)		(0.0581)
Voice	−8.772***		−8.992***	0.192***		0.174***	−0.175***		−0.196***
	(1.651)		(1.769)	(0.0360)		(0.0401)	(0.0410)		(0.0447)
First treatment	5.718***		5.289***	−0.162***		−0.146***	0.170***		0.166***
	(1.005)		(1.035)	(0.0234)		(0.0229)	(0.0252)		(0.0255)
Constant	24.98***	29.11***	25.34***	0.364***	0.229***	0.353***	0.486***	0.632***	0.490***
	(0.978)	(0.377)	(0.993)	(0.0241)	(0.00933)	(0.0218)	(0.0242)	(0.0132)	(0.0259)
Observations	2653	1333	2056	2653	1333	2056	2653	1333	2056
Number of subjects	443	410	306	443	410	306	443	410	306
R-squared	0.065	0.007	0.065	0.081	0.013	0.074	0.088	0.014	0.091

Standard errors clustered at the session level reported in parentheses. All specifications include subject fixed effects. The sample is restricted to observations where there was an agreement, and for the Unknown Agent and Consistency treatments only observations where the principal had the opportunity to control 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. The dependent variable in columns (1)–(3) is the transfer of the agent, in columns (4)–(6) it is an dummy variable that equals one if the transfer was less than or equal to 4, in columns (7)–(9) it is a dummy variable that equals one if the transfer was greater than or equal to 40.

***$p < 0.01$,

**$p < 0.05$,

*$p < 0.10$.

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 control between treatments is not driven by a selection effect.[18] Overall, we find that imposing control is detrimental to the principal in the *Baseline Treatments*, has no effect in the *Consistency Treatment*, and is beneficial in the *Voice Treatment*, highlighting the importance of procedural fairness in moderating the efficacy of control.

5.2 Who Responds Negatively to Control

Because we observe all subjects playing the role of the principal in the *Baseline Treatment*,[19] we can use a subject's frequency of imposing control when a principal in the *Baseline Treatment* as a measure of their attitude toward control. This attitude toward 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 on others and may react more negatively, with a larger behavioral response, when they are controlled. Conversely, subjects who see control as a reasonable precaution may prefer to control others and may not respond negatively to being controlled. Consequently, the decision to impose control as a principal can indicate whether a subject cares about procedural fairness or can indicate the extent to which a particular subject feels unilateral imposition of control is unjust.

In the *Baseline Treatment*, the median subject imposed control in two-third 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 *Other Subject Controlled* in each treatment for subjects above and below the median usage. The results are reported in Table 4.

18. We also run a specification in the *Voice Treatment* in which we separately control for the agent asking for control, the principal asking for control, and both asking for control (full regression results are available from the authors on request). We find a marginally significant negative effect of the agent asking for control if the other subject did not ($\beta = -5.57$, $p = 0.053$), suggesting that the agent may be punishing the principal, possibly interpreting the principal's failure to ask for control 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 asking for control ($\beta = -2.61$, $p = 0.279$), but a strong positive effect if the principal joined the agent in asking for control ($\beta = 10.80$, $p = 0.006$). 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 control himself. Additionally, this "punishment" effect by agents helps explain the low average transfer shown in the "No Minimum" bar in Panel A of Figure 1. If neither subject requests the restriction, the average transfer is 26.97, which is comparable to the other treatments.

19. Here, we exclude data from the Unknown Agent Treatment and only look at the data from the Baseline Treatment, which everyone played either first or second in the session.

Table 4. Effect of Subject Behavior as Principal in Baseline Treatment

	Average transfer		
	Baseline	Consistency	Voice
Variables	(1)	(2)	(3)
Other controlled and used control < two-third in baseline	−5.917***	−2.697	0.796
	(1.397)	(3.097)	(1.880)
Other controlled and used control ≥ two-third in baseline	0.787	−2.537	6.967**
	(1.280)	(2.109)	(2.113)
Constant	27.57***	27.43***	19.01***
	(0.524)	(1.310)	(0.908)
Observations	1880	255	518
Number of subjects	0.022	127	140
R-squared	0.022	0.013	0.026
	Transfer ≤ 4		
Variables	(4)	(5)	(6)
Other controlled and used control < two-third in baseline	0.141***	0.0611	−0.0547
	(0.0283)	(0.0760)	(0.0456)
Other controlled and used Control ≥ two-third in baseline	−0.0120	0.0741	−0.172**
	(0.0302)	(0.0910)	(0.0531)
Constant	0.297***	0.264***	0.438***
	(0.0107)	(0.0419)	(0.0227)
Observations	1880	255	518
Number of subjects	436	127	140
R-squared	0.028	0.010	0.025
	Transfer ≥ 40		
Variables	(7)	(8)	(9)
Other controlled and used control < two-third in baseline	−0.203***	−0.124	−0.00531
	(0.0387)	(0.0993)	(0.0549)
Other controlled and used control ≥ two-third in baseline	−0.00581	−0.117	0.101
	(0.0310)	(0.0768)	(0.0574)
Constant	0.572***	0.640***	0.361***
	(0.0124)	(0.0408)	(0.0281)
Observations	1880	255	518
Number of subjects	436	127	140
R-squared	0.047	0.032	0.008

Standard errors clustered at the session level reported in parentheses. All specifications include subject fixed effects. The sample is restricted to observations where there was an agreement, and for the Unknown Agent and Consistency treatments only observations where the principal had the opportunity to control the agent are included. The dependent variable in the top panel is the transfer of the agent, in the middle panel it is a dummy variable that equals one if the transfer was less than or equal to 4, in the bottom panel it is a dummy variable that equals one if the transfer was greater than or equal to 40.

***$p < 0.01$,

**$p < 0.05$,

*$p < 0.10$.

In the *Baseline Treatment*, we find a behavioral response only among agents who used control infrequently as principals. For this group, being controlled as an agent led to an estimated transfer decrease of 6 units, a 14 percentage point increase in the likelihood of making a transfer of 4 units or less, and a 20 percentage point decrease in the likelihood of transferring 40 units or more. By contrast, subjects in the *Baseline Treatment* who used control frequently as a principal had essentially zero response to being controlled as an agent. In the *Consistency 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 *Voice Treatment*, the positive effect of the other subject asking for control was only observed among subjects who asked for control frequently themselves—for these subjects transfers increased by an estimated 7 EUs and the frequency of transfers of 4 or less decreased by 17 percentage points. Subjects who used control infrequently have essentially a zero response to control in the *Voice Treatment*. Overall, the pattern of results suggests that there is important heterogeneity in how subjects perceived control, with usage of control as a principal being correlated with more positive reactions to being controlled as an agent. This pattern could reflect differences in intrinsic concern for procedural fairness or differences across individuals in whether the more unilateral control processes are interpreted as procedurally unfair.

One concern with interpreting these heterogeneity results in the *Baseline Treatment* is that subjects switch between playing as a principal and as an agent over the course of the 10 rounds in the *Baseline Treatment*. We are tempted to interpret these results as supportive of a story in which subjects who are inclined to impose control as a principal respond less negatively (or more positively) to control as an agent (due to differences in concern for—or attribution of—procedural unfairness). This interpretation would allow for a prescription that principals should feel comfortable controlling agents who themselves use control in settings where they are a principal (e.g., a CEO would be unlikely to observe a backlash control middle managers who are observed to control their front-line employees). However, an alternative explanation for this pattern of results 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 two-third of the time in the *Baseline Treatment*) and 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.

6. Discussion

In this article, we investigate how procedural fairness affects whether control backfires. As hypothesized, the procedure by which control is imposed affects how subjects respond. We find a large behavioral response, in which agents withhold effort, when control is imposed unilaterally and has an asymmetric effect on the agent, as in the *Baseline Treatments*. In the *Baseline Treatments*, the behavioral response is so large that we observe an average cost of control in which principals are worse off when they impose control than when they leave agents unconstrained. However, when control is imposed symmetrically, the behavioral response is mitigated; and when control is imposed bilaterally—giving the agent voice in the imposition of control—there is no behavioral response and principals are strictly better off from asking for control.

Beyond demonstrating the importance of procedural fairness in the efficiency of control, the results in this article offer three additional insights. First, our results can help guide parties deciding whether or not to impose control. Second, our results speak to a class of models that aim to explain why a cost of control might arise. Third, our study can help reconcile results from other experimental papers that sometimes fail to replicate the average cost of control result from Falk and Kosfeld (2006). We address each of these three in turn.

First, from the perspective of parties deciding whether or not to impose control, our results provide insight about when agents will respond negatively to control and when the contractual tools available can be expected to work as standard theory predicts. First, the main results of the paper highlight that when control is implemented in a procedurally fair way, the costs of control are mitigated and indeed imposing control is beneficial to the principal. Second, we do not observe a behavioral response, nor an average cost of control, among agents who frequently controlled as principals. This result suggests there is important heterogeneity in intrinsic concern for procedural fairness or in whether unilateral control processes are interpreted as procedurally unfair and suggests that principals can more safely impose control upon agents who themselves use control as managers.

Additional insights on this topic come from the additional experimental treatments discussed in Appendix C and Appendix D. We do not observe an average cost of control when the average action in the absence of control is too low (i.e., we only observe a cost of control when we allow parties to make an agreement that raises average actions, see results in Appendix C). This suggests that the procedural fairness concerns are most important in otherwise positive relationships where there is the potential for a large behavioral response to control. Fourth, we do not observe an average cost of control if control is strong enough (i.e., an average cost of control only arose with a minimum of 4 not with a minimum of 10, see results in Appendix D). Hence, sufficiently effective control can be beneficial even if it is imposed in a procedurally unfair way.

To summarize, our results suggest that principals and firms should be most concerned about an average cost of control when they have otherwise high performing agents, 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 control their own bad actions or if they can allow agents to have a voice in the imposition of control.

Second, while our experiment was not designed to test any of the models described below, our results speak to models aiming to explain the cost of control results that helped to motivate our paper. Some of the previously proposed models for the cost of control cannot explain the results in our setting. The model in Sliwka (2007), in which the principal's use of control signals a low norm of behavior, cannot explain behavior in our setup. First, the model depends on the principal having greater information about the norm than the agent, which is not true in our setting. Additionally, the model cannot explain why an agent consenting to control would affect behavior, since it adds no information about the norm. von Siemens' (2013) model of intentions-based reciprocity also cannot explain the pattern of results in our setting. We do not generally model an individual's actions toward himself as representing kindness or unkindness, and hence the *Consistency* treatment should not differ from the *Baseline*.

The Ellingsen and Johannesson (2008) model of esteem is potentially consistent with our data. In this model, individuals care about others' beliefs of their prosociality (esteem) and care more about esteem from prosocial individuals. Hence, a prosocial agent may choose high effort to signal his type to a prosocial principal. A key assumption of that model is that an individual's beliefs are correlated with their type, and hence controlling is a signal of selfishness. In the *Consistency* and *Voice Treatments*, however, a controlling subject who is selfish imposes a cost on himself if he is chosen as the agent, while a prosocial subject is not harmed by controlling himself, since the minimum would not bind on him. This difference can disrupt the signaling equilibrium (depending on parameter values), which would explain the reduction of the cost of control. As we show in Appendix E, however, this is not a sharp prediction of the model as there are parameter values that consistent with a separating equilibrium under all treatments. Additionally, the treatment differences are most likely when the number of high types is small (less than 25% of the population), however 40–60% of our subjects are high types (i.e., transfer 40

EUs). We therefore prefer our procedural fairness explanation to this model.

Third, our experimental results speak to an ongoing debate about the robustness of results in the literature on costs of control. In previous experimental work, principals were shown to be worse off on average when they imposed control on agents rather than leaving them unconstrained (Falk and Kosfeld 2006; Masella et al. 2014). But this average cost of control is not always observed, even in somewhat similar experimental settings (Hagemann 2007[20]; Schnedler and Vadovic 2011; Ziegelmeyer et al. 2012[21]). While these papers fail to find an average cost of control, they generally do replicate the behavioral response or "hidden cost of control" result in which a subset of agents contribute less when they are controlled than when they are not controlled. We help explain these divergent results by showing that for control to actually make principals worse off, it must not only be imposed unilaterally and asymmetrically on the agent, but it must also be sufficiently weak (i.e., not force agents to provide high effort itself), and average effort in the absence of control must be sufficiently high.

By focusing on procedural fairness and investigating how control is imposed, we have shown one way in which the behavioral response to control can be mitigated and eliminated. There are certainly other factors that will influence the behavioral response that arises due to control, which future work can and should address. More generally, demonstrations that control *can* undermine effort (or that incentives can undermine intrinsic motivation) are important first steps in improving our models by including behavioral phenomenon. The next step in developing these models is to understand when such perverse effects of control and incentives will arise. Identifying the boundaries of the cost of control, as well as other behavioral phenomenon,[22] will help us to write better models of behavior and provide better guidance to practitioners deciding how to motivate workers and contract with counterparties.

Funding

Funding for this research came from the authors' research funds provided by the University of Pennsylvania and the University of Michigan.

Conflict of interest statement. None declared.

20. 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 Kosfeld (2006), which has 72 agents and identifies the effect as significant.

21. Ziegelmeyer et al. (2012) find both directionally negative and directionally positive effects for the principal of imposing control, depending on the subject pool.

22. We see identifying boundaries on behavioral phenomena as a generally useful activity that pushes the field toward richer theories that incorporate these phenomena. Papers such as Ariely et al. (2009) pursue a similar approach.

Appendix A: Comparing Baseline Treatment with Unknown Agent Treatment

Table A1. Comparing Baseline to Unknown Agent

			No Control			Control		
	Want agreement	Want control	Transfer	% at min	% 40+	Transfer	% at min	% 40+
Baseline	0.840	0.592	25.27	0.323	0.514	18.93	0.511	0.306
Unknown agent	0.845	0.553	26.62	0.386	0.503	17.24	0.587	0.291
p-value	0.744	0.291	0.586	0.123	0.817	0.409	0.179	0.727
Pooled	0.841	0.583	25.46	0.332	0.512	18.73	0.520	0.305

p-value row reports the *p*-value of a test of whether Baseline and Unknown agent are different in a regression that clusters by the 25 sessions run in the main experiment.

Appendix B: Tables A2 and A3 with Unknown Agent Shown Separately from Baseline Treatment

Table A2. Transfers by Treatment (all pairs) with Unknown Agent Shown Separately from Baseline Treatment

	Transfer		Transfer $\leq$ 4		Transfer $\geq$ 40	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Other controlled in baseline	−1.780**	−1.910**	0.0602***	0.0814***	−0.0849***	−0.100***
	(0.800)	(0.711)	(0.0175)	(0.0217)	(0.0228)	(0.0258)
Other controlled in unknown agent	−5.032***	−3.580	0.150***	0.0969***	−0.155***	−0.0636
	(1.211)	(2.184)	(0.0345)	(0.0294)	(0.0325)	(0.0598)
Other controlled in consistency	0.253	2.742**	0.00315	−0.0325	−0.0338	0.0564*
	(1.469)	(1.276)	(0.0422)	(0.0537)	(0.0412)	(0.0280)
Other controlled in voice	4.146***	6.507***	−0.128***	−0.162***	0.0371	0.0761*
	(1.044)	(2.095)	(0.0260)	(0.0387)	(0.0289)	(0.0427)
Had agreement	14.57***	16.71***	−0.360***	−0.378***	0.397***	0.472***
	(0.959)	(1.451)	(0.0245)	(0.0345)	(0.0252)	(0.0368)
Unknown agent	1.493		−0.0327		0.0235	
	(1.688)		(0.0368)		(0.0434)	
Consistency	−0.929		0.00485		−0.00573	
	(1.961)		(0.0556)		(0.0490)	
Voice	−6.285***		0.178***		−0.121***	
	(1.644)		(0.0294)		(0.0355)	
First treatment	5.764***		−0.173***		0.148***	
	(0.818)		(0.0205)		(0.0192)	
Constant	9.177***	11.69***	0.738***	0.616***	0.0817***	0.147***
	(1.092)	(1.086)	(0.0285)	(0.0265)	(0.0269)	(0.0275)
Observations	3837	1908	3837	1908	3837	1908
Number of subjects	464	458	464	458	464	458
R-squared	0.188	0.198	0.219	0.203	0.247	0.266

Standard errors clustered at the session level reported in parentheses. All specifications include subject fixed effects. The sample is restricted for the unknown agent and consistency treatments to include only observations where the principal had the opportunity to restrict the agent. In columns (2), (4), and (6) the sample is further restricted to only the first treatment of a session, for these specifications treatment controls are dropped as they are collinear with the fixed effects. The dependent variable in columns (1) and (2) is the transfer of the agent, in columns (3) and (4) it is a dummy variable that equals one if the transfer was less than or equal to 4, in columns (5) and (6) it is a dummy variable that equals one if the transfer was greater than or equal to 40.

***$p < 0.01$,

**$p < 0.05$,

*$p < 0.10$.

Table A3. Transfers by Treatment (pairs with an agreement) with Unknown Agent Shown Separately from Baseline Treatment

	Transfer			Transfer $\leq$ 4			Transfer $\geq$ 40		
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Other controlled in baseline	−2.901**	−2.245**	−3.152**	0.0641***	0.0761***	0.0721**	−0.101***	−0.0934**	−0.101***
	(1.056)	(1.037)	(1.345)	(0.0221)	(0.0260)	(0.0273)	(0.0270)	(0.0375)	(0.0323)
Other controlled in	−4.763*	−2.979	−4.628*	0.132**	0.0257	0.139**	−0.174***	−0.0643	−0.179***
unknown agent	(2.588)	(4.372)	(2.469)	(0.0518)	(0.0609)	(0.0571)	(0.0467)	(0.0619)	(0.0585)
Other controlled in consistency	−2.334*	1.541**	−2.321	0.0474	−0.0349	0.0467	−0.0943**	0.0174	−0.0902
	(1.303)	(0.690)	(1.580)	(0.0380)	(0.0715)	(0.0358)	(0.0436)	(0.0221)	(0.0565)
Other controlled in voice	3.874***	2.008	4.330***	−0.125***	−0.0493	−0.145***	0.0404	0.0290	0.0631
	(0.984)	(2.504)	(1.305)	(0.0331)	(0.0513)	(0.0307)	(0.0337)	(0.0909)	(0.0423)
Unknown agent	2.724*		3.001*	−0.0425		−0.0532	0.0644		0.0703
	(1.532)		(1.662)	(0.0298)		(0.0328)	(0.0531)		(0.0518)
Consistency	0.160		−0.957	−0.00953		0.0158	0.0380		0.0116
	(1.897)		(1.947)	(0.0447)		(0.0444)	(0.0536)		(0.0590)
Voice	−8.639***		−8.892***	0.187***		0.169***	−0.169***		−0.190***
	(1.657)		(1.753)	(0.0362)		(0.0398)	(0.0415)		(0.0448)
First treatment	5.779***		5.353***	−0.164***		−0.148***	0.173***		0.169***
	(1.003)		(1.024)	(0.0237)		(0.0232)	(0.0252)		(0.0250)
Constant	24.66***	29.14***	24.99***	0.371***	0.231***	0.361***	0.477***	0.631***	0.480***
	(1.035)	(0.417)	(1.037)	(0.0252)	(0.00925)	(0.0227)	(0.0262)	(0.0127)	(0.0266)
	−2.901**	−2.245**	−3.152**	0.0641***	0.0761***	0.0721**	−0.101***	−0.0934**	−0.101***
Observations	(1.056)	(1.037)	(1.345)	(0.0221)	(0.0260)	(0.0273)	(0.0270)	(0.0375)	(0.0323)
Number of subjects	−4.763*	−2.979	−4.628*	0.132**	0.0257	0.139**	−0.174***	−0.0643	−0.179***
R-squared	(2.588)	(4.372)	(2.469)	(0.0518)	(0.0609)	(0.0571)	(0.0467)	(0.0619)	(0.0585)

Standard errors clustered at the session level reported in parentheses. All specifications include subject fixed effects. The sample is restricted to observations where there was an agreement, and for the unknown agent and consistency 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. The dependent variable in columns (1)–(3) is the transfer of the agent, in columns (4)–(6) it is a dummy variable that equals one if the transfer was less than or equal to 4, in columns (7)–(9) it is a dummy variable that equals one if the transfer was greater than or equal to 40.

***$p < 0.01$,

**$p < 0.05$,

*$p < 0.10$.

Appendix C: Control When There Is No Opportunity for an Agreement

In Sections 1 and 3 of the main text, we explained that we introduced the agreement in order to raise average actions in the absence of control so that we would have a better opportunity to observe a behavioral response and an average cost of control. We focused on this case because we expected procedural fairness to matter more when the relationship was otherwise positive and collaborative. We now show results demonstrating that giving subjects the opportunity to make the agreement had the intended effect. To do this test, we ran additional sessions with a *No Agreement Allowed Treatment* in which subjects did not have the opportunity to make an agreement. 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 five sessions, with a total of 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 A1 shows the average transfer with and without control in the *No Agreement Allowed Treatment* and compares it to behavior in the *Baseline Treatment* (including *Baseline* data from the main experiment and these new sessions).[23] There is a significant increase in transfers associated with giving subjects the opportunity to make the agreement. In addition, the opportunity to make an agreement allows us to more cleanly identify the average cost of control that arises when the principal imposes control in the *Baseline Treatment* in a setting that is procedurally unfair.

In the *No Agreement Allowed Treatment*, transfers decrease slightly from 16.41 when control is not imposed to 14.91 when control is imposed, and the difference is not significant ($p > 0.20$ for both subject-level and session-level permutation tests). Similarly, in the same data, the fraction of subjects transferring 4 or less increases from 30% to 36% in response to control, but the difference is only marginally significant (subject-level: $p = 0.06$, session-level: $p > 0.20$). It is worth pointing out that these small and insignificant differences contrast with the results in Falk and Kosfeld (2006), which has a design very similar to our *No Agreement Allowed* except that in Falk and Kosfeld (2006), subjects play a one-shot game via the strategy method and control imposes a minimum transfer of 5 rather than 4. They find that imposing a minimum transfer of 5 leads to a decrease in average transfer from 25.1 to 12.2, and an increase in the

23. We obtain essentially the same results whether or not we include *Unknown Agent Treatment* data in with *Baseline Treatment* data and whether or not we exclude any observations that took place in the second half of the experiment (i.e., comparing the *No Agreement Allowed Treatment*, which was always played first in a session, to data from the *Baseline Treatment* when it was played first in a session).

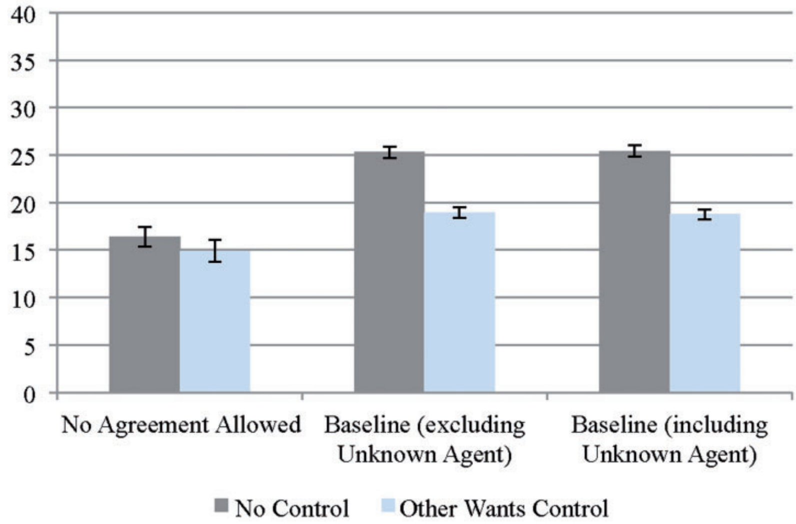


Figure A1. Transfers and Effect of Control by Whether Agreements Are Allowed. Figure A1 shows the average transfer as a function of whether the other player asks for control (right bar in each treatment) or does not ask for control (left bar in each treatment) when agreements are not allowed (left pair of bars) and in the *Baseline Treatments* from the original experiment and the additional sessions. Standard error bars (clustered by session) are shown around each mean.

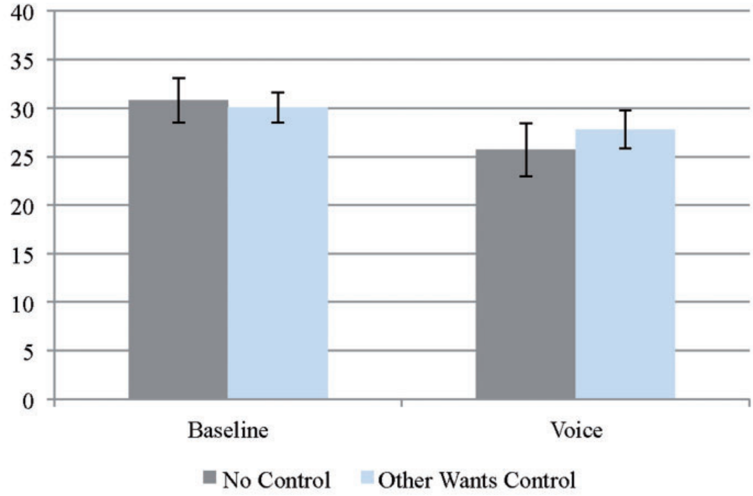
fraction of subjects transferring 5 or less from approximately 20% to approximately 50%.

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 willingness to transfer in the absence of control for their subject pool when compared with our subject pool. We are only able to get subjects to transfer an average of 25.1 units in the absence of control when subjects establish an agreement in our data (and only there do we see an average cost of control). Taken together, we only expect procedural fairness concerns to generate an average cost of control in settings where there is a strong willingness to transfer units in the absence of control, either by default as in Falk and Kosfeld (2006) data, or due to a specific agreement as in our data.

Appendix D: Restrictions with a Higher Minimum Transfer

In the *Baseline Treatment* of our main experiment, we observe a behavioral response among agents who are controlled. We find that agents are much less likely to transfer 40 units and much more likely to transfer the minimum of 4 units when they are controlled. This leads to an average cost of control, since the behavioral response is large relative to the benefit from raising transfers that were below 4 units to the minimum of 4.

Panel A: Average Transfer with an Agreement



Panel B: Average Transfer without an Agreement

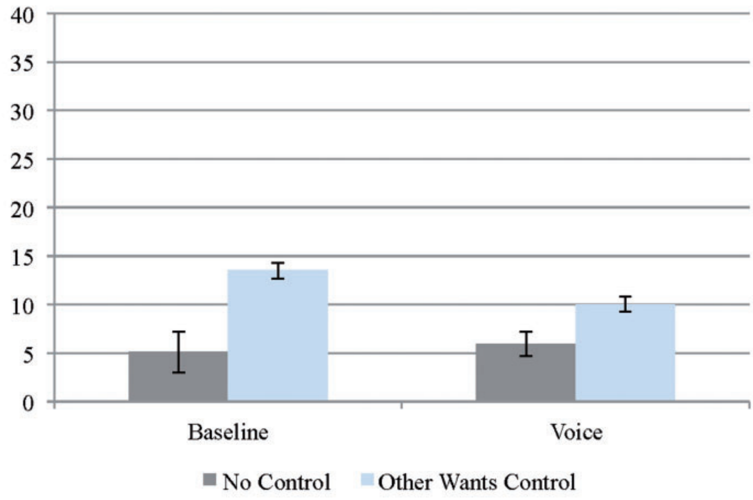


Figure A2. Effect of Control on Transfers by Treatment when Minimum is 10.
Figure A2 shows the average transfer as a function of whether the other player asks for control (right bar in each treatment) or does not ask for control (left bar in each treatment) when the minimum is 10 EUs. Panel A shows transfers when subjects have an agreement to play 40 EUs. Panel B shows transfers when no agreement is in place. Standard error bars (clustered by session) are shown around each mean.

Principals who know that control will lead to a large behavioral response may decide to avoid using control unless it can be implemented in a way that is more procedurally fair, which, as we showed in the main text, can eliminate (or reverse) the cost of control. Alternatively, however, principals may decide to invest in more powerful control, for example by finding a way to require a greater perfunctory performance (e.g., a better monitoring technology).

To test the impact of more powerful control, we ran five additional sessions with 94 subjects in the *Baseline* and *Voice* treatments in which control required a minimum transfer of 10 units rather than 4 units. Figure A2 shows the average transfer in each treatment as a function of whether an agreement was in place.

When there is an agreement, we no longer find an average 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 permutation tests). As expected, 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: $p = 0.05$; session-level: $p > 0.20$) while the fraction of transfers greater than or equal to 40 decreases from 57% to 46% (subject-level: $p = 0.02$; session-level: $p = 0.06$). In this case, however, the benefit of the increase due to the binding minimum counteracts the behavioral response. Therefore, sufficiently strong control can still be useful even when imposed in a procedurally unfair way. In the *Voice Treatment*, we find that the minimum is somewhat beneficial for the Principal, directionally increasing average transfers from 25.0 to 26.0 (subject-level: $p = 0.10$; session-level: $p > 0.20$). There is a directional decrease in the fraction of subjects transferring 10 or less with a restriction from 42% to 37% ($p > 0.20$ for both), and a directional increase in the number of subjects transferring 40 or more from 47% to 52% (subject-level: $p = 0.06$; session-level: $p > 0.20$).

As in our main experiment, we find that when there is no agreement, imposing powerful control (or asking for control in the Voice treatment) leads to higher average transfers from the agent. Overall, these results suggest that the cost of control arising from a procedurally unfair implementation of control can be overcome when control is powerful enough to compensate for any behavioral response.

Appendix E: Our Results and the Ellingsen and Johannesson (2008) Model of Esteem

In Ellingsen and Johannesson (2008), individuals have three components to their utility: material payoffs, inequity aversion (captured by parameter θ), and esteem (denoted by v). For simplicity, we assume there are two type: $\theta_L = 0$ and $\theta_H > 0$. The esteem that player i would feel from player j believing he was type θ_i is $v = g_i{*}s_j{*}\theta_i$, where g_i denotes player i's general concern for esteem, and s_j denotes the weight placed on the opinion of someone of player j's type. Individuals are assumed to care more about the opinions of high types than low types ($s_H > s_L$).

The principal's utility is then $U_P = 2x - |2x - (120 - x)|\theta_P + v_P$. The agent's utility is similarly $U_A = 120 - x - |(120 - x) - 2x|\theta_A + v_A$. Additionally, low types have belief p_L that their counterpart is a high type, while high types have believe $p_H > p_L$. Following Ellingsen and Johannesson (2008), we consider for the *Baseline Treatment* a simple separating equilibrium that follows the basic cost of control result,[24] namely low types control when possible and choose the minimum action as agents, while high types trust, choose 40 if trusted and choose 20 if controlled.[25] If principals who control are systematically more selfish than principals who trust, then a high-type agent will have a stronger incentive to choose a high action in order to signal their type (since esteem from a high type is more valuable than esteem from a low type). In a pooling equilibrium, esteem will be the same in the controlling and trusting cases, so we would expect the same action (and therefore no cost of control).

For the equilibrium to hold, we need the following conditions:

(a) Low type agent controlled does not pool with high types: $116 \geq 100 + g_L s_L \theta_H$ or $16 \geq g_L s_L \theta_H$.
(b) Low type agent trusted does not pool with high types: $120 \geq 80 + g_L s_H \theta_H$ or $40 \geq g_L s_H \theta_H$.
(c) High type agent controlled does 20 instead of $x > 20$: $100 - 60\theta_H \geq 120 - x - |120 - 3x|\theta_H$ or $1/3 \geq \theta_H$.
(d) High type agent controlled does 20 instead of 4: $100 - 60\theta_H + g_H s_L \theta_H \geq 116 - 108\theta_H$ or $g_H s_L \theta_H \geq 16 - 48\theta_H$.
(e) High type agent trusted does 40: $80 + g_H s_H \theta_H \geq 120 - 120\theta_H$ or $g_H s_H \theta_H \geq 40 - 120\theta_H$.
(f) Low type controls: $p_L(40) + (1-p_L)(8) \geq p_L(80 + g_L s_H \theta_H) + (1-p_L)(g_L s_L \theta_H)$ or $p_L \leq (8 - g_L s_L \theta_H)/(48 + g_L \theta_H [s_H - s_L])$.
(g) High type trusts: $p_H(80 + g_H s_H \theta_H) + (1-p_H)(-120\theta_H + g_H s_L \theta_H) \geq p_H(40 - 60\theta_H) + (1-p_H)(8 - 108\theta_H)$ or $p_H \geq (8 + 12\theta_H - g_H s_L \theta_H)/(48 + 72\theta_H + g_H \theta_H [s_H - s_L])$.

Therefore, we need inequality aversion to be not too strong, low types to be sufficiently pessimistic and not too esteem concerned, and high types to be sufficiently optimistic and sufficiently esteem concerned.

24. For simplicity we set aside here the role of agreements. Since the majority of subjects make the agreement, it should not have a signaling effect. We can also adapt the Elligsen and Johannesson (2008) model to include norm sensitivity instead of inequity aversion to explicitly allow for the effect of an agreement—results are similar.

25. These are the same action levels that Elligsen and Johannesson (2008) uses in their discussion of the costs of control.

For the *Consistency* treatment, conditions (a)–(e) remain the same. However, condition (f) is now $\frac{1}{2}[p_L(40)+(1-p_L)(8)]+\frac{1}{2}[116] \geq \frac{1}{2}[p_L(80+g_L s_H \theta_H)+(1-p_L)(g_L s_L \theta_H)]+\frac{1}{2}[120]$ or $p_L \leq (4-g_L s_L \theta_H)/(48+g_L \theta_H[s_H-s_L])$, which is a stricter criterion. Similarly, condition (g) is now $\frac{1}{2}[p_H(80+g_H s_H \theta_H)+(1-p_H)(-120\theta_H+g_H s_L \theta_H)]+\frac{1}{2}[80+p_H(g_H s_H \theta_H)+(1-p_H)(g_H s_L \theta_H)] \geq \frac{1}{2}[p_H(40-60\theta_H)+(1-p_H)(8-108\theta_H)]+\frac{1}{2}[116-108\theta_H]$ or $p_H \geq (22-48\theta_H-g_H s_L \theta_H)/(24+36\theta_H+g_H\theta_H[s_H-s_L])$. For this to be a stricter condition than the Baseline we need $432(2-3\theta_H-9\theta_H^2)+(14-60\theta_H)(g_H\theta_H[s_H-s_L])-(24+36\theta_H)g_H s_L \theta_H \geq 0$. Having $\theta_H < 7/30$ is a sufficient condition for this to be true. However, we note that this is a restrictive assumption, as we observed between 40% and 60% of subjects who made an agreement in each treatment transferring 40.

If the condition is stricter, then the separating equilibrium is less likely to be sustainable in the *Consistency Treatment* than in the *Baseline Treatment*. If instead a pooling equilibrium holds, then we would expect similar agent choices between controlling and trusting (since esteem will be equal), and hence a smaller cost of control.

The *Voice Treatment* further changes the criteria for the decisions to control or trust. The condition for the low types to control is slightly relaxed compared with the *Consistency Treatment* (since in some cases when he will be the agent the other player trusted): $\frac{1}{2}[p_L(40)+(1-p_L)(8)]+\frac{1}{2}[p_L(120)+(1-p_L)(116)] \geq \frac{1}{2}[p_L(80+g_L s_H \theta_H)+(1-p_L)(g_L s_L \theta_H)]+\frac{1}{2}[120]$ or $p_L \leq (4-g_L s_L \theta_H)/(44+g_L\theta_H[s_H-s_L])$. However, this condition is still stricter than the *Baseline Treatment*. The condition for the high types to trust is now $\frac{1}{2}[p_H(80+g_H s_H \theta_H)+(1-p_H)(-120\theta_H+g_H s_L \theta_H)]+\frac{1}{2}[p_H(80+g_H s_H \theta_H)+(1-p_H)(100-60\theta_H+g_H s_L \theta_H)] \geq \frac{1}{2}[p_H(40-60\theta_H)+(1-p_H)(8-108\theta_H)]+\frac{1}{2}[p_H(120-120\theta_H)+(1-p_H)(116-108\theta_H)]$ or $p_H \geq (12-18\theta_H-g_H s_L \theta_H)/(12+72\theta_H+g_H\theta_H[s_H-s_L])$. For this to be a stricter condition than in the *Baseline Treatment*, we need: $480(2+3\theta_H)(1-3\theta_H)+(14-60\theta_H)(g_H\theta_H[s_H-s_L])-36 g_H s_L \theta_H \geq 0$. Having $\theta_H < 7/30$ is again a sufficient condition for this to be true. Additionally, this is always a stricter condition than the *Consistency* condition.

Hence, there is a range of parameters for the Ellingsen and Johannesson (2008) model that are consistent our primary results: that there is a cost of control in the *Baseline Treatments* and this cost is reduced or eliminated in the *Consistency Treatment* and *Voice Treatment*. However, we note that the difference in equilibrium conditions between treatments is largest when θ_H is small, which is not consistent with our data. Additionally, there are many parameter values where the separating equilibrium obtains in all three treatments. Therefore, while the observed treatment differences are potentially consistent with the model, it is not a sharp prediction and is not particularly consistent with our data.

References

Alexander, Sheldon, and Marian Ruderman. 1987. "The Role of Procedural and Distributive Justice in Organizational Behavior," 1 *Social Justice Research* 177–98.

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

Ariely, Dan, Anat Bracha, and Stephan Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially," 99 *American Economic Review* 544–55.

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

Benz, Matthias, and Alois Stutzer. 2003. "Do Workers Enjoy Procedural Utility?," 49 *Applied Economics Quarterly* 149–172.

Benz, Matthias, and Bruno S. Frey. 2004. "Being Independent Raises Happiness at Work," 11 *Swedish Economic Policy Review* 95–134.

Bernheim, B. Douglas, and Michael Whinston. 1998. "Incomplete Contracts and Strategic Ambiguity," 88 *American Economic Review* 902–32.

Bies, Robert J., and Debra L. Shapiro. 1988. "Voice and Justification: Their Influence on Procedural Fairness Judgments," 31 *Academy of Management Journal* 676–85.

Bolton, Gary E., Jordi Brandts, and Axel Ockenfels. 2005. "Fair Procedures: Evidence from Games Involving Lotteries," 115 *The Economic Journal* 1054–76.

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

Cohen-Charash, Yochi, and Paul Spector. 2001. "The Role of Justice in Organizations: A Meta-Analysis," 86 *Organizational Behavior and Human Decision Processess* 278–321.

Charness, Gary, and Martin Dufwenberg. 2006. "Promises and Partnerships," 74 *Econometrica* 1579–601.

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

Deci, Edward. 1971. "Effects of Externally Mediated Rewards on Intrinsic Motivation," 18 *Journal of Personality and Social Psychology* 105–15.

———. 1975. *Intrinsic Motivation*. New York and London: Plenum Press.

Deci, Edward, Richard Koestner, and Richard Ryan. 1999. "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation," 125 *Psychological Bulletin* 627–68.

De Cremer, David, and Daan van Knippenberg. 2002. "How Do Leaders Promote Cooperation? The Effects of Charisma and Procedural Fairness," 87 *Journal of Applied Psychology* 858–66.

———. 2003. "Cooperation with Leaders in Social Dilemmas: On the Effects of Procedural Fairness and Outcome Favorability in Structural Cooperation," 91 *Organizational Behavior and Human Decision Processes* 1–11.

Dickinson, David, and Marie Claire Villeval. 2008. "Does Monitoring Decrease Work Effort?," 63 *Games and Economic Behavior* 56–76.

Dufwenberg, Martin, Simon Gächter, and Heike Hennig-Schmidt. 2011. "The Framing of Games and the Psychology of Play," 73 *Games and Economic Behavior* 459–78.

Ellingsen, Tore, and Magnus Johannesson. 2004. "Promises, Threats and Fairness," 114 *Economic Journal* 397–420.

———. 2008. "Price and Prejudice: The Human Side of Incentive Theory," 98 *American Economic Review* 990–1008.

Falk, Armin, and Michael Kosfeld. 2006. "The Hidden Cost of Control," 96 *American Economic Review* 1611–30.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic experiments," 10 *Experimental Economics* 171–8.

Frey, Bruno S. 1993. "Does Monitoring Increase Work Effort? The Rivalry with Trust and Loyalty," 31 *Economic Inquiry* 663–70.

———. 1994. “How Intrinsic Motivation Is Crowded In and Out,” 6 *Rationality and Society* 334–52.

Frey, Bruno S., and Felix Oberholzer-Gee. 1997. “The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out,” 87 *American Economic Review* 746–55.

Frey, Bruno S., Felix Oberholzer-Gee, and Reiner Eichenberger. 1996. “The Old Lady Visits Your Backyard: A Tale of Morals and Markets,” 104 *Journal of Political Economy* 1297–313.

Frey, Bruno S., Matthias Benz, and Alois Stutzer. 2004. “Introducing Procedural Utility: Not Only What, But Also How Matters,” 160 *Journal of Institutional and Theoretical Economics* 377–401.

Frohlich, Norman, and Joe A. Oppenheimer. 1990. “Choosing Justice in Experimental Democracies with Production,” 84 *American Political Science Review* 461–77.

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

———. 2000b. “A Fine Is a Price,” 29 *Journal of Legal Studies* 1–18.

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

Greenberg, Jerald. 1990. “Employee Theft as a Reaction to Underpayment Inequity: The Hidden Cost of Pay Cuts,” 75 *Journal of Applied Psychology* 561–70.

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

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

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1986. “Fairness and the Assumptions of Economics,” 59 *Journal of Business* S285–S300.

Kessler, Judd B., and Stephen Leider. 2012. “Norms and Contracting,” 58 *Management Science* 62–77.

Konovsky, Mary A., and Russell Cropanzano. 1991. “Perceived Fairness of Employee Drug Testing as a Predictor of Employee Attitudes and Job Performance,” 76 *Journal of Applied Psychology* 698–707.

Krawczyk, Michał. 2010. “A Glimpse Through the Veil of Ignorance: Equality of Opportunity and Support for Redistribution,” 94 *Journal of Public Economics* 131–41.

Kruglansky, Arie, Sarah Alon, and Tirtzah Lewis. 1972. “Retrospective Misattribution and Task Enjoyment,” 8 *Journal of Experimental Social Psychology* 493–501.

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

Leventhal, Gerald S. 1980. “What Should be Done with Equity Theory? New Approaches to the Study of Fairness Is Social Relationships,” in K.J. Gergen , M.S. Greenberg, and R.H. Willis, eds., Social Exchange: Advances in Theory and Research, 27–55. New York: Plenum.

Lind, E. Allan, and Tom R. Tyler. 1988. *The Social Psychology of Procedural Justice*. New York: Springer.

Lind, E. Allan, Ruth Kanfer, and P. Christopher Earley. 1990. “Voice, Control, and Procedural Justice: Instrumental and Noninstrumental Concerns in Fairness Judgments,” 59 *Journal of Personality and Social Psychology* 952–9.

Lind, E. Allan, Carol T. Kulik, Maureen Ambrose, and Maria V. de Vera Park. 1993. “Individual and Corporate Dispute Resolution: Using Procedural Fairness as a Decision Heuristic,” 38 *Administrative Science Quarterly* 224–51.

Masella, Paolo, Stephan Meier, and Philipp Zahn. 2014. “Incentives and Group Identity,” 86 *Games and Economic Behavior* 12–25.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard Press.

Schildberg-Hörisch, Hannah. 2010. “Is the Veil of Ignorance only a Concept about Risk? An Experiment,” 94 *Journal of Public Economics* 1062–6.

Schnedler, Wendelin, and Radovan Vadovic. 2011. "Legitimacy of Control," 20 *Journal of Economics and Management Strategy* 985–1009.

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

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

Spier, Kathryn E. 1992. "Incomplete Contracts in a Model with Adverse Selection and Exogenous Costs of Enforcement," 23 *RAND Journal of Economics* 432–43.

Sutter, Matthias, and Hannelore Weck-Hannemann. 2003. "Taxation and the Veil of Ignorance—A Real Effort Experiment on the Laffer Curve," 115 *Public Choice* 217–40.

Thibaut, John, and Laurens Walker. 1975. *Procedural Justice: A Psychological Analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Thibaut, John, and LaurensWalker. 1978. "A Theory of Procedure," 66 *California Law Review* 541–66.

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

Vanberg, Cristoph. 2008. "Why Do People Keep Their Promises? An Experimental Test of Two Explanations," 76 *Econometrica* 1467–80.

van Wanrooy, Brigid, Helen Bewley, Alex Bryson, John Forth, Stephanie Freeth, Lucy Stokes, and Stephen Wood. 2013. "The 2011 Workplace Employment Relations Study First Finding." Workplace Employment Relations Study Report.

Vernon, Guy. 2003. "Comparative Work Organization, Managerial Hierarchies and Occupational Classification," 25 *Employee Relations* 389–404.

von Siemens, Ferdinand A. 2013. "Intention-Based Reciprocity and the Hidden Costs of Control," 92 *Journal of Economic Behavior and Organizations* 55–65.

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

———. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.

Ziegelmeyer, Anthony, Katrin Schmelz, and Matteo Ploner. 2012. "Hidden Costs of Control: Four Repetitions and an Extension," 15 *Experimental Economics* 323–40.