Determinants of Relational Contract Performance: Experimental Evidence

Steven Y. Wu *

December 2020

Abstract

This study examines the determinants of relational contractual performance using data from a series of laboratory experiments. There is currently limited empirical evidence on the determinants of contractual performance, which includes contractual acceptance, and the delivery of promised quantity/quality under the terms of the contract. While theory predicts that the primary drivers of contractual performance are high discount factors, and contract designs that obey individual rationality and self-enforcement constraints, the empirical analysis suggests that other determinants such as a history of prior cooperation can matter as much, if not more, than the theoretical constraint conditions.

JEL Codes:

KEYWORDS: relational contracts, informal contracts, trust, cooperation, agricultural contracting

Running Head: *Determinants of contractual performance*

*Steven Y. Wu is Associate Professor, Purdue University, Krannert Building, 403 West State Street West Lafayette, IN 47907, sywu@purdue.edu. Funding from USDA-NIFA grant number 2010-65400-20430 and HATCH project IND010580 are gratefully acknowledged.

Contracts are important mechanisms for facilitating the trade of agricultural commodities when there are imperfections in spot markets. For example, moral hazard and adverse selection are the classic textbook information asymmetries that justify the use of contracts. Contracts can also be designed around other types of information issues such as the inability of outsiders to verify performance outcomes making it difficult or prohibitively costly for an arbitrator or a court to enforce contracts (Dixit 2007). This tends to be the rule rather than the exception because the set of performance outcomes that traders care about often exceed the set of outcomes that can be third-party enforced. This makes informal contracts, such as relational contracts, relevant for most agricultural contracting situations.[1]

This paper investigates the empirical determinants of relational contracting performance using data from a series of economic experiments. Experiments are well suited for studying contracting problems because the information imperfections that justify the use of contracts also pose problems for researchers who study contracts. That is, if there are important information gaps even among the actual participants, it is difficult to to imagine that researchers, who must rely on surveys or observation data, have sufficient information to rigorously study contracting. Additionally, Gil and Zanarone (2018) suggest that the empirical study of relational contracts pose a number of challenges given the difficulty of measuring some important variables needed to specify self-enforcement constraints. Finally, relational contracts may not even be observable to researchers. For example, under textbook contract theory, a fixed payment contract that lacks bonuses or piece rate payments may appear to lack sufficient incentives to motivate the agent. However, in a relational contract, performance incentives may tacitly exist because it is implicitly understood by the parties to the contract that the prospect of future business or even informal discretionary bonuses depend on good performance. Thus, it is thus not surprising that empirical work on relational contracts has lagged behind theoretical developments (Michler and Wu 2020).

A major advantage of laboratory experiments is that they permit researchers to design the information environment and control for other nuisance factors that might confound results. While laboratory experiments are sometimes criticized for their external validity due to the use of student subjects, Cason and Wu (2019) point out that laboratory experiments and students subjects may be the preferred way of testing general theories.[2]

As a starting point, I primarily focus on theoretical determinants of contractual acceptance and compliance. Specifically, within a principal-agent model of informal contracting, I investigate the factors that drive agents to (1) accept contracts offered by principals; and (2) to deliver the quantity or quality level agreed upon in the contract. These issues are important to applied economists and practitioners who worry about how to use contracts to induce farmers to participate in contract production or marketing, while ensuring quality compliance. For example, development agencies may want farmers to adopt contracts as a means of facilitating market access to smallholders (FAO 2017). Biomass companies may want to know how to use contracts to induce farmers to adopt new energy crops (Alexander et al. 2012). Food processors may want to induce farmers to adopt organic practices while ensuring quality compliance.

According to theory, the most important determinants of contract acceptance and compliance are whether a contract satisfies self-enforcement (incentive compatibility) and participation (individual rationality) constraints. I also investigate secondary theoretical factors such as trading history and a history of cooperation as these may be related to game theoretic *best response dynamics*. While contracts in practice may involve other determinants, such as access to inputs or financing, the theoretical determinants are grounded in the first principles of contract design and may potentially serve as the foundation on top of which more situation specific determinants can be stacked.

The experimental design maps a stylized canonical relational contracting model into an experimental trading platform where human subjects can conduct economic transactions using informal agreements. Subjects assigned to be principals can design and offer informal contracts to subjects assigned to be agents. Thus, the structure of each relational contract is endogenous. Allowing for endogeneity is an important aspect of testing contract theory as contract terms enable a principal to optimally deal with certain types of incentive problems. Studies that exogenously impose contractual form often invite criticisms from theorists (e.g. Carmichael (1985); Tirole (1999); Schmitz (2001)) because they involve ad hoc assumptions about solutions without justification.[3] Moreover, given that I control the trading environment and parameterizations, it is easy to determine whether each contract satisfies the theoretical self-enforcement and participatio constraints, and to determine how agents react to these contracts. Additionally, I ran two treatments involving different size discount factors. Theoretically, a larger discount factor should cause future payoffs to matter more and strengthen the potential for self-enforcement, which should affect both endogenous contract structure and agents' willingness to accept and deliver on contractual promises. Finally, given that the experiment is designed around repeat trading, additional determinants related to trading history can also be tested.

The results suggest that contracts that satisfy the self-enforcement and participation constraints robustly predict agents' willingness to accept/reject contracts. Moreover, prior cooperation and trading also appear to be important determinants of whether agents accept or reject contracts. The implication here is that contracts that are individually rational and are self-enforcing tend to be more credible and increase the odds of farmer acceptance. Moreover, if farmers have a history of trading with a contractor or the contractor has a generally good reputation, the probability of farmer adoption is higher. Surprisingly, higher discount factors did not consistently predict a decrease in contract rejection rates.

With regard to post-contractual performance; i.e. agents' willingness to deliver on quality/quantity specified in the contracts, things appear to be more nuanced. There is tentative evidence that higher discount factors and/or contracts that satisfy the theoretical constraints reduce agent shirking on quantity/quality obigations but results appear to be sensitive to the control variables used and the treatment. However, a history of cooperation and an increase in the promised share of surplus to agents seem to robustly predict whether agents will deliver on promised quality/quantity.

Background and Literature

Michler and Wu (2020) provide an appraisal of relational contracting in agricultural and resource economics. Their main points are that (1) researchers examining agricultural contracting often do not sufficiently distinguish between formal contracts and informal contracts; (2) there does not appear to be a consistent body of high quality work on relational contracting in agriculture; and (3) empirical studies, whether they be randomized control trials or based on observational data, tend to focus on the reasons for farmer participation and farmer welfare. Few studies examine the endogenous formation of contracts or how contract structure affects outcomes. For the most part, there are few empirical studies that are well grounded in theory, which limits the generalizability of specific studies.

Empirical work that tests and is grounded in theory is important for understanding the general forces that underlie most contracting relationships. Just as the rigorous development and testing of production theory provided an unifying framework for production economics and farm management, a coherent and consistent model of contracting that nests both formal and informal agreements can be used as a basis for studying applied contracting problems. This is particularly important going forward as contracts are frequently used by farmers and manufacturers to manage risks, and maintain efficiency and consistent product quality (MacDonald and Burns 2019), and by development agencies as a poverty mitigation mechanism (Bellemare and Bloem 2018).

Empirical testing of relational contracting is currently quite active in general economics. Gil and Zanarone (2018)(GZ) provide an overview of empirical work on relational contracting and highlights some challenges facing researchers using observational data. In particular, in order to precisely test relational contract theory, researchers would have to have reasonable measures or proxies for intertemporal discount factors, reservation utilities, cost savings from shirking, etc. One advantage of lab experiments is that these important variables can be controlled and parameterized by the researcher so clean tests can be undertaken.

In the following section, I outline a canonical relational contracting model that is flexible and yet simple enough to encompass a wide range of theoretical relational contracts observed in the literature from efficiency wage style fixed price contracts (e.g. Klein and Leffler (1981), Shapiro and Stiglitz (1984), Brown et al. (2004)) to discretionary bonus contracts(e.g. MacLeod and Malcomson (1989), Levin (2003)). This model should make clear GZ's point about the need to specify important variables in order to test relational contract theory. The model will also serve as the blueprint for the experimental design.

The Model

The model is a basic principal-agent model where a principal contracts with an agent to produce a commodity or provide some service denoted by q. For example, q could be the quantity of the fruit that the farmer (agent) delivers to a processor (principal). In a production contract, q might represent the quality of service because the integrator is essentially hiring labor, specialized knowledge, diligence, flexibility, etc. Let $q \in [\underline{q}, \overline{q}] \subseteq [0, ..., +\infty]$. Assume that the principal prefers higher q but higher q is costly to the agent. That is, if $r(q)$ is the principal's revenue, then $r'(q) > 0$, and if $c(q)$ is the agent's cost, then $c'(q) >$. For completeness, curvature assumptions, $r''(0) \leq 0$ and $c''(q) \geq 0$, are also assummed to facilitate well-behaved solutions. The reservation payoffs are $\overline{\pi}$ (principal) and $\overline{u}$ (agent). Finally, assume that $r(\underline{q}) - c(\underline{q}) \leq \overline{u} + \overline{\pi}$ so that contracting under minimal performance results in a loss in efficiency relative to next best alternatives.

The crucial assumption that distinguishes formal contracts from relational contracts is the degree to which q can be measured and verified by a third-party, such as a court or arbitrator. For some commodities, quality grading systems are well established so that if q represents quality, it can be included in a formal contract. However, in other situations where grading technology is not available, no credible third-party can verify q or specialized knowledge is needed to assess performance, third-party enforcement is prohibative so that it would be impossible to determine whether $q \geq \hat{q}$ where $\hat{q}$ denotes the performance level specified in the contractual agreement. Since this study is about relational contracts, it is assumed that q cannot be enforced by a third-party. However, to keep things simple, q is assumed to be observable to the "insiders"; i.e. the traders themselves. Thus, there is no moral hazard or adverse selection. The only informational asymmetry relates to the ability of third-parties to observe q.[4]

The generic form of the contract that is used to engage in trade is $w(q) = f + b(q)$ where $w(q)$ is total compensation, f is the fixed payment component that is independent of q, and $b(q)$ is a q dependent bonus. Note that in the textbook principal-agent setup, f and $b(q)$ are determined by the principal (in a take-it-or-leave-it situation) or jointly (when there is ex ante bargaining). The key point here is that contractual form is not exogeneously determined - it is determined endogenously. Moreover, if q is not third-party enforceable, then $b(q)$ would be a *discretionary bonus* because if q cannot be third-party verified, then neither can any payment that is conditioned on q. Thus, $b(q)$ must be discretionary in the sense that the principal can always claim poor performance by the agent as a justification to pay $b(q) < b(\hat{q})$ even if $q \geq \hat{q}$. Consequently, a credible contract that inspires performance from both parties must be self-enforcing to ensure that both parties carry out their obligations under the agreement.

The contracting timeline follows the typical textbook principal-agent sequence:

1. Principal designs and offers a contract to the agent where the contract includes a specification of terms $(\hat{q},\ f,\ b(\hat{q}))$.[5]

2. The agent decides to accepts or reject the contract. If rejected, the parties realize their reservation payoffs.

3. If accepted, the agent chooses the level of q.

4. The principal observes q and chooses $b(q)$.

Given the assumption that q is not third-party enforceable, a method for self-enforcement must be outlined. The received literature on relational contracts generally focuses on repeat trading as a means of self-enforcement. Note that the above contracting time-line represents what occurs within a single period (or stage-game) of a repeated game. In a relational contract, the stage-game is repeated an indefinite number of periods. Hence, future rewards and punishments can be conditioned on current performance.

More specifically, a relational contract is self-enforcing if it is a subgame perfect equilibrium of the indefinitely repeated game. In order to simplify the contract design problem, Levin (2003) shows that one can focus on stationary contracts; i.e. the same optimal contract is offered in every period, t. I assume that the principle designs a contract that is offered on a take-it-or-leave-it basis to the agent. Let δ be the common discount factor. Then the contract design problem is:

$$\max_{(\hat{q},f,b(\hat{q}))} \quad [r(\hat{q}) - f - b(\hat{q})] + \frac{\delta}{1-\delta}V(c), \tag{1}$$

$$\text{s.t.} \quad [r(\hat{q}) - f - b(\hat{q})] + \frac{\delta}{(1-\delta)}V(c) \geq \frac{\overline{\pi}}{(1-\delta)}, \tag{2}$$

$$[f + b(\hat{q}) - c(\hat{q})] + \frac{\delta}{1-\delta}U(c) \geq \frac{\overline{u}}{1-\delta}, \tag{3}$$

$$[r(\hat{q}) - f - b(\hat{q})] + \frac{\delta}{(1-\delta)}V(c) \geq [r(\hat{q}) - f] + \frac{\delta}{(1-\delta)}\overline{\pi}, \tag{4}$$

$$[f + b(\hat{q}) - c(\hat{q})] + \frac{\delta}{(1-\delta)}U(c) \geq \left[f - c(\underline{q})\right] + \frac{\delta}{(1-\delta)}\overline{u}, \tag{5}$$

Constraints (2) and (3) are the the participation constraints (PC) and (4) is the self-enforcement (SE) constraint for the principal to ensure that the principal will pay $b(q) \geq b(\hat{q})$

if $q \geq \hat{q}$. In words, this constraint says the long-term payoffs from paying the bonus in the current period exceeds the long-term payoffs for shirking on the bonus. Constraint (5) is the self-enforcement constraint for the agent which makes delivering $q \geq \hat{q}$ incentive compatible. Note that $\underline{q}$ is on the r.h.s. of (5) because if the agent shirks, so that the principal will not pay the bonus, then $\underline{q}$ is the most profitable shirk. The payoffs $V(c)$ and $U(c)$) are the continuation payoffs from the cooperative state. If at least one party fails to cooperate, then the parties transition to the non-cooperative state and each party earns only reservation payoffs.

An important point to note is that the principal's "control" variables for ensuring that the parties remain in the cooperative state are the contract variables ($\hat{q}$, f, $b(\hat{q})$). This is because if these contracting variables are chosen in such a way that all PC and IC constraints are satisfied, then, in princple, both parties will honor the agreement.

Formally, for each period t, a *relational contract* is a complete plan of action that describes (a) the requested quality, $\hat{q}$ along with the compensation f and $b(\hat{q})$; (b) the agent's acceptance decision; (c) the agent's choice of q for all periods from $t = 0$ to period $t - 1$. Moreove, the contract is self-enforcing if honoring the agreement is part of a sub-game perfect equilibrium. The contract must also describe what happens off the equilibrium path; i.e. what happens if at least one party shirks. Abreu (1988) suggest that in repeated games, the optimal punishment is for the parties to revert to the worst equilibrium outcome which would mean termination of trade. However, Levin (2003) suggests that termination may not be renegotiation proof; instead the parties can continue to trade but with contract terms adjusted so that the deviator is held to his reservation payoff.

Solving the constrained optimization problem (1) will yield the optimal $\hat{q}^*$ and self-enforcing payment terms. But for the purposes of this study, the main concern is not the specific quantative values of $\hat{q}^*$ or the payment terms but rather what drives agent performance, which includes contract acceptance and delivery of promised $\hat{q}$. Analyzing the constraints can provide some predictions about the determinants of agent performance.

Constraint analysis

An analysis of constraints (3), (4) and (5) can provide insights into the key determinants of sellers' willingness to accept/reject contracts and to deliver on contractually specified levels of q. Starting with the agent's contract acceptance decision, we have the following proposition.

Proposition 1. *Under the assumption of stationarity, the agent will accept the contract if (i)* $f + b(\hat{q}^*) - c(\hat{q}^*) \geq \overline{u}$*; and (ii) if* $\delta[r(\hat{q}^*) - f - \overline{\pi}] \geq b(\hat{q}^*)$*.*

Proof. Let $\hat{q}^*$ be the optimal $\hat{q}$ from solving the constrained problem (1). Then the agent's participation constraint (3) can be written (after some re-arranging) as $(1-\delta)\,[f + b(\hat{q}) - c(\hat{q})] + \delta U(c) \geq \overline{u}$. By stationarity, $U(c) = f + b(\hat{q}^*) - c(\hat{q}^*)$ so (3) simplifies to $f + b(\hat{q}^*) - c(\hat{q}^*) \geq \overline{u}$. Thus, so long as the combination of f and $b(\hat{q}^*)$ are chosen to ensure that the inequality holds, then the agent would earn more from accepting than rejecting the contract so long as the principal pays $b(\hat{q}^*)$ when $q \geq \hat{q}^*$. To ensure credibility of $b(\hat{q}^*)$, constraint (5) must hold. Simplifying (5) yields $\delta[r(\hat{q}^*) - f - \overline{\pi}] \geq b(\hat{q}^*)$. $\square$

In words, if the principal wants to contract for a performance level $\hat{q}^*$, it must structure the payments so that the agent is promised higher profit from delivering $\hat{q}^*$ than the next best alternative. Moreover, the portion of the payment that is at the discretion of the principal,

$b(\hat{q}^*)$, cannot be excessively large or the agent will not believe that the principal will actually make the payment.[6]

The second prediction has to do with the agent's choice of q given that the agent accepts the contract. The relevant constraint is the agent's self-enforcement constraint (5), which facilitates incentive compatibility for the agent to deliver $q \geq \hat{q}$. An analysis of this constraint yields the following proposition.

Proposition 2. *Under the assumption of stationarity, the agent will choose $q \geq \hat{q}^*$ only if* $b(\hat{q}^*) \geq (1-\delta)[c(\hat{q}^*) - c(\underline{q})] - \delta[f - c(\hat{q}^*) - \underline{u}]$

Proof. Recall that constraint (5) ensures that the agent's payoff from choosing $q \geq \hat{q}^*$ exceeds short-term payoffs from shirking. The constraint can be rewritten as $b(\hat{q}^*) \geq c(\hat{q}^*) - c(\underline{q}) - \frac{\delta}{1-\delta}[U(c) - \overline{u}]$. By stationarity, the constratint becomes $b(\hat{q}^*) \geq c(\hat{q}^*) - c(\underline{q}) - \frac{\delta}{1-\delta}[f + b(\hat{q}^*) - c(\hat{q}^*) - \overline{u}]$. Solving for $b(\hat{q}^*)$ yields $b(\hat{q}^*) \geq (1-\delta)[c(\hat{q}^*) - c(\underline{q})] - \delta[f - c(\hat{q}^*) - \overline{u}]$. □

Proposition 2 specifies the lower bound on $b(\hat{q}^*)$ to ensure incentive compatible for the agent to deliver $q \geq \hat{q}^*$. Note that, in combination with Proposition 1, this implies that the discretionary bonus must be bounded as follows:

$$\delta[r(\hat{q}^*) - f - \overline{\pi}] \geq b(\hat{q}^*) \geq (1-\delta)[c(\hat{q}^*) - c(\underline{q})] - \delta[f - c(\hat{q}^*) - \overline{u}] \tag{6}$$

Furthermore, note that as $\hat{q}^*$ increases, the r.h.s. of (6) increases faster than the l.h.s. due to the convexity of $c(q)$ and the concavity of $r(q)$. Consequently, the largest $\hat{q}^*$ is one that forces (6) to just bind with equality.

Another way to interpret (6), is that the optimal discretionary payment has to be sufficiently high to ensure that the agent finds it self-enforcing to honor the agreement but not so high that it is not credible for the principal to pay the bonus. This can also be related to the discount factor, δ, by solving (6) for δ which yields,

$$\delta \geq \underline{\delta} = \frac{c(\hat{q}^*) - c(\underline{q})}{r(\hat{q}^*) - c(\underline{q}) - \underline{u} - \underline{\pi}} \tag{7}$$

where $\underline{\delta}$ is the critical threshold for mutual self-enforcement of $b(\hat{q}^*)$.

Corollary 1. *$\hat{q}^*$ is implementable only if $\delta \geq \underline{\delta}$.*

Proof. The proof is by contrapositive. Suppose that $\delta < \underline{\delta}$. Then it must be the case that $\delta[r(\hat{q}^*) - f - \overline{\pi}] \leq b(\hat{q}^*) \leq (1-\delta)[c(\hat{q}^*) - c(\underline{q})] - \delta[f - c(\hat{q}^*) - \overline{u}]$ with at least one inequality being strict. Hence, if $b(\hat{q}^*) > \delta[r(\hat{q}^*) - f - \overline{\pi}]$ then by Proposition 1, the agent will reject the contract. If $b(\hat{q}^*) < (1-\delta)[c(\hat{q}^*) - c(\underline{q})] - \delta[f - c(\hat{q}^*) - \overline{u}]$, then by Proposition 2, the agent will not deliver $q \geq \hat{q}^*$. $\square$

The main implication of Corollary 1 is that a larger δ facilitates the self-enforcement of a relational contract. Therefore, exogenous increases in δ should reduce contract rejection and shirking on relational contracts.

Experimental Design

The model outlined in the previous section forms the basis for the experiment design. The main difference is that specific parameters and functional forms had to be chosen though all curvature assumptions were maintained so that there is minimum loss of generality. Before these parameterizations are discussed, the experimental context is described.

In each experimental session, subjects were randomly assigned to be either "buyers" (principals) or "sellers" (agents). The buyer-seller wording minimizes technical jargon for the subjects and adds salience to the trading intentions of the experiment. The subjects also were separated by dividers in the laboratory and they only knew each other by assigned ID numbers that were not associated with actual identities. After these assignments, subjects were read instructions and took a control questionnaire to facilitate understanding. Next, two non-paying trial periods with suppressed ID numbers were conducted to familiarize subjects with the trading screens. When the actual, paying experiments began, exogenously matched buyer-seller pairs play a sequence of repeated stage-games, where each stage-game mirrors the stage-game described in the theoretical section.

At the end of each stage-game or period, there is a δ probability that the matched pair will play another stage-game. Probablistic continuation is a common method of implementing indefinitely repeated games in the laboratory (e.g. Murnighan and Roth (1983); Bó (2005); Bo and Fréchette (2011); Fudenberg et al. (2012); Bo and Fréchette (2019)). The probability of continuation varied between treatments as follows:

1. **0.8 treatment**: $\delta = 0.8$ – implies an expected five period repeated game (supergame).[7]

2. **0.5 treatment**: $\delta = 0.5$ – implies an expected two period repeated game (supergame).

Trading for matched buyer-sellers pairs continue until randomly terminated which concludes a *supergame*. And then each buyer is rematched with another seller and a new supergame continues until it is randomly terminated and so on.[8]

The overall experiment, which can consist of multiple supergames, ended if one of the following occurred: (1) all buyer-seller matchings have been exhausted so there is no way to avoid repeat supergame matching when the current supergame ends; (2) if at least 18 periods (across all supergames) in the $\delta = 0.8$ treatment or at least 20 periods in the $\delta = 0.50$ treatment have occurred, then the session ends when the current supergame randomly terminates. Condition (1) ensures stranger matching so no group reputation effects would confound results and condition (2) is a time management condition. These conditions ensured long enough sessions to allow learning effects to mitigate early period noise from learning how to play. Each session had between 16 to 22 subjects with more subjects recruited for the $\delta = 0.5$ sessions because the supergames were shorter and therefore there were more of them per session, which means more subjects are needed to avoid repeat supergame matching.

The experimental design also allows buyers to endogenously choose contractual form; i.e. all contractual variables ($\hat{q}^*$, f, $b(\hat{q}^*)$) are chosen by a buyer and then the contract is offered on a take-it-or-leave-it basis to the seller.[9] This flexibility allows for the experiment to nest a number of popular contracts seen in the theoretical literature, including discretionary bonus contracts ($f > 0$, $b(\hat{q}) > 0$), efficiency wage contracts ($f > 0$, $b(\hat{q}) = 0$), and even pure bonus contracts ($f = 0$, $b(\hat{q}) > 0$).

With regard to specific parameterizations for experimental implementation, q can only take natural numbers in the set $Q_A = \{1, 2, ..., 15\}$. While a continuous space can be used, natural numbers facilitates experimental implementation by decreasing confusion and reducing the complexity of calculations. The stage-game payoff functions are $\pi = 12q - f - b(q)$ and $u = f + b(q) - (q^2)/2$ where q is the realized quality/quantity rather than the contracted quality/quantity $\hat{q}$. All subjects were provided with Table 1 so that the seller's cost of delivering a specific q did not require calculations.

The reservations payoffs are $\overline{\pi} = \overline{u} = 15$, which are realized if either the buyer does not

Table 1: Seller's Cost

Quality	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Cost	1	2	5	8	13	18	25	32	41	50	61	72	85	98	113

offer a contract or the seller rejects an offer.

Under these parameters, joint surplus is maximied at $q = 12$ which yields surplus of 72. If parties do not contract, the joint reservation payoffs is 30 so contracting at high q is efficient. However, if $q < 3$ then joint surplus from contracting is less efficient than not contracting. During the experiment, subjects accumulate payoff points, which are converted into USD at the rate of 30 points=$1. This is a standard approach in the experimental economics literature (e.g. see Bó (2005)).

The stage-game timeline in the experiment matches the theoretical principal-agent sequence described earlier:

1. **Proposal phase**-buyer offers a single contract with terms $(\hat{q},\ f,\ b(\hat{q}))$ to seller.

2. **Acceptance phase** - seller decides whether to accept or reject the offer.

3. **Quality phase**-seller chooses q.

4. **Payment phase**-if the buyer offered $b(\hat{q})$, then the buyer can choose actual $b(q)$ to pay. The payment f is guaranteed by the computer.

The experiments were conducted at a dedicated experimental economics lab with a no deception policy at a major U.S. state university under an approved IRB. Six sessions (three sessions of each treatment) involving 110 subjects were conducted. Average pay was approximately $28.5 U.S. dollars per-session, with a range from $21 to $59. These payouts translated into hourly pay that were consistent with average hourly rates of other experiments conducted in the same laboratory. All experiments were programmed with Z-tree (Fischbacher 2007). Prior to the live periods, considerable time was spent in each session familiarizing subjects with the trading platform via instructions, control questions, and practice rounds.[10] The average experiment lasted about three hours, including instructions, questionnaire, trial periods, post experimental payouts and post experimental questionnaire.

Results

Table 2 provides key summary statistics from all the experimental sessions. There were slightly more subjects across the $\delta = 0.50$ sessions because more subjects were needed to ensure stranger matching within each session in light of the shorter (on average) supergames. However, the average number of periods of each session was longer in the $\delta = 0.80$ sessions given the longer supergames.

Table 2: Summary statistics

		$\delta = 0.50$	$\delta = 0.80$
# of sessions		3	3
# of total subjects		60	50
	# of buyers	30	25
	# of sellers	30	25
Averge # of periods per session		24	28.3
# of total trading opportunities		724	707
# of total offers		542	586
	# self-enforcing	13	285
Total # accepted		295	351

The total number of trading opportunities, which is the total number of periods across all subjects and sessions, is roughly equal across the two treatments (724 versus 707). Note that buyers were not required to make offers but the overwhelming majority did in both treatments (542 and 586). A self-enforcing offer is one where a buyer's offered contract ($\hat{q}$, f, $b(\hat{q})$) satisfies the all four constraints (the two PC and SE constraints) under the experimental model parameterizations discussed earlier. One can see that only 13 offers were self-enforcing in the $\delta = 0.50$ sessions but a substantial 285 offers were self-enforcing in the $\delta = 0.80$ sessions. Overall, 295 and 292 offers were accepted and these equal the number of completed trades.

Hypotheses

Before moving on to the next set of statistics, the two main hypotheses tested in this paper are described. These hypotheses are related to the determinants of relational contractual performance and follow naturally from the propositions in the theoretical section.

Hypothesis 1. *Contract offers that satisfy the self-enforcement and participation constraints lead to reduced rejection rates and reduced rates of $q < \hat{q}$.*

Hypothesis (1) follows naturally from propositions (1) and (2). Indeed if a single constraint is violated, the theory predicts that the contract may violate the equilibrium conditions for a self-enforcing relational contract which could compromise parties' willngness to either accept or honor the requested performance level.

The next hypothesis is related to Corollary 1.

Hypothesis 2. *An exogenous increase in δ reduces contract rejection rates and the rate at* *which $q < \hat{q}$.*

An increase in δ implies that the future becomes relatively more important so that self-enforcement should be theoretically stronger. This hypothesis can be directly tested because the treatment variation involved $\delta = 0.50$ and $\delta = 0.80$.

Descriptive statistics related to the hypotheses.

Before proceeding to the formal hypothesis testing, some plots across periods can provide a visualization of the evolution of seller rejection rates and seller shirk rates (i.e. $q < \hat{q}$) as subjects gained experience.

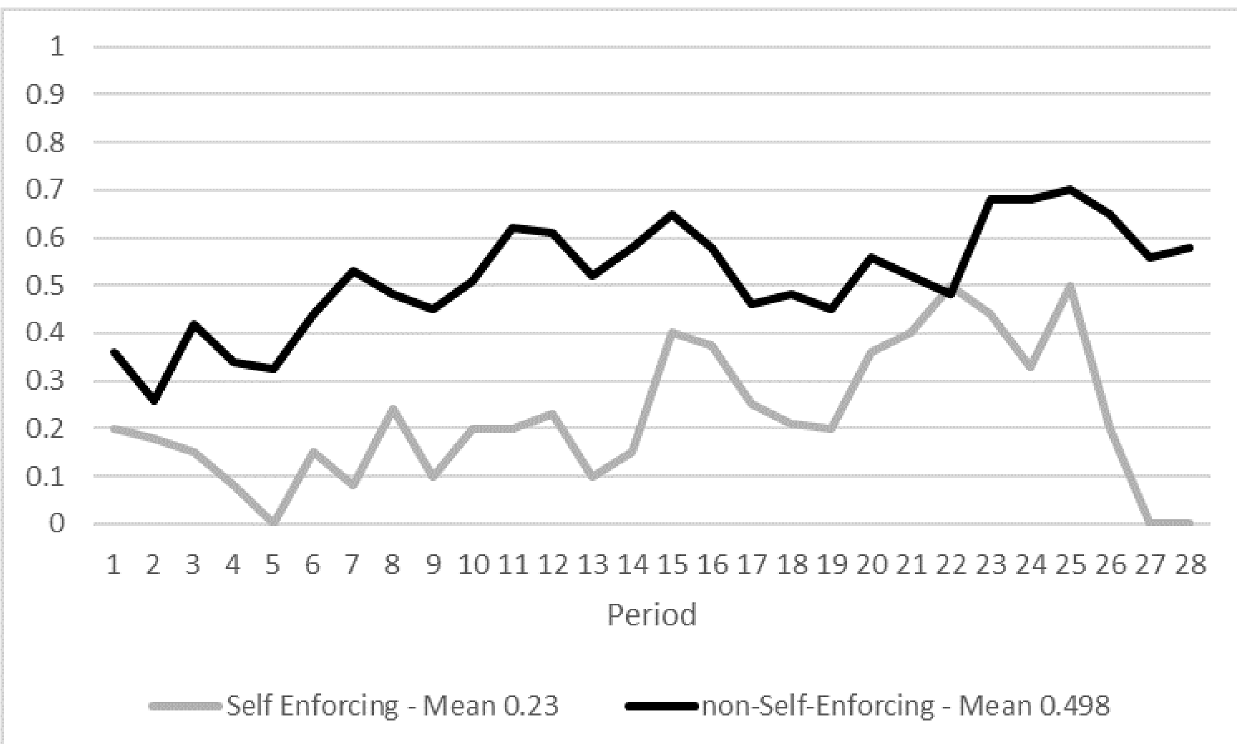

Figure 1: *Seller contract rejection rate for self-enforcing and non-self-enforcing contract offers.*

Figure 1 plots seller rejection rates (for all sessions) across periods for offers that are self-enforcing versus those that are not self-enforcing. For all periods, the rejection rate is 0.23

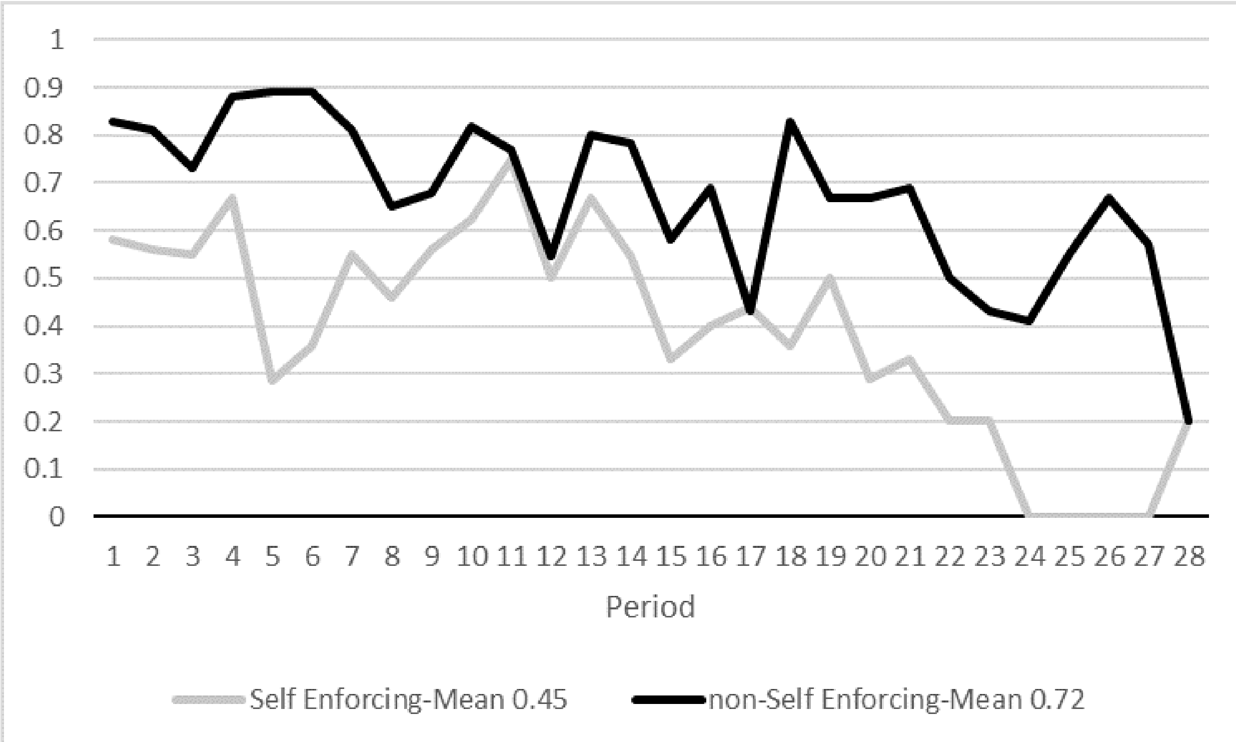

Figure 2: *Seller shirk rate for self-enforcing and non-self-enforcing offers.*

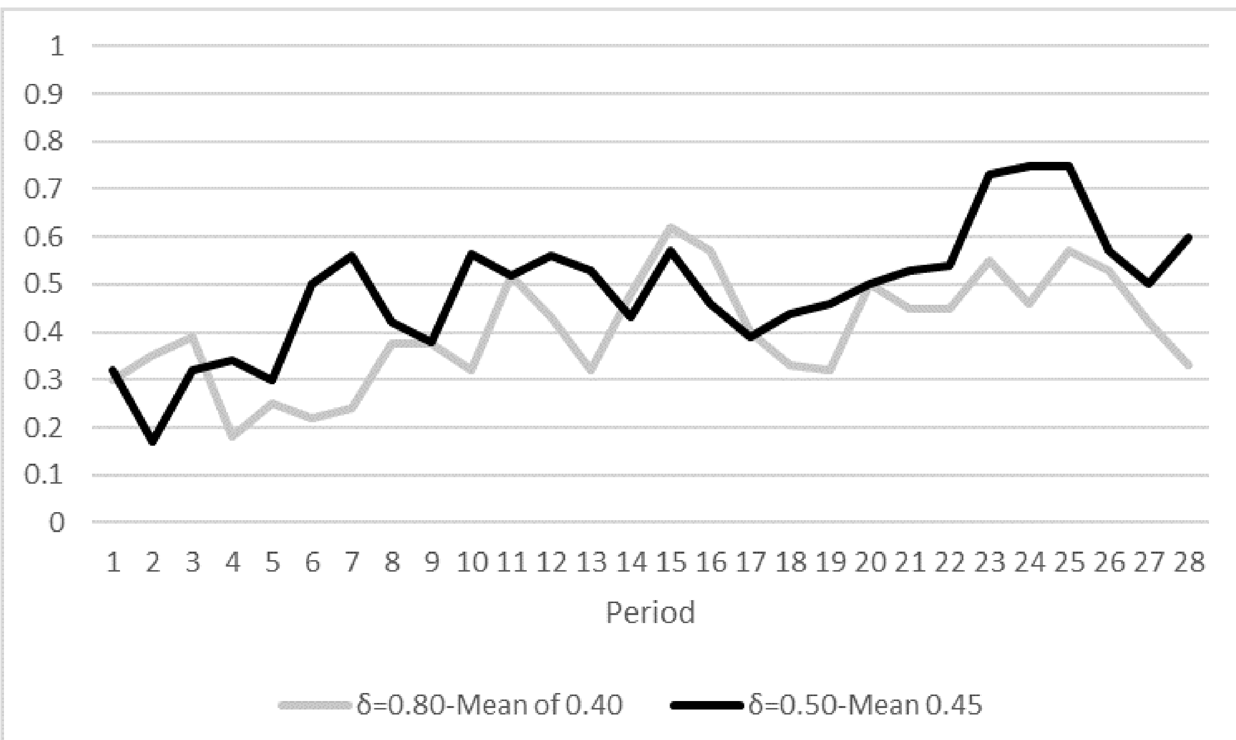

Figure 3: *Seller contract rejection rate for* $\delta = 0.80$ *and* $\delta = 0.50$ *treatments.*

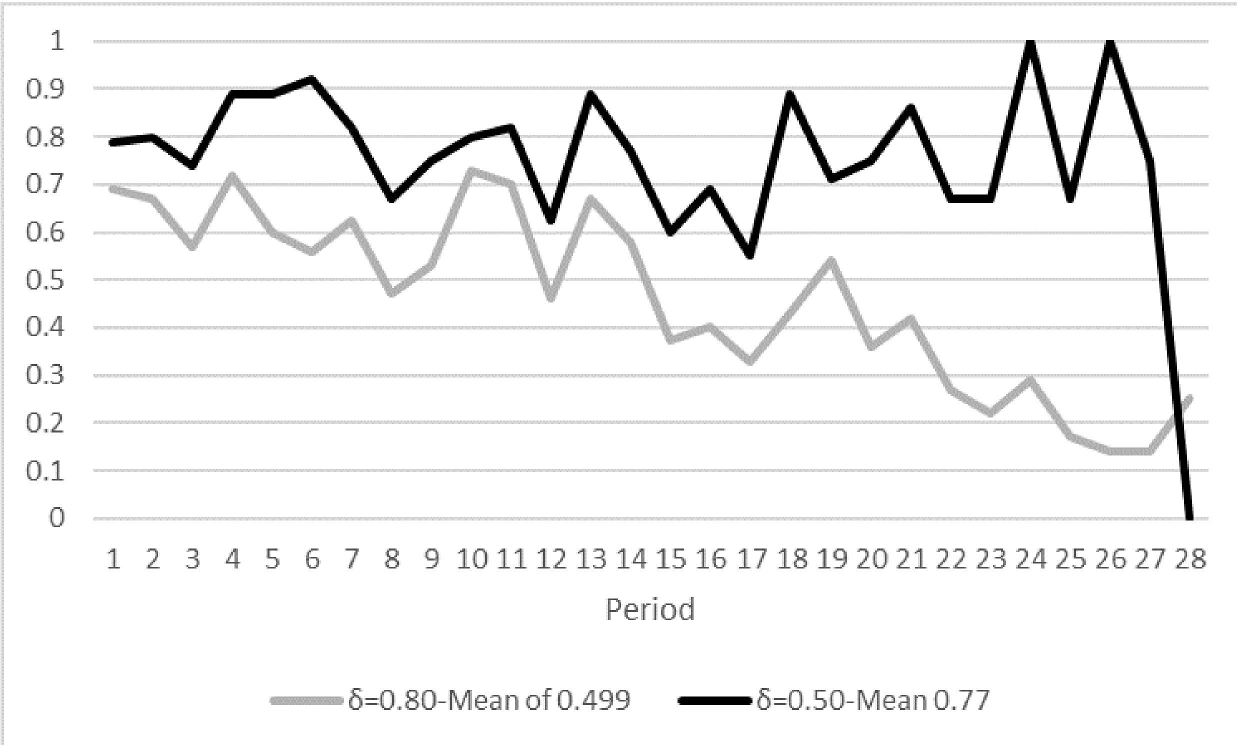

Figure 4: *Seller shirk rate for* $\delta = 0.80$ *and* $\delta = 0.50$ *treatments.*

for self-enforcing contracts but more than double at 0.498 for non-self-enforcing contracts which is consistent with Hypothesis 1. Morever, the gap mostly persists across periods.

Figure 2 plots seller shirk rate of contracts that have been accepted. The overall shirk rate is 0.45 for self-enforcing contracts and a much higher 0.72 for non-self-enforcing contracts. This is again consistent with theory and Hypothesis 1.

Figure 3 plots rejection rates for the $\delta = 0.80$ versus the $\delta = 0.50$ sessions. The overall all rejection rates were similar across the two treatments (0.40 versus 0.45) and, as seen in the graph, there is no obvious separation in rejection rates across periods. Thus, one cannot draw conclusions about Hypothesis 2 simply through visualization of this graph. Formal hypothesis tests will be presented later in this section.

Figure 4 plots seller shirk rate of the $\delta = 0.80$ and $\delta = 0.50$ treatments. The overall shirk rate is 0.499 for the $\delta = 0.80$ treatment but substantially higher at 0.77 for the $\delta = 0.50$ treatment. Moreover, the gap appears to persist across nearly every period. These patterns are consistent with Hypothesis 2, though formal hypothesis tests will be conducted.

Hypothesis tests

This subsection describes the results of the formal hypothesis tests conducted using regression analysis. Table 3 contains the results from four linear probability model (LPM) regressions examining the determinants of what drives sellers to reject contracts. The dependent variable takes a value of "1" if the seller rejected a contract offer. Four different specifications were used to provide some insights into the robustness of the results.[11] All regressions also include the control variables, *period*, which is a count of periods in the session to capture subject learning, and *period*-squared to capture nonlinearities. Seller fixed effects were also included in regressions (3) and (4).

With respect to the independent variables, the "$\delta = 0.80$ dummy" takes a value of "1" if the observation came from the $\delta = 0.80$ treatment and "0" if it belong to the $\delta = 0.50$ treatment. The "Self-enforcement-participation constraint dummy"(SEPC) takes a value of "1" if the observation contains a contract offer that satisfies the SE and PC constraints. The $\hat{q}$ variable is just the requested quality level that the buyer specified in the contract offer. This variable is important given that q is the primary driver of the joint surplus of the contracting relationship and if q falls below a certain level under the experimental parameterizations, surplus becomes negative. These three independent variables are largely motivated by the theoretical model described earlier.

Under Hypotheses 1 and 2, the coefficients for the $\delta = 0.80$ dummy and the SEPC

Table 3: LPM Estimates

	Binary Dependent Variable - Seller Reject=1			
	(1)	(2)	(3)	(4)
$\delta = 0.80$ *dummy*	0.10*	0.077	0.52***	-0.31***
	(0.08)	(0.05)	(0.04)	(0.07)
Self-enforcement-participation constraint dummy	-0.36***	-0.24**	-0.27***	-0.27***
	(0.04)	(0.04)	(0.03)	(0.03)
$\hat{q}$	-0.02	-0.01***	-0.016	-0.015**
	(0.01)	(0.004)	(0.009)	(0.005)
1-memory cooperation dummy		-0.285***		-0.21*
		(0.06)		(0.09)
Promised seller share		-0.009		-0.03
		(0.02)		(0.03)
1-memory no-trade dummy		0.33***		0.25***
		(0.04)		(0.04)
Period	0.02**	0.02	0.02**	0.019
	(0.006)	(0.01)	(0.005)	(0.01)
Period^2	-0.0005**]	-0.0004	-0.0004**	-0.0005
	(0.0001)	(0.0003)	(0.0001)	(0.0003)
Constant	0.50***	0.37***	0.26**	0.70***
	(0.14)	(0.07)	(0.095)	(0.09)
Seller fixed effects	No	No	Yes	Yes
Observations	1128	789	1128	789
R^2	0.11	0.33	0.24	0.41

-Robust standard errors clustered on sessions are reported in parentheses.

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

dummy should be negative. That is, higher discount factors and contracts that satisfy all constraints should increase contractual acceptance. Regression (1) in Table 3 suggests that only the SEPC has the correct sign and is statistically significant ($p < 0.01$). The $\delta = 0.80$ is particularly interesting as it is significant at the 10% level of significance but has the wrong sign. One possible explanation is that the impact of δ operates through the contract variables so that including the SEPC dummy and $\hat{q}$ may dampen the direct effect of δ. But I ran other specifications with the SEPC and $\hat{q}$ variables omitted and the $\delta = 0.80$ dummy remained non-significant.

Regression (2) adds a 1-memory cooperation dummy variable, which takes a value of "1" if the same buyer-seller pair engaged and honored a relational contract in the previous period (i.e. $q \geq \hat{q}$ and $b \geq b(\hat{q})$). A 1-memory no-trade dummy was also added which takes a value of "1" if the buyer-seller pair did not trade at all in the previous period. Finally, the share of the total surplus promised to the seller under the contract terms was also included as a regressor. While these variables do not follow directly from the model presented earlier, the game-theoretic concept of *best response dynamics* is based on the idea that each player believes that the other player will choose the actions s/he chose in the previous period. Thus, the 1-memory cooperation dummy coefficient is expected to be negative while the 1-memory no-trade coefficient is expected to be positive. The promised seller share is expected to decrease rejection rates because higher share relaxes the participation constraint.

The results of regression (2) suggest that the both the 1-memory cooperation and no-trade coefficients are statistically significant ($p < 0.01$) and have the expected signs. That is, prior period cooperation reduces rejection probability but prior period no-trade increases rejection probability. The promised seller share coefficient is not statistically significant. The SEPC dummy coefficient continues to have the correct sign and is statistically significant ($p < 0.05$). The $\delta = 0.80$ dummy coefficient is not statistically significant.

Regressions (3) and (4) are identical to (1) and (2), respectively, with the exception that seller fixed effects are added since unobserved seller heterogeneity could create selection effects into certain contract types so that the error term might be correlated with the SEPC dummy and $\hat{q}$. Once these fixed effects are added, the SEPC dummy coefficients remain robust and statistically significant ($p < 0.001$). The 1-memory cooperation dummy coefficient and the 1-memory no-trade coefficient also retain the correct signs and statistical significance at at least the 10% level. The other coefficients of interest either are not significant or not robust to other variables or seller fixed effects. The $\delta = 0.80$ treatment coefficient is particularly problematic as it also switches signs from regressions (3) to (4). Recall that visually, it was hard to detect that there was a difference in rejection rates across the $\delta = 0.80$ and $\delta = 0.50$ in Figure 3, and the regressions fail to yield additional clarity.

Overall, it appears that contracts that satisfy the theoretically important self-enforcement and participation constraints are robustly important in predicting contract rejection behavior. This is consistent with Hypothesis 1. However, we find no evidence that an increase in δ from 0.50 to 0.80 reduces rejection rates which contradicts Hypothesis 2. The best response dynamics proxies, 1-memory cooperation and 1-memory no-trade, also seem to predict contract rejection by sellers.

Table 4 reports the results from six LPM regressions concerning seller shirking where the dependent variable equals "1" if $q < \hat{q}$ and "0" otherwise. For similar theoretical reasons, the set of regressors is the same as those found in Table 3 for reject decisions, with the exception that an interaction term between the $\delta = 0.80$ dummy and the SEPC dummy was included in regressions (3) and (6). Recall from Table 2 that the number of self-enforcing contracts depends heavily on the δ treatment. This interaction term is meant to capture the possibility that the impact of δ and SEPC on seller shirk probability is interralated.[12]

Under Hypothesis 1, the sign of the SEPC coefficient should be negative. Looking at regressions (1), (2), (4), and (5) that do not include the interaction term, the estimated SEPC coefficients are negative but statistically significant only in (1) and (5). In regression (3), which includes the interaction term, it appears that satisfying the SEPC constraints

Table 4: LPM Estimates

	Binary Dependent Variable - Seller Shirk=1					
	(1)	(2)	(3)	(4)	(5)	(6)
$\delta = 0.80$ *dummy*	-0.15	-0.01	0.01	-0.56***	-0.28***	0.28**
	(0.09)	(0.03)	(0.03)	(0.05)	(0.06)	(0.08)
Self-enforcement-participation constraint dummy	-0.16**	-0.05	0.22***	-0.07	-0.06**	0.05
	(0.05)	(0.05)	(0.04)	(0.06)	(0.02)	(0.03)
$\hat{q}$	0.008	0.02**	0.017**	0.02**	0.02**	0.02**
	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
1-memory cooperation dummy		-0.69***	-0.69***		-0.38**	-0.37**
		(0.06)	(0.06)		(0.1)	(0.1)
Promised seller share		-0.13**	-0.13**		-0.26***	-0.26***
		(0.03)	(0.03)		(0.04)	(0.04)
1-memory no-trade dummy		-0.09	-0.08		-0.04	-0.04
		(0.06)	(0.06)		(0.06)	(0.06)
$\delta = .80 \times SEPC$			-0.30***			-0.12***
			(0.05)			(0.02)
$(\delta = .80) + (\delta = .80 \times SEPC)$			-0.29***			0.17
			(0.06)			(0.09)
$(SEPC) + (\delta = .80 \times SEPC)$			-0.08*			-0.07**
			(0.035)			(0.02)
Period	0.001	-0.003	-0.002	0.0007	-0.005	-0.005
	(0.005)	(0.005)	(0.004)	(0.008)	(0.004)	(0.004)
Period^2	-0.0005**	-00008	-0.0001	-0.0003	0.00007	0.00007
	(0.00016)	(0.0001)	(0.0001)	(0.0003)	(0.0001)	(0.0001)
Constant	0.78***	0.79***	0.77***	0.82***	0.93***	0.90***
	(0.08)	(0.06)	(0.06)	(0.08)	(0.14)	(0.08)
Seller fixed effects	No	No	No	Yes	Yes	Yes
Observations	646	407	407	646	407	407
R^2	0.14	0.46	0.46	0.50	0.62	0.62

-Robust standard errors clustered on sessions are reported in parentheses.

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

in the $\delta = 0.50$ treatment does not decrease seller shirking but it does in the $\delta = 0.80$ treatment. This can be seen from the negative interaction coefficient (-0.30, $p < 0.001$) and the negative estimated sum of the interaction coefficient and the SEPC coefficient (-0.08,

$p < 0.10$). Similar qualitative results are seen in regression (6), which includes seller fixed effects. These results mostly support Hypothesis 1 with regard to seller shirking at least in the $\delta = 0.80$ treatment. However, it is important to note that only 13 out of 543 offers in the $\delta = 0.50$ treatment satisfied SEPC. Moreover, when sellers delivered on the contractually obligated $\hat{q}$ in the $\delta = 0.80$ treatment, only 7% of the time did buyers fail to pay the promised $b(\hat{q})$. However, in the $\delta = 0.50$ treatment, nearly 22% of trades for which $q \geq \hat{q}$ resulted in the buyer failing to pay the promised bonus. The combination of the low frequency of credible contracts combined with the higher frequency of buyers failing to honor their bonuses in the $\delta = 0.50$ treatment could mean that seller trust was largely eroded even when SEPC conditions were satisfied.

Under Hypothesis 2, the coefficient for the $\delta = 0.80$ dummy is expected to be negative. In the regressions without the interaction term, all estimated coefficients had the expected sign but were statistically significant only in regressions (4) and (5), which include seller fixed effects. Regressions (3) and (6) include the interaction term and the estimated coefficients for this term are both negative, suggesting that the impact of an increase in δ in reducing seller shirking depends on whether contracts satisfy the SEPC constraints. Overall, there is only tentative evidence that an increase to $\delta = 0.80$ reduces seller shirking. It appears that much of the impact of δ on seller shirking operates indirectly through δ's impact on the propensity of buyer's to offer SEPC contracts.

With regard to the 1-memory dummies, the 1-memory cooperation coefficients have the expected sign of reducing seller shirk both with and without seller fixed effects. This variable appears to be a robust predictor of both reject and shirk decisions. Moreover, the absolute values of the 1-memory cooperation coefficients are larger than other coefficients, underscoring the value of prior cooperation even further. The 1-memory no-trade coefficients, however, are not statistically significant. However, this variable was important in explaining reject decisions. Thus, it appears that whether a seller traded or not with the same buyer in the previous period only drives acceptance decisions but not shirk decisions conditional on the contract being accepted.

Finally, promised seller share does appear to reduce shirking. This makes intuitive sense because if the payments are designed in such a way so as to compensate the seller relatively more if the contract is honored, then sellers may be more motivated to satisfy the terms of the contract.

Conclusion and discussion

This study uses economic experiments to examine the factors that drive agents (e.g. sellers or producers) to (1) adopt contracts and (2) to deliver on contractual obligations. The results mostly support the hypothesis that contracts that satisfy self-enforcement and individual rationality constraints reduce contract rejection rates by sellers. Surprisingly, there is no robust evidence that an increase in the discount factor reduces contract rejection by sellers. There is tentative evidence that satisfying the self-enforcement and individual rationality constraints reduces seller shirking, but primarily in the treatment with the higher discount factor. There is also tentative evidence that an increase in the discount factor reduces seller shirking but only if contracts satisfy the aforementioned constraints. It is possible that a higher discount factor reduces seller shirking primarily through its impact on the number of contracts that satisfy the constraints. The results also point to the importance of some variables that are not explicitly accounted for by the bare bones relational contracting model. In particular, if a principal and agent cooperated in the previous trading period, this appears to robustly predict both rejection and shirk decisions by agents.

The implications of these results are that, if agricultural economists or development agencies would like to facilitate contract adoption by farmers, it is important as a first step to ensure that contracts satisfy self-enforcement and participation constraints. Moreover, convincing farmers to adopt new crops or technology may be easier if the farmers have had a history of cooperation with the contractor, which might mean that self-enforcement could be relaxed. Of course, contractors are often not just concerned about adoption but also about whether farmers will deliver on the quality or volume requirements specified in the contracts. Our study suggest that a prior history of cooperation is the most robust predictor of contract compliance, though ensuring that the contract satisfies self-enforcement and participation costraints could help.

A qualification of the study is that it does not account for all important factors that may lead to contractual adoption and performance. However, given the paucity of empirical studies on this topic, this study can be seen as a starting point for testing the first principles of contract design; i.e. for identifying the necessary ingredients of any successful contract. Future applied studies or program evaluations that focus on specific crops or industries should incorporate situation specific determinants in addition to the determinants identified in this study.

Footnotes

[1]Baker et al. (1994) suggest that, even when formal contracts are used, relational contracts based on tacit expectations might still be layered on top of formal contracts to complete the governance of the full set of outcomes that the trading parties care about. One example is that many agricultural contracts contain detailed provisions on how quality or efficient input use are measured but are silent about quantity obligations. This would be an example where a formal contract governs quality but an informal agreement governs the volume.

[2]Cason and Wu (2019) provide a detailed discussion of the advantages and disadvantages of both laboratory and field experiments. Two points that are relevant for this study are the following. First, the "external validity" of laboratory experiments do not come from any specific laboratory session but rather from the theory that is being tested. In other words, lab experiments can be used to stress test theory through various interventions. Theories that survive multiple stress tests are more likely to generalize. Second, field experiments with field professionals are important for program evaluation where the main goal is to strive for internal validity of a study for a specific program for a specific population. But in the absence of theoretical grounding, there is no guarantee that the results of the field experiment is any more external valid in other settings as a lab experiment. These points are relevant because this is a study about a general theory of contracting not a program evaluation of a specific contracting situation.

[3]The contractual form is actually a solution to a principal's optimization problem.

[4]Within the mechanism design literature, there is the possibility that cleverly designed contracts, such as options contracts, can resolve third-party verifiability problems. However, strong assumptions must be made about renegotiation (Edlin and Hermalin 2000). Rather than get bogged down in the implementation literature, this study adheres to standard assumptions within the relational contracting literature.

[5]In practice, contracts can contain numerous terms beyond these such as act of God clauses, dispute resolution clauses, etc. However, a model is a parsimonious abstraction meant to capture the most critical components of an economic situation. A parsimonious description of a contract is that it is an enforceable (self or third-party) agreement that includes performance objectives along with payment terms to incent

those objectives.

[6]Note the participation constraint (3) still needs to hold even in the absence of stationarity in that the discounted long-run payoffs from honoring the contracts exceeds discounted reservation payoffs. Moreover, the discretionary payment needs to be credible. Thus, the assumption of stationarity simplies the modeling by allowing us to focus on single period payoffs rather than work with a model awash with summation notation, but the fundamental insights and intuition are unaffected.

[7]The expected number of periods is $\frac{1}{1-\delta}$.

[8]Note that a supergame is essentially a repeated game in the context of the theoretical model. Thus each of the sessions might involve multiple supergames or repeated games with different buyer-seller pairs.

[9]Take-it-or-leave-it offers are common in come agricultural sectors in the U.S. like boiler and hog contracting. It is also a realistic assumption in situations involving smallholders and large contractors. Even when bargaining is involved, the predictions from the take-it-or-leave-it are robust with the only change being that f should increase to allocate more surplus to the seller.

[10]For the most part, subjects appeared to have learned how to trade because, with the exception of one or two outliers per session, most were making quick decisions after the first few periods.

[11]Logit regressions were also estimated, but for the most part, the qualitative results were consistent with the LPM results with the logit results being slightly stronger in terms of p-values. The LPMs were also estimated with clustering at the session level. By clustering at such a high level and using the LPM rather than the logit results, I took a conservative approach to hypothesis testing by decreasing the odds of finding statstical significance.

[12]This interaction variable was also included in some initial regressions for contract rejection. However, the addition of this interaction term had little qualitative impact so only the simpler LPMs without the interaction term were included in Table 3.

References

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica*, pages 383–396.

Alexander, C., Ivanic, R., Rosch, S., Tyner, W., Wu, S., and Yoder, J. (2012). Contract theory and implications for perennial energy crop contracting. *Energy Economics*, 34(4):970–979.

Baker, G., Gibbons, R., and Murphy, K. (1994). Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics*, 109:1125–1156.

Bellemare, M. F. and Bloem, J. R. (2018). Does contract farming improve welfare? a review. *World Development*, 112:259–71.

Bo, P. and Fréchette, G. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, 101:411–429.

Bo, P. and Fréchette, G. (2019). Strategy choice in the infinitely repeated prisoner's dilemma. *American Economic Review*, 109:3929–3952.

Bó, P. D. (2005). Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *American Economic Review*, pages 1591–1604.

Brown, M., Falk, A., and Fehr, E. (2004). Relational contracts and the nature of market interactions. *Econometrica*, 72(3):747–780.

Carmichael, L. (1985). Can unemployment be involuntary? comment [equilibrium unemployment as a worker discipline device]. *American Economic Review*, 75(5):1213–1214.

Cason, T. and Wu, S. (2019). Subject pools and deception in agricultural and resource economics experiments. *Environmental and Resource Economics*, 73(3):743–758.

Dixit, A. K. (2007). *Lawlessness and Economics: Alternative Modes of Governance*. Princeton University Press.

Edlin, A. and Hermalin, B. (2000). Contract renegotiation and options in agency problems. *Journal of Law, Economics, and Organization*, 16(2):395–423.

FAO (2017). Contract farming and the law: What do farmers need to know? http://www.fao.org/3/a-i7581e.pdf. (Accessed on 09/06/2019).

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Fudenberg, D., Rand, D. G., and Dreber, A. (2012). Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review*, 102(2):720–49.

Gil, R. and Zanarone, G. (2018). On the determinants and consequences of informal contracting. *Journal of Economics and Management Science*, 27(4):726–41.

Klein, B. and Leffler, K. (1981). The role of market forces in assuring contractual performance. *The Journal of Political Economy*, pages 615–641.

Levin, J. (2003). Relational incentive contracts. *The American Economic Review*, 93(3):835–857.

MacDonald, J. and Burns, C. (2019). Marketing and production contracts are widely used in u.s. agriculture, *Amber Waves*, USDA, ERS, Washington, D.C., July 1. https://www.ers.usda.gov/amber-waves/2019/july/marketing-and-production-contracts-are-widely-used-in-us-agriculture/.

MacLeod, W. and Malcomson, J. (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica*, pages 447–480.

Michler, J. D. and Wu, S. Y. (2020). Relational contracts in agriculture: Theory and evidence. *Annual Review of Resource Economics*, 12:111–127.

Murnighan, J. K. and Roth, A. E. (1983). Expecting continued play in prisoner's dilemma games: A test of several models. *Journal of Conflict Resolution*, 27(2):279–300.

Schmitz, P. (2001). The hold-up problem and incomplete contracts: A survey of recent topics in contract theory. *Bulletin of Economic Research*, 53:1–17.

Shapiro, C. and Stiglitz, J. (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review*, 74:433–444.

Tirole, J. (1999). Incomplete contracts: Where do we stand? *Econometrica*, 67(4):741–781.

APPENDIX-Instructions for Treatment $\delta = 0.80$

Note: The $\delta = 0.50$ instructions are identical except for the change in continuation probability from 80% to 50%.

You can earn money during this experiment, with the exact amount depending on the decisions you make during the experiment. Your experimental income is calculated in points, which will be converted into cash at the rate of: $1 = 30 points. We will start you off with a balance of 150 points ($5).

All written information you received from us is for your private use only. You are not allowed to pass over any information to other participants in the experiment. Talking during the experiment is not permitted. Violations of these rules may force us to stop the experiment.

General Information

This experiment is about how people buy and sell goods for which quality matters. Participants are divided into two groups: half will be buyers and the other half sellers. And then a trading period will start in which a buyer and seller will trade one unit of a good that can vary in quality. The price agreed upon between the buyer and seller and the quality of the good traded will determine how much money each party makes in that period. There will many trading periods throughout the course of this experiment.

Who will you trade with? At the beginning of the experiment, the computer will randomly match each participant in the room with another participant to form a buyer-seller pairing. You will be informed whether you are the buyer or seller in your pairing. You will trade with your pair-member. You will *not* be informed of the actual identity of the other person (and s/he will not be informed of your identity). *All sellers and buyers are assigned a numeric ID which is not associated with their real identity.* You will also retain your ID and role (e.g. buyer or seller) through the entire experiment.

For how many periods will you trade with the same person? All participants will remain matched with their pair-member for a random number of periods. **How is this determined?** At the end of each period, the computer will determine randomly whether the same pairings will continue for the next period or whether new pairings will be formed. In any given period, there is an **80% chance** that the same pairings will continue for the next period. In other words, in any given period, there is a 80% chance that you will continue to trade with the same person in the next period. To help you understand this, imagine that the computer has been programmed to spin a roulette wheel. If it lands on 1,2,3,4, 5, 6, 7, or 8 then you will continue to trade with the same person the next period. But if it lands on 9 or 10 the current pairings are immediately terminated. And then for the next period, the computer will randomly match you with a different person in the room to form a new pairing. This process will repeat for every new pairing. *At the beginning of each period, you will be notified on-screen whether the random matching process has kept you with the same person or matched you with a new person.*

When does the entire experiment end? If one of two conditions hold: (1) The experiment will end if all participants have already been matched with all possible trading partners. *This is because no participant will be matched with the same person more than once during this experiment.* For example, if there are 10 buyers and 10 sellers, then no buyer or seller will have more than 10 unique pairings. After 10 unique pairings, the experiment ends. (2) Even if all unique pairings have not been exhausted, the last pairing will occur once the experiment has lasted at least 18 periods. In other words, if you have traded at least 18 periods for the experiment, then your current pairing is your last one. **This does not mean the experiment stops at 18 rounds exactly;** it only means that when your last pairing **randomly** ends, you will not be paired with a new partner.

To summarize , if you have had less than 10 different trading partners during the experiment, but the experiment has not lasted at least 18 total periods, then when your current match is randomly terminated, the computer will match you with a new person and the experiment would continue. However, if the experiment has lasted at least 18 total periods, then the experiment will end once your current pairing is randomly terminated.

CONDUCTING TRADES

Each trade occurs within a trading period. Each trading period is then divided into a *proposal phase* followed by *a quality determination phase* and then followed by a *payment determination phase***.**

a) During the *proposal phase*, the buyer can make a proposal on the terms of trade to the seller. The seller can either accept or reject the proposal.
b) If the seller accepts the proposal, then during the *quality determination phase*, the seller chooses the actual quality level to supply.
c) After quality is observed, comes the *payment determination phase*. During this phase, the buyer can make final adjustments in payment depending on the initial terms of the proposal.

During each phase, you can take as much time as you need to make a good decision, but the faster you make your decision, the faster the experiment will move.

Specific details of each phase are given below:

1. The Proposal Phase

Each period starts with a proposal phase. A proposal allows the parties to agree to the terms of trade by including a list of promises and obligations of both parties (see below for details). *The buyer can submit a single proposal during the proposal phase. Once a proposal is submitted, the seller will decide to accept or reject the proposal.*

How does a buyer make a proposal? A proposal screen will appear that will require the buyer to enter values for the following terms: *desired quality*, *price*, and a *performance bonus***.** These terms are described below.

a) **Desired quality** - A buyer can ask the seller to deliver a specific level of quality. Quality can range from 1 to 15, where higher numbers indicate higher quality (whole numbers only). To specify quality, the buyer enters a number in the "*Desired quality* " field.
   Important: during the quality determination phase to come later, the seller can choose any quality level s/he wishes. In other words, the Desired quality allows the buyer to request a specific quality level but it is not binding on the seller to deliver this quality level.

b) **Price** – This allows the buyer to state the price she will pay for the good. The buyer enters a price in the "*Price*" field. The price ranges from 0 to 200 (whole numbers).

   The price the buyer specifies will be *binding*. That is, this is similar to an upfront payment or a legally binding obligation – once the proposal is agreed upon, the computer will ensure that the price is paid to the seller.

c) **Performance bonus**– For the case when desired quality is discretionary, the buyer can state that s/he will pay a bonus that might be linked to quality. *To enter a bonus, click on the "yes" box next to "would you like to offer a bonus?"* Then enter a number in in the "*Bonus*" field to specify the size of the bonus (enter a whole number from 0 to 200). *If the buyer does not wish to offer a bonus, simply click "no" next to "would you like to offer a bonus?"* The total payment is *price* plus *bonus*.
   Important: The stated bonus is *not binding***.** During the payment determination phase to come later, the buyer can choose any bonus level s/he wishes. Thus, this is a discretionary bonus. However, if the buyer clicked "no" to offering bonus, then there will be no payment determination phase for the buyer in this period. The **Price** then becomes the final payment.

After the buyer has specified desired quality, price and performance bonus, s/he needs to click "OK" to submit it. Next comes the quality determination phase.

1. Quality Determination Phase

Following the proposal phase, all sellers who accepted an agreement will determine the level of quality that they will supply to their buyers. A seller can choose any quality s/he wants to from 1 to 15. The *Quality Determination* Screen will appear and a seller can enter his/her quality choice in the "*Actual Quality"* field. Nothing restricts the seller from choosing a quality level that is different from the "desired quality" level specified in the proposal.

2. Payment Determination Phase

Following the quality determination phase, all buyers who offered a bonus will determine the level of actual bonus that s/he will pay to the seller. During this phase, **after quality is observed by the buyer**, the buyer will choose actual bonus to be paid to the seller. The *Payment Determination* screen will appear and the buyer will enter his/her bonus choice in the "*Actual Bonus"* field. **Nothing restricts the buyer from choosing a bonus level that is different from the bonus that was specified in the proposal**. The actual bonus can range from 0 to 200 at the buyer's discretion.

How Are Points (Income) Calculated?

How do Buyers Make Money?

- If the **buyer does not make an offer or the seller rejects the offer**, the **buyer will receive 15** points for that period.
- If the buyer's proposal is accepted, the buyer's points for the period depend on the actual quality, the price and the actual bonus paid. That is,

Buyer Points = 12*Actual Quality – Price –Actual Bonus

- As you can see, the higher the actual quality, the more points the buyer earns. At the same time, the lower total payments (price plus actual bonus), the more points the buyer earns.
- **In summary, higher quality at lower payments means more points for the buyer.**

How do Sellers Make Money?

- If the **seller rejects the proposal or the buyer does not make an offer**, the **seller will receive 15** points for that period.
- If the seller has accepted an offer, then the seller's points depends on the price, actual bonus, and production costs s/he incurs. The points of a seller is determined as follows:

Seller Points = Price +Actual Bonus– Production Costs

- As you can see, **the higher the actual payments, the more points a seller earns. At the same time, the higher the quality, the higher the production costs, which reduces points.**
- **How are production costs calculated**? The higher the quality the seller supplies, the higher the costs. Roughly speaking, the cost is determined by the following formula: $Cost = \frac{q^2}{2}$. We say "roughly speaking" because we will round the cost number to the nearest whole number. The following table gives you the exact cost in whole numbers of producing each quality level.

Quality	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Cost	**1**	**2**	**5**	**8**	**13**	**18**	**25**	**32**	**41**	**50**	**61**	**72**	**85**	**98**	**113**

Points for all buyers and sellers are determined in the same way. **Each buyer can therefore calculate the income of his/her seller and each seller can calculate the income of his/her buyer.** Note that buyers and sellers can incur losses in each period. These losses are subtracted from your points balance.

At the end of each period, the buyer and seller will be shown an "**income screen**." The following information is displayed on this screen:

- the ID number of your trading partner.
- the **Price** the buyer offered.
- the **Proposed Bonus**
- the **Actual bonus** granted
- the buyer's **Desired Quality**
- the **Actual quality** delivered by the seller.
- the points earned (lost) by both parties in this period.

Please enter all the information on the screen in the documentation sheet supplied to you. This will help you keep track of your performance across periods so that you can learn from your past results.

At the beginning of the next period, the computer will inform you if you have been randomly matched with the same trading partner or with a different partner.

Before we begin the experiment, we ask all participants to complete a questionnaire which will test familiarity with the procedures. The experiment will not begin until all participants are completely familiar with all procedures. In addition, we will conduct **2 trial periods of the proposal phase** so that you can get accustomed to the computer. During the trial periods, no money can be earned. Your ID numbers will also be suppressed on the screen during the trial periods.