Discretionary Latitude and the Nature of Relational Contracting

by
Steven Y. Wu and Brian E. Roe[*]

Prepared for the American Economics Association Session: *Contracts, Incentives and Cooperation Between and Within Firms*.
ASSA Meetings, Chicago, Illinois, January 2007

Preliminary Draft: Do Not Cite Without Authors' Permission[**]

Abstract

Relational contracts typically combine explicit terms that are third-party enforceable with implicit components that can be adjusted at the parties' discretion. While discretionary *ex post* adjustments lies at the heart of relational contracting, previous empirical research has been primarily motivated by the *repeat purchase mechanism* (RPM) of Klein and Leffler (1981), which is a relational contract that permits only limited discretionary latitude. We use economic experiments to compare the RPM against more general forms of relational contracts, such as those that allow for discretionary bonuses/deducts. Our results suggest that trading outcomes are profoundly impacted by simple alterations in discretionary latitude. In experiments where maximum discretion is allowed (i.e. the explicit contract is highly incomplete), trading patterns more closely resemble those observed under complete contracts rather than those observed under the RPM. Thus, efficiency wages, generous rent sharing, and/or reputation based trading, which occur under the RPM, may not be general features of relational trading. Our results also provide support for the theory of *strategic ambiguity* of Bernheim and Whinston (1998); i.e. if there are barriers to complete contracting, then increasing the degree of contractual incompleteness may increase efficiency.

Keywords: relational contracts, implicit incentives, experimental economics, reputation.
JEL Classification: C91 D23 D84 D86 J33 K12

1. INTRODUCTION

Self enforcing relational contracts are important mechanisms for exchange in environments characterized by weak contract enforcement institutions or where performance is difficult to describe *ex ante* and difficult to verify *ex post* by a third-party. While much theoretical progress has been made in recent years concerning the nature and structure of relational contracts (e.g. Dixit 2003; Levin, 2003; Baker, Gibbons and Murphy 2002; Bernheim and Whinston 1999; MacLeod and Malcomson 1989 and 1998, among others), empirical knowledge of relational trading is still limited. Most of the empirical or experimental literature on relational trading tends to focus on implications of the *repeat purchase mechanism* of Klein and Leffler (1981), which is a relational contract of a very specific form (MacLeod 2006). Under the repeat purchase mechanism, a buyer and seller have the opportunity to trade repeatedly across many periods. In each period, the parties agree to trade a good at a fixed price, after which the seller chooses quality which is not third-party enforceable. Hence, buyers must create self-enforcement by threatening to cancel future trade with the seller if low quality is observed. However, for the threat of termination to carry weight, the buyer must pay the seller an above market-clearing "efficiency-wage" so that the seller has something to lose from termination.

Some recent empirical papers have either directly tested or studied some implication of the repeat purchase mechanism (henceforth RPM), including, but not limited to, the articles by Cabral and Hortacsu (2006), List (2006), Brown, Falk and Fehr (2004), and Jin and Leslie (2003). One general finding is that, with repeat purchase opportunities and sufficient information about quality, reputation effects do seem to discipline sellers.[1] A second set of more specific findings from the Brown, Falk and Fehr experiments are remarkably consistent with the theoretical predictions of the RPM. Specifically, in the absence of third-party quality enforcement, reputation is important as buyers seek out specific sellers with whom they had success in the past and then form long-term relationships that involve efficiency wages and generous rent sharing. The market then resembles a collection of bilateral trading islands that are relatively insulated

---

[1] While Cabral and Hortacsu (2006) find that reputation effects matter, they point out that it is difficult to link this finding to a specific theoretical model such as the repeat purchase mechanism.

from competitive pressures. This is in stark contrast to the case where quality is third-party enforceable, in which case trade is no longer characterized by reputation based trading, rent sharing, or efficiency wages.

While the RPM has received considerable attention in the literature, MacLeod (2006) points out that the RPM may not be the most efficient trading mechanism even in the absence of third-party enforcement of contracts. In general, relational contracts involve explicit terms, based on a formal contract, that are third-party enforceable along with implicit components, based on mutually self-enforcing expectations, that can be adjusted at the parties' discretion. Thus, the discretion to make *ex post* adjustments to contract terms to punish and reward trading partners lies at the heart of relational contracting. Under the RPM, the formal term is the price, but the seller has discretion on quality choice and can thus behave opportunistically. The only discretionary instrument available to buyers for disciplining sellers is termination of the relationship. Thus, a mutually productive self-enforcing relationship can only emerge under the expectation that buyers will pay rents to sellers so long as the seller delivers high quality, but will terminate the relationship otherwise. If indeed, the RPM is a sub-efficient trading mechanism, then it may not be widely used in actual markets, making it important to investigate other forms of relational contracts.[2]

The purpose of the present study is to use economic experiments to compare and contrast relational trading under the RPM against more general forms of relational contracts that allow for greater discretionary latitude. This paper is, to the best of our knowledge, the first to experimentally study relational trading when parties have a broader range of discretionary powers that go beyond those that are available under the RPM. In practice, termination is rarely the only discretionary instrument available for disciplining sellers. For example, Banerjee and Duflo (2000) show that sellers in the Indian software industry can take corrective remedial actions after low quality is observed in order to preserve their reputation with buyers. An even simpler form of a

---

[2] Even though papers by List (2006), Jin and Leslie (2003), and Resnick and Zeckhauser (2002) study reputational trading using market data, it is not clear that the observed reputation mechanisms can be rationalized by the RPM. As Cabral and Hortacsu (2006) point out, reputation mechanisms can be quite complex and difficult to link with any specific theory.

remedial action is for buyers to demand or for sellers to provide discretionary deduct in prices, which is an assumed part of many relational contracting models (e.g. Levin 2003), and allowable under the Uniform Commercial Code. At a more general level, specific organizational forms or institutions can also affect the degree of discretionary latitude available to trading partners. For instance, Baker, Gibbons, and Murphy (2002) suggest that make-or-buy decisions in the theory of the firm can affect the amount of discretion available to parties, which alters the set of feasible relational contracts. On a related theme, supply contracts tend to provide different discretionary powers relative to labor contracts which can affect whether downstream firms would prefer to integrate with an upstream firm (making the upstream entity an "employee") or remain separate entities (making the upstream firm a "contractor"). In particular, discretionary *ex post* deducts are often acceptable to suppliers when the buyer receives a low quality shipment of goods. On the other hand, due to minimum wage rules or social norms, discretionary deducts may violate self-enforcement constraints in labor agreements so that discretionary bonuses rather than deducts tend to be the norm. Different legal regimes may also affect the feasibility of certain relational contracts. Williamson (1991) suggests that contract law is such that courts will tend to adjudicate on disputes over prices and other obligations between firms but courts will often refuse to hear disputes over similar matters when the parties in dispute are different divisions of the same firm. Thus, the formal enforcement regime within firms is weaker than between firms leading to different degrees of discretion in the underlying relational contracts.

To ensure methodological integrity, our basic experimental platform is based on the experimental marketplace designed by Brown, Falk and Fehr (2004) (henceforth BFF), which is a proven design for contrasting trading outcomes under the RPM to trading outcomes under a "complete contract" involving exogenous third-party enforcement of quality. To this platform, we add two additional treatments that are unique to our study and allow us to compare and contrast the RPM to relational contracts that expand the range of discretionary latitude. Specifically, our two treatments provide buyers with discretionary power to make *ex post* price adjustments thereby increasing the

number of instruments that buyers have to discipline sellers. The ability to make discretionary price adjustments is a crucial element of relational contract theory and is consistent with well known models proposed by MacLeod and Malcomson (1989) and Levin (2003). Thus, our experiments involve four total treatments. In one treatment (ICF-2), we provide buyers with the *ex post* discretionary to offer any price she chooses regardless of what the *ex ante* written offer was. This permits discretionary bonuses and deducts. Moreover, because neither price nor quality is third-party enforced, the underlying written contract is extremely incomplete. In another treatment (ICF-B), buyers have the latitude to offer discretionary bonuses only. This might mimic an employee contract where the price (or wage) is enforceable but the discretionary bonus is not. Thus, the underlying written contract is still incomplete, but less so than under ICF-2. A third treatment (ICF-1) that allows only for discretionary termination is identical to the ICF treatment used in the BFF experiments and serves as the RPM control. Our final treatment is based on BFF's "C" condition, which involves perfect third-party enforcement of the written contract and serves as the "complete contract" experiment. Note that the regimes can be ranked in terms of the degree of contractual incompleteness as follows: ICF-2>ICF-B>ICF-1>C. Thus, ICF-2 represents the most "incomplete" of the incomplete contracts, and ICF-1 (i.e. the RPM) is the most "complete" of the incomplete contracts (price is fully enforceable). In all treatments, we maintain the assumption that discretionary termination is a feasible option.

Our data is based on 1,313 trades from 18 experimental sessions involving 216 subjects. Even though our experiments took place in a different country and had minor design changes relative to BFF's, we were able to replicate many of BFF's results under ICF-1 which suggests that efficiency wages, generous rent sharing, and reputation based trading (i.e. identify of trading parties matter) are fairly robust outcomes of the RPM. In addition, our experiments yielded the following new insights:

1) Efficiency was highest with complete contracting, as expected. However, ICF-2 was the most efficient of the incomplete contracting regimes, while the RPM was

the least efficient. Thus, when comparing just incomplete contracting regimes, higher degrees of incompleteness yielded higher efficiency.

2) Bilateral trading islands are important only under ICF-1 and ICF-B. That is, reputation based trading where buyers prefer to trade with specific-sellers may not be a general feature of incomplete contracts with repeat trading opportunities. Under ICF-2, the regime with an extreme degree of incompleteness, the identity of trading partners appear to matter only slightly more than under the C regime.

3) Generous rent-sharing is not a general feature of incomplete contracts. While generous rent-sharing exists under the RPM, the degree of rent-sharing observed under ICF-2 is similar to what is observed under C.

4) When buyers had discretionary latitude, they used it. In ICF-2 experiments, 58% of all trades involved buyers using discretionary deducts. Discretionary bonuses were offered in only 12% of the trades, which suggests that buyers preferred sticks to carrots. In treatment ICF-B, where deducts were ruled out, bonuses were observed in 27% of trades. This suggests that even though buyers preferred to use deducts, they will increase the use of bonuses when it is their only option.

In addition to the results above, we also tested a number of additional predictions from a theoretical model we discuss in Section 2. The majority of these theoretical predictions were surprisingly consistent with our experimental outcomes.

The main implication of our results is that many of the relational trading outcomes that emerge under the RPM may be considerably weakened under other forms of relational contracts. In fact, in ICF-2 where maximum discretion is allowable (i.e. contract is extremely incomplete), the trading patterns that emerge are closer to the patterns observed under perfect third-party enforcement (i.e. complete contract) than under the RPM (e.g. a contract with a moderate degree of incompleteness). This novel finding suggests that efficiency wages, generous rent sharing, and reputation based trading where agents' identities matters may not be as strong as previously thought even in the absence of third-party enforcement of quality. On a related theme, our findings can also contribute to an understanding of degree to which bonuses and deducts are part of a

relational contract and how seller performance responds to these payments, which is an important topic of investigation according to MacLeod (2006).

Our results also provide some support for the theory of *strategic ambiguity* discussed by Bernheim and Whinston (1998). That is, if there are barriers to complete contracting, perhaps due incomplete law or missing institutions for third-party enforcement, then it may be optimal for parties to increase the degree of incompleteness. Thus, endogenous incompleteness may be a rational response to barriers to complete contracting. Finally, our results support Acemoglu and Johnson's (2005) claim that weak contracting institutions appear to have only second order impacts on long-run economic growth. Indeed many of our experimental subjects were able to execute trades that approached full efficiency even in regime ICF-2, which mimics an environment completely devoid of contracting institutions. Thus, our results may provide some insights into the importance of legal institutions for enforcing contracts.

2. THEORY AND PREDICTIONS

The degree of discretionary latitude under each relational contract depends on the degree to which the underlying formal contract is third-party enforceable. When a formal contract can be perfectly enforced by a third-party, then it is a complete contract and the contracting parties have no discretionary latitude to deviate from the contract. Because few real world contracts are perfectly enforced by a third-party, incomplete contracts tend to be the norm and at least one party to a contract typically has some discretionary latitude (e.g. Bernheim and Whinston 1998; Scott 2003). Nonetheless, even in the realm of incomplete contracts, contracts will typically range from highly "complete" incomplete contracts (i.e. those that specify traders' obligations to the maximal extent possible given limits to third-party enforceability), to those that are fully incomplete (i.e. contracts are good on paper only – nothing is legally binding). In a world of incomplete contracts, parties have to find ways to self-enforce certain obligations, either through the design of option contracts using mechanism design principles (e.g. Noldeke and Schmidt 1995) or by engaging in relational contracting. We focus on relational contracting and

assume a trading environment that rules out option contracts (more on this later). Even when parties rely on relational contracting, the degree of incompleteness of the formal contract can have profound effects on the set of feasible equilibrium paths available to parties engaged in a relational contract (Bernheim and Whinston 1998; Dixit 2003).   We now analyze how the structure of the underlying formal contract (e.g. degree of contractual incompleteness) affects the nature of the relational contract.

Suppose that a buyer and seller can potentially trade one unit of a good with a quality index $q \in [\underline{q}, \overline{q}]$, where q is observable and may or may not be enforceable by a third-party depending on the contract enforcement regime. If trade occurs at some price, p, the payoffs to the buyer and seller are $\pi = R(q) - p$ and $U = p - c(q)$, respectively, where the revenue function, $R(q)$, obeys $R(\underline{q}) = 0$, $R'(q) > 0$ and $R''(q) \leq 0$. The cost of producing a good of quality q is given by the function $c(q)$, where $c(\underline{q}) = 0$, $c'(q) > 0$ and $c''(q) \geq 0$. Hence, the buyer and seller's profits from exchange are functions of q. If no trade occurs, then the buyer earns $\overline{\pi}$ and the seller earns a reservation payoff of $\overline{u}$. Social surplus is then given by $S = R(q) - c(q) - \overline{u} - \overline{\pi}$. Assume that $S \geq 0$ and $R'(q) \geq c'(q)$, $\forall q \in [\underline{q}, \overline{q}]$, so that trade at $q = \overline{q}$ results in social efficiency.

The timing of a one-shot trading (stage) game is as follows. At time 0, the buyer may make a take-it-or-leave-it contract to the seller. The contract specifies a contract price, *P*, and a contract quality, *Q*. At time 1, the seller decides whether to accept or reject the contract. If the contract is rejected, the game ends and each party earns reservation payoffs $\overline{\pi}$ and $\overline{u}$. If the contract is accepted, the parties move to time 2 where the seller chooses actual *q,* which may or may not equal *Q*, depending on whether quality is third-party enforceable. At time 3, after *q* is observed, the buyer chooses actual price, *p* which may differ from *P* if price is not third-party enforceable. We assume that the parties cannot renegotiate the trading decision after performance is observed so that option contracts in the spirit of Noldeke and Schmit (1995) can be ruled out. Thus,

remedies to noncontractibility must depend on repeated interactions rather than on renegotiation and options.[3]

   To be consistent with our experimental design, we now consider four contract enforcement regimes where the amount of discretionary latitude available to parties (i.e. the degree of contractual incompleteness) varies across the regimes. We assume that the enforcement regime is exogenous to the trading parties who cannot influence the legal structure of the economy in which they operate. Our analysis also focuses on what *would* be the trading outcomes *if* the agents configure written contracts that are enforceable to the maximum extent allowable under the enforcement regime. That is, we focus on *conditionally complete* contracts where contracts are enforceable to the fullest extent possible given limits to enforcement, although the contracts are still incomplete when compared to the contract with perfect third-party enforcement. Agents, of course, can always endogenously select less complete contracts in any given regime. Our analysis would then provide insights into when agents may choose to write conditionally complete contracts versus conditionally incomplete contracts. When there are no barriers to third-party enforcement, then the qualifier "conditionally" can be eliminated and parties can write "complete contracts." The four regimes we examine are:

1. Complete Contract (C): In this regime, we assume that agents can structure a contract that is perfectly third-party enforceable so that the parties have no *ex post* discretionary latitude. Here, both P and Q are exogenously enforced so that the only possible *ex post* outcome is $p=P$ and $q=Q$. This regime corresponds to the "C" condition of BFF. This is the only regime in which agents can write complete contracts.

2. Incomplete Contract 1 (ICF-1): This regime is identical to regime C with the exception that quality is not third-party enforceable. Thus, *ex post* the seller has discretionary latitude to choose any quality it wants to so that $q \neq Q$ is possible. This regime is identical to BFF's ICF condition. This regime also corresponds to the RPM, which would be the "conditionally complete" contract in this regime. That is, with a

---

[3] This assumption enables us to focus on relational contracting environments such as those studied by Levin (2003).

barrier to third-party enforcement of *q,* the most complete contract in this regime would entail full enforcement of *P*.

3. Incomplete Contract B (ICF-B): This regime is identical to ICF-1 with the exception that the buyer can adjust prices upward *ex post.* In other words, the buyer has the discretionary latitude to offer an unenforceable bonus to the seller. Employee contracts and informal contracts between restaurant waiters/waitresses and customers are examples of contracts that include unenforceable discretionary bonuses or tips. The conditionally complete contract in this regime is "less complete" than in ICF-1 because *P* is not fully enforceable, as the buyer always has some latitude to offer some *p>P*, *ex post.*

4. Fully Incomplete Contract (ICF-2): In this regime, neither *P* nor *Q* are third party enforceable so that full discretionary latitude exists for both parties. Thus, *ex post,* the seller can choose any *q* he desires and the buyer can choose any *p* she desires without restriction. One can think of this treatment as one that is completely devoid of formal contracting institutions or good faith laws to protect the contracting parties. Here, buyers have no legal recourse for enforcing quality and sellers cannot collect payment from buyers, except through informal means. This treatment also allows for the use of discretionary bonus or discretionary deducts as the buyer can choose to either increase *p* above *P* or reduce *p* below *P*, *ex post.*

2.1. *Static Contracting*

In one-shot interactions, it is fairly straightforward to rank the various contracting regimes in terms of efficiency (assuming contracts are conditionally complete) as some regimes will result in higher quality, and hence, higher efficiency, than other regimes.

PROPOSITION 1: *In a one-shot interaction, the trading outcomes are*

i) $q = \bar{q}$ *and* $p = c(\bar{q}) + \bar{u}$ *in regime C,*

ii) $q = \underline{q}$ *and* $p = c(\underline{q}) + \bar{u}$ *in regimes ICF-1 and ICF-B, and*

iii) *no trade takes place in regime ICF-2.*

PROOF: See appendix.

Proposition 1 predicts that full efficiency occurs only under regime C, with full third-party enforcement of contracts. With only partial enforcement (regimes ICF-1 and

ICF-B), trade occurs but only at the minimal quality level. When third-party enforcement is completely absent (ICF-2), the market collapses and no trade occurs. This proposition is consistent with common intuition amongst economists and legal scholars that a greater degree of contractual completeness improves efficiency. Hence, improving third-party enforcement of contracts would enhance welfare.

2.2. *Repeat Trading*

In contrast to the one-shot setting where predicted outcomes are stark for ICF-1, ICF-B and ICF-2, higher efficiency and cooperation can be sustained via informal or relational incentives when repeat trading is possible. In repeated interactions, the promise of future rewards and punishments can discipline short-run opportunism and promote cooperation. However, we show that the strength of cooperation and the magnitude of future rewards and punishments are dependent on the contract regime. Each contract regime provides trading parties with different degrees of discretionary latitude and many of the predictions from the one-shot environment are reversed in the repeat trading environment. The key message of this section is that, when agents can rely on informal incentives, incremental improvements in third-party enforcement, which permits agents to write more "complete" contracts, no longer necessarily improves efficiency.

To show that the repeat trading environment can reverse the prediction of the one-shot environment, we need to show that the shadow of the future can provide sufficient incentives for traders to not to shirk on agreements. It is easiest to show this under the assumption that buyers and sellers interact an infinitely number of times, although it would not be difficult conceptually to show this in a finitely repeated setting under the assumption that there exists in the population "cooperative" types. Kreps, Milgrom, Roberts, and Wilson (1982), and BFF have both shown that if some fraction of traders are "cooperative" or not purely self-interested and that this is common knowledge, then positive rents can exist even in the last stage of a finitely repeated game. The existence of such rents creates a mechanism for sustaining cooperation in earlier periods even among selfish traders. We focus on the infinitely repeated game because it is

mathematically less cumbersome and should yield similar qualitative predictions as a finitely repeated game with some cooperative types. We begin our modeling by stating the following assumptions.

A.1. INFORMATION: *The buyer and seller only know their own past actions or the past actions of parties with whom they have interacted. The buyer and seller do not know the past actions of parties with whom they have not traded with.*

A.2. STRATEGIES: *In each period, t, a buyer offers a contract (P,Q) to a seller, followed by the seller's reject/accept decision. Upon acceptance, the seller then chooses q, which may differ from Q. After observing q, the buyer may or may not choose $p \neq P$. Finally, the buyer decides whether to renew the contract the next period. Continuation of a relationship occurs if and only if a buyer offers a contract and the seller accepts the contract. These decisions depend on the history of play through period t-1.*

A.3. PAYOFFS: *Each party's payoff is given by the discounted sum of her stage-game payoff. The common discount factor is $\delta \in (0,1]$.*

A.4. EQUILIBRIUM: *Like most repeated games, there are multiple equilibria. The equilibrium outcomes we focus on are subgame perfect under the assumption that buyers and sellers cooperate if cooperation results in higher present value of payoffs than non-cooperation. Cooperation occurs in each stage t if the history of play through t-1 has been cooperation and the parties break off trade in response to any deviation. We assume that no renegotiation of either the explicit contract or the implicit agreement occurs after a deviation. The following deviations may be observed in each regime:*

i) Regime C – if the history of trade has been high price and high quality (P*,Q*), then, the buyer deviates in period t by offering some $P \neq P^*$.

ii) ICF-1 – same as regime C with the addition that, after acceptance, the seller can further deviate by choosing $q < Q^*$.

iii) ICF-B – same as regime ICF-1 with the addition that if the seller supplies $q \geq Q^*$ the buyer can further deviate by choosing $p = P^*$ (buyer withholds bonus payment for good performance).

iv) ICF-2 – same as regime ICF-B with the addition that if the seller chooses
$q \geq Q^*$, *the buyer can further deviate by choosing* $p \leq P^*$ *(buyer deducts the*
price or fails to pay bonus even when performance is good).[4]

Note that by using the term "may be observed" in describing the deviations, we introduce
some vagueness into A.4. However, this vagueness is unavoidable, especially in an
experimental setting, as implicit components of relational contracts depend heavily on the
expectations held by the parties as a relationship evolves. For example, in regime ICF-B,
rather than pay a bonus, the buyer may set a high *P** relative to the market so that the
seller may continue to cooperate even when no bonus is paid; i.e. there is no expectation
of a bonus. Similarly, in ICF-2, it is unclear whether buyers and sellers expect deducts or
bonuses (or both) to be used a priori. In our experiments, we observed a range of
behaviors as some paid bonuses, others used deducts and some used neither.

The only possible deviation under C is in the offer and acceptance/rejection
decisions that occur in between the stage-games. Thus, so long as the buyer makes an
offer (*P*,Q**) in each period that covers a seller's reservation payoff, the seller will
accept the contract. If the buyer deviates by offering some (P', Q^*) where $P' \neq P^*$, then
the seller may accept the contract and exert $q = \underline{q}$ or reject the contract depending on
whether P' covers the seller's reservation utility or not. Given full enforcement, $(P^* Q^*)$
should be identical to the one-shot contract in each period. Repeated trading should not
matter as the shadow of the future is not needed to enforce the contract.

Another important point to note is that, while our assumption that no
renegotiation is possible is less than theoretically satisfying, the purpose of our model is
to provide a heuristic framework for generating testable hypotheses for our experiments.[5]
Thus, we make the assumption for modeling convenience as it allows us to avoid
complex issues associated with the renegotiation of infinitely repeated games. Even if we

---

[4] With regard to case (iv), it should be noted that it is also possible that, if $q<Q^*$, the buyer and seller can
still "settle-up" by the buyer paying some price $p < P^*$; i.e. the buyer gets a rebate for low quality. If this
is the case, the buyer and seller may let bygones be bygones and continue to cooperate in the future.
[5] See Bernheim and Whinston (1998) for an exhaustive theoretical treatment of renegotiation in repeated
trading environments with incomplete contracts.

relax this assumption, what is important in our context is whether experimental subjects believe that punishments for deviation (i.e. breaking off of trade) will be enforced or whether parties will renegotiate. If subjects believe that renegotiation is likely to occur, then incentives for cooperation will be undermined. We believe that there are two reasons why our subjects would find punishments for deviation credible. First, our experimental design does not permit explicit contracts to be renegotiated at any time in a stage-game. Second, even if subjects can renegotiate agreements across periods (across stage-games), the fact that there are more sellers and buyers (more details later) in our experimental marketplace means that there will always be some unemployed sellers in each period. Thus, when a deviation occurs, buyers know that there will always be another seller waiting in line to receive a contract which weakens buyers' incentives to renegotiate with any one seller. Moreover, because sellers know this, they also know that there is a nontrivial probability that they will be unemployed in the future if they deviate.

Under ICF-1 (RPM), if the seller accepts a contract (P, Q), then she can either cooperate by choosing $q \geq Q$ or shirk by choosing $q < Q$ in each period. We assume that any shirking by the seller will trigger non-cooperation in the future by the buyer so we consider only the seller's most profitable deviation from the contract, which is $q = \underline{q}$ and yields a stage game payoff of $P - c(\underline{q})$. Then in the future, the seller earns, $\overline{u}$ so that if it is optimal for the seller to shirk, then the present value of the seller's payoffs is:

(1) $V_{ICF-1}^{s} = P - c(\underline{q}) + \dfrac{\delta \overline{u}}{1 - \delta}$

On the other hand, if it is optimal for the seller to cooperate, then the present value of the seller's payoff is,

(2) $V_{ICF-1}^{c} = \dfrac{P - c(q)}{1 - \delta}$ where $q \geq Q$

Equations (1) and (2) imply that it is optimal for the seller to honor the agreement if and only if,

(3) $\dfrac{P - c(q)}{1 - \delta} \geq P - c(\underline{q}) + \dfrac{\delta \overline{u}}{1 - \delta}$

Expression (3) can be restated as,

(4) $P \geq c(\underline{q}) + \bar{u} + \dfrac{c(q) - c(\underline{q})}{\delta}$

which gives the lower bound on P for inducing seller cooperation. Therefore, at the beginning of the stage game, if the buyer wants to induce the seller supply some $Q > \underline{q}$, it must offer at least the one-shot price plus a premium of $\dfrac{c(Q) - c(\underline{q})}{\delta}$. A profit-maximizing buyer would offer $P = c(\underline{q}) + \bar{u} + \dfrac{c(Q) - c(\underline{q})}{\delta}$. This yields stage-game profits of $U_{ICF-1} = P - c(q) = \bar{u} + \dfrac{(1 - \delta)\left[c(Q) - c(\underline{q})\right]}{\delta}$ for the seller, which exceeds the seller's outside payoff when $\delta < 1$. We must also ensure that the buyer is willing to participate, which requires that $\pi_{ICF-1} = R(Q) - P \geq \bar{\pi}$ or

(5) $\pi_{ICF-1} = R(Q) - c(\underline{q}) - \bar{u} - \left[\dfrac{c(Q) - c(\underline{q})}{\delta}\right] \geq \bar{\pi}$.

Turning now to ICF-B, note that the key difference between this regime and ICF-1 is that after the buyer observes the seller's choice of q, the buyer has the discretion to choose some $p(q) \geq P$; that is, the buyer can pay the seller a discretionary bonus. We can denote this bonus by $b(q) = p(q) - P > 0$. The predictions under this regime are highly sensitive to the expectations of the parties. For example, the buyer may offer a low P and have an informal understanding with the seller that a bonus will be paid whenever $q \geq Q$, which raises the sellers pay *ex post*. A failure by the buyer to pay the bonus constitutes a deviation and a signal of future non-cooperation which can cause the relationship to unravel. In this case, a relational agreement enforces the payment of the bonus. However, it is also plausible that the buyer offers a high initial P that is attractive enough to where a seller continues to cooperate even if no bonus is paid.

To analyze ICF-B more formally, we suppose that the expectation is that a bonus is paid for good performance and withheld for poor performance. Consider the last mover of the stage-game where the buyer has observed the choice of q by the seller. If

the seller has chosen $q \geq Q$ (the case of $q < Q$ is trivial), the buyer can pay a "bonus" by choosing $p(q) > P$ or not pay a bonus by choosing $p(q) = P$ thereby triggering future non-cooperation by the seller. If it is optimal for the buyer to shirk on the bonus, then the present value of the buyer's payoffs is:

(6) $\quad \Pi^{s}_{ICF-B} = R(q) - P + \dfrac{\delta \tilde{\pi}}{1 - \delta}$

On the other hand, if it is optimal for the buyer to pay the bonus, then we have,

(7) $\quad \Pi^{c}_{ICF-B} = \dfrac{R(q) - P - b(q)}{1 - \delta}$

Equations (6) and (7) imply that it is optimal for the buyer to pay the bonus if and only if,

(8) $\quad \dfrac{R(q) - P - b(q)}{1 - \delta} \geq R(q) - P + \dfrac{\delta \tilde{\pi}}{1 - \delta}$

Backward induction within the stage game brings us to the choice of q by the seller. The seller's problem is analogous to his problem under ICF-1 with the exception that now shirking to increase his stage game payoff will cost him both the bonus $b(q)$ and future cooperation with the buyer. Hence, it is optimal for the seller to honor the agreement if and only if,

(9) $\quad \dfrac{P + b(q) - c(q)}{1 - \delta} \geq P - c(\underline{q}) + \dfrac{\delta \tilde{u}}{1 - \delta}$

This inequality (9) can be expressed as,

(10) $\quad P \geq c(\underline{q}) + \bar{u} + \dfrac{c(q) - c(\underline{q}) - b(q)}{\delta}$

Comparing (10) to (4), one can see that one of the benefits of a bonus is that it allows the buyer to lower P. Thus, at the contract formation stage, the buyer must offer (P, Q) such that both parties are willing to participate. That is, the contract must satisfy,

(11) $\quad U_{ICF-B} = P + b(q) - c(Q) \geq \bar{u}$

(12) $\quad \pi_{ICF-B} = R(Q) - P - b(q) \geq \bar{\pi}$

We now look at ICF-2 where neither P nor Q are enforced by a third-party. Within each trading stage, after the buyer observes the seller's choice of q, the buyer can

choose any $p(q) \geq 0$ she wants including not paying for the good ($p(q) = 0$). The buyer can also condition price on quality so as to create discretionary bonus and deducts to motivate the seller. Then we would have a relational contract similar to those studied by Levin (2003). Whether opportunism occurs (buyer refuses payment) or relational contracting emerges depends on the evolution of expectations and norms of trading during the course of repeated transactions. To conduct the formal analysis, we assume that the expectations are such that if the seller supplies $q \geq Q$, and the buyer chooses $p \geq P$, then cooperation is achieved and the parties are willing to continue to cooperate in the future. If $q < Q$, then the buyer pays $p < P$ and the parties no longer cooperate in the future. It is also possible that, by deducting the price in response to low quality, the parties have "settled-up" and can cooperate in the future. But we saw little evidence of this in our experiments. Non-cooperation can also be triggered if $q \geq Q$ and $p < P$, which is a situation where the seller has honored the agreement but the buyer does not.

The sequence of steps in a stage-game of ICF-2 is similar to the steps in ICF-B with the exception that, in the final step, after the buyer observes q and chooses $p(q)$, there is no restriction on the range of $p(q)$, except it cannot be negative. We define a "bonus" as $w(q) = p(q) - P > 0$ and a "deduct" as $w(\underline{q}) = p(q) - P < 0$. But if a deduct triggers future non-cooperation, then the buyer may as well choose the most profitable deduct which would be to set $p = 0$ which implies that $w(\underline{q}) = -P$. Thus, if the seller has chosen $q \geq Q$, the buyer can either honor the agreement by choosing $w(q)$ or behave opportunistically and choose $w(\underline{q})$, which also triggers future non-cooperation. It is optimal for the buyer to honor the agreement if and only if,

(13) $\dfrac{R(q) - P - w(q)}{1 - \delta} \geq R(q) - P - w(\underline{q}) + \dfrac{\delta \overline{\pi}}{1 - \delta}$

Backward induction within the stage game brings us to the seller's choice of q. It is optimal for the seller to honor the agreement if and only if,

(14) $\dfrac{P + w(q) - c(q)}{1 - \delta} \geq P + w(\underline{q}) - c(\underline{q}) + \dfrac{\delta \overline{u}}{1 - \delta}$

Inequality (14) can be expressed as,

$$(15) \quad P \geq c(\underline{q}) + \bar{u} + \frac{c(q) - c(\underline{q}) - [w(q) - w(\underline{q})]}{\delta} - w(\underline{q})$$

By letting $w(\underline{q}) = -P$ in (15), it is straightforward to show that P is indeterminate in this case. What is important is that,

$$(16) \quad U_{ICF-2} = P + w(q) - c(q) \geq \bar{u}$$

which implies that the seller's expected total pay under cooperation must be $P + w(q) = p \geq \bar{u} + c(q)$ or that *ex post* price chosen by the buyer cannot fall below $p \geq \bar{u} + c(q)$ in a cooperative relationship. Thus, at the contract formation stage, the contract (P,Q) the buyer offers must satisfy,

$$(17) \quad \pi_{ICF-2} = R(Q) - P - w(Q) \geq \bar{\pi}$$

Equations (1)-(17) are useful for deriving the major results that arise under repeated trading and are summarized in the following proposition.

> PROPOSITION 2: *When the buyer and seller can trade repeatedly, outcomes under regime C are identical to the one-shot case. For the incomplete contract regimes, equilibrium outcomes are:*
>
> i) *contracted quality is such that* $Q_{ICF-B} = Q_{ICF-2} = \bar{q} \geq Q_{ICF-1} > \underline{q}$,
>
> ii) *sellers cooperate by producing* $q \geq Q$ *and shirking of* q *by sellers is most likely to occur under ICF-1, followed by ICF-B and ICF-2,*
>
> iii) *contracted prices under ICF-B and ICF-1 are*
>
> $$P_{ICF-B} = c(\underline{q}) + \bar{u} < c(\underline{q}) + \bar{u} + \frac{c(Q_{ICF-1}) - c(\underline{q})}{\delta} = P_{ICF-1}$$
>
> *and* P_{ICF-2} *is indeterminant,*
>
> iv) *buyers' offer contingent pay* $b(q) = c(Q_{ICF-B}) - c(\underline{q})$ *in ICF-B and*
>
> $$w(q) - w(\underline{q}) = c(Q_{ICF-2}) - c(\underline{q}) \text{ in ICF-2.}$$

Proposition 2 offers some testable predictions that are useful for guiding the experimental data analysis. Part (i) predicts that buyers' should request the efficient level of quality in

ICF-B and ICF-2, which should be weakly greater than quality requests under ICF-1. The weak inequality comes from the fact that buyers *may* request the efficient quality level if the discount factor is sufficiently high (see proof of Proposition 2). Intuitively, the cost of implementation (total payments made to sellers) to induce each quality level is higher under ICF-1 than under the other regimes, although the difference in cost depends on δ. This is because buyers must ensure that sellers earn rents under ICF-1 in order to induce them not to shirk. Part (ii) predicts that shirking by sellers should be observed most frequently under ICF-1 and least frequently under ICF-2. This is due to the fact that the seller's dynamic incentive constraints – equations (3), (9), and (14) – are most relaxed under ICF-2 and least relaxed under ICF-1. Part (iii) predicts that average contracted price should be higher under ICF-1 than under ICF-B. No predictions emerge for ICF-2 regarding contracted price. Part (iv) predicts that the range of contingent pay – that is, the size of the bonus or the spread between the bonus and deduct – should be equal to the cost difference between producing Q and producing minimal quality.[6]

Repeat trading also has some implications for the distribution of profits. Corollary 1 summarizes the payoffs to sellers under each regime.

COROLLARY 1: *In equilibrium, the seller earns rents under ICF-1. Under C, ICF-B and ICF-2, the seller earns profits equal to his reservation payoff.*

Intuitively, the repeat trading mechanism (ICF-1) relies on efficiency wages and the threat of termination to motivate sellers to deliver high quality. Because the other regimes include additional incentive instruments such as discretionary bonus, deducts, or formal enforcement, efficiency wages are less important for incentivizing sellers.

Another issue has to do with the propensity for buyers and sellers to cooperate under each regime. A cooperative outcome where both the buyer and seller honors the contract is a sub-game perfect Nash equilibrium if δ is sufficiently high.

PROPOSITION 3: *Let $\underline{\delta} > 0$ such that $\forall \delta \in [\underline{\delta}, 1)$, cooperation is achievable. Then $\underline{\delta}_{ICF-1} = \underline{\delta}_{ICF-B} < \underline{\delta}_{ICF-2}$.*

---

[6] Recall that producing minimum quality is the seller's most profitable deviation.

Proposition 3 predicts that, in the incomplete contract regimes, the range of discount factors that can support a cooperative equilibrium is the same under ICF-1 and ICF-B, and larger than under ICF-2. Thus, one might expect to observe more cooperative outcomes under ICF-1 and ICF-B than under ICF-2. The intuition for this result is that, given that buyers can choose to pay any price, including prices that fall below the contracted price P, they can behave opportunistically by withholding payment and reaping short-term gains. The pressure to behave opportunistically in this regime means that higher discount factors are required to prevent buyer's from shirking on price.

A problem that may be of interest is how buyers may structure a relational contract with sellers in the ICF-B regime. Buyers can either choose a low P and offer a large $b(q)$ when $q \geq Q$ or offer a high P and a small bonus. It would be tempting to conclude that the two are equivalent when parties are risk neutral and sellers hold the correct expectations concerning the buyer's willingness to pay the bonus. We show, however, that increasing P can expand (weakly) the set of δ that would support a cooperative sub-game perfect Nash equilibrium and would never decrease the set of δ that supports cooperation.

COROLLARY 2: *Under ICF-B, $\underline{\delta}$ is weakly decreasing P.*

That is, buyers can induce greater cooperation by choosing higher P which suggests that it is preferably not to use pure bonus contracts with $P = 0$. However, raising P too high means that the buyer would overpay. The optimal P is specified in Proposition 1 (iii).

3. EXPERIMENTAL DESIGN

Our basic experimental platform is based on the design of BFF; indeed, our C and ICF-1 regimes are nearly identical to theirs. Our regimes ICF-B and ICF-2 are unique to our study and allow us to examine how exogenously altering the contract enforcement regime would affect efficiency, distribution and the nature of trading if agents wrote conditionally complete contracts.

We ran a total of eighteen experiments where six of the experiments were regime C, four were ICF-1, three ICF-B, and five were ICF-2. For each experiment, twelve

student subjects were recruited at a major university in the Midwest. The twelve subjects are partitioned into a group of five buyers and another group of seven sellers. The absolute number of buyers and sellers were smaller than BFF's experiments (seven buyers and ten sellers) due to limitations in laboratory size, but the ratio of buyer to sellers is nearly the same. A total of 216 subjects participated in the experiments. Each experiment has 17 trading rounds – two practice rounds and 15 'live' rounds that may determine eventual cash payment. Given five buyers, the total number of possible trades per-round is five. This translates into seventy-five possible trades per experiment.

An important point to note is that, while our theoretical model is based on an infinitely repeated game, our experiment is a finitely repeated game. In theory, when the ending round is common knowledge and if it is common knowledge that all subjects are strictly self interested, then it is straightforward to show, using backward induction, that cooperation should not occur in any round. In this case, the one-shot predictions outlined in Proposition 1 should hold in all fifteen rounds. Nevertheless, a number of past studies have shown that cooperation still occurs in the early to middle rounds of finitely repeated games and only begin to breakdown near the end of an experiment (e.g. Axelrod 1981; Andreoni and Miller 1993; Cooper 1996, among others). Moreover, BFF's ICF experiments, which were nearly identical to our ICF-1, show that cooperation does indeed occur and only begins to decline in rounds close to the end. BFF suggest that the presence of "fair" types makes it possible for cooperation to be achieved because fair types will honor contracts even in the final round. Thus, rents may exist even in the final round and the possibility of capturing these rents serves to discipline even selfish workers in early rounds. Along similar lines, Kreps, Milgrom, Roberts and Wilson's (1982) famous model shows that if it is common knowledge that some people are cooperative, then it is possible to support a cooperative perfect Bayesian equilibrium in finitely repeated games. The upshot is that the predictions of section 2.2. are driven by the possibility that agents' actions can affect future payoffs under repeat trading. As long as mechanisms, such as cooperative types, exist for generating rents even in the final round of a finitely repeated game, then all qualitative results of section 2.2. should still hold.

An infinitely repeated game then becomes a useful parable for thinking about how people strategically trade-off short term gains for long term payoffs (Rubinstein 1991). Moreover, Dal Bó (2005) used experiments to compare the degree of cooperation that occurs between finitely and infinitely repeated games. His results suggest the level of cooperation is lower under finitely repeated games. Hence, if our results from a finitely repeated game are consistent with the theory outlined in section 2.2, then our conclusions should only be strengthened in an infinitely repeated trading environment.

During the experiments, all trading takes place on networked computers enclosed in cubicles to eliminate between-subject visual contact. Anonymity is further preserved by assigning all subjects identification (ID) numbers. ID numbers are fixed across rounds allowing subjects to develop and track reputations. While there are fifteen rounds in each experiment, each individual round is sub-divided into various "phases". In phase 1 ("the trading phase"), buyers offer contracts specifying a price-quality combination for a unit of an abstract good; that is, buyers offer a contract consisting of the pair (P, Q). Sellers can only accept or reject offers. A buyer can make as many offers as desired in the trading phase, but once one offer is accepted, all other offers are withdrawn and no additional offers can be made. Similarly, once a seller accepts an offer, no other offers can be entertained. The trading phase lasted a minute and a half, which differed from BFF's in that BFF allowed three minutes.[7] In short, each buyer and seller can conclude at most one trade per round. No buyers (sellers) are obligated to make (accept) offers during the trading phase. Because the contract (P,Q) is perfectly enforced under C, once the trading phase ends, earnings for buyers and sellers are calculated, and the round ends. Each buyer knew what she and her seller made during the round, but did not know the earnings of other buyers and sellers in the market. However, under ICF-1, ICF-B and ICF-2, the round continues into phase 2 ("quality determination phase"). In this phase, if a seller has agreed to a contract, then s/he can choose actual quality, q, that differs from the quality level s/he agreed to in the contract; i.e. q may differ from Q. While sellers

---

[7] Pilot tests were conducted by allowing for 2.5 minutes (150 seconds). However, we observed that most of the offers were completed within a minute and a half so we shortened the trading phase to reduce the length of the experiment. The shortened time period did not seem to affect results as our qualitative results under C and ICF-1 are very similar to BFF's results.

were deciding on *q,* buyers were asked to specify what quality level s/he *expected* the seller to supply and how certain s/he was that these expectations would be fulfilled. Under ICF-1, after *q* is chosen, income is calculated and the round ends. However, under ICF-B and ICF-2, the round continues into phase 3 ("price determination phase"). In this phase, after buyers observe *q* chosen by sellers, buyers choose the actual price they pay. Under ICF-B, the restriction $p \geq P$ is imposed. Under ICF-2, there are no restrictions on *p* except that it is in the interval [0, 100].[8] While buyers were making their decision on *p,* sellers specify what price s/he *expects* the buyer to choose and how certain s/he is that the expectations will be fulfilled. Finally, income is calculated and the round ends.

During the trading phase, buyers can extend two types of offers: public and private. Public offers are displayed on the computer screens of all sellers and buyers; any seller can accept any public offer. Private offers are extended by entering a specific seller's ID number into the computer. Only the seller identified sees the offer and only s/he can choose to accept it. Private offers enable reputation based trading and long-term relationships, which lie at the core of relational contracting theory. For example, if a buyer predicts benefits from contracting with a specific seller, the buyer can make a single, private offer to that seller in each round rather than venturing into the open market and hoping that that seller is the first to accept the offer.[9] Moreover, renewing private offers across rounds permits long-term relationships to form.

Every round features the same five buyers and seven sellers. Fewer buyers than sellers creates buyer concentration because at least two sellers do not trade in each round. This forces sellers to compete for a limited number of contracts, which tilts bargaining power in favor of buyers.

In order to implement experiments, we parameterize our model as follows: $R(q) = 10q$, $\bar{\pi} = 0$, $\bar{u}$ equals 5 or 10, $\underline{q} = 1$ and $\bar{q} = 10$. Moreover, we assume that $c(q)$ is represented by the following cost schedule:

---

[8] *P* was also restricted to be in this interval.
[9] Firms often establish "private trades" by contacting specific suppliers with whom they have good relationships in order to avoid costly public solicitations when the desired supplier is already known.

Quality	1	2	3	4	5	6	7	8	9	10
Cost	0	1	2	4	6	8	10	12	15	18

Note that marginal cost never exceeds "3" and the buyer's marginal revenue is always "10". Thus, marginal revenue always exceeds marginal cost, as was assumed in the theoretical model, so it would be socially efficient for parties to trade at $q = \bar{q} = 10$.

Round specific payouts are determined for buyers as follows:

$$(18) \quad \pi = \begin{cases} 10q - p & \text{if an agreement is reached,} \\ 0 & \text{if no agreement is reached,} \end{cases}$$

All payments are given in experimental points where subjects earn one dollar for 70 points. The seller's profit is:

$$(19) \quad U = \begin{cases} p - c(q) & \text{if an agreement is reached,} \\ \bar{u} & \text{if no agreement is reached,} \end{cases}$$

where $\bar{u}$ is a reservation payoff in the absence of trade. This reservation payoff was equal to 10 in two of the six C experiments, in all three ICF-B experiments, and in two of the ICF-2 experiments. The reservation payoff was set to $\bar{u} = 5$ in the eleven remaining experiments. The variation in reservation payoffs should only induce buyers to change their price offers to ensure that sellers' reservation payoffs are covered, but efficiency should not be affected. All subjects were told that they would earn experimental "profits" based on the payoff functions (18) and (19). Experimental profits were then converted into dollars at the rate of $1 = 70$ profit points.

The experiment was programmed using "z-tree" software (Fischbacher 1999). Subjects were also asked to fill out short questionnaires, which took anywhere from five to twenty minutes to complete, to test subjects' understanding of experimental instructions and to obtain information about subject characteristics (e.g. demographics, social preferences, GPA, etc.).[10] Subjects were informed that actual earnings depend

---

[10] The questionnaires for social preferences were based on a small set of games similar to the Charness and Rabin (2002) games, which took about 20 minutes and were administered before the contracting game. Questions to test subjects' understanding of the contracting instructions were also administered prior to the

upon the rules of the game and the participant's and other participants' actions. In addition, subjects started each contracting experiment with $5 in their account balance. Average earnings were in the neighborhood of $23 per subject per experiment.[11] Each contracting experiment took between 40 to 60 minutes to complete.

4. RESULTS

For the six C experiments, subjects executed 436 out of 450 possible trades. Of the 436 trades executed, 94 (21.5%) were private trades. For the four ICF-1 experiments, subjects executed 290 out of 300 possible trades. Of the 290 trades, 164 (56.5%) were private trades. For the three ICF-B experiments, 213 out of 225 possible trades were executed with 115 of the 213 trades being private trades (54%). Finally, for the five ICF-2 experiments, 374 out of the possible 375 trades were executed with 105 of the 374 being private trades (28%). Recall that Proposition 1(iii) predicts that no trade should take place under ICF-2 in a one-shot setting, which should theoretically also hold for finitely repeated games. But the fact that 374 out of 375 possible trades occurred suggests unequivocally that subjects may not believe that the experimental marketplace is populated with only strictly selfish types who would shirk on agreements in the final round of the finitely repeated game. This suggests that subjects do cooperate and may trade-off short term and long term gains when making decisions.

4.1. *Efficiency*

Recall that under our model parameters, full efficiency is achieved if buyers and sellers trade at $q = \overline{q} = 10$. Thus, higher quality implies higher efficiency. The repeat trading model of section 2.2 makes several predictions concerning the level of quality offered

---

contracting games and took between 5 to 10 minutes to complete. We do not believe that administering these questionnairess before the contracting games had any significant impact on results for two reasons. First, the same questionnaires were administered before all experiments so differences in results across contracting regimes would not be driven by these questionnaires. Second, our results from our C and ICF-1 experiments were qualitatively very similar to BFF's C and ICF conditions. While there were minor changes in level effects, the qualitative results and evolution in the pattern of trade were remarkably similar. Thus, results appear to be robust to any design/questionnaires differences.

[11] The going rate for a two-hour experiment at the host campus was about $20.

and the degree to which shirking occurs under each regime. Proposition 2 (i) and (ii)

allows us to test the following hypotheses using our experimental data.

> H.1: *There should be no difference in the level of quality specified*
>
> *in the contracts offered by buyers under ICF-B, and ICF-2. Moreover, Q=10.*
>
> H.2: *The level of contracted quality under ICF-1 should be no*
>
> *greater than the level of quality requested under ICF-B or ICF-2.*
>
> H.3: *Seller deviation from contracted quality should be greatest*
>
> *under ICF-1, followed by ICF-B and then ICF-2.*

These three hypotheses lead to a corollary hypothesis.

> H.4: *Actual quality chosen should be greatest under C, followed by*
>
> *ICF-2, ICF-B and then ICF-1.*

TABLE 1
Quality and Quality Deviation Summary Statistics across Contracting Regimes

	Obs.	Avg. Contracted Quality Q	Avg. Actual Quality q	% of trades where q<Q	Avg. size of shortfall, Q-q
Regime C	436	8.95	8.95	0	0
Regime ICF-1	290	8.06	5.82	61%	2.31
Regime ICF-B	213	8.22	6.51	56%	2.03
Regime ICF-2	374	8.22	7.12	45%	1.36

Table 1 reports summary statistics for quality. The average Q requested under

contract is highest under C and lowest under ICF-1. With regard to hypothesis H.1, it

appears that the average Q requested under ICF-B and ICF-2 are nearly identical, which

is consistent with the hypothesis. This is confirmed using a non-parametric Kruskal-

Wallis (KW) test ($p=0.18$) so that H.1 cannot be rejected by the data. H.1. also suggests

that buyers should request the efficient level of quality under ICF-B and ICF-2; i.e.

$Q_{ICF-B} = Q_{ICF-2} = 10$. To test this hypothesis, we used a Wilcoxon test to determine

whether the pooled data from regimes ICF-B and ICF-2 are significantly different from

10, which yielded a p-value=0.00.[12] One concern that may arise is that buyers did not ask sellers to supply the efficient level of quality because a finitely repeated game does not induce enough cooperation to make it feasible to support this high level of quality. However, there are two reasons to doubt that a finitely repeated game was the reason why efficient quality was not requested. First, we conducted the same test using data from regime C, where cooperation should not matter and also found that buyers requested $Q<10$ ($p=0.00$). Second, an average of $Q=8.22$ under ICF-B and ICF-2 is still relatively high quality request and far exceeds the draconian predictions of one-shot model.

H.2 predicts that Q under ICF-1 will be lower (weakly) than under ICF-B and ICF-2. While an average Q of 8.06 under ICF-1 is lower than averages of 8.22 under ICF-B and ICF2, a KW test rejected any statistical difference between ($p=0.91$). Because H.2 only predicts that Q_{ICF-1} should be *weakly* less, the results of the test does not allow us to reject H.3. Nonetheless, this conclusion is rather uninteresting without providing some economic intuition about when $Q_{ICF-1} < Q_{ICF-B\&2}$ (subscript indicates ICF-B and ICF-2 data are pooled) might occur. Note from Proposition 2 that the total pay-out the buyer would have to make to the seller for delivering some Q under ICF-B or ICF-2 is $P_{ICF-B} + b(q) = P_{ICF-2} + w(q) = \bar{u} + c(Q)$. On the other hand, total pay under ICF-1 is $P_{ICF-1} = c(\underline{q}) + \bar{u} + \dfrac{c(Q) - c(\underline{q})}{\delta}$. Thus, the marginal cost of implementing Q is $c'(Q)$ in ICF-B or ICF-2 and $\dfrac{c'(Q)}{\delta}$ under ICF-1. In other words, the marginal cost of implementing a given Q depends on the discount factor δ. When $\delta = 1$, there is no difference in marginal costs under ICF-1, ICF-B and ICF-2, but as δ gets smaller, marginal cost under ICF-1 increases over marginal costs under ICF-B and ICF-2 and we would also expect to see the buyer request lower Q under ICF-1. In the context of our experiments, a parameter such δ is rather meaningless if interpreted literally as it is difficult to imagine that our experimental subjects discount the "future" in our finitely repeated experiment, which lasted no longer than 45 minutes. Nevertheless, δ may be

---

[12] We pooled the data from these two regimes because the earlier KW test did not reject H.1.

seen as a parable for the degree to which subjects might be willing to trade short term gains for future round gains. If subjects have the belief that there are many cooperative types in the population, then they expect large payoffs from taking cooperative actions so this would be analogues to having a "large" δ. In this case, the marginal cost of implementing a given Q under ICF-1 would approach the marginal costs under ICF-B and ICF-2. The fact that we could not reject the conclusion that $Q_{ICF-1} = Q_{ICF-B\&2}$ under H.2 may thus have the economic interpretation that buyers have the belief that that there are sufficient number of cooperative sellers to support a high Q.

Before formally testing hypothesis H.3, it may instructive to first examine H.4 and the evolution of actual quality chosen by sellers. Note that hypothesis H.4 is a prediction about the ranking of actual quality chosen by sellers under various regimes. One of the interesting implications of this hypothesis, which is consistent with the arguments raised by Bernheim and Whinston (1998), is that when there are barriers to third-party enforcement, more complete contracts are not necessarily more efficient. For example, ICF-1 is the most "complete" of the incomplete contracts whereas ICF-2 is the most "incomplete" of the incomplete contracts. Yet, in a repeated trading environment, the less complete incomplete contracts appear to facilitate more powerful informal incentives that should lead to higher efficiency. Intuitively, when complete contracts are impossible to achieve, at least one party will have some discretionary latitude during the course of the transaction. Thus, it may be beneficial to relax parts of the formal contract (i.e. make it more incomplete) to provide the other party with some latitude in order to achieve balanced discretionary powers. This creates a deterrent, which prevents parties from abusing their latitude. H.4, however, does not discount the value of third-party enforcement if third-party enforcement is perfect. With perfect third-party enforcement, parties can write fully complete contracts that perfectly specify all obligations of all parties so that parties have no discretionary latitude to behave opportunistically. In this case, the problem of ensuring balanced discretionary powers is no longer an issue.

Figure 1 illustrates the evolution of average quality chosen by sellers across all 15 rounds. The results appear to be consistent with H.4 in that chosen quality is highest

under C, followed by ICF-2, ICF-B and then ICF-1, although the gap between ICF-B and ICF-1 appears to be small. Also note that as the contracting game approaches the final round, shirking occurs under ICF-1 and ICF-B, which is consistent with Kreps, Milgrom, Roberts, and Wilson's (1982) model of finitely repeated games with cooperative agents. This is also consistent with previous experiments of finitely repeated games, mostly notably the ICF treatment of BFF. Moreover, the evolution of quality under C and ICF-1 looks remarkably similar to BFF's results, which suggests that the minor differences between our experimental design and theirs did not substantially influence the results. What is also remarkable is that shirking did not seem to occur under ICF-2 even as the experiment approached the final rounds. One possible explanation for this is that if sellers hold the belief that some buyers will honor contracts so long as sellers honor contracts even in the final round (i.e. there exists cooperative buyers), then sellers may be reluctant to shirk for fear of triggering a low price response from buyers.
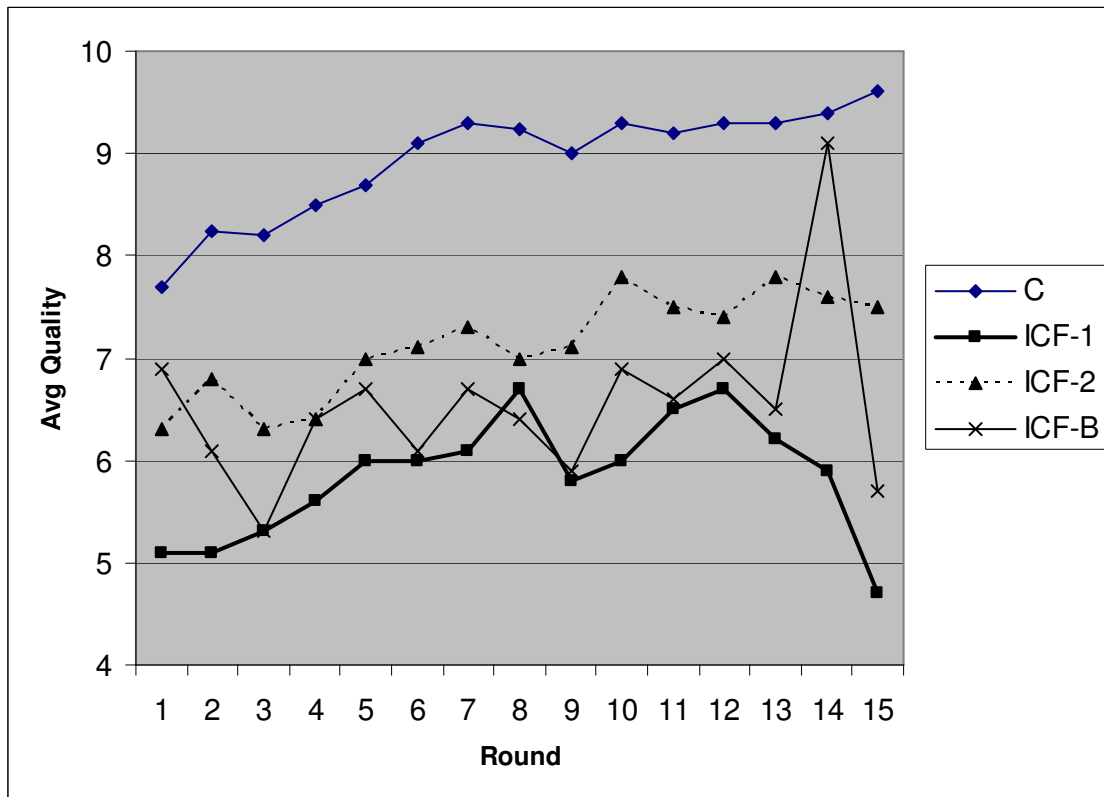

FIGURE 1. – Average quality chosen by sellers across all rounds.

To test H.4, we first test the null hypothesis that $q_C = q_{ICF-2}$, which was rejected by a KW test (p=0.0001), suggesting that $q_C > q_{ICF-2}$ held in our experiments. We next test $q_{ICF-2} = q_{ICF-B}$, which yielded a p-value of 0.104 so that there is tentative evidence that $q_{ICF-2} > q_{ICF-B}$. Finally a test of $q_{ICF-B} = q_{ICF-1}$ yielded a p-value of 0.02, which allows us to reject the null hypothesis at the 5% level of significance. We therefore have moderately strong evidence that $q_{ICF-B} > q_{ICF-1}$. For the most part, the evidence seems to support H.4 so that the evidence is consistent with the earlier claim that making an already incomplete contract more incomplete can improve efficiency.

We now return to hypothesis H.3, which predicts that the frequency to which $q<Q$ should be greatest under ICF-1 and least under ICF-2 (degree of shirking under ICF-B should lie in between). In Table 1, we also reported the percentage of trades for which actual quality fell short of contracted quality and the average absolute size of the quality shortfall. Using either measure, the results seem to be consistent with H.3 as the percentage of trades for which $q < Q$ is highest under ICF-1 (61%) and lowest under ICF-2 (45%). Similarly, the absolute size of the shortfall averaged 2.31 quality units under ICF-1 and only 1.36 under ICF-2. A KW test rejects the null hypothesis that the percentage of trades for which $q < Q$ is the same under ICF-B and ICF-2 (p=0.01). However, we could not reject the null hypothesis that the percentage of trades for which $q < Q$ is the same under ICF-1 and ICF-B (p=0.28). These results are consistent with the conclusion that the least amount of shirking occurred under ICF-2 which is consistent with H.3. However, we could not conclude that the most amount of shirking occurred under ICF-1 since there appeared to be no statistical difference in the amount of shirking that occurred under ICF-1 and ICF-B.

We also examined the absolute size of the quality shortfall (Q - q) and the results seem to be consistent with the results obtained from using percentage of trades for which $q < Q$. The null hypothesis that $(Q-q)_{ICF-1} = (Q-q)_{ICF-B}$ could not be rejected (p=0.346), while the null that $(Q-q)_{ICF-B} = (Q-q)_{ICF-2}$ is rejected (p=0.0005). Thus,

our data suggests that the intensity of shirking is equally strong under ICF-1 and ICF-B, and weaker under ICF-2.

A summary of the results in this section are that (1) we could not reject the conclusions that buyers requested the same level of Q under ICF-1, ICF-B and ICF-2 and we rejected the conclusion that $Q=10$; (2) we have statistical support for the conclusion that actual q chosen by sellers under the various regimes can be ranked $q_C > q_{ICF-2} \geq q_{ICF-B} > q_{ICF-1}$; and (3) our data suggests that the frequency and intensity of shirking was strongest under ICF-1 and ICF-B and weakest under ICF-2.

4.2. *Prices, Bonuses, and Deducts*

The repeat trading model also makes several predictions concerning the level of contracted price, P, offered by buyers to sellers. A qualitative implication of Proposition (iii) is that the contracted price, P, will be lower under ICF-B than under ICF-1 since ICF-1 relies on efficiency wages to motivate quality whereas ICF-B relies on a discretionary bonus. In particular, Proposition (iii) and (iv) yield the prediction that, under ICF-B, buyers will guarantee a price P that is equal to the sum of the seller's reservation utility plus the cost of producing the lowest quality level, $\underline{q}$. Then if the seller honors the contract by producing $q \geq Q$, then the buyer pays a bonus equal to the cost to the seller of honoring the contract. In contrast, buyers under ICF-1 attempt to induce the seller to honor the contrast by paying a rent equal to $\dfrac{c(Q_{ICF-1}) - c(\underline{q})}{\delta}$, which increases when δ decreases. As mentioned earlier, within the context of our experiments, we can instead think of δ is a metaphor for buyers' beliefs about how cooperative sellers are. If buyers believe that sellers are mostly uncooperative, then it would require large rents to induce cooperation. Finally, Proposition 1 (i) suggest that, under regime C where Q is exogenously enforced, the buyer can simply offer a P that just induces participation so that P will equal the seller's reservation payoff plus the cost of producing the contractually specified quality level. Seller's do not earn rents under C so we can expect $P_C < P_{ICF-1}$. We also have $P_{ICF-B} < P_C$ because it is the combination of the contract price

and the discretionary bonus that allows a seller to satisfy his participation constraint from honoring the contract. This discussion leads to the following testable hypothesis.

H.5: *Contracted price under ICF-B is lower than contracted price under C and contracted price under C is lower than contract price under ICF-1.*

Note that we can say little about contract price under ICF-2 because buyers can structure contracts in many different ways; i.e. they can offer a high P and then use a discretionary deduct to punish underperforming sellers or they can offer a low P combined with a discretionary bonus to mimic ICF-B contracts. It would be interesting to examine what buyers actually did using our experimental data.

TABLE 2
Price and Price Deviations Summary Statistics across Enforcement Regimes

	Obs.	*Avg. Contracted Price, P*	*Avg. Actual price, p*	*% of trades where p<P*	*% of trades where p>P*	*Avg. Difference p-P*
Regime C	436	30.66	30.66	0	0	0
Regime ICF-1	290	39.78	39.78	0	0	0
Regime ICF-B	213	34.43	37.11	0	27%	2.68
Regime ICF-2	374	41.32	31.69	58%	12%	-9.63

Table 2 reports summary statistics for prices and price adjustments. H.5 predicts that P should be lowest under ICF-B and highest under ICF-1. Consistent with H.5, the average $P_C = 30.66$, which is lower than average $P_{ICF-1} = 39.78$. However, average $P_{ICF-B} = 34.43$ is larger than average P_C which is inconsistent with H.5. To formally test H.5, we conducted a KW test of $P_{ICF-1} = P_C$, which was rejected (p=0.0001), which is consistent with the conclusion that $P_{ICF-1} > P_C$. We also tested $P_{ICF-1} = P_{ICF-B}$, which was rejected at the 5% level of significance (p=0.012), which provides some support for the conclusion that $P_{ICF-1} > P_{ICF-B}$. Finally, we tested $P_{ICF-B} = P_C$, which was

rejected ($p=0.0004$). Since average P_{ICF-B} exceeds average P_C, a rejection of $P_{ICF-B} = P_C$ is evidence in support of $P_{ICF-B} > P_C$, which contradicts H.5.[13]

One possible explanation for our result $P_{ICF-B} > P_C$ is that, because a bonus is discretionary in ICF-B, sellers may be reluctant to accept contracts with a low P when there is no third-party enforceable guarantee that the bonus will be paid. A typical seller would have to be very trusting to accept a contract offering a low P and then deliver high quality without a promise that her costs will be covered. To explore this issue further, recall that after our seller subjects chose quality in the ICF-B experiments, they were asked to specify what level of p they expect buyers' to choose before they observe what buyers' actually choose. For the subset of trades for which contracts $q \geq Q$ in ICF-B, only 51% of sellers expected a bonus to be paid. Moreover, when examining the expectations data only for early rounds 1 thru 3, before long term relationships had a chance to take root and before sellers had much experience trading with buyers, only 33% of sellers who honored their agreements expected bonuses to be paid. Finally, the average contract price of rejected offers was 25.44, which is substantially lower than the average of 34.43 of accepted offers ($p=0.0001$).

An interesting result is that average $P_{ICF-2} = 41.32$, which suggests that buyers offer relative high prices in the ICF-2 regime. In fact, the statistical evidence for a difference between average P under ICF-1 and ICF-2 is weak ($p=0.12$). However, it should be noted that, in the ICF-2 experiments, buyers imposed deducts on sellers in 58% of trades and average contracted price exceeded actual prices by an average of 9.63. Hence, although buyers offered nearly the same contracted price in our ICF-1 and ICF-2 experiments, actual prices paid by buyers were much lower under ICF-2.

---

[13] It is possible that this result is driven by the fact that the reservation payoff to the seller was set at 10 in all three ICF-B experiments, but set at 10 in only two of the six C experiments (it was set at 5 in the remaining four C experiments). Thus, the average reservation payoff is lower under C and buyers may respond accordingly by offering a lower P. We tested for this possibility by running a censored regression (constant omitted) with P as the dependent variable. The right-hand-side variables included dummies for each of the four regimes, and a "reservation" variable to control for variation in reservation payoffs. However, it appears that controlling for the reservation payoffs had no impact on the qualitative conclusions so we did not report the results from this regression.

Given that deducts were so prevalent and so large in ICF-2, it might be worth exploring whether buyers were behaving opportunistically by failing to honor promised prices to sellers or whether buyers were legitimately using deducts to provide incentives. If buyers are using deducts for incentive provision, then we ought to observe a correlation between the incidence of deducts and seller performance.

TABLE 3
Probability of Deducts and Bonuses[a]

	(1) Deduct dF/dX ICF-2 data only	(2) Bonus dF/dX ICF-2 and ICF-B data
Shirking dummy (1if $q<Q^*$)	0.77*** (0.037)	-0.22*** (0.03)
Private trade dummy	-0.18* (0.097	0.13*** (0.04)
P^*	0.01*** (0.003)	-0.005*** (0.009)
Length of private relationship.	-0.07*** (0.025)	0.006 (0.004)
ICF-B dummy		0.12*** (0.05)
Experiment Fixed Effects	Yes	Yes
Obs.	349	547
Wald Chi-sq	$\chi(8)=116.4$ $p=0.00$	$\chi(11)=140.6$ $p=0.00$
Pseudo R^2	0.567	0.385

[a]Regression are probits with robust standard errors (in parentheses). Reported coefficients are marginal effects (Δ probability for small change regressor)
***,**,*Indicates that the estimate is significantly different from 0 at the 1%, 5%, and 10% levels, respectively.

Table 3 reports the marginal effects from probit regressions that estimate the probability of deducts and bonuses. Column (1) contains the marginal effects estimates for deducts. Marginal effects give us the change in probability of a deduct given small

changes in the independent variables. Note that when suppliers shirk on quality, it increases the probability of deducts by 0.77, which strongly suggests that deducts are used for incentive provision. Moreover, the probability of deducts decreases when trades are private and when the length of the relationship between a buyer and seller is long. Thus, buyers engaged in cooperative, self-enforcing relationships with sellers are less likely to use deducts. Finally, small increases in contracted price P increases the probability of a deduct. While these results do suggest that deducts are used for incentive provision, it should be noted that opportunism cannot be ruled out as deducts were still imposed in 28% of trades (not reported in Table 3) for which sellers honored contracts.

TABLE 3
Probability of Deducts and Bonuses[a]

	(1) Deduct dF/dX ICF-2 data only	(2) Bonus dF/dX ICF-2 and ICF-B data
Shirking dummy	0.77***	-0.22***
(1if $q<Q^*$)	(0.037)	(0.03)
Private trade dummy	-0.18*	0.13***
	(0.097	(0.04)
P^*	0.01***	-0.005***
	(0.003)	(0.009)
Length of private	-0.07***	0.006
relationship.	(0.025)	(0.004)
ICF-B dummy		0.12***
		(0.05)
Experiment Fixed	Yes	Yes
Effects		
Obs.	349	547
Wald Chi-sq	$\chi(8) = 116.4$	$\chi(11) = 140.6$
	$p=0.00$	$p=0.00$
Pseudo R^2	0.567	0.385

[a]Regression is a probit with robust standard errors (in parentheses). Reported coefficients are marginal effects (Δ probability for small change regressor)
***,**,*Indicates that the estimate is significantly different from 0 at the 1%, 5%, and 10% levels, respectively.

Column (2) reports marginal effects for bonuses. The results are intuitive in that shirking reduces the probability of bonuses by 0.22, private trading increases the probability of bonuses by 0.13, and small increases in P reduces the probability of bonuses. The positive ICF-B marginal effect suggests that buyers are more likely to use bonuses in the ICF-B experiments relative to the ICF-2 experiments, which is also intuitive since ICF-2 buyers can use deducts as substitutes for bonuses for incentive provision. This result is consistent with the summary statistics in Table 2, which show that bonuses occurred in 27% of trades in ICF-B but in only 12% of trades in ICF-2.

In summary, the key conclusions of this section are that (1) contracted prices are highest under ICF-1 and ICF-2, lowest under ICF-C, and contracted price is higher than expected under ICF-B, which may be because it is difficult to entice sellers to accept contracts with low guaranteed prices when bonuses are only discretionary; (2) the probability that buyers will impose deducts increases when sellers shirk, decreases under private trading, decreases in relationships that have a long history, and increases with contracted price; and (3) the probability of bonuses decreases with seller shirking, increases with private trading, decreases with contracted price, and bonuses were more prevalent in ICF-B experiments than in ICF-2 experiments.

4.3. *Distribution and Rent Sharing*

Corollary 1 following Proposition 2 predicts that sellers should earn rents over their reservation payoffs under ICF-1, but not under any other contracting regime. Recall that ICF-1 is essentially the RPM, which enforces quality through the use of an efficiency wage which guarantees sellers rents when there is cooperation. Our next hypothesis is:

H.6: *Buyers share rents generously with sellers under ICF-1 but hold sellers close to their reservation payoffs in the other regimes.*

One implication of H.6 is that the ratio of sellers' surplus to total surplus, $\dfrac{U - \bar{u}}{\pi + U - \bar{u}}$,

should be larger under ICF-1 than under the other regimes.[14] Figure 2 illustrates the distribution of this ratio across different scenarios.
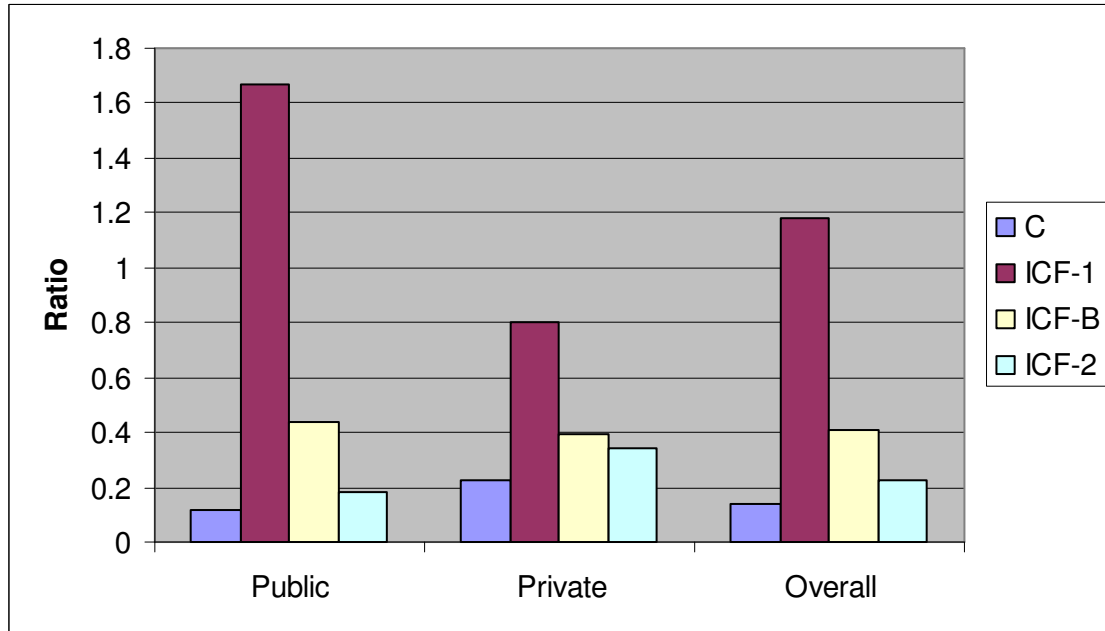

FIGURE 2. - Ratio of sellers' surplus to total surplus.

In a qualitative sense, the ratio of sellers' surplus to total surplus is consistent with H.6 since this ratio is clearly highest under ICF-1 both overall and across public and private trades. This was verified by the fact that KW tests applied to the overall data rejected the equality of the ratio across the following regimes: (1) ICF-1 vs. all other regimes pooled ($p=0.0001$), (2) ICF-1 vs. C (rejected: $p=0.0001$), ICF-1 vs. ICF-5 (rejected: $p=0.0001$), and ICF-1 vs. ICF-2 (rejected: $p=0.0001$).

Overall, sellers earned over 100% of total surplus under ICF-1, close to 80% under private trading, and over 160% in public trading.[15] One logical explanation for high seller surplus under ICF-1 is that Q is unenforceable while P is enforceable; thus, the power to engage in opportunism belongs only to sellers. The fact that the ratio was

---

[14] Note that $\bar{\pi} = 0$ in our experiments.

[15] Sellers can earn more than 100% (ratio greater than 1) when buyers make negative profits.

highest in public trading is consistent with an opportunism explanation given that shirking is most likely to occur in public trading where sellers were not involved in self-enforcing agreements. Nonetheless, even if opportunism did not exist, buyers clearly intend to share rents generously with sellers in ICF-1. Based on the results of earlier sections, the average contract offered by buyers to sellers under ICF-1 is $(P, Q) \approx (40, 8)$, which would yield total profits of 63 with sellers receiving a surplus of 23 so that the ratio would be 0.365. In both C and ICF-2, sellers captured less than 25% of total surplus overall, although the situation improved for sellers under private trading.

Figure 3 tracks the ratio of seller surplus to total surplus over time. The evolution of the ratio of sellers' to total surplus is consistent with Figure 2 in that seller surplus is highest under ICF-1 in every round and seller surplus under C and ICF-2 is quite low converging toward zero over time.
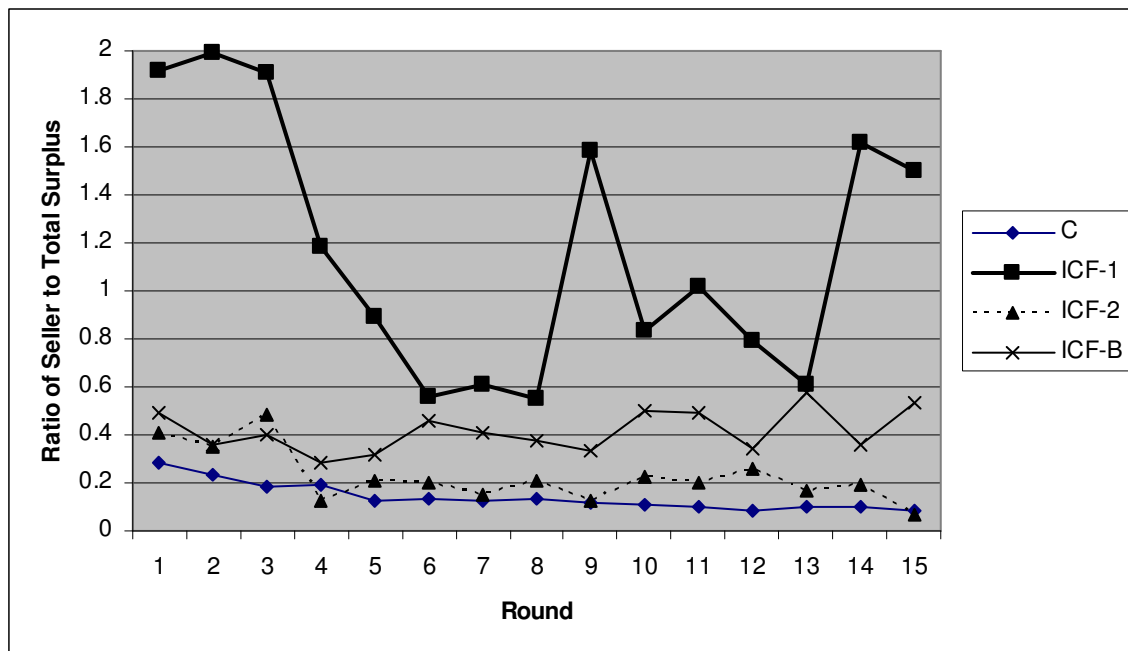

FIGURE 3: Evolution of the ratio of sellers' surplus to total surplus.

Under ICF-B sellers faired better than under C and ICF-2 which is inconsistent with H.6. Nonetheless, this result is consistent with our earlier finding that contracted price P under ICF-B is higher than predicted. It may be that in order to induce sellers to

accept contracts, buyers had to offer a *P* that provided some rents to sellers. Since actual *p* can never fall below *P* in the ICF-B regime, an increase in *P* is correlated with increased rent to sellers.

The key result of this section that generous rent-sharing is not a necessary outcome of incomplete contracts or of market interaction in the absence of third-party enforcement of contracts. Indeed, the patterns that emerge in Figures 2 and 3 suggest the counterintuitive result that, as third-party enforcement is relaxed and as the *degree* of contractual incompleteness *increases* from ICF-1, to ICF-B and then to ICF-2, rent sharing with sellers actually *decreases.* The pattern of rent sharing under ICF-2, where third-party enforcement is completely missing, more closely resembles the complete contract, C than other types of incomplete contracts. It appears that generous rent sharing is a specific feature of the RPM, and to a lesser extent, the relational contract that promises only discretionary bonuses.

An interesting pattern to note is that quality progressively increases as we move from more stringent to less stringent enforcement regimes (ICF-1 to ICF-B to ICF-2) while rent-sharing with sellers declines. To understand what is going on, it would be useful to refer to the sellers' dynamic incentive constraints under each regime given be equations (3), (9), and (14). It is clear from equation (3), which is the seller's dynamic enforcement constraint under ICF-1, that the only way for this constraint to be satisfied to prevent seller shirking is for *P* to be high enough to ensure rents to the seller for not shirking. Note that when $\delta = 1$, which implies that the highest degree of cooperation is expected, then $P \geq \bar{u} + c(q)$ in which case the buyer would hold the seller right at her reservation payoff if she supplies quality level, *q*. However, as soon as $\delta < 1$ so that sellers are expected to be less cooperative, then a rent has to be paid to sellers to induce cooperation. In contrast, (9) is the seller's dynamic incentive constraint under ICF-B. Note from Proposition 2 that the optimal contract under ICF-B implies $P = \bar{u} + c(\underline{q})$ and $b(q) = c(q) - c(\underline{q})$. The significance of $P = \bar{u} + c(\underline{q})$ is that if it is substituted into the r.h.s. of (9), the r.h.s. becomes $\dfrac{\bar{u}}{1-\delta}$. In other words, the optimal contract ensures that

the seller can do no better than her outside payoff if she shirks. Then the bonus only has to be high enough to ensure that the l.h.s. of (9) is at least as great as her outside payoff, which is accomplished by letting $b(q) = c(q) - c(\underline{q})$, which just covers the cost of producing high quality. The ICF-B regime provides the buyer with enough discretionary latitude to reduce the seller's payoff from shirking which reduces rents needed to motivate the seller. This argument is magnified under ICF-2 because the buyer has so much discretion that she can force the seller's payoff from shirking to fall below the seller's outside payoff. To see this, consider the r.h.s. of (14) which is the seller's payoff from shirking after accepting a contract that requests a quality level, *q*. Because the buyer has the latitude to completely withhold payment, she can impose a severe punishment on the seller for shirking thereby leaving the seller with only a short-term gain of $-c(\underline{q})$, which is clearly less than her outside payoff of $\bar{u}$. Hence, in equilibrium, the seller never has an incentive to shirk after having accepted a contract; in fact, once a contract is accepted by the seller, the buyer can hold the seller hostage and ensure that she doesn't even earn her outside payoff, which provides powerful incentives for the seller to deliver high quality. The buyer only has to structure the contract in such a way that the seller's total pay from cooperating restores her to her outside payoff. This is accomplished with the contract the specifies a total pay of $P + w(q) = c(q) + \bar{u}$, which pays no rents to sellers. In short, discretionary latitude for buyers allows them to alter sellers' payoffs from shirking and alter the feasible continuation equilibria.

4.4. *Identity and Reputation Based Trading*

A key implication of Proposition 3 is that cooperation should be easier to achieve under ICF-1 and ICF-B relative to ICF-2. In addition, because the role of cooperation is to provide self-enforcement of contracts, it should play no role in regime C because contracts are exogenously enforced. Within the context of our experiments, private offers enable reputation based trading and long-term relationships, which lie at the core of relational contracting. For example, if a buyer predicts benefits from contracting with a specific seller (i.e. can earn relationship-specific rents), the buyer can make a private

offer to that seller in each round rather than venturing into the open market and hoping that that seller is the first to accept the offer. Had we not incorporated private trading, it would have been difficult for parties to establish relational agreements based on identity and reputation as buyers would have had to hope that their targeted sellers were first to accept their public contracts. Let the fraction of trades conducted via private offers be denoted by the symbol, ρ. Then our next hypothesis is,

H7: *The fraction of trades conducted via private offers are ranked as follows:*

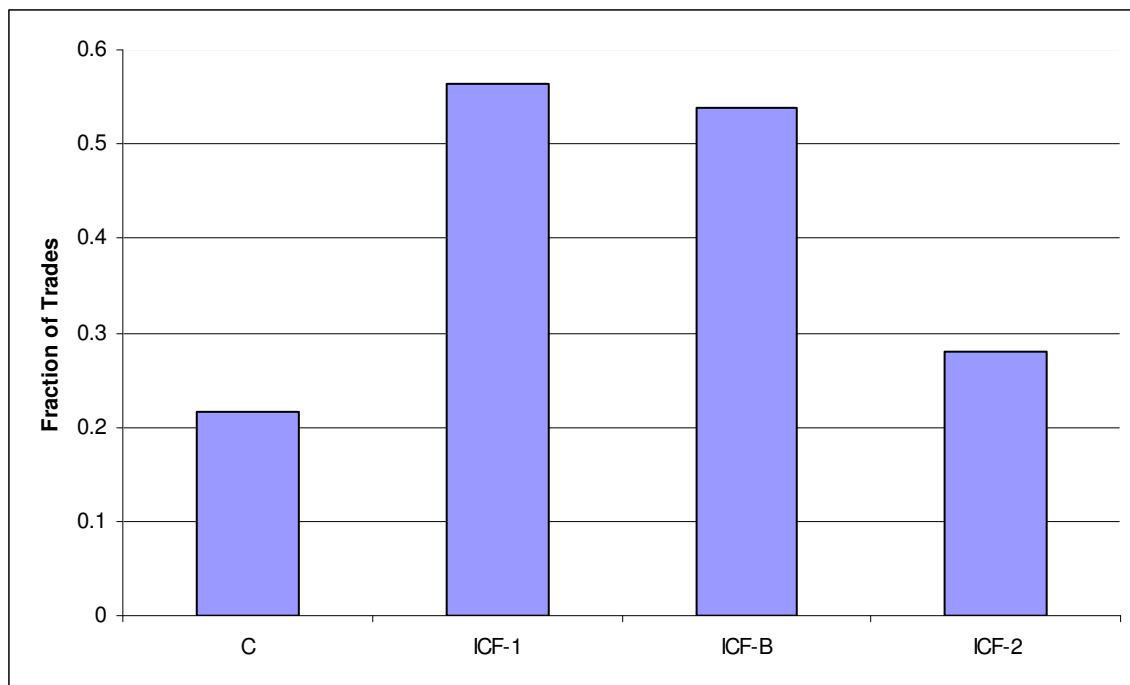$$\rho_{ICF-1} = \rho_{ICF-B} > \rho_{ICF-2} > \rho_C$$

FIGURE 4: Fraction of private trades in each enforcement regime.

Figure 4 provides an overview of the fraction of private trading that occurred under each of the contracting regimes. The pattern of private trading appears consistent with H.7 in that the incidence of private trading is highest under ICF-1 and ICF-B, and lowest under C. A KW test revealed that $\rho_{ICF-1} = \rho_{ICF-B}$ could not be rejected (p=0.568) so that there appears to be no statistical difference between the fraction of private trades that occurred under ICF-1 and ICF-B which is consistent with H.7. We also tested

40

ρ_{ICF-1} & $\rho_{ICF-B} = \rho_{ICF-2}$, as well as ρ_{ICF-1} & $\rho_{ICF-B} = \rho_C$, both of which were rejected (p=0.0001 for both tests). Finally, a test of $\rho_{ICF-2} = \rho_C$ was rejected at the 5% level of significance (p=0.03). These results are remarkably consistent with H.7.
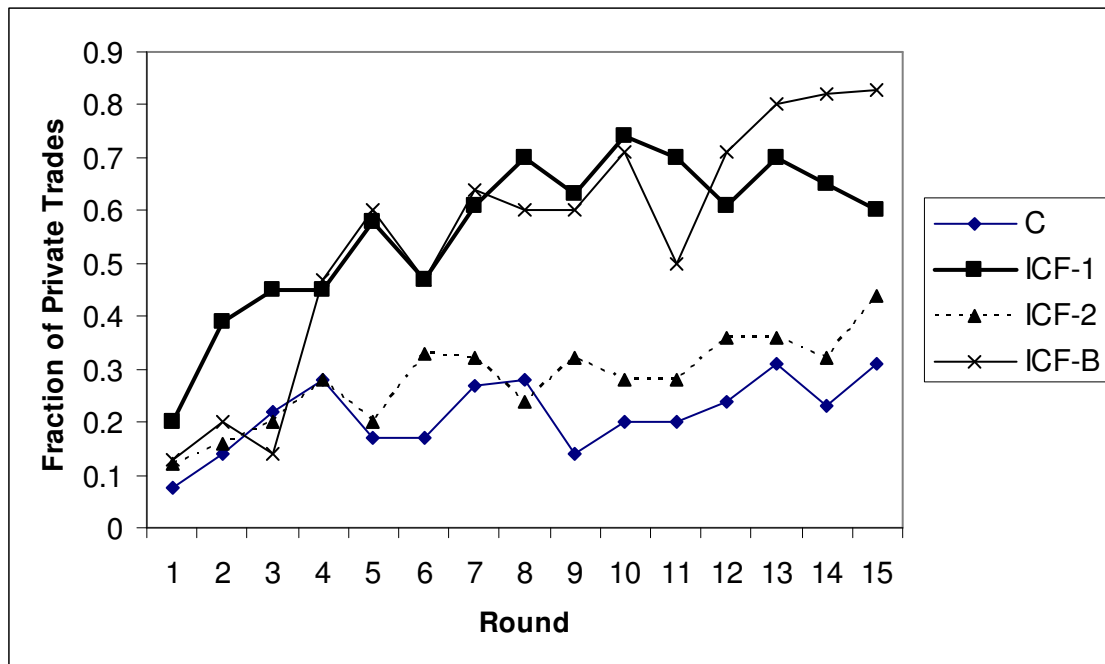

FIGURE 5. Evolution of the fraction of private trades over time.

It would also be useful to assess the pattern of trading over time and Figure 5 allows us to do this. It is interesting to note that the pattern of evolution between ICF-2 and C is very similar even though C is the complete contract and ICF-2 is the most incomplete contract. Meanwhile, the two contracts with intermediate levels of incompleteness evolve very differently. Private trading increases fairly steeply under ICF-1 and ICF-B in the early rounds topping out at about 0.75 under ICF-1 and at around 0.85 under ICF-B. The trajectory of ICF-2 is much flatter and never exceeds 0.45 in any round. Private trading is consistently lowest under C and never tops 0.31 in any round.

BFF suggest that, in the absence of third-party enforcement, markets resemble a collection of bilateral trading islands formed via private trades. Competitive pricing is

less important than reputation and the identity of the trading partner. This continued to be true in our ICF-1 experiments, as well as our new ICF-B experiments. However, in the case where third-party enforcement was completely absent, bilateral trading islands begin to weaken and trading again resembles the complete contracting regime. Our results suggest that it is difficult to draw general conclusion about the nature of market interactions in the absence of third-party enforcement.

An obvious puzzle is if the fraction of private trades is so low under ICF-2, how is it that average quality in this regime is higher than under either ICF-1 and ICF-B? In the absence of third-party enforcement, parties must find ways to self-enforce contracts and one of the primarily mechanisms for facilitating self-enforcement is to establish cooperative agreements through private trading. Yet, the vast majority of trades in ICF-2 are public, so how is high quality achieved? One possible explanation is that, because price is not third-party enforceable, buyers have strong incentives to renege on payments. By accepting the good and then not paying for it, the short term gains to buyers can be substantial. Thus, the typical buyer must see large benefits from continuation with a specific seller to continue to cooperate. This then puts tremendous pressure on sellers to deliver high quality to enhance buyers' revenues so that buyers will not renege on price. Even in public trading, the pressure to deliver high quality exists for sellers for two reasons. First, if a seller wants to establish credibility with a buyer so that the buyer will trade with him privately in the future, he must entice the buyer with high quality. Second, a close examination of sellers' dynamic incentive constraint suggests that sellers never have an incentive to renege once the seller has accepted a contract. Recall that once a contract has been accepted, buyers can hold sellers hostage by withholding payment so that sellers would earn even less than her outside payoff. Sellers therefore must perform to prevent a retaliatory response from buyers within the stage-game.

To examine these issues further, Table 4 presents probit results of the determinants of sellers' *expectations* of whether buyers will shirk i.e. choose $p<P$ (deducts). Recall that during our ICF-2 experiments after sellers had chosen quality and before sellers had observed actual p chosen by buyers, sellers were asked to state what p

they expected sellers to choose. If expected p was less than contracted price P, then sellers expect buyers to shirk or use deducts. Our probits allows us to examine the determinants of sellers' expectations of buyer shirking for public and private trading.

TABLE 4
Determinants of Sellers' Expectations of Buyer Shirking in ICF-2 (Expected Price $< P$)[a]

	(1) dF/dX ICF-2 Public data only	(2) dF/dX ICF-2 Private data only
q chosen by seller	0.002	-0.13**
	(0.02)	(0.06)
Seller shirked (dummy = 1 if $q<Q$)	0.41***	0.38**
	(0.07)	(0.17)
Contracted price, P	0.007***	0.017***
	(0.003)	(0.005)
Length of relationship up to current round		-0.003
		(0.02)
Experiment Fixed Effects	Yes	Yes
Obs.	269	90
Wald Chi-sq	$\chi(7) = 49.90$ $p=0.00$	$\chi(7) = 21.23$ $p=0.00$
Pseudo R^2	0.138	0.415

[a]Regression is a probit with robust standard errors (in parentheses). Reported coefficients are marginal effects (Δ probability for small change regressor)
***,**,*Indicates that the estimate is significantly different from 0 at the 1%, 5%, and 10% levels, respectively.

It is interesting to note that the strongest determinant (largest estimated marginal effect) is whether sellers held up their end of the bargaining by honoring the contracted quality. The marginal effects were large, significantly different from zero, and consistent across both public and private trading (0.41 and 0.38, respectively). Sellers thus have strong incentives to honor contracts as they believed that failing to do so would significantly increase the probability that buyers would deduct price. The absolute level of quality only seemed to matter in private trading where sellers' anticipated that a one-unit increase in quality would reduce the probability of buyer shirking by 0.13. It is also interesting to note that sellers' expected an increase in the probability of buyer shirking

when contracted *P* was higher. Perhaps sellers had some intuitive sense that high promised prices might be "too good to be true" because it would be more costly for buyers to honor contracts. Under private trading, the length of the relationship had little impact on sellers' expectations concerning buyer shirking. Thus, it appears that once a seller has accepted a contracted, the history of the relationship did little to influence sellers' expectation of buyer behavior within the stage game. Overall, the probit results suggest that having the discretion to adjust prices offered buyers a powerful incentive instrument even within a stage-game.[16] This could explain why high quality is achievable in ICF-2 even though the incidence of private trading was lower.

5. CONCLUSION

In this study, we use economic experiments to compare and contrast relational trading under the *repeat purchase mechanism* of Klein and Leffler (1981) against more general forms of relational contracts, such as those that allow for discretionary bonuses and deducts (e.g. MacLeod and Malcomson 1989; Levin 2003, among others). The major finding of this study is that the efficiency and the nature of relational trading are profoundly impacted by seemingly simple alterations in the degree of discretionary latitude available to traders. In one treatment where maximum discretion is allowable (i.e. no terms of the underlying written contract are third-party enforceable so the contract is extremely incomplete), the trading patterns that emerge are closer to the patterns observed under perfect third-party enforcement (i.e. complete contract) than under the repeat purchase mechanism. Moreover, among the relational contracts that we study, the repeat purchase mechanism proved to be the least efficient. This finding, combined with the fact that some theorists have claimed that the repeat purchase mechanism is a relatively inefficient trading mechanism, suggests that it may not be widely used in competitive environments even without third-party enforcement of contracts. An

---

[16] Of course, if we focus only on a stage-game independent of repeated game effects, then the question arises as to why sellers would accept a contract in the first place. Clearly, sellers have to believe that there are some "cooperative" buyers in the population who would honor the contracted price so long as the seller delivered promised quality. If all buyers were strictly selfish, or if sellers believed that buyers were strictly selfish, then sellers would never accept a contract in the first place.

implication of this is that efficiency wages, generous rent sharing, and reputation based trading, which are outcomes of the repeat trading mechanism, may not generalize to all environments that lack third-party enforcement. If indeed efficiency wages and generous rent sharing are not trading outcomes under more general relational contracts, then this may provide a possible resolution to the debate concerning whether rent sharing and efficiency wages can persist in equilibrium (see Carmichael 1985). It is possible that parties can easily use discretionary adjustments and side-payments that reduce the need to use rents as a motivational device.

Our findings also offer a possible explanation to an issue that has puzzled some legal scholars, which is: Why do so many real world contracts appear to be deliberately incomplete? Scott (2003) collected information on a large number of cases that have been dismissed by American courts and finds that a surprising number of them appear to be endogenously incomplete in the sense that many contracting parties decline to condition performance on verifiable performance measures that can be specified in the contract at low cost. Our results, combined with Bernheim and Whinston (1998) theoretical conclusions regarding *strategic ambiguity,* suggest that endogenous incompleteness might be a rational response to barriers to complete contracting.

A possible direction for future research might involve a more detailed examination of how explicit, enforceable contracts interact with informal incentives. For example, one can conduct experiments based on the assumption that buyers care about two performance outcomes, such as quality and quantity, where only one of these measures is third-party enforceable. It would be interesting to see whether an increase in contractual incompleteness would still enhance efficiency.

REFERENCES

Acemoglu, D. and S. Johnson. (2005): "Unbundling Institutions," *Journal of Political Economy,* 113, 949-995.

Andreoni, J.A., and J.H. Miller. (1993): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," *Economic Journal,* 103, 570-85.

Axelrod, R. (1981): "The Emergence of Cooperation Among Egoists," *American Journal of Political Science,* 34, 306-18.

Baker, G., R. Gibbons, and K.J. Murphy. (2002): "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics,* 117, 39-84.

Banerjee, A.V., and E. Duflo. (2000): "Reputation Effects and the Limits to Contracting: A Study of the Indian Software Industry," *Quarterly Journal of Economics,* 115, 989-1017.

Bernheim, B., and M.D. Whinston. (1998): "Incomplete Contracts and Strategic Ambiguity," *American Economic Review,* 88, 902-932.

Brown, M., A. Falk, and E. Fehr. (2004): "Relational Contracts and the Nature of Market Interactions," *Econometrica,* 72, 747-780.

Cabral, L., and A. Hortacsu. (2006): "The Dynamics of Seller Reputation: Theory and Evidence from eBay," Mimeo, New York University and University of Chicago.

Carmichael, H.L. (1985): "Can Unemployment be Voluntary?: Comment," *American Economic Review,* 75, 1213-1214.

Cooper, R. (1996): "Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games," *Games and Economic Behavior,* 12, 187-218.

Dal Bó, P. (2005): "Cooperation Under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games," *American Economic Review,* 95, 1591-1604.

Dixit, A. (2003): "On Modes of Economic Governance," *Econometrica*, 71, 449-481.

Dixit, A.K. (2004): *Lawlessness and Economics,* Princeton: Princeton University Press.

Fischbacher, U. (1999): "Z-Tree: Experimental Software." University of Zurich.

Jin, G.Z., and P. Leslie. (2003): "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics,* 118,

409-451.

Klein, B., and K. Leffler. (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy,* 89, 615-641.

Kreps, D.M., P. Milgrom, J. Roberts, and R. Wilson: (1982): "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory, 27, 245-252.*

Levin, J. (2003): "Relational Incentive Contracts," *American Economic Review,* 93, 835-857.

List, J.A. (2006): "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions," *Journal of Political Economy,* 114, 1-37.

MacLeod, B.W. (2006): "Reputations, Relationships and the Enforcement of Incomplete Contracts," IZA Discussion Paper 1978, Institute for the Study of Labor (IZA).

MacLeod, B.W., and J.M. Malcomson. (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 57, 447-480.

MacLeod, B.W., and J.M. Malcomson. (1998): "Motivation and Markets," *American Economic Review*, 88, 388-411.

Noldeke, G., and K. Schmidt. (1995): "Option Contracts and Renegotiation: A Solution to the Hold-Up Problem," *Rand Journal of Economics,* 26, 163-79.

Scott, R.E. (2003): "A Theory of Self-Enforcing Indefinite Agreements," *Columbia Law Review,* 103, 1641-1699.

Resnick, P., and R. Zeckhauser. (2002): "Trust Among Strangers in Internet Transactions: Empirical Analysis of Ebay's Reputation System," In *The Economics of the Internet and E-Commerce,* ed. By M.R. Baye, Volume 11 of *Advances in Microeconomics,* Amsterdam: Elsevier Science.

Rubinstein, A. (1991): "Comments on the Interpretation of Game Theory," *Econometrica,* 59, 909-924.

Williamson, O. (1991): "Comparative Economic Organization: Analysis of Discrete Structural Alternatives," *Administrative Science Quarterly,* 36, 269-96.