NBER WORKING PAPER SERIES

CONTRACTS AS REFERENCE POINTS - EXPERIMENTAL EVIDENCE

Ernst Fehr
Oliver D. Hart
Christian Zehnder

Working Paper 14501
http://www.nber.org/papers/w14501

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2008

We would like to thank Philippe Aghion, Bjorn Bartling, Tore Ellingsen, Ray Fisman, Bob Gibbons, Steve Leider, Bentley MacLeod, Al Roth, Klaus Schmidt, and Birger Wernerfelt for helpful comments. We would like to acknowledge financial support from the U.S. National Science Foundation through the National Bureau for Economic Research and the Research Priority Program of the University of Zurich on the "Foundations of Human Social Behavior." The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Contracts as Reference Points - Experimental Evidence
Ernst Fehr, Oliver D. Hart, and Christian Zehnder
NBER Working Paper No. 14501
November 2008
JEL No. D0,K0,C9

ABSTRACT

In a recent paper, Hart and Moore (2008) introduce new behavioral assumptions that can explain long-term contracts and important aspects of the employment relation. However, so far there exists no direct evidence that supports these assumptions and, in particular, Hart and Moore's notion that contracts provide reference points. In this paper, we examine experimentally the behavioral forces stipulated in their theory. The evidence confirms the model's prediction that there is a tradeoff between rigidity and flexibility in a trading environment with incomplete contracts and ex ante uncertainty about the state of nature. Flexible contracts - which would dominate rigid contracts under standard assumptions - cause a significant amount of shading on ex post performance, while under rigid contracts, much less shading occurs. Thus, although rigid contracts rule out trading in some states of the world, parties frequently implement them. While our results are broadly consistent with established behavioral concepts, they cannot easily be explained by existing theories. The experiment appears to reveal a new behavioral force: ex ante competition legitimizes the terms of a contract, and aggrievement and shading occur mainly about outcomes within the contract.

Ernst Fehr
Institute for Empirical Economics Research
University of Zurich
Blumlisalpstrasse 10
8006 Zurich
Switzerland
efehr@iew.unizh.ch

Christian Zehnder
Faculty of Business and Economics
University of Lausanne
Quartier UNIL-Dorigny
Internef 612
CH-1015 Lausanne
Switzerland
zehnder@iew.uzh.ch

Oliver D. Hart
Department of Economics
Littauer Center 220
Harvard University
Cambridge, MA 02138
and NBER
ohart@harvard.edu

I. Introduction

In a recent paper Hart and Moore (2008) develop a theory which provides a basis for long-term contracts in the absence of noncontractible investments. The theory is also capable of rationalizing the employment contract, which fixes wages in advance and leaves discretion to the employer. The theory rests, however, on strong behavioral assumptions that lack direct empirical support and deviate in important ways from the assumptions made in standard contract theory. For this reason, and because of the potentially far-reaching implications of the new theory, it seems particularly important to provide empirical support for the theory. In this paper, we provide an experimental test of some of the key predictions of the theory of Hart and Moore (2008). In doing so, we identify new behavioral forces that can neither be explained by traditional contract theory nor by currently existing behavioral models. These forces are, however, predicted by the Hart-Moore notion that competitively determined contracts perform the function of reference points.

It is useful to start with some background and motivation. According to the standard incomplete contracts literature, trading parties find it difficult to write a long-term contract because the future is hard to foresee. As time passes and uncertainty is resolved, the parties can complete their contract through renegotiation. The typical model supposes symmetric information and no wealth constraints, so that Coasian bargaining ensures ex post efficiency. However, there is a hold-up problem: as a consequence of renegotiation, each party shares some of the fruits of prior (non-contractible) investments with the other party. Anticipating this, each party under-invests.

While this approach has been useful for studying asset ownership (see Grossman and Hart (1986), Hart and Moore (1990)), it has been less useful for studying the employment relationship and the internal organization of large firms. In order to broaden the approach, Hart and Moore (2008) develop a theory based on the idea that an ex ante contract, negotiated under competitive conditions, shapes parties' entitlements regarding ex post outcomes. A party compares the ex post outcome to other outcomes permitted by the contract, and if he does not get what he feels entitled to he is aggrieved and shades on non-contractible aspects of performance. This yields a tradeoff between contractual rigidity and flexibility. A flexible contract is good in that parties can adjust to the state of nature, but bad in that there is a lot of shading. In contrast, a rigid contract is good in that there is little shading, but bad in that parties cannot adjust to the state of nature.

Although some of the assumptions underlying the model are broadly consistent with well-established behavioral concepts such as reference-dependent preferences (e.g., Kahnemann and Tversky (1979), Köszegi and Rabin (2006)), self-serving biases (e.g., Babcock and Loewenstein (1997)), and social preferences (e.g., Rabin (1993), Fehr and Schmidt (1999)), there is not yet any empirical evidence that directly supports the idea that contracts are reference points for trading relationships. It is the aim of this study to close this gap with evidence from a controlled laboratory experiment.

Our experiment is based on the payoff uncertainty model in Hart and Moore (2008). In this model, a buyer and a seller trade one unit of a standard good, but there is uncertainty ex ante about the buyer's value and the seller's cost. This uncertainty is resolved ex post, and there is symmetric information throughout. However, value and cost are not verifiable, and so state-contingent contracts cannot be written. The model assumes that ex post trade is voluntary. Given that value and cost are uncertain, there may be no single price such that both parties gain from trade at this price whenever value exceeds cost. Thus, to ensure trade, a range of possible prices may be required in the ex ante contract. However, under the assumptions of the model, this leads to ex post aggrievement and shading.

In Hart and Moore (2008), the first best can be achieved if either value or cost is certain, given that lump sum transfers are possible. In the experiment, we rule out lump-sum transfers. A consequence is that the first-best result does not apply, and we can simplify matters by assuming that only the seller's cost is uncertain. We also make the simplifying assumption that only the seller can shade.

In the experiment buyers and sellers meet and contract at date 0. Trade of zero or one unit of a good occurs at date 1. There are two states of nature: a good state, in which a seller's cost is low, and a bad state, in which cost is high. At date 0, the state of nature is uncertain. Contracting involves two steps. In the first step, a buyer determines what type of contract he wants: a flexible contract or a rigid contract. In a rigid contract, the price is fixed; a flexible contract defines a price range out of which the buyer will pick the price after the state of nature has been realized. In the second step, the contract is auctioned off to sellers. The auction not only determines which of the sellers gets the contract but also defines the contract terms (i.e., the fixed price in rigid contracts and the lower bound of the price range in flexible contracts). The buyer and seller then move to date 1. They are now in a situation of bilateral monopoly, and the state of nature is determined. Trade takes place only if the date 0 contract allows for a mutually beneficial outcome in the realized state. Competition ensures that the

price in the rigid contract is sufficiently low that trade is possible only in the good state, while trade is always possible in the flexible contract. In the flexible contract, the buyer picks a price above the seller's cost from the price range. In the rigid contract, only the fixed price is available. After trade occurs (if it is possible), the seller decides whether to shade. Shading takes the form of costly sabotage: the seller spends resources to lower her quality from normal to low. Shading has a small cost for the seller but greatly reduces the buyer's value.

Under the assumptions of the standard economic model (rationality, selfishness and sub-game perfection), the prediction for this experiment is straightforward. Since shading is costly, sellers should always provide normal quality, irrespective of the contract type and the price. Buyers should anticipate the sellers' behavior and therefore always choose the lowest price permitted by their contract. The competitive auction used to assign contracts to sellers should ensure that the fixed price in rigid contracts and the lower bound of the price range in flexible contracts is at the competitive level. Accordingly, both contract types yield the same outcome in the good state of nature, while only the flexible contract allows for trade in the bad state. This implies that buyers should always choose the more profitable flexible contract.

However, if the behavioral assumptions of Hart and Moore (2008) apply, the prediction is different. The assumption that contracts are reference points does not affect the prediction concerning the competitive auction outcomes. But if contracts define reference points, the contract type may affect the sellers' quality choice. Since rigid contracts pin down outcomes, sellers get what they expect and should not be aggrieved. Accordingly, shading should not occur in rigid contracts. In flexible contracts, in contrast, sellers may be aggrieved if they get a lower price than they had hoped for. This may trigger shading. In response, buyers may either increase the price in flexible contracts to avoid shading or accept the possibility of getting low quality. Either way, the reference dependent behavior of sellers has a negative impact on the buyers' profit in flexible contracts. Thus, if the willingness to engage in shading is strong enough, buyers may find switching to rigid contracts profitable.

What are the results of the experiment? They are largely in line with Hart and Moore (2008). The auction process indeed induces strong competition for contracts. Both the fixed price in rigid contracts and the lower bound of the price range in flexible contracts converge to the competitive level over time. However, despite the fact that, in principle, buyers have the possibility to pay the same prices in both types of contracts when the good state is realized, we observe that buyers pay significantly higher prices in flexible contracts. Moreover, depending on the price paid, there is considerable seller shading in flexible

contracts in the good state. In contrast, there is almost no shading in rigid contracts. Under the parameter values of the experiment, the rigid contract is more profitable than the flexible contract even though it precludes trade in the bad state. Furthermore, a substantial fraction of buyers choose the rigid contract.

It is worth noting that these results not only provide empirical support for the model of Hart and Moore but they also constitute new insights into the behavioral economics of fairness. To see this in more detail it is important to note that rigid contracts typically lead to very low earnings for the seller and a very uneven distribution of the gains from trade. Thus, by proposing a rigid contract, a buyer makes an unfair proposal so that one would expect the sellers to shade a lot under rigid contracts. In fact, theories of inequity aversion (Fehr and Schmidt 1999, Bolton and Ockenfels 2000) suggest that there should be considerable shading in the rigid contract since the surplus is very unevenly distributed. Likewise intention-based fairness theories (Rabin 1993, Charness and Rabin 2002, Dufwenberg and Kirchsteiger 2004, Falk and Fischbacher 2006) also suggest that there should be shading in the rigid contract since the choice of the rigid contract signals rather ungenerous intentions (the rigid contract lowers the seller's payoff in the good state and prevents trade in the bad state, and so it would be generous of the buyer to choose the flexible contract). However, despite the very uneven distribution of the gains from trade the sellers rarely shade in rigid contracts.

Our evidence becomes even more puzzling – when viewed through the lens of traditional theories of fairness – if we compare the low frequency of shading in rigid contracts with what happens under flexible contracts. In the latter we observe a lot of shading even though the sellers receive a higher share of the gains from trade than under rigid contracts. This pattern of shading across contract types makes perfect sense, however, if competitively determined contracts provide reference points which demand special normative status. We thus believe that our experiment reveals a new behavioral force: ex ante competition legitimizes the terms of the contract, and aggrievement occurs mainly about outcomes within the contract and not about the contract itself.

The paper is structured as follows: In section II, we describe the design of our experiment and provide details on procedures. Section III contains the behavioral predictions. We present our results in section IV and discuss them in section V. Section VI concludes.

II. Experimental Design

In this section we present our experimental design. Section II.A roughly summarizes the main features of the implemented market. We then specify the details of the experimental game in Section II.B, followed by a description of the experimental procedures in Section II.C.

II.A. Overview

We implement an experimental market where buyers and sellers meet to trade. Each transaction involves two dates. At date 0 the trading parties meet and conclude a contract. At date 1 they trade zero or one unit of a product. Traders do not yet know the state of nature at date 0. Accordingly, a flexible contract that allows adjustments to the realized state of nature may be desirable. The market setup implies that supply exceeds demand. Accordingly, sellers compete for contracts with buyers at date 0. The allocation of contracts takes place in two steps: First, each buyer determines whether he wants to conclude a flexible or a rigid contract. Then a competitive auction determines which seller gets the contract. All contracts are incomplete, because quality cannot be perfectly specified, i.e. a contract cannot prevent a seller from choosing low quality instead of normal quality. The uncertainty about the state of nature is resolved at date 1. However, while the state of nature is observable to the trading parties, it is not verifiable. This implies that state contingent contracts are infeasible. Trade occurs only if the contract terms are such that both parties can benefit from trading in the realized state of nature. Otherwise the parties walk away and realize an outside option. In case of trade the buyer refines the terms within the limits of the contract and the seller determines the quality.

II.B. Experimental Game and Parameters

There are 28 market participants in each experimental session, 14 in the role of buyers and 14 in the role of sellers. In each of the 15 periods of the experiment sellers and buyers interact in groups of two buyers and two sellers. To minimize the role of reputational considerations, these interaction groups are randomly reconstituted in every period.

In each period buyers and sellers have the possibility to trade a product. While every buyer can buy at most one unit of the product per period, each seller can sell up to two units. Since there is an equal number of buyers and sellers, this implies that the supply of the product is twice as large as the demand. Thus, sellers face competition for buyers. When a buyer purchases a unit of the product from a seller his payoff is given by his valuation for the product v minus the price p. The payoff of the seller is calculated as the difference between

the price p and the production cost c. While the buyer's valuation for the product depends only on the seller's ex-post quality choice q, the seller's production cost also depends on the realized state of nature s. There are two states of the world: a good state ($s = g$), in which the seller's production costs are low and a bad state ($s = b$), in which the production costs are high. The good state occurs with probability $w^g = 0.8$.

The payoffs of buyers and sellers can be summarized as follows:

Buyer's payoff: $\pi_B = v(q) - p$.

Seller's payoff: $\pi_S = p - c(q, s)$.

When trade takes place sellers can choose between two quality levels: normal quality ($q = q^n$) or low quality ($q = q^l$). The production costs for low quality are slightly higher than the production costs for normal quality: $c(q^l, s) > c(q^n, s)$. The idea is that it is most convenient for sellers if they simply do their job. They can, however, sabotage output (at the expense of a small cost) if they want to.[4] For each unit of the product which a seller cannot sell – either because he lost the contract to the other seller in his trading group at date 0 or because his contract does not allow for a mutually beneficial trade at date 1 – he realizes an outside option of $x_S = 10$. When a buyer is unable to trade a unit of the product at date 1, he also realizes an outside option of $x_B = 10$.

Table I summarizes the cost and value parameters of the experiment:

Insert Table I here

Each period of the experiment is structured as follows:

Date 0: Contracting

Step 1: Random formation of interaction groups

At the beginning of every period the interaction groups consisting of two buyers and two sellers are randomly determined. The rematching of participants at the beginning of every period makes sure that reputation effects cannot play a substantial role in our experiment.

Step 2: The buyer's contract choice

Before buyers' contracts are auctioned off to sellers, each buyer has to decide which contract type t he wants to offer in this period. It is important to note that the buyer can choose only the type of the contract, while the actual terms of the contract are defined by

[4] An alternative interpretation is that there is a black-boxed incentive technology for good quality (e.g., the seller is punished in case of shading).

the sellers in a competitive auction process later on. Specifically, the buyer can choose between two types of contracts: rigid contracts ($t = r$) and flexible contracts ($t = f$). A rigid contract fixes the price at date 0. The level of the fixed price p^r is endogenously determined in the subsequent contract auction. The auction is set up in such a way that the fixed price lies always in the following interval: $p^r \in [c(q^l,g) + x_s, m] = [35, 75]$, where m is the exogenously set maximum for the fixed price.[5] A flexible contract, in contrast, specifies a price range $[p^l, p^h]$ at date 0 out of which the buyer can choose the actual price at date 1. The upper bound of the price range is exogenously fixed and identical to the buyer's valuation of the product when the seller provides normal quality: $p^h = v(q^h) = 140$. The lower bound of the price range is endogenously determined in the subsequent contract auction. The interval of potential lower price bounds is identical to the one for fixed prices in rigid contracts: $p^l \in [35, 75]$.

Step 3: The sellers' contract auction

When both buyers in an interaction group have chosen their contract type, the two contracts are auctioned off to the sellers. The sequence of the auctions is randomly determined within each group. In case of a rigid contract, the auction directly determines the fixed price p^r. In case of a flexible contract, the auction determines the lower bound of the price range p^l. In both cases the auction starts off at 35 and then increases by one unit every half second. Each of the two sellers has a button that allows him to accept the contract at any time during the auction. Thus, the first seller who is willing to accept the displayed fixed price or the displayed lower price bound respectively gets the contract. The seller who loses the auction and does not get the contract realizes the outside option x_S.

Date 1: Trade

Step 4: Determination of the state of nature

After the contracts have been auctioned off to the sellers, a computerized random device determines the state of nature for each contract independently. Both sellers and buyers observe the realized state for their contracts and are informed whether a mutually beneficial transaction can take place or not. Trade can always take place when the buyer

[5] The minimum of 35 for the fixed price ensures that the seller cannot make losses relative to his outside option in the good state even if he provides low quality. This feature ensures that sellers do not refrain from choosing low quality, just because they want to avoid losses (loss aversion). The maximum of 75 for the fixed price makes sure that the price is always below the seller's cost in the bad state of nature. This guarantees that trade cannot occur if the bad state is realized. However, as we will see later on, in the experiment the upper bound was never binding.

has chosen a flexible contract, because the price range allows the buyer to choose prices that cover the seller's cost in both states of nature. In the case of a rigid contract, in contrast, trade occurs only in the good state. In the bad state the fixed price is always lower than the seller's cost, such that mutually beneficial transactions are not feasible. If trade does not occur, the buyer and the seller realize their outside options (x_B and x_S).

Step 5: The buyer's price choice

Once the state has been revealed, the buyer can choose the actual trading price. In a rigid contract the buyer does not have a choice, since the price has already been fixed at date 0 and cannot be changed. In a flexible contract, however, the buyer can determine his price. If the good state has been realized the buyer can choose any price $p \in [p^l, p^h]$. In the bad state the buyer has to make sure that the price is such that the seller cannot make losses, i.e., he must choose a price that satisfies $p \in [c(q^l,b) + x_S, v(q^h)] = [95, 140]$.[6]

Step 6: The seller's quality choice

Sellers observe the price choice of their buyer and then determine their quality. In both types of contracts the sellers have the choice between normal (q^n) and low (q^l) quality. Remember choosing low instead of normal quality increases the seller's cost by 5 units irrespective of the contract type and realized state of nature (see Table I).

Step 7: Profit calculations

After the quality choice of sellers all decisions have been made. Profits are calculated and displayed on subjects' screens.

Step 8: Market information for the buyers

Subsequent to viewing the profit screen buyers also get some aggregated information about the market outcome. Specifically, they are informed about profits of buyers in both contract types averaged over all past periods. Furthermore, they learn how many buyers have chosen the rigid contract and the flexible contract in the current period.[7]

The screen with the market information for buyers ends the period. After this a new period begins and the participants are randomly reassigned to a new interaction group.

[6] Again we do not allow prices to be such that the seller can make losses by choosing low quality, because we want to avoid the possibility that people refrain from shading due to loss aversion (see also Footnote 5).

[7] The aim of the provision of this information was to make learning easier for buyers. Since our setup allows for many possible constellations (two contract types, two states of nature, two quality levels, many prices), learning from individual experience is rather difficult.

II.C. Subjects, Payments and Procedures

All subjects were students of the University of Zurich or the Swiss Federal Institute of Technology Zurich (ETH). Economists and psychologists were excluded from the subject pool. We used the recruitment system ORSEE (Greiner 2004). Each subject participated in only one session. Subjects were randomly subdivided into two groups before the start of the experiment; some were assigned the role of buyers and others the role of sellers. The subjects' roles remained fixed for the whole session. All interactions were anonymous, i.e., the subjects did not know the personal identities of their trading partners.

To make sure that subjects fully understood the procedures and the payoff consequences of the available actions, each subject had to read a detailed set of instructions before the session started. Participants then had to answer several questions about the feasible actions and the payoff consequences of different actions. We started a session only after all subjects had correctly answered all questions. The exchange rate between experimental currency units ("points") and real money was 15 Points = 1 Swiss Franc (~US $ 0.83, in summer 2007).

In order to make the sellers familiar with the auction procedure we implemented two trial auctions – one with a rigid contract and one with a flexible contract – before we started the actual experiment. In the trial phase each seller had his own auction, i.e., they did not compete with another seller and no money could be earned.

The experiment was programmed and conducted with z-Tree (Fischbacher 2007). We had 28 subjects (14 buyers and 14 sellers) in four of our five sessions and – due to no-shows – 24 subjects (12 buyers and 12 sellers) in the remaining session. This yields a total number of 136 participants in the experiment. A session lasted approximately two hours and subjects earned on average 49 Swiss Francs (CHF 49 ~ US $ 41, in summer 2007).

III. Behavioral Predictions

In this section we derive the predictions for our experiment and discuss some design features.

III.A. Predictions under Pure Self-Interest

If we assume common knowledge of rationality and money-maximizing behavior, the predictions for this experiment are straightforward. Since shading on performance is costly, purely selfish sellers provide normal quality irrespective of the realized price in both types of contracts. Buyers anticipate sellers' behavior and choose the lowest price allowed by the

contract. In the contract auctions trade rivalry between sellers implies that the fixed price in rigid contracts, respectively the lower price bound in flexible contracts, ends up at the competitive level, i.e. $p^r = 35$ and $p^l = 35$.[8] Accordingly, when buyers choose their contract types they anticipate the following outcomes: in the good state of nature both contract types deliver the same outcome ($\pi_B = v(q^n) - p = 140 - 35 = 105$), but in the bad state of nature the flexible contract is more attractive, because it allows for trade ($\pi_B = v(q^n) - p = 140 - 95 = 45$), while the rigid contract leads to the realization of the outside-option ($\pi_B = x_B = 10$). This implies that buyers always choose the flexible contract. We summarize the prediction of the standard economic model as the

Standard Hypothesis:

a) Market forces imply that the fixed price in rigid contracts and the lower bound of the price range in flexible contracts end up at the competitive level, i.e., $p^r = p^l = 35$.
b) Sellers never choose low quality irrespective of the contract type and price level. Buyers always choose the lowest price available in flexible contracts.
c) Buyer's profits are higher in flexible contracts than in rigid contracts. Therefore, buyers prefer flexible contracts.

III.B. Predictions if Contracts are Reference Points

According to Hart and Moore (2008), an ex ante contract, negotiated under competitive conditions, shapes parties' entitlements regarding ex post outcomes. In Hart and Moore (2008), a party compares the ex post outcome to the best outcome permitted by the contract, and if he does not get what he feels entitled to he is aggrieved and shades on non-contractible aspects of performance. In the Appendix we extend Hart and Moore (2008) to allow for the case where parties may feel entitled to an outcome other than the best outcome. The model's predictions are broadly similar to Hart and Moore (2008). Rigid contracts pin down outcomes, sellers get what they expect, and so sellers are not aggrieved. Accordingly, shading should not occur in rigid contracts. However, in a flexible contract, the seller may be aggrieved and shade if he gets a lower price than he had hoped for. We show that the heterogeneity in seller entitlements implies that the frequency of shading is decreasing in price. Given this the buyer will either increase the price in flexible contracts to avoid shading or accept the possibility of getting low quality. Thus, although the flexible contract guarantees trade in both states, the

[8] Remember: Since $p = 35$ corresponds to $p = c(q^l,g) + x_S$ and the seller must offer at least $p = c(q^l,b) + x_S = 95$ in the bad state of nature, a seller can never be worse off if he accepts a contract than if he accepts his outside option.

reference dependent behavior of sellers has a negative impact on the buyers' profit. Hence flexible contracts may be less profitable than rigid contracts. Reference dependent behavior does not, however, change the auction outcome: rivalry between sellers still ensures that the lower bound of the price range in flexible contracts, and the fixed price in rigid contracts, is 35. (Recall that we do not allow the price to fall below 35.)

These considerations lead to the

Reference Point Hypothesis:

a) Market forces imply that the fixed price in rigid contracts and the lower bound of the price range in flexible contracts end up at the competitive level, i.e., $p^r = p^l = 35$.
b) In rigid contracts sellers never choose low quality irrespective of the price level. In flexible contracts sellers' quality provision is price dependent. Heterogeneity in seller entitlements implies that the frequency of shading is decreasing in the price. Given the price dependence of quality, buyers may not choose the lowest price available in flexible contracts.
c) Buyer's profits in flexible contracts are lower than predicted by the standard model. If the impact of the reference dependent preferences is strong, buyers may even make higher profits in rigid contracts than in flexible contracts.

III. C. Discussion of Design Features

It is important to emphasize that the aim of this paper in not to determine whether people succeed in choosing the optimal contract structure. We are interested rather in the more fundamental question of whether the underlying behavioral assumptions of the Hart-Moore model turn out to be of empirical relevance. To study this question in a clean and controlled way, we have intentionally abstracted from many theoretically relevant features of the model. In this section we discuss to what extent the simplifications and specific decision structures in our experimental setup affect our predictions.

While Hart and Moore (2008) assume that shading opportunities of buyers and seller are symmetric, we simplify matters by restricting shading to sellers. For the theoretical investigation of optimal contracts this simplification would have serious consequences, because it implies that a flexible contract can always achieve the first best if the seller is given the right to choose price.[9] In other words, the trade-off between contractual flexibility and

[9] In equilibrium buyers would choose the flexible contract, sellers would accept the competitive lower price bound at date 0, choose a price ex post equal to the buyers' valuation ($p = v(q^n)$), and provide normal quality ($q =$

rigidity would be destroyed.[10] However, since optimality of contracts is not the focus of this study, we avoid this problem by restricting the set of feasible contracts. Specifically, we exclude side payments and consider only structures in which the buyer decides about the contract type and chooses the ex-post price. Furthermore, since we aim at examining whether and how the contract type affects the sellers' performance choice, we maximize the contrast between contractual rigidity and flexibility by implementing extreme forms of contracts. In the rigid contract the price is entirely fixed. The flexible contract, in contrast, determines only the lower bound of the price range and allows for any reasonable price above this limit (i.e., the upper bound is exogenously fixed at the buyers' maximal willingness to pay: $p^h = v(q^n)$).[11] Studying these extreme cases gives us the best chance to illustrate the trade-off between rigidity and flexibility in contracts.

Another important simplification and limitation is that we have not allowed the parties to write informal state-contingent contracts. For example, the buyer and seller could agree that price will depend on the seller's realized cost, which is observed by both parties. Hart and Moore (2008) discuss informal contracts of this kind. They argue that such contracts may be problematic in situations where there is a little bit of asymmetric information and the parties exhibit self-serving biases. Under these conditions, each party may be able to convince himself that the state is favorable to him. This is likely to lead to aggrievement and shading, as in the flexible contracts studied in this paper. Obviously, a considerably more complicated experiment would be required to test the role of informal state-contingent contracts in the presence of asymmetric information and self-serving biases. We leave this for future work.

Hart and Moore (2008) assume that trading parties can choose the amount of shading in a continuous way and that there is no upper bound on shading. For the sake of simplicity we implement shading as a binary choice. While a more continuous action space would change details of our predictions, it would not affect our main hypothesis that price dependent shading should occur only in flexible contracts. In addition, Hart and Moore (2008) assume that trading parties are indifferent between perfunctory and consummate performance, i.e., shading is neither costly nor beneficial. However, they emphasize that assuming indifference

---

q^n). Redistribution of surplus from sellers to buyers (because of competition) could be achieved through lump sum payments conditional on winning the contract auction.

[10] Hart and Moore (2008, p. 34/35) discuss how the predictions of the model can be rescued even if buyers cannot engage in shading.

[11] Hart and Moore (2008) predict that buyers would prefer the flexible contract to be less extreme. Given the decision structure and the restriction on price setting in our environment, the optimal flexible contract would set the upper bound at $c(q^l, b) + x_s$. This would still allow for trade in the bad state of nature, but it would eliminate the room for aggrievement and shading in the bad state of nature.

is just a technically convenient way to capture the idea that the cost of providing low quality is not substantially higher or lower than the cost of providing normal quality. With regard to the aim of our paper, implementing strict indifference between shading and normal performance in the experiment would have been problematic. The reason is that indifference would not rule out equilibrium shading under standard economic assumptions in our setup. In order to make sure that shading cannot be explained if people are motivated by pure self-interest, we implemented costly shading. However, since the increase in the sellers' costs is low ($c(q^l, s) - v(q^n, s) = 5$) relative to the damage which shading imposes on the buyer ($v(q^n) - v(q^l) = 40$), our setup is still in line with the spirit of the model.[12]

It is obvious that the probabilities with which the two states of nature occur are decisive for the relative attractiveness of rigid and flexible contracts. Since we intend to study the impact of contract types on behavior, we need a sufficient number of observations for flexible and rigid contracts. The rigid contract is interesting only if the disadvantage due to the non-existence of trade in the bad state of nature is not too large. We therefore decided that the good state of nature should occur with a high probability ($w^g = 0.8$).[13]

IV. Results

In this section we present and discuss our results. The analysis of our data at the aggregate level in Section IV.A reveals that the outcomes of our experiment largely confirm the reference point hypothesis and contradict the predictions of the standard economic theory. We observe that there are significant differences in average prices and quality levels between rigid and flexible contracts in the good state of nature. In Section IV.B we demonstrate how these findings can be explained in light of the different price dependence of sellers' performance choices across the two contract types. Sections IV.C and IV.D illustrate that the predictions of the reference point theory are also relevant for individual behavior of sellers and buyer.

[12] Alternatively we could also have chosen to make the provision of low quality slightly less costly than the provision of normal quality. However, the case of costly shading probably leads to stronger effects. It seems more likely that aggrievement triggers costly shading than that the absence of aggrievement causes people to engage in costly voluntary cooperation. The reason is that aggrievement certainly causes a negative sentiment, while the absence of aggrievement may be completely neutral and may not imply the positive sentiment necessary to induce costly cooperation. Of course, this remains an empirical question that should be addressed in future work.

[13] Another way to make sure that we have a sufficient number of observations in both contract types would have been to assign contract types exogenously. However, this would have changed the spirit of the experiment in a fundamental way. From the perspective of the seller, it certainly makes a big difference whether the buyer himself chooses to limit his ability to adjust the price ex-post or whether this is imposed by the experimenter (see the discussion in Section V).

IV.A. Aggregate Findings

We find that our data are largely in line with all three parts of the reference point hypothesis stated at the end of Section III. Table II and Figure I summarize our main results. Table II presents averages of prices, quality choices, auction outcomes, profits and contract choices for rigid and flexible contracts in the good and bad state of nature. Figure I displays the development of prices and quality choices over time.

Insert Table II here

IV.A.1. Auction Outcomes in Rigid and Flexible Contracts

We start with the competition part of the reference point hypothesis. Figure I illustrates the power of competition in the auction phase of our experiment. The fixed price in rigid contracts and the lower price bound in flexible contracts converge to the competitive price of 35 over time. In the final period the auctions deliver an average fixed price of 35.7 and an average lower bound of 35.2. Because auction outcomes are somewhat higher in the early period of the experiment the overall averages of both the fixed price in rigid contracts and the lower bound of the price range in flexible contracts are slightly above the predicted level of 35. Both averages turn out to be about 40 (see Table II). As expected a non-parametric signed-rank test using session averages as observations confirms that the auction outcomes for rigid contracts and flexible contracts are not significantly different.[14]

Insert Figure I here

IV.A.2. Prices and Quality in Rigid and Flexible Contracts

The fact that auction outcomes do not differ across contract types implies that, in principle, the buyers would have the possibility to pay the same prices in both types of contracts when the good state of nature is realized. However, if the reference point hypothesis is correct, buyers in flexible contracts may have incentives to increase their prices above the lower price bound, because low prices may aggrieve sellers and lead to shading. This is in fact what we observe. In 75 percent of the flexible contracts in which the good state has been realized, buyers pay a price which is strictly above the lower bound of the price range determined in the auction. Although the lower price bound is only about 40 on average, the average price level is 51 (see Table II). This difference between the actual price paid by the buyer and the lower bound of the price range is very stable and does not disappear over time (see Figure I).

[14] The session averages for the fixed price in rigid contracts are: 37.3, 40.7, 40.5, 41.2 and 43.3. The session averages for the lower bound of the price range in flexible contracts are: 37.5, 41.6, 40.5, 38.7 and 43.4.

In rigid contracts, in contrast, the final price at date 1 is equal to the fixed price which has been determined in the auction at date 0. We have already shown above that these prices are around 40 on average and converge to the competitive level of 35 over time.[15] This implies that in the good state buyers pay on average substantially higher prices in flexible contracts than in rigid contracts. A non-parametric signed rank test confirms that the price difference between rigid and flexible contracts is statistically significant (p-value = 0.031 (one-sided)).[16]

Although prices are close to the competitive level, shading is almost absent in rigid contracts. Sellers provide normal quality in 94 percent of the cases in which the good state is realized. In flexible contracts, however, the higher prices are not always sufficient to prohibit sellers from shading. In the good state sellers provide normal quality in only 75 percent of the cases (see Table II). The difference in the frequency of shading between the two contract types is statistically significant (non-parametric signed rank test, p-value = 0.031 (one-sided)) and very stable over time (see Figure I).[17]

IV.A.3. Payoffs and Contract Choice

The differences in price and quality levels have important implications for payoffs of buyers and sellers in the good state of nature. Since prices are higher and quality is lower in flexible contracts, buyers earn, on average, considerably lower payoffs in flexible contracts (78.9) than in rigid contracts (96.8). The opposite is true for sellers. Although the higher frequency of shading increases sellers' costs in flexible contracts, the price difference is large enough to offset this. While the average payoff of sellers in flexible contracts is 29.8, their payoff in rigid contracts is 20.4 (see Table II). Both payoff differences are highly significant according to non-parametric signed rank tests (Sellers: p-value = 0.031 (one-sided), Buyers: p-value = 0.031 (one-sided)).

In the bad state of nature rigid contracts do not allow for trade. Accordingly, buyers and sellers realize their outside-option. In flexible contracts trade takes place and buyers must offer them at least a price of 95. We observe that buyers pay on average a price of 98.4 (see Table II). Thus, while average prices are substantially higher than the lower price bound in the good state, prices in the bad state are very close to the minimal price buyers can offer. In

[15] The "auction outcome" in Table II is the average of the fixed prices in all rigid contracts . The "price" in Table II and Figure I, in contrast, is the average of the fixed prices in all rigid contracts in which trade occurred. However, since the state of nature is randomly determined, this does not cause a systematic difference.

[16] The session averages for the price in rigid contracts are: 37.3, 40.8, 40.8, 41.0 and 43.4. The session averages for the price in flexible contracts are: 51.7, 49.7, 49.0, 50.2 and 54.0.

[17] In the good state the session level frequencies of high quality in rigid contracts are 89%, 97%, 95%, 91% and 96%. The corresponding numbers for flexible contracts are 78%, 76%, 79%, 67% and 75%.

response shading is slightly more frequent in the bad state than in the good state. However, a non-parametric signed rank test shows that this difference is not statistically significant.[18] We will later investigate whether the price setting strategies of buyers in flexible contracts reflect profit-maximizing behavior. Since outside options generate only a payoff of 10, sellers and buyers are better off with a flexible contract in the bad state of nature. Average payoffs are 29.7 for buyers and 16.9 for sellers, respectively.

We have established that buyers indeed face a trade-off between rigidity and flexibility. Given the stronger tendency for shading in flexible contracts, rigid contracts are more attractive in the good state of nature. However, since fixed prices prohibit trade when costs are high, having a flexible contract is of advantage in the bad state of nature. But which contract is more profitable in total? It turns out that overall the need to pay higher prices and the higher frequency of shading are strong enough to render flexible contracts less profitable for buyers than rigid contracts. While the average buyer payoff is 77.9 in rigid contracts, it is only 68.9 in flexible contracts. This difference is statistically significant (non-parametric signed rank, p-value = 0.031 (one-sided)).[19] Sellers, in contrast, are better off in flexible contracts. Average seller payoffs in rigid contracts are 18.1, compared to 27.9 in flexible contracts. Also this difference is statistically significant (non-parametric signed rank, p-value = 0.031 (one-sided)).[20] The finding that rigid contracts yield higher profits for buyers than flexible contracts is, of course, highly dependent on the choice of parameters. It is certainly easy to find other parameter constellations which yield the opposite results (e.g., higher probability for bad state, weaker impact of shading on buyer's value, etc.). However, our findings not only illustrates that a trade-off between contractual flexibility and rigidity exists, but also that there are parameters under which this trade-off has strong consequences for economic outcomes.

In contrast to the standard prediction buyers have chosen rigid contracts in 50 percent of the cases (see Table II). If we look at the development over time, we observe that the share of rigid contracts has an upward tendency. It starts off at 38 percent in period 1 and ends up at

[18] In the bad state of nature the session level frequencies of high quality in flexible contracts are 70%, 46%, 83%, 58% and 83%.
[19] The session averages for buyer payoffs in rigid contracts are: 80.0, 80.1, 78.8, 75.0 and 75.3. The session averages for buyer payoffs in flexible contracts are: 71.0, 69.9, 71.4, 65.3 and 67.2.
[20] The session averages for buyer payoffs in rigid contracts are: 15.4, 18.4, 18.3, 18.1 and 20.1. The session averages for buyer payoffs in flexible contracts are: 28.1, 25.8, 25.7, 25.8 and 29.8.

55 percent in period 15. An OLS regression of the fraction of rigid contracts on periods indicates that this positive time trend is statistically significant.[21]

IV.B. Sellers' Quality Choice: Contract-Dependent Price-Quality Relationships

The discussion of Table II and Figure I has shown that in the aggregate our findings are supportive for the reference point hypothesis. Next, we analyze whether the underlying behavioral patterns are also in line with the assumption that contracts provide reference points for trading relationships. We start with sellers' performance choices. Figure II displays the relative frequency of normal quality conditional on the price paid by the buyer for each contract type and both states of nature. In addition, the figure also shows the relative frequency with which each price level is chosen by buyers. Notice that prices on the horizontal axis are rounded to the nearest multiple of ten. The figure provides strong support for the seller behavior predicted by the reference point hypothesis. In rigid contracts sellers almost always choose normal quality even if prices are very close to the competitive level. There is no noteworthy correlation between prices and the frequency of normal quality.[22] For flexible contracts, in contrast, the figure suggests a strong positive correlation between prices and the willingness to provide normal quality in both states of the world. If prices are close to the competitive level in the good state of nature, normal quality is chosen in less than 60 percent of the contracts. The frequency of normal quality is clearly increasing in price, but in order to reach the same average quality as observed in rigid contracts, buyers must raise their price to a level of at least 75. In the bad state of nature prices close to the lowest possible level also trigger a lot of shading. At prices between 95 and 104 sellers provide normal quality in less than 70 percent of the contracts. Also in the bad state substantial price increases are necessary to reach a high quality level on average.

Insert Figure II here

We provide statistical backup for our observations on price dependence of quality with a regression analysis in Table III. In the first column of the table we investigate the good state of nature. We regress an indicator variable for choosing normal quality on price increments, an indicator variable for flexible contracts, and the interaction term of the two. We define price increments as the difference between the actual price and the competitive price of 35.

[21] The regression uses one observation per period and session. The dependent variable is the fraction of rigid contracts, the explanatory variable is period. The estimated coefficients are as follows: Constant = 0.42, p-value < 0.001; Period = 0.010, p-value = 0.002 (p-values are based on robust standard errors).

[22] The low frequency of normal quality when prices are around 60 should be ignored, because it is based on very small number of outliers (see price distribution in Figure II).

Using price increments instead of prices allows us to interpret the constant as the frequency with which sellers provide high quality when buyers offers the competitive price of 35 in rigid contracts. The constant of 0.94 therefore reflects that prices close to the competitive level do not trigger much shading in rigid contracts. Furthermore, the coefficient of price increments is close to zero and insignificant indicating that sellers' quality choices in rigid contracts do not depend on prices in a statistically significant way. The situation is very different in flexible contracts. The significantly negative coefficient of the dummy for flexible contracts shows that if prices are at the competitive level sellers are much more likely to choose low quality in flexible contracts than in rigid contracts (–0.34, p-value < 0.01). The regression also confirms the statistical significance of the positive impact of higher price increments on sellers' quality choices in flexible contracts (F-Test: price increment + price incr. x flex. Contr. = 0, p-value < 0.01).[23] In column (2) we show that a probit estimation (marginal effects reported) using the same set of variables yields results similar to the ones of the linear probability model used in column (1). Column (3) investigates the bad state of nature. We regress the indicator variable for choosing normal quality on price increments (now defined as the difference between price and the lowest possible price of 95). The constant indicates that the frequency of normal quality is only 0.66 when buyers pay the lowest possible price in the bad state of nature. In addition, the significant coefficient confirms that there is also a significant impact of price increments on quality in the bad state of nature. Column (4) documents that a probit estimation yields similar results (marginal effects reported).

Insert Table III here

IV.C. Individual Seller Behavior: Quality Choice Within and Across Contracts

The analysis of sellers' behavior in Figure II and Table III is based on pooled data from all sellers in the experiment. However, the reference point hypothesis relies on behavioral assumptions about preferences and makes specific predictions regarding individual behavior of sellers in rigid and flexible contracts. Since contract assignment is endogenous in the experiment, our analysis hitherto does not provide evidence that our aggregate findings are the consequence of different behavior of the same sellers in different types of contracts. It could also be that the aggregate effects are the consequence of self-selection of distinct seller

[23] Since the lower bound of the price range is not always equal to the competitive price, one might suspect that the realized lower bound could be relevant for the seller's quality choice. For example, the seller could evaluate the generosity of the price paid by the buyer relative to the lower bound of the available price range. However, a regression of quality on price increments and the lower bound of the price range reveals that this does not seem to be the case. The coefficient for the lower bound of the price range is close to zero and not significant.

groups into different contract types. For example, the result that shading is more frequent in flexible contracts than in rigid contracts would also be observed if those sellers who self-select into flexible contracts are systematically more likely to provide low quality than those sellers who self-select into rigid contracts. In the following we dig deeper and examine whether sellers accept both types of contracts, and if so, how their behavior differs across contract types.

In Figure III we show the distribution of rigid and flexible contracts over individual sellers in the experiment. We observe that most sellers do not self-select into a specific type of contract, i.e., most sellers conclude several rigid as well as several flexible contracts. Specifically, the figure reveals that every seller has experienced each contract type at least once and for most sellers there are multiple observations for each contract type (84 percent of sellers have experienced at least four contracts of each type). Furthermore, even if we consider only contracts in which trade actually occurred, we still have at least one observation for each seller and contract type. This implies that each seller has made at least one quality decision in each type of contract, such that we can compare sellers' performance choices across contract types.

Insert Figure III here

In Table IV we analyze individual behavior in detail. According to the reference point hypothesis sellers may shade on performance in flexible contracts, but never in rigid contracts. We find that 51 of the 68 sellers in our experiment exhibit a behavioral pattern which is consistent with this prediction. 27 of these 51 sellers do not provide low quality in either contract type (first column), while the other 24 sellers provide low quality in some of their flexible contracts (second column). Notice: while the behavior of sellers who do not shade on performance at all can also be explained by standard economic theory (see the standard prediction in section III.A), this behavior does not contradict the reference point hypothesis. If a seller happens to receive offers above his threshold price whenever he concludes a flexible contract, it is plausible that he never shades on performance. Since sellers do not indicate their threshold price in our experiment, we cannot compare the threshold prices of sellers across differently behaving groups. However, Table IV shows that sellers who never provide low quality have concluded a lower number of flexible contracts and receive, on average, higher price offers in these contracts (especially in the good state of nature). These two factors make it less likely that a seller with a given intensity of reference dependence engages in shading.

The remaining 17 sellers in our experiment exhibit behavior which is not consistent with the reference point hypothesis, i.e., they provide at least once low quality in a rigid contract (see third and fourth column in Table IV). A closer look reveals that 7 of these sellers show behavioral patterns which are "almost in line" with the prediction of the reference point hypothesis: They provide exactly once low quality in a rigid contract and they shade more often in flexible contracts than in rigid contracts, both in absolute and relative terms. Only 10 of our 68 sellers make decisions which are clearly not in line with the reference point hypothesis.

In addition, the reference point hypothesis also suggest a positive (or zero) correlation of prices and quality in flexible contracts for each individual seller. However, the limited number of observations per individual combined with the fact that average prices (and therewith possibly also threshold prices) change over time make this a rather tough test. Nevertheless, for the good state of nature we observe positive correlations of prices and quality for 18 of the 24 subjects who exclusively shade in flexible contracts (column 2 of Table IV). In 7 of these 18 cases the correlation is statistically significant at the 10% level. The remaining 6 sellers have either never provided low quality in the good state (2 cases) or the correlation is insignificantly negative. If we do the same analysis for the 7 sellers whose quality choices are "almost in line" with the reference point hypothesis (column 3 of Table IV) we find a positive correlation of prices and quality for each individual. In 5 cases the correlation is significant at the 10% level.[24]

Insert Table IV here

IV.D. Individual Buyer Behavior: Price Setting and Contract Choice

Our analysis has shown that seller behavior is by and large in line with the reference point hypothesis. It is interesting to examine how buyers react to the behavioral patterns of sellers when they choose a contract type and set prices in flexible contracts. From Table II we know that overall buyers choose the rigid contract in 50 percent of the cases. Although we have shown that there is a positive time trend (see the regression results in Footnote 21), it may appear rather astonishing that the significantly higher average payoffs have not motivated more buyers to choose the rigid contract. However, since the higher average payoff in rigid contracts comes with a higher variance (see the payoffs per state in Table II) buyers' risk aversion (or loss aversion) may play an important role for this finding. The buyers may

[24] Given the even smaller number of observations, a similar analysis is not feasible for the bad state of nature.

perceive the possibility to end-up with the no-trade outcome of rigid contracts as too risky. Our post-experimental questionnaires confirm this presumption. About 50 percent of buyers indicate that they didn't want to take the chance of getting only the outside option of 10 in case of the realization of the bad state and therefore went for flexible contracts.

A second important issue regarding the buyers' behavior is their response to the price dependence of quality in flexible contracts. Are buyers able to figure out the profit-maximizing price or do they lose money in flexible contracts, because they fail to adopt the price properly? Figure IV displays average buyer profits conditional on the realized price for both contract types and both states of nature. Notice that prices are rounded to the nearest multiple of ten. In addition, the figure also displays the relative frequency with which each price has been realized. In flexible contracts in which the good state has been realized, it makes apparently most sense to set a price between 35 and 54. These prices yield average payoffs between 83 and 84 to buyers. These most profitable prices are chosen in 68 percent of the contracts. Buyers who set prices between 55 and 64 still get an average payoff of 77. This happened in 16 percent of the contracts. Increasing prices above 65 has strongly negative consequences for average payoffs. However, buyers picked such high prices in only 16 percent of the contracts, implying that buyers are rather successful in finding the most profitable price range. The same is true for flexible contracts in which the bad state of nature has been realized. In this case Figure IV shows that the most profitable prices are between 95 and 104. These prices result in average payoffs of about 30. Buyers choose prices in this range in 90 percent of the cases. In another 7 percent of contracts buyers pick prices in the range from 105 to 115. These prices yield only slightly lower average payoffs of 28. Higher prices, which yield much lower payoffs, are chosen only in 3 percent of the contracts. In rigid contracts buyers cannot influence the price. Since quality is not price dependent in rigid contracts, average profits of buyers are highest when prices are low.

Insert Figure IV here

While we find that the majority of buyers choose prices close to the profit-maximizing level, there is still some room for improvement. However, even if buyers had chosen optimal prices in every flexible contract average payoffs would still be below average payoffs of rigid contracts: $0.8 \times 84 + 0.2 \times 30 = 73.2 < 77.9$ (see Table II). This illustrates that in our setup the impact of sellers' reference dependence is strong enough to change the optimal structure of the contract relative to the standard economic model.

V. Discussion

Part of our interpretation of the data relies on the observation that many people seem to be willing to punish their trading partners for inappropriate behavior (in our case the payment of low prices in flexible contracts), even if this is costly for them and yields no material gain. Since such behavior has been extensively studied in theoretical and empirical work on social preferences, it is important to understand how our study relates to this literature.

Theories of social preferences assume that people are not solely motivated by their material self-interest but also take social considerations, especially fairness concerns, into account. There are two main classes of fairness theories: theories based on inequity aversion and reciprocity models. Models of inequity aversion (Fehr and Schmidt 1999, Bolton and Ockenfels 2000) suppose that people dislike inequitable outcomes and are therefore willing to forgo material payoff in order to prevent these outcomes from occurring. Despite their simplicity models of inequity aversion can explain many experimental results which seem puzzling from the perspective of the standard self-interest model (see e.g., Fehr and Schmidt 2002). Most important for our purpose is the evidence from the ultimatum game. In the ultimatum game the first mover (proposer) has to decide how to divide a certain amount of money between himself and the second mover (responder). The responder can either accept or reject the suggested division of pie. In case of acceptance each party receives the share suggested by the proposer, in case of rejection both parties get zero. The typical result in this experiment is that proposers offer a considerable share of the pie because they rightly fear that responders reject low amounts (see, e.g., Güth et al. 1982, Camerer and Thaler 1995). These findings are relevant for our study because they are similar to what we observe in flexible contracts, where buyers have to offer a high price in order to deter sellers from shading on performance. Since paying low prices generates unequal payoffs of buyers and sellers in our setup, inequity aversion implies that sellers should shade on performance when buyers pay low prices, because the provision of low quality reduces the payoff inequality. However, while inequity aversion is in line with the rejection of low offers in ultimatum games and the price dependent shading in flexible contracts, it cannot explain why sellers almost never provide low quality in rigid contracts. The reason is that in models of inequity aversion the allocation of final payoffs alone determines the perceived unfairness of an outcome. Accordingly, these models predict the same price dependency of quality in both types of

contracts. As prices are close to the competitive level in rigid contracts, these models predict a high frequency of shading.[25]

The more sophisticated reciprocity models take into account that the same outcome may trigger different fairness perceptions depending on how the outcome came about. Specifically, these models assume that the evaluation of an outcome depends not only on the realized allocation of payoffs but also on people's belief about the intentions of their trading partner (Rabin 1993, Dufwenberg and Kirchsteiger 2004, Charness and Rabin 2004, Falk and Fischbacher 2006). The intuition is that people perceive an unfavorable outcome as more unfair if another player intentionally chooses the outcome than if the payoff distribution is beyond the other player's control. The main difference between the two contract types in our experiment is that rigid contracts restrict the buyer's choice set to one predetermined outcome which is usually not very favorable to the seller, while flexible contracts enable the buyer to choose between many possible outcomes, including ones which are very attractive to the seller. If we translate this into the framework of the ultimatum game, the relevant comparison would be between one setup where the proposer is forced to offer a small fraction of the pie and a setup where the proposer can choose between offering small and large fractions of the pie. Falk et al. (2003) implement two treatments in an ultimatum game experiment which make such a comparison possible. In one of their treatments the proposer, who is endowed with 10 points, can choose between the equal split (5 points to each player) and an unequal allocation in which the proposer gets 8 points while the responder is left with 2 points. In another treatment the proposer has no choice but to offer the unequal allocation (8 for himself and 2 for the responder). The evidence shows that the rejection rate for the unequal allocation strongly depends on whether the equal split is available or not. When the proposer could have chosen the equal split, the unequal offer is rejected in roughly 44 percent of the cases. When the unequal allocation was the only alternative available to the proposer, the rejection frequency was only 18 percent.[26] While these findings are not in line with inequity aversion, they can be explained by reciprocity. If the proposer chooses the unequal allocation when the equal split would have been available, it is clear that the proposer intentionally created an

[25] Using the models of inequity aversion to derive full-fledged predictions for our experiment is not trivial. One problem is that the models do not specify precisely how the reference group relative to which individuals evaluate their payoffs is determined (e.g., sellers could compare their payoff to the (expected) payoff of both buyers and the other seller in their group or only to the buyer and the other seller involved in the trade or even only to their direct trading partner). However, while different assumptions with respect to the reference group may influence the predicted behavior in auctions and contracts, they do not affect the prediction that the price-quality relation should be independent of the contract type.

[26] Charness and Rabin (2002) report evidence from similar games (see Berk23and Berk27, page 844). Their results show the same pattern, although both rejection rates are much smaller (9 percent if the equal split was available and 0% otherwise).

"unfair" outcome. As this is perceived as an unkind action, it is likely to trigger an unkind response by a reciprocally motivated responder. If the unequal allocation is realized when the proposer did not have a choice, responders may still not like it, but they do not blame the proposer and are therefore less likely to retaliate.

However, there is a fundamental difference between the experiment of Falk et al. (2003) and our experiment. In the ultimatum games of Falk et al. (2003) the proposer's strategy space is chosen by the experimenter, i.e., the proposers is assigned to either a game with only one option or a game with two options. In our experiment, in contrast, the contract types are endogenously determined within the experiment, i.e., the buyers themselves choose whether they want to interact within a rigid or a flexible contract. This difference is important, because it is likely to change how outcomes are attributed to intentions. If the proposers in Falk et al. (2003) had had the choice between the two games, reciprocity theories would no longer predict a difference in rejection rates for the unequal allocation across games. The reason is that choosing the game with the restricted choice set would be perceived as an intentionally unfair action which calls for punishment. Similarly, in the context of our experiment intention based reciprocity models would predict that sellers punish buyers who choose the rigid contract. Under the assumption that the buyer anticipates the competitive outcomes of the contract auctions, choosing the rigid contract means that the buyer intentionally chooses a low payoff for the seller in both states of the world. In the good state trade will take place at the competitive price and in the bad state the seller is forced to take his unattractive outside option. Thus, a generous buyer would rather choose the flexible contract and pay a high price in both states of the world.

A crucial difference between the fairness models discussed above and Hart and Moore (2008) is that the latter explicitly assume that people perceive outcomes of a competitive bargaining process as acceptable or fair. The idea is that competition adds an objective dimension to the bargaining process which provides the outcome with a certain justification. This assumption is important in our experiment. In contrast to the experiments in Falk et al. (2003) where the payoff structure of the ultimatum game is exogenously given, the contract terms in our experiment are endogenously determined in competitive auctions. The low frequency of shading in rigid contracts suggests that sellers do indeed not blame the buyer for the unequal outcome in a rigid contract, but rather view it as the natural and justified outcome of a competitive market. This implies that buyers can circumvent the punishment for unequal

outcomes by delegating the determination of the outcome to the forces of a competitive market.[27]

The result that sellers earn higher average payoffs in flexible contracts than in rigid contracts is to a certain extent an artifact of our simplified design. In the more general theoretical setting of Hart and Moore (2008) the availability of ex-ante side payments would imply that the differences in seller payoffs across contracts should disappear. However, the fact that sellers shade significantly less in rigid contracts although they get substantially lower profits than in flexible contracts makes our results only stronger. The reason is that people's fairness consideration work against the reference point hypothesis. Reciprocity suggests that sellers may want to punish buyers for choosing the contract which gives them lower payoffs. Since the introduction of side payments would equalize sellers' payoffs across contracts, such fairness concerns would become less relevant and there would probably be even less shading in rigid contracts.

VI. Conclusions

In this paper we provide empirical support for the behavioral assumptions underlying the theory of Hart and Moore (2008). Our experimental evidence is in line with the idea that competitively determined contract terms constitute a reference point for a trading relationship. When buyers implement flexible contracts, which allow for many different outcomes, sellers seem to be disappointed if the buyer chooses an outcome which is not attractive to them. In response to an unfavorable outcome sellers are willing to engage in costly shading activities which reduce buyers' payoffs. However, when the buyers implement a contract with very rigid terms such that outcomes are pinned down from the outset and sellers know exactly what to expect, the same unfavorable outcomes do not trigger shading. Given uncertainty about the state of nature, these behavioral regularities imply a trade-off between contractual rigidity and flexibility. While flexible contracts are desirable because they allow the buyer to adjust the contract to the state of nature, rigid contracts have the advantage that they avoid inefficient shading activities.

[27] Recent experimental work shows that people can also avoid punishment for unequal outcomes by shifting the relevant decision to another person (see Bartling and Fischbacher 2008 and Coffman 2008). In Coffman (2008) player A has two options. He can either play a dictator game with player C or he can sell the dictator game to player B, who then plays the game with player C. If player A decides to sell the game, the price is determined in a competitive double auction. In both cases player C observes the outcome and has the possibility to attribute punishment points to player A. It turns out that player A is less punished for the same final outcome if he sells the game to player B than if he plays the dictator game himself. In future work it would be interesting to look more closely at how these findings are related to our work.

There are several ways in which our work could be extended. First, it would be interesting to know how important the assumption of ex ante perfect competition is. Suppose that there is a single seller from the start and the buyer picks the price at date 0 in the fixed price contract. At date 1, the seller may well regard the price as unfair since it was not negotiated competitively, i.e., the contract no longer serves as a reference point. One might therefore now expect shading in the fixed price contract. Preliminary results of an experiment that we have carried out suggest that this is indeed the case (see Fehr et al. (2008)). Second, we have tested a simplified version of Hart-Moore (2008) in which, although contract choice is endogenous, the structure of contracts and the degree of flexibility are predetermined. In future experiments it would be desirable to investigate truly endogenous contract structures, including the possibility of lump sum transfers.

Finally, we have ignored renegotiation. Suppose that a fixed price contract is written at date 0 and the bad state occurs. Assume that the parties have the chance to renegotiate to enjoy the gains from trade. Will they do so? Will such a mutually beneficial renegotiation create bad feeling, given that the price was already fixed, and hence lead to shading? Does the possibility of renegotiation open the door to hold-up; that is, will the seller try to force a renegotiation in order to grab more of the surplus, even if trade would occur in the absence of renegotiation? Experiments can clarify the answers to these important questions, and thereby provide new understanding of the foundations of long-term trading agreements.

Appendix

In this Appendix we formally derive the reference point hypothesis described in section III.B. In order to formalize the role of aggrievement we define the reference price $p^R(t, s)$ as the price the seller feels entitled to, given the contract type t and the realized state of nature s. While Hart and Moore (2008) assume that the seller always feels entitled to the best outcome admitted by the contract, we make a more general and less extreme assumption. We suppose that the seller's reference price p^R can be any price permitted by the terms of the contract. In addition, we also allow for the possibility that the reference price depends on the realized state. In a flexible contract the reference price therefore satisfies $p^R(f, s) \in [p^l, p^u]$. In a rigid contract, in contrast, the reference price of the seller must be equal to the fixed price: $p^R(r, s) = p^r$. Loosely following the formulation in Hart and Moore (2008) we assume that the seller's utility can be written as follows:

$u_S = \pi_S - \theta \max[(p^R(t, s) - p), 0]\, I(q)$,

where $\theta \geq 0$ and $I(q)$ is an indicator function, which is unity if $q = q^n$ and zero otherwise. The second term captures the psychic costs of aggrievement, which become relevant if the realized price p is smaller than the seller's reference price p^R. The parameter θ measures the intensity of the seller's aggrievement if he feels shortchanged. The indicator function $I(q)$ implies that a seller can completely offset his aggrievement if he shades on performance and thereby hurts the buyer by lowering his valuation for the delivered product.[28]

Rigid contracts never lead to aggrievement. The buyer has no other possibility than to pay the fixed price which has been determined in the competitive auction. Since the fixed price is also the seller's reference price, the seller gets exactly what he feels entitled to. Thus, whenever the good state is realized, there is trade and the seller delivers the normal quality. If the bad state is realized, trade does not take place and the buyer realizes his outside option. Accordingly the buyer's expected payoff from choosing the rigid contract can be expressed as follows:

$$\mathrm{E}\pi_B^r = w^g[\mathrm{v}(q^n) - p^r] + (1 - w^g)x^B.$$

In flexible contracts, in contrast, realized prices can be different from reference prices and aggrievement may play a role. This implies that the seller's performance is no longer independent of the price. Specifically, if the buyer offers a low price, the seller may be willing to bear the cost of shading in order to avoid the psychic cost of aggrievement. Formally, we

[28] The assumption that shading completely offsets aggrievement may seem strong. In the original model of Hart and Moore (2008) the reference dependent disutility of traders consists of the difference between the money equivalent of aggrievement and the monetary loss which shading imposes on the other party. Since shading is continuous and unbounded, traders can always avoid this disutility. As long as the reference dependent part of sellers' utility function is not too important, our formulation is in line with the original model. Specifically, it needs to be true that $\theta(p^R(t, s) - p) \leq v(q^n) - v(q^l) = 40$. This is not extremely restrictive. For example, even if the reference dependent part of utility and the monetary part are equally important (i.e., $\theta = 1$), the seller can still have a reference price of 75 in the good state and 135 in the bad state. An alternative way to formalize the impact of shading on aggrievement would be to write the intensity of aggrievement as a function of quality: $\theta(q)$, where $\theta(q^n) > \theta(q^l)$. If $\theta(q^l) > 0$, this implies that shading only partly offsets aggrievement. While this formulation leads to identical results concerning the quality choice in flexible contracts (the threshold price would be defined as: $p^T = p^R - [c(q^l, s) - c(q^n, s) / \theta(q^n) - \theta(q^l)]$), it changes the prediction for the acceptance choices of sellers in auctions of flexible contracts. As long as shading completely offsets aggrievement, sellers are always willing to accept the contract at any available price in the auction (see also Footnote 8). This is no longer true if shading only partly offsets aggrievement. Since sellers cannot completely avoid suffering from aggrievement, it may now be the case that some sellers prefer to realize the outside option if they anticipate that the buyer will set a low price in a flexible contract. This implies that the lower price bound of flexible contracts would no longer converge to the competitive level, but would depend on the intensity of the reference dependent preferences of the two sellers involved in the auction. As a consequence the auction outcome would provide the buyer with information about the preferences of the winning seller. This information would influence the optimal price in both states of nature. Since sellers anticipate that buyers use information conveyed by the auction outcome to adjust their price, some types of sellers may have strategic incentives to hide their true preferences and signaling may become an issue. In section IV we illustrate that both the fixed price in rigid contracts and the lower price bound in flexible contracts converge to the competitive level and there is no statistically significant difference between the auction outcomes in the two contract types. Therefore we do not further pursue the implications of this alternative model.

can define a threshold price p^T which the buyer needs to pay in order to motivate the seller to provide normal quality in a flexible contract:

$$p^T(s) = p^R(f, s) - [c(q^l, s) - c(q^n, s) / \theta].$$

Sellers may be heterogeneous with regard to both the reference prices p^R and the intensity of aggrievement θ. This implies that different sellers may have different threshold prices p^T. Let $F_s(\cdot)$ be the distribution function of threshold prices in the state s. After observing the realization of the state at date 1 a buyer who has concluded a flexible contract chooses his price as follows:

$$p^f_s = \arg\max v(q^n)F_s(p) + v(q^l)[1 - F_s(p)] - p.$$

This shows that in flexible contracts the optimal price in each state of nature depends on the characteristics of the distribution function of threshold prices in this state. For example, if many sellers have rather low reference prices and moderate price increases substantially lower the frequency of shading, buyers have strong incentives to increase the price above the lower price bound. On the other hand, if most reference prices are high and only large price increases help to prevent shading, the buyers may prefer to pay a low price and accept the consequences of a high frequency of shading. In any case, however, the fact that the sellers' quality choices depend on the price offered by the buyer has a negative impact on buyers' profits in flexible contracts relative to the standard prediction. The buyer's expected payoff if he chooses a flexible contract is:

$$E\pi_B^f = w^g[v(q^n)F_s(p^f_g) + v(q^l)(1-F_s(p^f_g)) - p^f_g] + (1 - w^g)[v(q^n)F_s(p^f_b) + v(q^l)(1-F_s(p^f_b)) - p^f_b].$$

This establishes the trade-off between contractual flexibility and rigidity described in the reference point hypothesis in section III.B: i) Our assumption that sellers can completely offset aggrievement by shading implies that they are willing to accept a contract at any available price (see Footnote 28 for details). Accordingly, the competitive auctions will drive the fixed price in rigid contracts and the lower bound of the price range in flexible contracts down to the competitive level ($p^r = p^l$ =35). ii) Rigid contracts have the advantage that buyers can pay low prices and do not suffer from shading in the good state. The downside of rigid contracts is that there is no trade in the bad state. iii) Flexible contracts guarantee trade in both states of nature. However, buyers have to increase the price above the competitive level and/or bear the consequences of shading in both states of nature. This reduces buyers' payoffs.

References:

Babcock, L., and Loewenstein, G., (1997): "Explaining Bargaining Impasse: The Role of Self-Serving Biases", Journal of Economic Perspectives 11(1), 109-126.

Bartling, B., and Fischbacher, U. (2008): "Shifting the Blame: On Delegation and Responsibility", mimeo, University of Zurich.

Bolton, G., and Ockenfels A. (2000): "ERC - A Theory of Equity, Reciprocity and Competition", American Economic Review, 90, 166-193.

Camerer, C., and Thaler, R. H. (1995): "Anomalies: Ultimatums, dictators and manners", The Journal of Economic Perspectives 9(2), 209–219.

Charness, G. and Rabin, M. (2002): "Understanding Social Preferences with Simple Tests", Quarterly Journal of Economics 117(3), 817–869.

Coffman, L. (2008): "Ethics and Strategic Intermediation", mimeo, Harvard Business School.

Dufwenberg, M., and Kirchsteiger, G. (2004): "A Theory of Sequential Reciprocity", Games and Economic Behavior 47, 268-98.

Falk, A., Fehr, E., and Fischbacher, U. (2003): "On the Nature of Fair Behavior", Economic Inquiry 41(1), 20-26.

Falk, A., and Fischbacher, U. (2006): "A Theory of Reciprocity", Games and Economic Behavior 54(2), 293-315.

Fehr, E., Hart, O. and Zehnder, C. (2008): "Contracts, Reference Points, and Competition", Mimeo, University of Lausanne.

Fehr, E., and Schmidt, K. (1999): "A Theory of Fairness, Competition and Cooperation", Quarterly Journal of Economics 114, 817-68.

Fehr, E., and Schmidt, K. (2002): "Theories of Fairness and Reciprocity - Evidence and Economic Applications", in Dewatripont, L. Hansen and St. Turnovsky (Eds.), Advances in Economics and Econometrics - 8th World Congress, Econometric Society Monographs, Cambridge, Cambridge University Press.

Fischbacher U. (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments", Experimental Economics 10(2), 171-178.

Greiner, B. (2004): "An Online Recruiting System for Economic Experiments", in Kurt Kremer, Volker Macho (Eds.): Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Göttingen: Ges. für Wiss. Datenverarbeitung, 79-93.

Grossman, S., and Hart, O. (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration", Journal of Political Economy 94, 691-719.

Güth, W., Schmittberger, R., and Schwarze, B. (1982): "An experimental analysis of ultimatum bargaining", Journal of Economic Behavior and Organization 3(4), 367–388.

Hart, O., and Moore J. (1990): "Property Rights and the Nature of the Firm", Journal of Political Economy 98, 1119-58.

Hart, O., and Moore, J. (2008): "Contracts as Reference Points", Quarterly Journal of Economics 123(1), 1-48.

Kahneman, D., and Tversky, A. (1979): "Prospect theory: An analysis of decisions under risk", Econometrica 47(2), 263–291.

Köszegi, B., and Rabin, M. (2006): “A model of reference-dependent preferences”, Quarterly Journal of Economics 121(4), 1133–1166.

Rabin, M. (1993): “Incorporating Fairness into Game Theory and Economics”, American Economic Review 83, 1281-1302.

Table I: Experimental Parameters

State of nature	Good [Prob(s = g) = 0.8]		Bad [Prob(s = b) = 0.2]	
Seller's quality	normal ($q = q^n$)	low ($q = q^l$)	normal ($q = q^n$)	low ($q = q^l$)
Seller's costs	20	25	80	85
Buyer's valuation	140	100	140	100

Notes: The table summarizes the main parameters of the experiment. Buyers' valuation for the product and sellers' production costs are displayed for both states of nature and both quality levels available to the seller.

Table II: Summary of Outcomes in Rigid and Flexible Contracts

Contract Type	Rigid Contract		Flexible Contract	
State of Nature	Good	Bad	Good	Bad
Average Price	40.7	-	51.1	98.4
Rel. Freq. of Normal Quality	0.94	-	0.75	0.70
Average Auction Outcome	40.7		40.2	
Average Profit Buyer (per state)	96.8	10	78.9	29.7
Average Profit Seller (per state)	20.4	10	29.8	16.9
Average Profit Buyer (over both states)	77.9		68.9	
Average Profit Seller (over both states)	18.1		27.2	
Rel. Freq. of Contract	0.50		0.50	

Notes: The table summarizes the outcomes for rigid and flexible contracts in both states of nature. All numbers are based on the data of all 5 sessions. *Average Price* is the average of the trading price and *Relative Frequency of Normal Quality* measures how often the seller has chosen the normal quality. For rigid contracts this information is only available for the good state, because trade does not occur in the bad state. *Average Auction Outcome* is the average of the fixed price in case of rigid contracts and the lower bound of the price range in case of flexible contracts. *Average Profit Buyer (Seller) (per state)* measures the average payoff of buyers (sellers) for each state and contract. In rigid contracts the payoffs in the bad state of nature are the outside options of the market participants. *Average Profit Buyer (over both states)* is the overall average payoff of buyers (sellers) for each contract type. *Relative Frequency of Contract* is the share of the total number of contracts that corresponds to each contract type.

Table III: Price Dependence of Quality Across Contract Types

Dependent Variable	Quality [s=g]		Quality [s=b]	
	OLS	Probit [ME]	OLS	Probit [ME]
	(1)	(2)	(3)	(4)
Price increment	0.000	0.000	0.013*	0.023***
	[0.002]	[0.004]	[0.005]	[0.009]
Flexible contract	-0.335***	-0.298***		
	[0.060]	[0.060]		
Price inc. x Flex	0.009*	0.009*		
	[0.004]	[0.005]		
Constant	0.936***		0.657	
	[0.025]		[0.075]	
Observation	805	805	104	
R^2	0.13		0.03	

Notes: *Price increment* is defined as price minus 35 in columns (1) and (2) and as price minus 95 in columns (3) and (4). *Flexible contract* is an indicator variable which is unity if the contract is of the flexible type and zero otherwise. *Price inc. x Flex* is the interaction term of price increment and flexible contract. Columns (1) and (3) report coefficients of OLS estimations. Columns (2) and (4) report marginal effects based on probit estimations. Since observations within sessions may be dependent all reported standard errors are adjusted for clustering at the session level. *** indicates significance at the 1 percent level, ** indicates significance at the 5 percent level and * indicates significance at the 10 percent level.

Table IV: Sellers' Quality Choices at the Individual Level

	Consistent with RPH		Inconsistent with RPH		Total
Shading	No	Flex. only	Mostly flex.	Flex.&Rigid	
Number of sellers	27	24	7	10	68
Population fraction	0.40	0.35	0.10	0.15	1
Rigid contracts					
Average price [s=g]	40.6	41.3	41.6	39.8	40.8
Rel. freq. of q^n [s=g]	1.00	1.00	0.83	0.67	0.93
Av. number of contracts	8.2	6.5	8.1	7.3	7.5
Flexible contracts					
Average price [s=g]	54.2	49.4	50	53.4	51.9
Rel. freq. of q^n [s=g]	1.00	0.67	0.47	0.70	0.78
Average price [s=b]	98.7	97.7	98.5	99.6	98.4
Rel. freq. of q^n [s=b]	1.00	0.61	0.58	0.56	0.77
Av. number of contracts	6.3	8.5	9.3	7.1	7.5

Notes: The shading categories are defined as follows: "No" stands for sellers who always provide normal quality irrespective of the contract type. "Flex. only" are sellers who never provide low quality in rigid contracts, but provide low quality in flexible contracts at least once. "Mostly flex." are sellers who provide low quality exactly once in rigid contracts and at least once in flexible contracts. "Flex. & Rigid" are all sellers who do not fit into one of the other categories. Numbers in the last column ("Total") slightly differ from the numbers in Table II, because numbers reported in this table are averages over individual averages, while Table II directly averages over all contracts.

Figure I: Development of Quality and Prices over Time

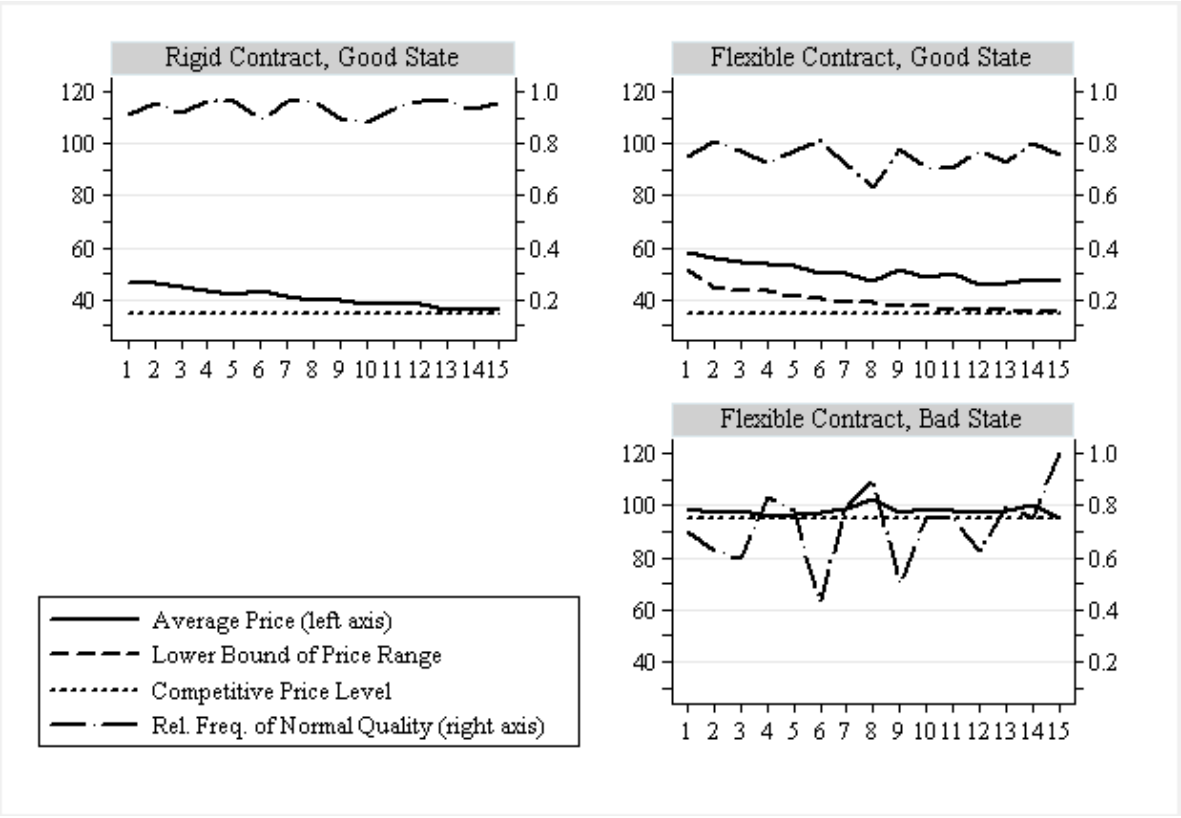

Figure II: Price Distribution and Frequency of Normal Quality Conditional on Price

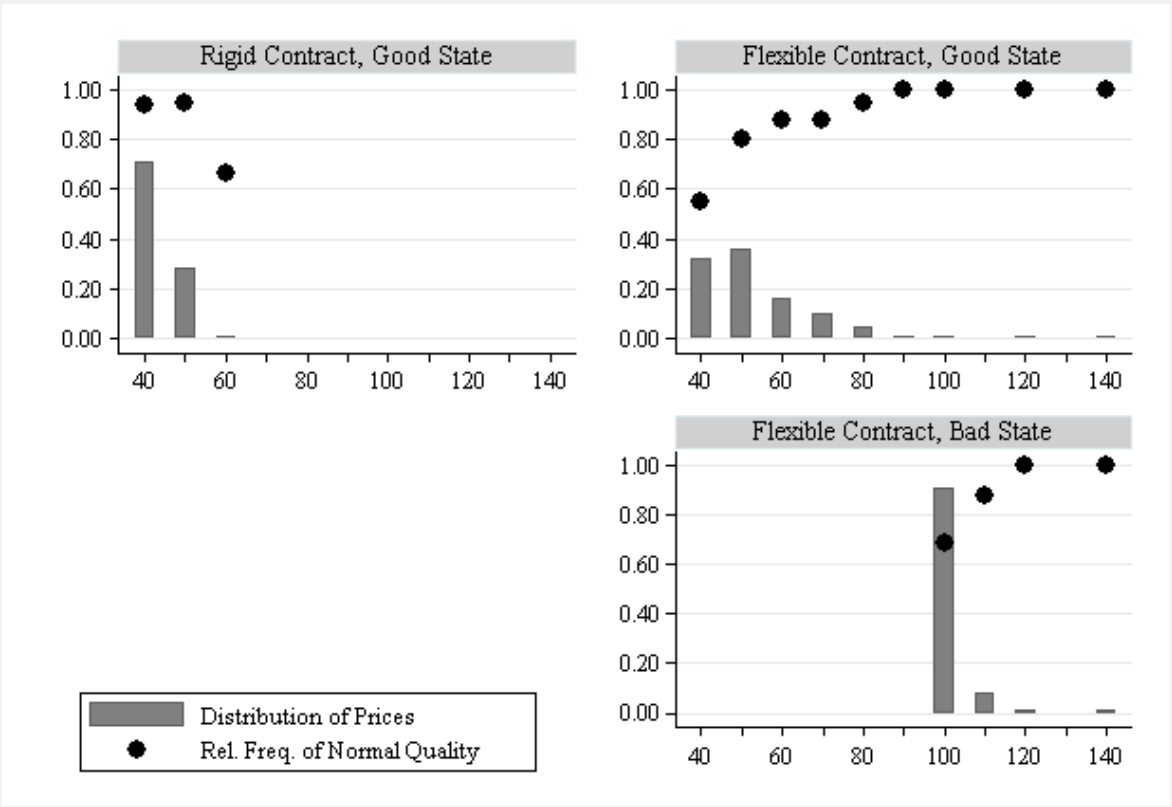

Figure III: Accepted Number of Rigid and Flexible Contracts per Individual Seller

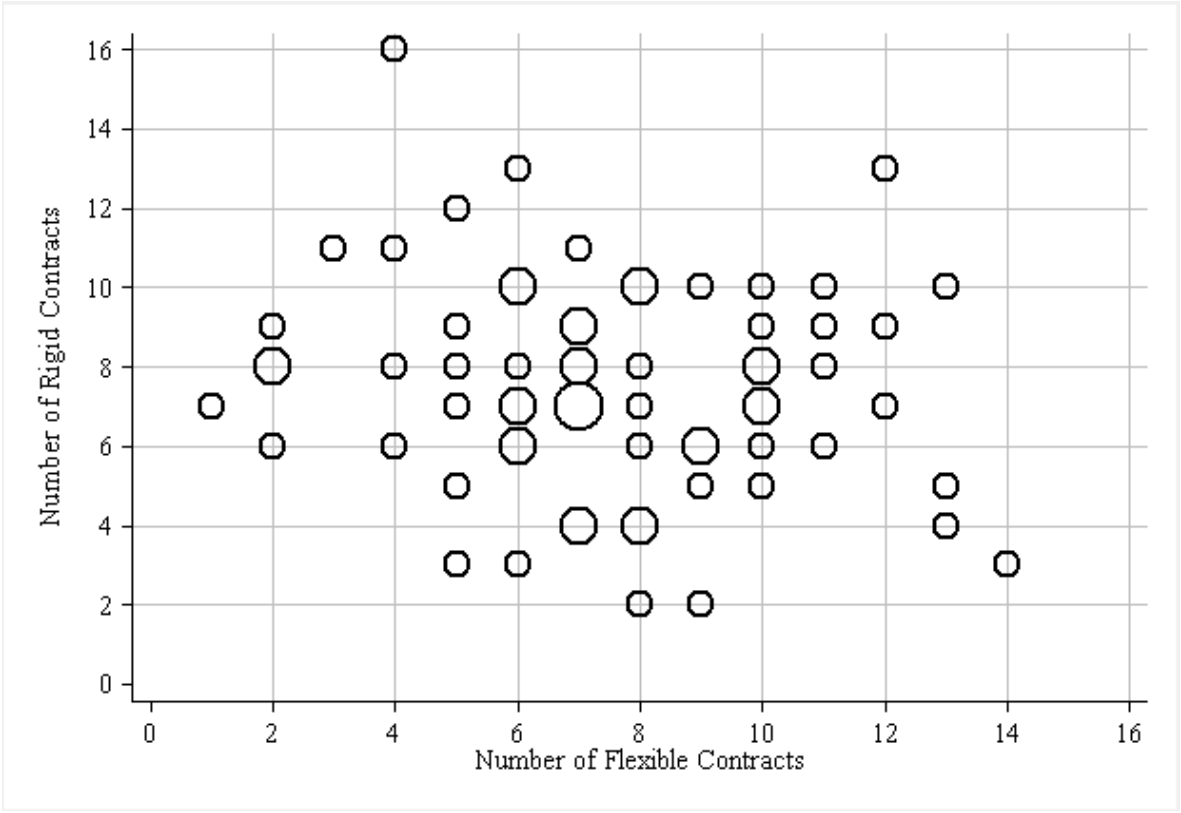

Figure IV: Price Distribution and Average Profit of Buyers Conditional on Price