

Erlei, Mathias; Reinhold, Christian

Conference Paper

To Choose or Not to Choose: Contracts, Reference Points, Reciprocity, and Signaling

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Incentives and Contracts, No. G10-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Erlei, Mathias; Reinhold, Christian (2010) : To Choose or Not to Choose: Contracts, Reference Points, Reciprocity, and Signaling, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Incentives and Contracts, No. G10-V1, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/37528>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

To Choose or Not to Choose: Contracts, Reference Points, Reciprocity, and Signaling^{*}

MATHIAS ERLEI AND CHRISTIAN REINHOLD^{**}

First Draft: February 2010

Abstract

Hart and Moore (2008) argue that varying degrees of flexibility in contracts induce differing reference points and aspiration levels for parties' shares of a transactions total surplus. As a consequence, a trade-off between adaptational flexibility and the prevention of distributional conflicts emerges. In a recent paper Fehr et al. (2009b) analyze a buyer-seller-relationship with incomplete contracts and ex ante uncertainty about the sellers' cost level to test these effects. We rerun their experiment and introduce another treatment with exogenously determined contract types. Like FHZ we find reference point effects in both treatments. However, uncooperative shading behavior in our treatments differs substantially from the one described in FHZ. Furthermore, it makes a significant difference whether contract types are determined by buyers or determined exogenously. We explain this by introducing two further effects, a reciprocity effect and a signaling effect.

Keywords: contracts, reference points, experiment

JEL Classification: C91, D03, D23

^{*} For outstanding research assistance we like to thank Eva Schulze Heuling.

^{**} Email: Mathias Erlei (m.erlei@tu-clausthal.de), Christian Reinhold (crein@tu-clausthal.de).

1 Introduction

Hart and Moore (2008) provide a new behavioral theory for the emergence of flexible vs. rigid contracts. They assume that contractual performance depends on whether the trading parties are able to realize the profits they believe to be entitled to. If they are unable to get this amount of profit, however, they feel aggrieved and are inclined to punish their trading partner by not delivering consummate performance. By doing so, they generate a welfare loss. Hart and Moore (2008) argue that through an irrevocable fixing of prices under competitive terms buyers' and sellers' aspiration levels become compatible so that neither aggrievement of any trading partner nor a delivery of bad quality ("perfunctory performance") will occur. In this way concluding contracts with rigid terms serves to reduce inefficiencies.

However, concluding rigid contracts has its own drawback. By fixing the terms of a contract irrevocably, it may become impossible to adapt to unforeseen events. This, however, decreases welfare. Consequently, there exists a trade-off between being able to effectively adapt to new circumstances and the ability to avoid distributional conflicts. By extending this simple logic, Hart and Moore provide a basis for long-term contracts in the absence of non-contractable investments and an explanation of rigid employment contracts.

Fehr et al. (2009b) present a first experimental investigation of Hart and Moore's theory. In a first stage, buyers determine whether to offer a rigid or a flexible contract. Next, contracts are auctioned off. The auction design ensures a competitive equilibrium so that sellers accept minimal prices for their products. In rigid contracts these auction prices are irrevocable and identical to final prices. In unfavorable circumstances, determined by a chance move, these prices do not allow the trade to be realized. According to Hart and Moore (2008) rigid contracts will ensure the delivery of high quality by the sellers. In contrast to this, flexible contracts always allow trade. The costs of flexible contracts consist of distributional conflicts that are induced by incompatible aspiration levels – subjective entitlements to shares of total surplus – of buyers and sellers. Buyers are allowed to increase prices above auction prices in order to appease sellers so that they do not provide low quality products. In the final stage sellers decide whether to provide normal or low quality. Note that a delivery of low quality increases the seller's costs so that no seller has a material incentive to do so.

By and large, Fehr et al. (2009b) find results in line with Hart and Moore (2008). There is less shading in rigid contracts although sellers use their potential for increasing prices. This definitely supports the view of rigid contract prices as reference points. Two robustness checks also support Hart and Moore (2008): (a) reducing the range of price increases that is feasible in flexible contracts diminishes the provision of low quality (perfunctory performance); (b) eliminating the competitive determination of (auction) prices induces a significant increase in the provision of low quality (Fehr et al. (2009a)).

Fehr et al. (2009b) admit that two features of their experimental results are quite surprising. Firstly, they find only very little perfunctory performance in rigid contracts. This is surprising because payoffs are highly asymmetric in these cases and standard measures for fairness or inequity aversion suggest much more perfunctory performance. Furthermore, the act

of choosing rigid contracts itself can be interpreted as an uncooperative behavior that deserves punishment by delivering low quality. Secondly, this oddity is enhanced by the fact that there is much more perfunctory performance in flexible contractual relations although the payoff asymmetry is significantly reduced by buyers' voluntary price increases. Fehr et al. (2009b) conjecture that this can be explained by 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." (Fehr et al. (2009b), p. 6) In our judgment this explanation looks rather artificial. Do sellers really make such a neat distinction between "within the contract" and "the contract itself" while at the same time they do not take into account that buyers' choice of rigid contracts causes their profits to decrease significantly. Furthermore, the choice of contract types is by no means determined under competitive conditions.

This is where our analysis begins. Our motivation for making an own investigation is based upon two open questions: (1) Why don't sellers punish buyers for choosing rigid contracts? (2) Why don't sellers appreciate buyers' voluntary price increases in flexible contracts? In order to test whether Fehr et al.'s interpretation of the data is appropriate, we replicate FHZ's main treatment and introduce a new treatment. In this new treatment contract types are not determined by players. If FHZ are right, this change in experimental design must not have an impact on subjects' behavior. If, however, players' behavior changes significantly, there must also be other forces at work than just reference point effects.

Our main findings are: (a) In the replication treatment with endogenous contract types we find much more cases of delivering low quality (perfunctory performance) than in FHZ. (b) In case of *exogenously* determined rigid contracts perfunctory performance decreases substantially. Our interpretation of (a) and (b) is that subjects in our experiment do indeed punish buyers for *choosing* rigid contracts. We call this a reciprocity effect. (c) There is significantly less perfunctory performance if flexible contracts are determined exogenously than if flexible contracts are actively chosen by buyers. Our interpretation of this phenomenon is that players' aspiration levels do not depend only on contract types but also on the signal that buyers send by choosing flexible contracts which are more favorable to sellers. Sellers then expect their trading partner to be more cooperative so that sellers' aspiration levels with respect to their share of surplus increases. This will, in turn, make sellers more inclined to provide perfunctory performance if buyers turn out to be less generous than expected. We call this the signaling effect. (d) We find, nevertheless, evidence for the existence of reference point effect in both treatments with endogenous and exogenous determination of contract types.

The paper is structured as follows: In Section 2 we describe our experiment's design and provide details on procedures. In Section 3 we present the behavioral predictions. Section 4 contains our results and their discussion. Section 5 concludes.

2 Experimental Design

In this section we present our experimental design and procedures. We consider a relationship between a buyer and a seller who are interacting in a two period relationship. At date 0 they

conclude an incomplete contract for the trade of one unit of a widget. Its incompleteness is due to the possibility that trade may not be feasible on date 1. The feasibility depends on the nature of the contract and on the ex ante unknown state of nature. The relationship's dynamic resembles the "fundamental transformation" as described by Williamson (1985). In such a setting transaction partners start in a competitive market structure which after some events, e.g. a specific investment, changes into a structure of a one-sided or bilateral monopoly. In all treatments we have two different kinds of contracts, rigid and flexible contracts, which brings our experiment in line with the model of Hart and Moore (2008).

We implemented two different treatments, an Endogenous Contract Treatment (EndCT) and an Exogenous Contract Treatment (ExCT). The EndCT is closely related to the design of Fehr et al. (2009b). In our ExCT we abolish the buyers' ability to choose between contract types. Each session consisted of 15 such periods. Figure 2 (given in the appendix) shows the structure of each period. To facilitate subjects' understanding of the experiments we subdivided each period into seven sequential stages:

Stage 1 – Formation of groups:

In order to avoid reputational effects, buyers and sellers were randomly divided into interaction groups of each two sellers and two buyers at the beginning of each period.

Stage 2 – Determination of contract types:

The determination of contract types varies between the two treatments. In the EndCT buyers decide whether to choose the rigid or the flexible contract design. In the ExCT the contract design is exogenously given to the participants. The difference between the designs relates to the process of determining the final product price. In rigid contracts the auction price (Stage 3) is binding for buyers and sellers. Furthermore, if the state of nature (stage 4) turns out to be bad, these binding contracts do not allow for trade to be accomplished. In flexible contracts buyers may increase product prices above the level given by the auction if they wish to do so and trade is always feasible. Table 1 shows the intervals of feasible final prices with respect to contract types and states of nature.

Stage 3 – Sellers compete for contracts in clock auctions:

After the two buyers in a group have chosen their contract types, contracts are auctioned off to the sellers. Just like FHZ we implemented an inverse clock auction with a starting price of 35, which is increased each second by one point. The auction ends when either the upper limit of 75 is reached or if one of the sellers accepts the current auction price by clicking a button on the computer screen. The first seller accepting the price gains the contract and enters into the bilateral contract performance stages. As described above, the final price is either fixed to the level of the auction price (rigid contracts) or may voluntarily be increased by the buyer. Sellers who don't gain any contracts realize an outside option payment of 10 points.

In order to design a truly competitive auction, the supply of widgets is twice as large as its demand. More precisely, each seller is able to produce up to two widgets per period while buyers can buy just one widget. This means that any seller may serve both buyers while any

buyer can only buy from one seller. Consequently, there is an excess supply which induces a highly competitive auction environment. In contrast to FHZ, the auctions for the two possible contracts in each group are conducted simultaneously. The auction boxes are randomly placed next to each other on the computer screen. Sellers are free to engage in any of the auctions as long as none of them has accepted an auction price. If one of the contract prices is accepted, the corresponding auction is immediately finished.

Table 1: Intervals of Feasible Final Prices

Contract type	State of nature	
	Good	Bad
Rigid contract	Auction price	No trade
Flexible contract	[Auction price, 140]	[95, 140]

Note: The auction price was determined by an inverse clock auction between the two sellers in one group. It could range from 35 to 75.

Stage 4 – Determination of the state of nature:

After sellers and buyers have concluded contracts, they are informed about the state of nature. This state determines the parameters for the remaining part of the period. Firstly, it has an impact on sellers' costs (see Table 2). In bad states sellers' costs increase by 60 points. In addition, minimum final prices are also increased to 95. Secondly, the state of nature determines whether sellers and buyers can trade at all. If buyers have chosen rigid contracts and a bad state occurs, trade becomes impossible. In all other cases trade is going to be accomplished (see Table 1). Good states occur with a probability of 80 percent and bad states with a probability of 20 percent. In the no-trade case buyers and sellers get an outside option of 10 points each.

Table 2: Experimental Parameter

State of nature	Good (prob = 0.8)		Bad (prob = 0.2)	
	normal	low	normal	low
Seller's cost	20	25	80	85
Buyer's value	140	100	140	100

Stage 5 – Determination of the final price:

The process of determining final prices depends on the chosen contract type. In case of rigid contracts participants do not have to make any decisions because the final price is given by the auction price. With flexible contracts buyers can voluntarily increase final prices above the auction price (or above 95 in bad states) and up to a maximum of 140 (see Table 1).

Stage 6 – Sellers choose quality:

If trade takes place, sellers have to determine product quality which can be either low or normal quality. Irrespective of the state of nature choosing low quality instead of normal increases sellers' costs by 5 points and decreases buyers' value by 40 points (see Table 2).

Stage 7 – Determination of profits:

At the end of each period profits are displayed to buyers and sellers. The payoffs were calculated as follows:

Seller profits: $\pi_S = \text{price} - \text{costs}$

Buyer profits: $\pi_B = \text{value} - \text{price}$

For the purpose of comparing our results with the results of FHZ, it is necessary to summarize in which ways our EndCT differs from the experiment of FHZ. First of all, we used different instructions which, of course, are somewhat similar to those of FHZ.¹ As mentioned, we used simultaneous auctions (Stage 3) instead of sequential auctions because we wanted to find out which contract type is preferred by the sellers. Finally, FHZ provided buyers with aggregate information about average buyer profits as an additional eighth stage of each period. Average profits were given separately for the two contract types over all past periods. Though giving this information facilitates learning, it also weakens the independency of individual decisions and fosters contagion effects. We felt that next to the speed of the learning process it might have an impact on the direction of learning.

All sessions have been conducted in December 2008 and January 2009. We conducted two sessions for each of the two treatments with 20 subjects each (10 buyers and 10 sellers). Subjects were students from the Clausthal University of Technology studying various majors like Business Administration or Industrial Engineering. In order to make them familiar with our experimental design we started every session with two training periods which did not get remunerated. In both treatments 45 experimental currency units (points) corresponded to one Euro. In addition each participant received a show-up fee of 5 Euro. On average, subjects earned 18.28 EUR (about 25.41 USD at the time of the experiment). A session lasted for about one and a half hour. The experiments were computerized using the software z-Tree (Fischbacher 2007).

3 Hypotheses

The main motivation for carrying out our experiment is that we presume that the process of determining contract types has an impact on subjects' behavior. Different modes of determining contract types may lead to different modes of behavior. In particular, sellers may be able to learn that rigid contracts decrease their expected profits and cause an unfair distribution of profits. Sellers realizing this will be inclined to interpret the choice of rigid contracts as unfriendly behavior that deserves punishment by providing low quality. On the other hand,

¹ Instructions are available upon request.

choosing flexible contracts may be understood as a signal for cooperative behavior. Such a signal will increase sellers' expectations of profits. Assuming this to increase sellers' aspiration levels, low price increments will not meet sellers' aspiration levels and provoke shading. Summarizing, according to our considerations the feasibility of actively choosing contract types increases the inclination to provide low quality. Consequently, replacing the active choice of contract types by a chance move will lead to less shading in both types of contractual relationships. Keeping this in mind, we can now present our central hypotheses.

After having formed the groups, the second stage in each period consists of determining contract types. In the EndCT we expect the same pattern of decisions as in Fehr et al. (2009b). Because there is no choice between contract types in the ExCT, it does not make sense to provide any behavioral hypotheses for this decision.

Hypothesis 1 (Contract Choice):

In the Endogenous Contract Treatment subjects choose rigid contracts in about 50 percent of the cases and flexible contracts in the remaining 50 percent of the cases.

In Stage 2 of each period contracts are auctioned off. Since there is always an excess supply, equilibrium prices are equal to minimum prices. Following FHZ we expect to see deviations in early rounds due to subjects' learning. In contrast to FHZ, we implemented simultaneous auctions for both buyers' contracts within one group. This renders possible to analyze which kind of contract sellers prefer. In FHZ sellers' profits in flexible contracts are significantly higher than in rigid contracts. Expecting a similar pattern of profits, sellers will prefer flexible contracts. In groups in which there is one auction for rigid contract and one auction for flexible contract, sellers will thus try to get a flexible contract first and only after this they will try to get the rigid contract. Due to the nature of the clock auction prices for rigid and for flexible contracts will differ in these cases.

Hypothesis 2 (Auction Prices):

- a) *In both treatments auction prices converge towards the competitive equilibrium, $p_A = 35$.*
- b) *On average, prices for flexible contracts are lower than prices for rigid contracts.*

Whereas rigid contracts leave no room for price adjustments, flexible contracts allow buyers to increase prices within the price range above the competitive lower price bound. This opportunity may affect the sellers' shading decision so that increasing the price may incline sellers not to provide low quality. This price decision at date 1 (Stage 5) is somewhat similar to an ultimatum game (Güth, Schmittberger, Schwarze (1982)) in which a proposer makes an offer how to share an amount of money between her and the responder. The responder then decides whether to accept the proposal or reject it. In this type of games rejections induce a payoff of zero. In our experiment, buyers act as a kind of proposer by offering a particular share of total surplus to the seller. After this the seller "accepts" the offer by providing normal quality or "rejects" it by providing low quality. In contrast to the ultimatum game, the seller can lower the buyer's profit by only 40 points in our experiment and his costs increase by

only 5 points. In line with Forsythe et al. (1994) we're expecting that buyers will pay significant higher prices than the lower price bound in order to avoid shading.

FHZ show that buyers are indeed willing to pay significant price increments in flexible contracts. Accordingly, we expected to find a similar pattern in the EndCT. However, behavior in the ExCT may differ. Being able to choose between contract types and choosing flexible contracts may have an impact on sellers' aspiration levels in the EndCT which is absent in the ExCT. In particular, we presume that a buyer choosing flexible contracts signals a preference for cooperative and fair behavior. Thus sellers' expectations raise. This signaling effect does not exist in ExCT because sellers do not choose contract types. Consequently, sellers' aspiration levels remain smaller so that buyers need to pay only smaller amounts of money to prevent sellers from shading.

Hypothesis 3 (Price Increment):

- a) *In both treatments buyers pay prices higher than the lower price bound in flexible contracts and consequently price increments above zero.*
- b) *In the exogenous treatment a lower aspiration level of the sellers causes lower price increments or higher average quality.*

In Stage 6 of each period sellers choose whether to provide low quality or normal quality. This decision is at the core of the papers by Hart and Moore (2008) and FHZ. According to the reference point hypothesis rigid contracts leaving no room for further adaptations provide a reference point to which both parties agree. Although profits may be highly asymmetric the existence of a reference point prevents shading. The evidence in FHZ confirms that there is fewer shading in rigid contracts than in flexible contracts.

Assuming that the process of determining contract types has an impact on subjects' behavior, we conjecture that other effects than just the reference point effect have to be at work. We argue that the possibility of choosing between contract types increases shading in both rigid and flexible contracts. It remains unclear, however, whether this treatment effect is larger in rigid contracts or in flexible contracts. If shading due to the choice of rigid contracts is smaller than shading due to unfulfilled expectations in flexible contracts, Fehr et al.'s and results can be explained without reference to reference point effects. Comparing shading in both treatments may clarify the nature of subjects' behavior. According to the theory of contracts as reference points, the magnitude of reference point effects does not depend on the process of determining the contract type but only the contract type itself. Consequently, there will be no differences in shading between the treatments. If, in contrast, shading decreases to zero in the ExCT, then the reference point hypothesis must be rejected. Finally, if shading in the ExCT is smaller than in the EndCT and if shading in flexible contracts is larger than in rigid contracts (ExCT) then all effects (reference point effects, punishment effects/reciprocity for choosing rigid contracts and signaling effects for choosing flexible contracts) are confirmed. This leads us to two alternative hypotheses:

Hypothesis 4a (Pure Reference Points):

In both treatments shading will be close to the evidence in FHZ, i.e. subjects will provide low quality in about

- a. 6 percent (rigid contracts)*
- b. 25 percent (good states in flexible contracts)*
- c. 30 percent (bad states in flexible contracts) of the cases.*

Hypothesis 4a is in perfect accordance to Hart and Moore (2008) and FHZ. Our alternative hypothesis refers to the existence of reciprocity and signaling effects:

Hypothesis 4b (Reciprocity and Signaling Effects):

(i) In the Endogenous Contract Treatment shading will be close to the evidence in FHZ, i.e. subjects will provide low quality in

- a. 6 percent (rigid contracts)*
- b. 25 percent (good states in flexible contracts)*
- c. 30 percent (bad states in flexible contracts) of the cases.*

(ii) In the Exogenous Contract Treatment shading will be lower than in the Endogenous Contract Treatment

- d. If shading decreases to zero, there are no reference point effects.*
- e. If shading remains above zero and if there is more shading in flexible than in rigid contracts, there are reference point effects, reciprocity effects and signaling effects.*

4 Results

In this section we present and discuss our results. In the EndCT we find surprisingly large deviations from the results presented in FHZ. Buyers choose flexible contracts more often and they pay smaller price increments. Sellers provide low quality more often than in FHZ. This difference in shading is largest with respect to rigid contracts. Our new treatment with contract determination by chance moves shows a large shift in behavior. There is less shading in both contract types. This confirms our hypothesis that the results presented in FHZ are driven also by other factors than pure reference point effects. We first present the results and then provide an explanation. Our aggregate results are presented in Table 8.

Insert Table 8 here

In FHZ each contract type is chosen in 50 percent of the cases. Table 8 shows that in our sessions with endogenous contracts buyers reveal a clear preference for flexible contracts. 72.3 percent of the contracts were flexible. Taking into account that average buyers' profits in rigid contracts (65.69) are (insignificantly) larger than in flexible contracts (61.2) this result becomes even more remarkable. The more or less equal distribution of contract types in the ExCT is irrelevant here, for contracts have not been chosen by buyers but have been determined exogenously.

Result 1 (Contract Choice):

Although average buyers' profits in flexible contracts are smaller than their profits in rigid contracts they choose flexible contracts in more than two out of three cases.

Obviously, subjects' behavior in our experiment is different from subjects' behavior in FHZ. We identify two possible explanations for this: (1) We, unlike FHZ, have not provided subjects with any statistical information about the aggregate profitability of rigid and flexible contracts in transactions of other players. Due to this lack of information subjects may have been unable to identify the more profitable contract type. (2) We had a different subject pool than FHZ. FHZ excluded all students of economics and psychology whereas more than 75 percent of our subjects study business administration or industrial engineering. Consequently, their education may have had an impact on their behavior in the laboratory. It is somewhat surprising, however, that these economically trained participants favoured the less profitable but – with respect to the distribution of profits – fairer contract type.

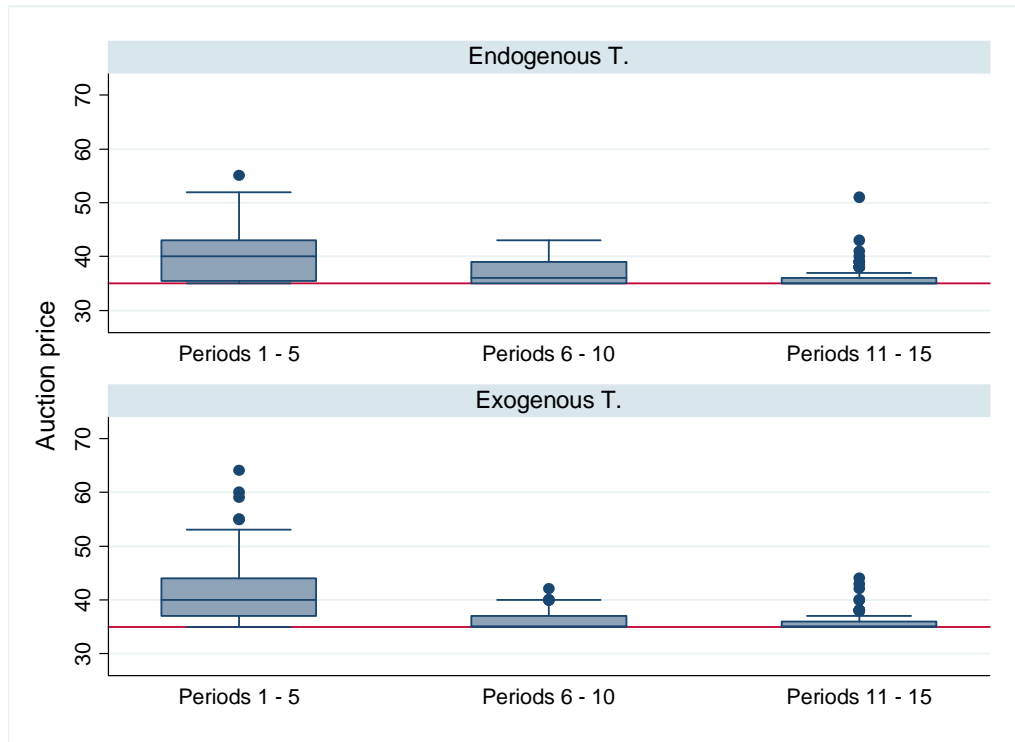
After buyers had chosen their contract types all contracts were auctioned off to the sellers. The results of the auctions are close to the results in FHZ and close to our expectations.

Result 2 (Auction Prices):

- (a) *Prices converge to the equilibrium value $p = 35$.*
- (b) *Prices for flexible contracts are lower than those for rigid contracts.*

Figure 1 shows a box plot for auction prices in groups of periods (1-5, 6-10, 11-15). It can easily be seen that auction prices in the first periods are well above 35 and converge over time to the equilibrium price. In the final five periods mean auction prices in EndCT and ExCT are 36.07 and 35.99 respectively. In contrast to FHZ we had simultaneous auctions of both contracts in each period and group. Consequently, while accepting one contract subjects lost time and could only accept the second contract after the corresponding time lag. This and normal reaction time of participants explain most of the divergence of prices from equilibrium values.

Figure 1: Auction Prices



The simultaneity of auctions has another nice feature. Excess supply in markets for contracts induces intense competition between sellers so that they will immediately accept contracts in equilibrium. If sellers prefer one type of contract to the other, we should find that sellers first try to get this contract type and only after that the other one. This implicates that prices for preferred contracts will be lower than prices for the other contracts.

Comparing sellers' profits shows that in flexible contracts sellers' profits are substantially larger: sellers' mean profits in flexible contracts are 22.44 (endogenous contracts) and 24.94 (exogenous contracts), in rigid contracts they earn only 15.10 (endogenous contracts) and 15.91 (exogenous contracts). Thus we expect prices to be lower for flexible contracts. Table 8 shows that part (b) of hypothesis 2 is also confirmed. In both treatments auction prices for flexible contracts are lower on average and the difference is statistically significant at the 5 percent level (signed rank test on individual means).

Result 3 (Price increment):

- (a) *In both treatments the mean price increments are above zero.*
- (b) *Mean price increments don't differ between the treatments.*

There are two different ways to define the price increment. First, the price increment (hereafter p35) consists of the additional points above the minimal feasible price of 35 in the clock auction.² Unfortunately this definition doesn't take into account that the contract / auc-

² FHZ use this definition.

tion price serves as a reference point in principle. Consequently we define the price increment as the difference between the final price (chosen by the buyer at date 1) and the auction price in good states or 95 in bad states of nature. Thus the price increment in rigid contracts is always zero. As long as result 2 holds, most of the difference between the two definitions of the price increment will disappear in the last periods of the experiment.

In our post-experimental questionnaire we asked the sellers to reveal their minimal price increment required for abstaining from shading in a flexible contract, good state of nature and an auction price of 35. On average, in the EndCT the sellers claim 20.25 points. In the ExCT we observe only a mean price increment of 10.95 points. So, in accordance with our conjectures, the removal of buyers' responsibility for contract choice reduces sellers' aspiration levels significantly (one-sided rank sum test $p < 0.01$). However, changing between treatments does not have an impact on buyers' willingness to pay. In the questionnaire we also asked how much buyers were willing to pay in order to foreclose shading in flexible contracts, good states of nature and an auction prices of 35. On average, buyers were willing to pay up to 14.35 in the EndCT and 14.25 in the ExCT.

Table 3 confirms our post-experimental observations. It shows the mean price increment with respect to the two different treatments. None of the differences between the treatments is statistically significant (one-sided rank sum test with subject means, $p > 0.10$). Price increments are, by and large, the same. Accordingly, buyers' behavior with respect to price increments does not depend on the endogeneity or exogeneity of contract types.

Table 3: Mean Price Increments in Both Treatments

	State of nature		
	Bad	Good	Total
<i>Fehr et al. (2009b)</i>	3.4	10.9	---
EndCT	2.80 (4.154)	10.04 (12.789)	7.51 (11.135)
ExCT	3.87 (4.932)	11.12 (13.587)	9.20 (12.326)

Notes: Standard deviations are in parentheses. A one-sided rank sum test with subject means shows that none of the differences between the treatments is statistically significant ($p > 0.10$).

The decision whether to provide normal or low quality is at the core of our experiment. Following Hart and Moore (2008) we refer to the provision of low quality as “shading”. Our main research question is whether the difference in shading behavior with respect to contract types may also be interpreted in a different way as reciprocity and signaling effects. If behavior in the ExCT is significantly different to behavior in the EndCT, there must exist some additional effects that FHZ subsume under reference point effects.

Table 4 summarizes shading behavior in our experiment. The first distinctive feature of our results is that we find much more shading in the EndCT than FHZ do. In particular, there is more shading in rigid contracts.

Table 4: Shading Behavior
with Respect to Treatment, Contract Type, and State of Nature

	Contracts types and states of nature			
	rigid		flexible	
	bad	good	bad	good
EndCT	---	0.32	0.37	0.33
(FHZ)	---	(0.06)	(0.30)	(0.25)
ExCT	---	0.15	0.08	0.22

Note: Numbers are relative frequencies of shading. Numbers in parentheses are the corresponding relative frequencies of shading in FHZ.

Even more important, there are only small differences in shading between rigid and flexible contracts. None of these small differences is statistically significant (signed rank test on subject means). In the EndCT shading in rigid contracts and good states is just a little smaller than in flexible contracts and good states. Compared to the results of FHZ reference point effects seem to be rather small in our sessions.

Things are somewhat different in the ExCT. Compared to endogenous contracts shading is substantially and significantly smaller in all combinations of states and contract types (one-sided rank sum test; $p < 0.05$). Furthermore, in good states there is less shading in rigid contracts than in flexible contracts. Our results clearly contradict Hypothesis 4a (pure reference points) and partially confirm Hypothesis 4b.

Result 4:

- (i) *In the Endogenous Contract Treatment shading behavior is not close to the evidence in FHZ. Subjects provide low quality in*
 - a. *32 percent (rigid contracts)*
 - b. *33 percent (good states in flexible contracts)*
 - c. *37 percent (bad states in flexible contracts) of the cases.*
- (ii) *In the Exogenous Contract Treatment shading is lower than in the Endogenous Contract Treatment. Furthermore, shading remains above zero and there is more shading in flexible than in rigid contracts.*

It is remarkable how much more shading we find in rigid contracts than FHZ do. This might raise doubts concerning the validity of the reference point hypothesis. However, it has to be taken into account that prices in flexible contracts are higher than in rigid contracts so

that there is less incentive for sellers to shade. In Table 5 we present relative frequencies of shading for different price intervals. In order to compare prices in different states price increments (p35) are defined by final prices minus minimal feasible prices, i.e. prices minus 35 in good states or prices minus 95 in bad states. Table 5 shows that in all but one intervals shading in rigid contracts and good states is below shading in flexible contracts and good states. This confirms the reference-point-effect hypothesis (RPE).

Table 5: Shading in Different Price Intervals

Price increments (p35)	Endogenous Contracts				Exogenous Contracts			
	rigid		flexible		rigid		flexible	
	bad	good	bad	good	bad	good	bad	good
p35 = 0		.333	.488	.409		.091	.125	.333
0 < p35 ≤ 5		.313	.200	.404		.191	.000	.308
5 < p35 ≤ 10		.385	.100	.412		.231	.167	.071
10 < p35 ≤ 15		.000	.667	.188		.000	.000	.200
15 < p35		.000		.205		.000		.156

Note: Numbers are relative frequencies of shading.

Our results indicate that it makes a difference whether contracts are determined endogenously or exogenously. Endogenously determined rigid contracts induce clearly more shading than exogenously determined contracts. When sellers had to decide whether to shade or not, the only difference between both treatments consisted in the fact that buyers had consciously determined the contract type so that they were responsible for it. Our interpretation of the increase in shading thus is that sellers punished buyers for selecting the wrong type of contracts. This, in turn, is nothing else than negative reciprocity (RecE) and affects shading behavior in opposite direction to reference point effects.

Endogenous choice of contract types increases shading in flexible contracts, too. Because we regard the decision in favor of a flexible contract as a signal to the seller that increases his aspiration level, low price increments disappoint sellers so that they are more inclined to punish buyers (SigE). Table 6 summarizes our theory about the three effects.

Table 6: Theoretical Effects

	Rigid Contracts	Flexible Contracts
Endogenous Contracts	RPE: –	RPE: +
	RecE: +	SigE: +
Exogenous Contracts	RPE: –	RPE: +

Note: RPE denotes the reference point effect, RecE the reciprocity effect and SigE the signaling effect. The signs “+” and “–” indicate that the corresponding effect increases (+) or decreases (–) shading.

According to Table 6 and as long as the reciprocity effect isn’t too large, shading is highest in case of flexible and endogenous contracts and lowest in case of rigid and exogenous contracts. The remaining two cases, rigid endogenous contracts and flexible exogenous contracts are somewhere in between. It is ex ante unclear in which of these two cases there will be more shading. These hypotheses are confirmed by subjects’ behavior that is shown in Table 5.

In order to conduct a more rigorous test, we conducted some econometric estimations of shading behavior. In our data set shading is given by the binomial variable shading. It equals 1 in case of shading and zero in case of normal quality. Final prices can be separated in auction prices (priceauc) plus price increments (priceinc).³ Table 6 shows that the reference point effect is the same for endogenous and exogenous contracts. It refers only to the choice of the contract type. FlexC is the corresponding variable being 1 for flexible contracts and zero for rigid contracts. The reciprocity effect applies only to rigid and endogenous contracts. The binomial variable endorigid covers this effect. This interactive variable is 1 for rigid contracts in the EndCT and zero otherwise. The signaling effect occurs only in flexible endogenous contracts. The corresponding variable (endoflex) equals 1 in these cases and zero otherwise. According to our theoretical considerations shading is increasing in FlexC, endorigid and endoflex.

Table 7 shows the results of our estimations. The dependent variable in all regressions is shading. The first estimation in column 2 serves as our benchmark estimation.⁴ Let us first look at the benchmark model, a random effects logit model with few control variables. Column 2 shows that the coefficients of FlexC, endorigid and endoflex are all significant at the 5 percent level. Their signs are all positive. Therefore, our results corroborate the existence of all three effects, the reference point effect, the reciprocity effect and the signaling effect. Fur-

³ Priceinc must not be mixed up with p35 in Table 5.

⁴ The qualitative results of our estimation do not depend on the method of estimation and the inclusion of further control variables like state of nature, sex, age, and period. We conducted probit estimations, population averaged estimations, pooled logit estimations and OLS estimations with similar results.

thermore, the coefficients of *endorigid* and *endoflex* are higher than the coefficient of *FlexC* indicating that the signaling and the reciprocity effect are as least as important as the reference point effect. The coefficient of *priceinc* is significantly below zero meaning that buyers can decrease the probability of shading by increasing the final price.

Table 7: Random Effect Logit Regressions and Social Preferences

Variable	Model I	Model II	Model III	Model IV	Model V	Model VI
<i>FlexC</i>	.819 (.031)	.828 (.031)	.767 (.039)	.744 (.047)	.642 (.067)	.756 (.045)
<i>endorigid</i>	1.11 (.046)	1.11 (.046)	1.113 (.044)	1.122 (.044)	1.127 (.043)	1.12 (.045)
<i>endoflex</i>	1.24 (.015)	1.24 (.016)	1.232 (.016)	1.22 (.017)	1.191 (.019)	1.219 (.017)
<i>priceinc</i>	-.0807 (.000)	-.087 (.012)	-.059 (.001)	-.060 (.003)	-.051 (.005)	-.070 (.041)
<i>good state</i>	.630 (.024)	.801 (.203)		.330 (.188)		.600 (.269)
<i>reci_unfair</i>		-.003 (.425)	.0099 (.035)			-.005 (.382)
<i>ia_shade</i>				.899 (.064)	1.157 (.012)	.911 (.062)
constant	-2.86 (.000)	-2.78 (.000)	-3.052 (.000)	-3.47 (.000)	-3.399 (.000)	-3.34 (.000)
N	525	525	525	525	525	525
LL	-249.3	-249.3	-249.7	-248.1	-248.5	-248.1
LL(0)	-280	-280	-280.1	-279	-279.3	-279
R ² ; R ^{2,ps}	.110	.110	.109	.111	.110	.111

Notes: dependent variable: shading; p-values (one-sided test) in parentheses; N: number of observations; LL: log-likelihood.

Finally, in good states we find significantly more shading. This may be explained by two facts. Firstly, there is simply less scope for price increments in bad states ($140 - 95$) than in good states ($140 - \text{auction price} \approx 140 - 35$). Consequently, comparatively modest price increments in bad states look more generous than larger price increments in good states. Secondly, the distribution of payoffs is more unequal in good states than in bad states. For example, if the buyer chooses a price increment of 15 in bad states, both players' payoffs are identical. In good states, however, the same price increment induces highly unequal payoffs, $\pi_B = 90$ and $\pi_S = 30$.⁵

Thus far our results are economically and statistically significant. We believe, however, that due to the highly asymmetric payoffs in our experiment it is necessary to test whether

⁵ Here we assume the auction price to be 35, the equilibrium price.

additional variables characterizing social preferences or fairness preferences have an impact on our main results. We measure such social preferences by two variables. The first one, *reci_unfair*, measures the difference between the buyer's and the seller's hypothetical profits in case of higher buyer's profit assuming that the seller doesn't shade. Our second variable, *ia_shade*, is a binomial variable with direct reference to the models of inequality aversion by Fehr and Schmidt (1999). Given the utility function of Fehr and Schmidt,

$$U_i = \pi_i - \alpha_i \max\{\pi_j - \pi_i, 0\} - \beta_i \max\{\pi_i - \pi_j, 0\},$$

and assuming that $\alpha_i = \beta_i = 0.5$ we calculated whether shading is utility maximizing in the sense of Fehr and Schmidt.⁶ The variable equals 1 whenever shading maximizes utility and it equals zero otherwise.

Table 7 shows the results of these estimations, too. Model I is the benchmark model, again. In Model II and Model IV we simply add one of the new variables for social preferences. Model VI includes both variables. Model II and Model V substitute the corresponding social preference variable for the state variable.

The introduction of the new variables has hardly any impact on the coefficients of *FlexC*, *endorigid* and *endoflex*, i.e. those variables being at the core of our paper. The magnitudes of coefficients remain, by and large, unchanged and statistically significant. Again, the existence of reference point effects, reciprocity effects and signaling effects is confirmed. Likewise, there is no substantial effect on the coefficient of *priceinc*. In contrast to this, the coefficient of the state variable is highly sensitive to the introduction of the variables for social preferences. The coefficients of all three variables, *good_state*, *reci_unfair* and *ia_shade*, are statistically significant at the five percent level when they are included exclusively. When they are jointly included with one or both other variables, then only *ia_shade* remains weakly significant (at the ten percent level). This sensitivity is due to the high correlation between these variables, in particular between *good_state* and *reci_unfair* ($\rho = .68$).

Summarizing the results of regression analysis, we can state our final result.

Result 5:

According to the results of panel regressions we find that

- *there is evidence for the existence of reference point effects, reciprocity effects and signaling effects;*
- *all three effects are of a similar order of magnitude;*
- *results are robust with respect to included variables and estimation methods;*
- *variables referring to distributional fairness, *reci_unfair* and *ia_shade*, contribute to the explanation of subjects' behavior.*

⁶ Because of the large payoff asymmetries in our experiment we had to select rather low values for α and β . Otherwise inequality averse players would shade in virtually all cases.

5 Conclusion

In this paper we provide a second experimental test of Hart and Moore's (2008) behavioral theory of contracts. Hart and Moore's main message is that rigid contract terms, in particular rigid prices, may serve as reference points that can prevent ex post inefficiencies caused by incompatible subjective entitlements. In this paper we find results deviating significantly from the first test, presented by Fehr et al. (2009b). FHZ provide remarkable support for Hart and Moore (2008). However, they admit to have found some other results surprising themselves. In particular, they find no evidence for negative reciprocity due to buyers' uncooperative choices of contract types and surprisingly little remuneration for buyers' voluntary increases of contract prices. Because we regard their explanation as rather unconvincing, we, after having performed a replication treatment, introduce a new treatment in which buyers cannot determine contract types anymore. In the new treatment contract types are exogenously determined by the experimenter. According to Fehr et al.'s interpretation of their data this change in experimental design must not have an impact on behavior. We find, however, that the new design induces a substantial change in behavior. Furthermore, subjects' behavior in replicating sessions differs significantly from the behavior of subjects in FHZ.

Firstly, we find more shading in endogenously chosen rigid contracts than FHZ. We regard this increased shading as the (negative) reciprocity effect that FHZ are surprised not to find: Sellers punish buyers for choosing rigid contracts decreasing sellers' payoffs. This effect is so large that shading in our sessions is approximately the same in endogenously rigid and endogenously flexible contracts.

Secondly, determining contract types exogenously, as we do in our new treatment, induces a decrease in shading by approximately 50 percent for both rigid and flexible contracts. Consequently, it does make a huge difference whether contracts types are chosen by the players themselves or not and Fehr et al.'s interpretation of the data seems to be incomplete. The decrease in shading with rigid contracts can be explained the omission of the reciprocity effect since buyers are not responsible for contract types anymore. We interpret the decrease in shading with flexible contracts as the omission of a signaling effect. This signaling effect that works in case of endogenous contracts stems from the possible interpretation of sellers that choosing flexible contracts may signal a more cooperative behavior of buyers that increases sellers' aspiration levels. At any rate, Fehr et al.'s conjecture that "aggrievement occurs mainly about outcomes within the contract and not about the contract itself" (Fehr et al. (2009b), p. 6) is not supported by our data.

Thirdly, we also find evidence for the existence of reference point effects. This effect is embodied in the fact that, given final prices and treatments, shading in rigid contracts is smaller than in flexible contracts. Fourthly, in our experiment reference point effects are of a smaller magnitude than in FHZ. In particular, they are smaller than the treatment effects. Fifthly, we find evidence for the relevance of social preferences and fairness considerations with respect to payoff asymmetries.

Summarizing, we find that not only contract types and their corresponding reference point effects determine sellers shading behavior but also the process by which contract types are selected. In our experiment sellers take into account buyers' behavior in all stages of the experimental game while Fehr et al. (2009b) find evidence only for the feasible price range that remains after the auction phase and may cause aggrievement and perfunctory performance.

Hart and Moore (2008) as well as FHZ emphasize that the existence of reference point effects is dependent on ex ante competition. According to both papers prices don't serve as reference points without ex ante competition. Unfortunately, the experimental design of FHZ establishes only a one-sided competition and gives all market power to the buyers. There isn't any competition between buyers at all. We believe that this does not really correspond to the fundamental transformation as put forward by Williamson (1985). Therefore, a Double Auction in place of the seller auction might be a better test of Hart and Moore's theoretical approach which has to be left for future research.

Finally, in many experiments we find that environments in which subjects have only few or no information about trading partners' payoffs induces behavior that is closer to the standard economic approach. Since information about trading partners' payoffs is often highly incomplete in real life, it remains to be seen whether the actual experimental results also carry over to such situations. Again, this question has to be left for future research.

References

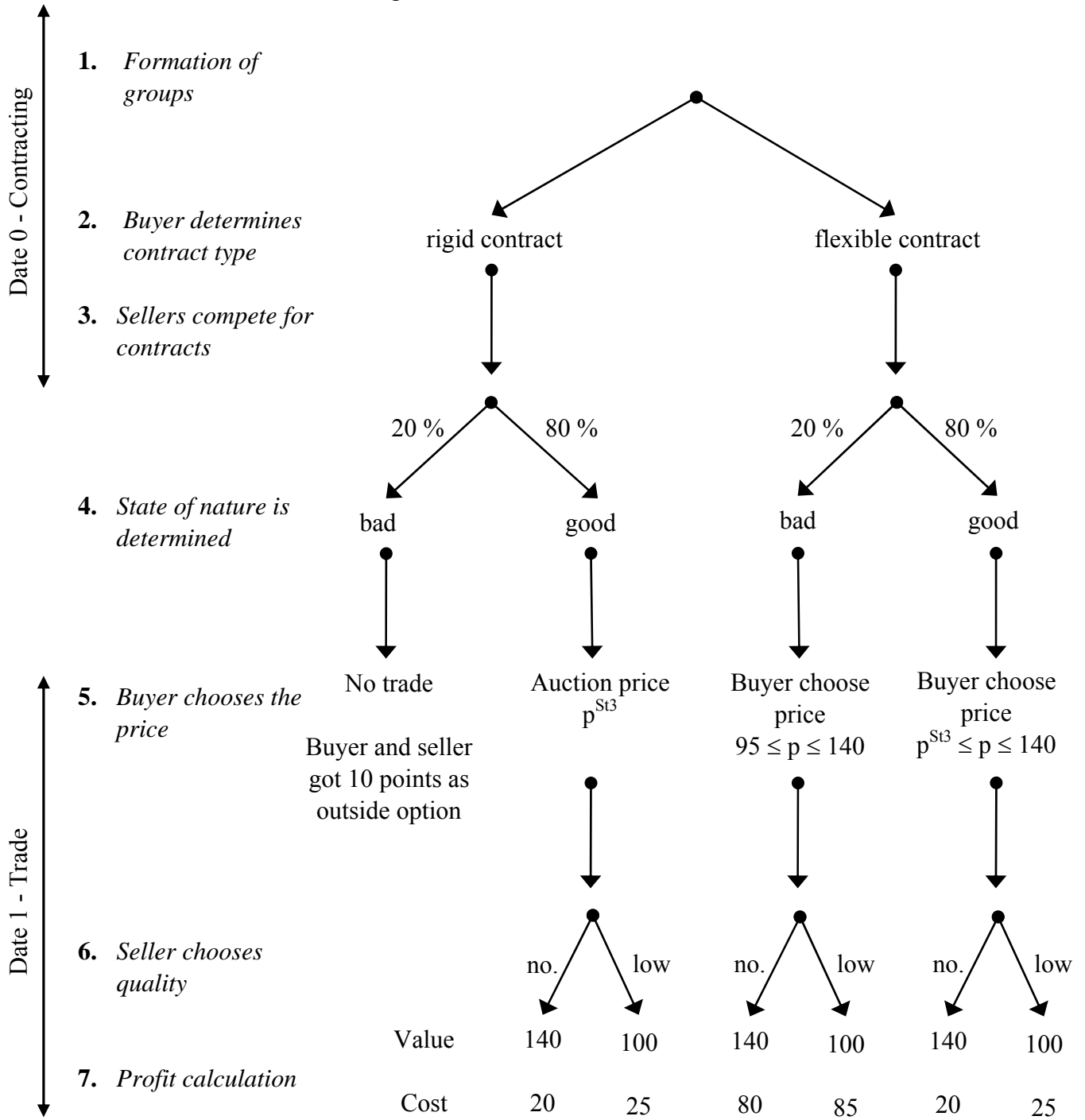
- Fehr, E., Hart, O., and Zehnder, C. (2009a): “Contracts, Reference Points, and Competition – Behavioral Effects of the Fundamental Transformation”, *Journal of the European Economic Association*, 7(2-3): 561-572.
- Fehr, E., Hart, O., and Zehnder, C. (2009b): “Contracts as Reference Points – Experimental Evidence”, mimeo, Harvard University.
- Fehr, E., and Schmidt, K. (1999): “A Theory of Fairness, Competition, and Cooperation”, *Quarterly Journal of Economics*, 114(3), 817-868.
- Fischbacher, U. (2007): “z-Tree: Zurich Toolbox for Ready-made Economic Experiments”, *Experimental Economics*, 10(2), 171-178.
- Forsythe, R. L., Horowitz, J., Savin, N. E., and Sefton, M. (1994): “Fairness in simple bargaining games”, *Games and Economic Behavior*, 6, 347-369.
- 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. (2008): “Contracts as Reference Points”, *Quarterly Journal of Economics*, 123(1), 1-48.
- Williamson, O. (1985): “The Economic Institutions of Capitalism”, New York: Free Press.

Appendix

Table 8: Aggregate Results

Treatment	Endogenous Contracts		Exogenous Contracts	
Contract Type	Rigid	Flexible	Rigid	Flexible
Rel. Freq. of Contracts	.277	.723	.48	.52
Auction Price (mean)	38.69	37.41	38.44	37.03
Final Price (mean)	38.69	47.57 / 97.8	38.44	48.38 / 98.87
Rel. Freq. of low quality	.322	.341	.147	.184
Profit Buyer (mean)	65.69	61.2	73.12	70.88
Profit Seller (mean)	15.1	22.44	15.91	24.94

Figure 2: Structure of Each Period



Legend:

- low – low quality
- no. – normal quality
- p^{st3} – auction price