LE
7/25/08
12:40 pm

Contracts as Reference Points – Experimental Evidence

Ernst Fehr*, Oliver Hart† and Christian Zehnder‡

First Draft: April 2008

Abstract

In this paper we test experimentally the empirical relevance of the behavioral forces suggested by Hart and Moore (2008). Our evidence confirms the model's prediction that there is a trade-off between rigidity and flexibility in a trading environment with incomplete contracts and ex ante uncertainty about the state of nature. Our findings support the hypothesis that a contract constitutes a reference point for a trading relationship. It appears that a party's ex post performance depends on whether he gets what he feels entitled to relative to outcomes permitted by the contract. Flexible contracts allow for adjustments of terms, such that trade is always feasible. However, since flexible contracts allow for many outcomes, discrepancies between expected and realized payoffs are likely. It seems that parties are often disappointed by the realized outcome, because the contract allowed for better terms. As a consequence flexible contracts cause a significant amount of shading. In rigid contracts, in contrast, the fixed terms of the contract imply that trade cannot occur in all states. But since outcomes are pinned down from the beginning, the trading parties know exactly what to expect and so there is not much room for disappointment. Accordingly, when trade takes place shading is much less frequent in rigid contracts than in flexible contracts.

Keywords: Contracts, Reference Points, Experiment

* University of Zurich, efehr@iew.uzh.ch.
† Harvard University, ohart@harvard.edu.
‡ Harvard Business School, czehnder@hbs.edu

I. Introduction

In this study we test experimentally the empirical relevance of the behavioral forces suggested by Hart and Moore (2008). Their theory is based on the assumption that contracts provide reference points for trading relationships. The idea is that the terms of the contract shape the trading parties' entitlements regarding ex-post outcomes. Specifically, they assume that each trader evaluates the ex-post outcome relative to the best outcome permitted by the contract. A flexible contract, which allows for many possible outcomes, is therefore likely to lead to conflicting expectations. If a party feels that he is not getting what he is entitled to, he may shade on performance and create a deadweight loss. These assumptions imply a trade-off between contractual flexibility and rigidity. The advantage of rigid contracts is that they pin down the outcomes from the outset. Accordingly, there is not much room for aggrievement because every trading party knows exactly what to expect. The downside of rigid contracts is that they do not allow the trading parties to adjust the contract terms to the realized state of nature. As a consequence, trade does not occur in all states. Flexible contracts, in contrast, enable the parties to refine the contract terms and thereby guarantee trade. At the same time, however, these contracts create misaligned entitlements, such that efficiency losses due to shading activities are more likely.

The theory of Hart and Moore (2008) adds an alternative and complementary perspective to the literature on incomplete contracts (see Grossman and Hart 1986, Hart and Moore 1990). While previous models rely on efficient ex-post renegotiation of contracts and identify ex-ante underinvestment as the source of inefficiency, the new theory investigates the consequences of imperfectly contractible ex-post trade. In addition, Hart and Moore (2008) introduce important behavioral elements. While some of the assumptions underlying the model are broadly consistent with well established and tested 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 reciprocity (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.

We implement an experimental market where buyers and sellers meet to trade. Each transaction takes place over a time span of two dates. The trading parties meet and contract at date 0 and they trade zero or one unit of a good at date 1. There are two states of nature: a good state, in which the seller's production costs are low, and a bad state, in which the costs are high. At date 0 there is uncertainty about the state of nature, so that a flexible contract that admits several outcomes may be desirable. Contracting involves two steps. In the first step the buyer determines what type of contract he wants to conclude, i.e., he chooses between a flexible and a rigid contract. In a rigid contract the price is fixed, a flexible contract defines a price range out of which the price can be picked after the state of nature has been realized. In the second step the contract is auctioned off to the sellers. The auction not only determines which of the sellers gets the contract, but it also defines the contract terms (i.e., the fixed price in rigid contracts and the price range in flexible contracts). Both flexible and rigid contracts are incomplete in the sense that the seller's performance is only partly contractible, i.e., the contracts cannot prevent the seller from ex-post shading on performance. Shading takes place in form of costly sabotage, i.e., the seller spends resources to lower his quality from normal to low. The uncertainty about the state of nature is resolved at date 1. However, while the state of nature is observable to the seller and the buyer, it is not verifiable to a third party like the court. As a consequence, the trading parties cannot write state contingent contracts at date 0. Trade takes place at date 1 only if the contract concluded at date 0 allows for a mutually beneficial outcome in the realized state of nature. Our parameters are such that the rigid contract allows for trade only in the good state, while trade is always feasible in the flexible contract. In the no-trade case the parties realize an outside option. If trade occurs the buyer refines the contract by picking a price admitted by the contract. In the flexible contract he can choose any price in the price range as long as it covers the seller's costs. In the rigid contract only the fixed price is available. After the buyer's price choice, trade takes place and the seller determines his quality.

Under the assumptions of the standard economic model 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 auctions 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 are 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 very 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. Thus, while sellers should always choose the normal quality in rigid contracts, their quality choice in flexible contracts may be dependent on the price. In response, buyers may either increase the price in flexible contracts to avoid shading, or accept the high probability of getting low quality. In any case, the reference dependent behavior of sellers has a negative impact on the buyer's profit in flexible contracts. Thus, if the willingness to engage in shading is strong enough, buyers may find switching to rigid contracts profitable.

Our results are largely in line with the predictions implied by the assumption that contracts are reference points for trading relationships. 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. Accordingly, there is no significant difference between the fixed price in rigid contracts and the lower bound of the price range in flexible contracts. However, despite the fact that, in principle, the buyers have the possibility to pay the same prices in both types of contracts when the good state of nature is realized, we observe that buyers pay significantly higher prices in flexible contracts. The reason is the following: while most sellers provide normal quality irrespective of the price in rigid contracts, a considerable fraction of sellers is willing to deliver high quality in flexible contracts only if the buyer pays a high price. This behavior is exactly in line with the prediction of the theory of Hart and Moore (2008). In rigid contracts the sellers are not disappointed by the low prices because these prices were already determined in the auction and the sellers knew what to expect. In flexible contract, in contrast, sellers may hope for a high price and they may be disappointed if they do not get it. Thus, the only way for buyers to reduce the probability of shading in a flexible contract is to pay higher prices. However, in our experiment the price increase in flexible contracts is not sufficient to prevent shading completely. While we observe almost no shading in the rigid contracts, the relative frequency of low quality is substantially higher in flexible contracts.

The fact that buyers who have chosen flexible contract pay higher prices but get lower quality in the good state of nature has a negative impact on the average profitability of flexible contracts for buyers. This negative effect is large enough to offset the advantage that flexible contracts guarantee trade in both states of nature, i.e., in contrast to the prediction of the standard economic model rigid contracts generate higher profits for buyers than flexible contracts. As a consequence many buyers choose the rigid contract at date 0.

The remainder of 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 and discuss our results in section IV. Section V 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 of widgets 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 to a third party. 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 of the widget.

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 sellers' 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 costs) 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 happens to be unable to trade 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.

[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).

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 conclude 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 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: $p^r \in [c(q^l,g) + x_s, m] = [35, 75]$.[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.

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

[5] The lower bound of this interval 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 upper bound of the interval makes sure that the price is always below 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.

beneficial transaction can take place or not. Trade can always take place when the buyer 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 contracts 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 setups 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). 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 $ 1, in March 2008).

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 computerized experiment was programmed and conducted with the experimental software 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 $ 49, in March 2008).

III. Behavioral Predictions

In this section we derive the predictions for our experiment. Section III.A contains a game theoretic analysis of our treatment conditions for the case where all market participants are rational profit-maximizers. Under this assumption sellers never shade on performance and therefore buyers unambiguously prefer flexible contracts, because they guarantee trade in both states of nature. The model of Hart and Moore (2008) suggests, however, that this standard economic analysis may be incomplete, because it ignores the fact that contracts may provide reference points for trading relationships. In Section III.B we therefore examine the implications of reference dependent preferences in our experimental market setup. We assume risk neutrality and rationality, but we suppose that the trading parties evaluate their ex post outcomes relative to the outcomes permitted by the contract. Parties who feel that they do not get what they are entitled to may be aggrieved and therefore shade on performance. We show that the presence of such preferences generates an important new trade-off between rigidity and flexibility. While rigid contracts still have the disadvantage that they prohibit trade in the bad state, a fixed price may now help to avoid inefficient shading. Accordingly, if contracts provide reference points, rigid contracts may be much more attractive to buyers than predicted by the standard economic model.

The model of Hart and Moore (2008) is very general in many dimensions. For the sake of feasibility our design adds a lot of structure and contains many simplifications. While we feel that our design is suitable to investigate the empirical relevance of the assumed behavioral mechanisms, we are aware that there are many other potentially interesting designs. In section III.C we therefore discuss in how far the predictions for the experiment depend on our specific design and we explain why we think that our design is especially useful to investigate the questions we are interested in.

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, end 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 under as 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.

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.
c) Buyers always choose the lowest price available in flexible contracts.
d) Buyer's profits are higher in flexible contracts than in rigid contracts. Therefore, buyers predominantly choose flexible contracts.

III.B. Predictions if Contracts are Reference Points

In the following we derive an alternative prediction based on the behavioral assumptions suggested by Hart and Moore (2008). In particular, while we stick to the assumptions of rationality and risk-neutrality, we also consider the possibility that contracts provide a reference point for trading relationships. The idea is that contract terms shape entitlements, i.e., the trading parties evaluate their ex post outcomes relative to the outcomes permitted by the contract. If a party feels that he is not getting what he is entitled to, he may shade on performance and create a deadweight loss. In the context of our setup the model of Hart and Moore (2008) suggests that the contracts concluded at date 0 determine reference points relative to which the trading parties evaluate the outcome obtained at date 1. Specifically, sellers may compare the realized outcome at date 1 to their preferred outcome among the outcomes admitted by the contract. If the realized outcome is less beneficial than their preferred outcome, they may be disappointed and suffer from a direct psychic loss. However, sellers may be able to offset this psychic cost by shading on performance and therewith lowering the payoff of the buyer. In other words, disappointed sellers can regain utility by transferring the hurt back to the buyer. Therefore, they may be willing to shade on performance in response to a disappointing outcome, even though shading is somewhat costly.

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 reference price p^R can be any price that is available to the buyer within the terms of the concluded contract. Since the state of nature determines the available gains from trade, 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 buyer can only pay the fixed price. Accordingly, 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. We suppose that sellers may be aggrieved if the realized price p is smaller than their reference price p^R. The parameter θ measures the intensity of their aggrievement. However, we also assume that a seller can offset his aggrievement if he shades on performance and thereby hurts the buyer by lowering his valuation for the delivered product.[9]

This behavioral assumption about sellers' behavior has important implications for the predicted outcome of our experiment. We will show that there is a trade-off between flexibility and rigidity if contracts provide reference points for trading relationships. We begin with the rigid contracts. The advantage of rigid contracts is that they never lead to aggrievement. The reason is that the seller always gets what he expects. Since the buyer has no other possibility than to pay the fixed price which has been determined in the competitive auction, the fixed price is also the seller's reference price. Thus, whenever the good state is realized, trade occurs and there is no shading. If the bad state is realized, trade does not take

[9] 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.

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, aggrievement may play a role implying that the seller's performance may no longer be 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 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:

$$p^f{}_s = \arg\max v(q^n)F_s(p) + \mathrm{v}(q^l)[1 - F_s(p)] - p.$$

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. In some cases it may be optimal that the buyer increases the price above the lower price bound in order to avoid shading, in other cases the buyer may prefer to pay a low price and accept the risk of suffering from shading. In any case, however, the fact that contracts provide reference points for trading relationships has a negative impact on the buyer's profit in flexible contracts relative to the standard prediction. The buyer's expected payoff if he chooses a flexible contract is:

$$\mathrm{E}\pi_B{}^f = w^g[v(q^n)F_s(p^f{}_g) + \mathrm{v}(q^l)(1\text{–}F_s(p^f{}_g)) - p^f{}_g] + (1 - w^g)[v(q^n)F_s(p^f{}_b) + \mathrm{v}(q^l)(1\text{–}F_s(p^f{}_b)) - p^f{}_b].$$

Sellers are willing to accept a contract at any available price (see Footnotes 8 and 9 for details). Competitive forces in the contract auctions will therefore drive the fixed price in rigid contracts and the lower bound of the price range in flexible contracts down to the competitive level ($p^r = p^l = 35$). When the buyer has to choose the contract type at date 0, he anticipates the competitive outcome of the auction. But unlike in the case of purely selfish preferences he now faces a trade-off: While a rigid contract prohibits trade in the bad state of nature, it allows the buyer to pay a low price without causing aggrievement and shading in the good state of nature. A flexible contract, in contrast, guarantees trade in both states of nature, but the buyer has to increase the price above the lowest available level and/or bear the consequences of shading in both states of nature. Which of the two contracts is more profitable depends on how strong the negative impact of shading on buyer profits in flexible contracts is relative to the costs caused by the absence of a trade opportunity in the good state of nature in rigid contracts. These considerations allow us to formulate 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 sellers' preferences implies that the frequency of shading is decreasing in the price.
c) Given the price dependence of quality, buyers may not choose the lowest price available in flexible contracts.
d) 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 and Predictions

Hart and Moore (2008) make strong behavioral assumptions. Specifically, they assume that contracts provide reference points for trading relationships and that trading parties engage in shading if they do not get what they felt entitled to. They theoretically investigate how the presence of these behavioral forces influences the optimal structure of contracts under various conditions. It is important to emphasize that the aim of this paper in not to investigate 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 model turn out to be of empirical relevance. To study this question in a clean and controlled way, we intentionally abstract 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 all decision rights are given to sellers.[10] In other words, the trade-off between contractual flexibility and rigidity would be destroyed.[11] 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 only consider 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, only determines 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: $p^h = v(q^n)$).[12] Studying these extreme cases gives us the best chance to illustrate the trade-off between rigidity and flexibility in contracts.

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 implemented 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 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

[10] Specifically, if our setup would allow for any type of contract, the first best could always be achieved by leaving both, the choice of the contract type and the price setting to the seller. In equilibrium sellers would choose the flexible contract, accept the competitive lower price bound at date 0, choose a price equal to the buyers valuation at date 1 ($p = v(q^n)$), and provide normal quality ($q = q^n$). Redistribution of surplus from sellers to buyers (because of competition) could be achieved through lump sum payments conditional on winning the contract auction.

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

[12] 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.

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.[13]

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 only interesting 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 ($w^g = 0.8$).[14]

IV. Results

In this section we present and discuss our results. We analyze the outcomes in rigid and flexible contracts in subsection IV.A. In strong contrast to the prediction of standard economic theory we observe that rigid and flexible contracts exhibit significant differences in average prices and quality levels when the good state of nature is realized. We demonstrate that these differences are largely in line with our reference point hypothesis and can be explained in light of sellers' performance choices in the two contract types. In Section IV.B we discuss our findings and relate them to the existing literature. We argue that some of our findings are related to previous findings in the experimental fairness literature. However, we also show that existing fairness models cannot explain our results. The reason is that these models do not account for the fact that contracts determine the reference points relative to which the trading parties evaluate realized outcomes. As a consequence, they fail to predict that the same seller may behave very differently in rigid and flexible contracts.

IV.A. Outcomes in Rigid and Flexible Contracts

In the following we investigate whether our reference point hypothesis accurately predicts the outcome of our experiment. The reference point hypothesis is based on the behavioral

[13] Alternatively we could also have chosen to make the provision of low quality slightly less costly than the provision of normal quality. However, we expect that 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 does not necessarily imply the positive sentiment necessary to induce costly cooperation. Of course, this remains an empirical question that should be addressed in future work.

[14] 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 (we will discuss this point in much more detail in section IV.B).

assumption that contracts serve as reference points for trading relationships (Hart and Moore, 2008). The reference point hypothesis provides three testable predictions: First, the auction process used to assign contracts should induce competition among sellers. As a consequence, the auction outcomes for both contract types – the fixed price in rigid contracts and the lower bound of the price range in flexible contracts – should reflect the competitive price of 35. Second, if contracts define reference points, the contract type may affect sellers' performance choices. Since rigid contracts pin down outcomes, sellers should not be aggrieved, because they always get what they expect. Accordingly, we should not observe shading in rigid contracts. In flexible contracts, in contrast, sellers may be aggrieved if they get a lower price than they had hoped for and this may trigger costly shading. Thus, we would expect that sellers always choose the normal quality in rigid contracts, while their quality choice in flexible contracts may depend on prices. Third, in response to the price dependence of a seller's quality choice buyers may pay prices above the lower price bound in flexible contracts. If pronounced enough, the difference in sellers' behavior across contract types may also influence the buyers' contract choices. Specifically, if the willingness to engage in shading is strong, buyers should switch to rigid contracts.

As a first step we present an overview of our experimental results, i.e., we compare the outcomes in rigid and flexible contracts at the aggregate level. We find that our data are largely in line with all three parts of the reference point hypothesis. As a second step we show that not only the aggregate outcomes, but also the underlying behavioral regularities are supportive of the view that contracts indeed provide reference points for trading relationships.

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

We start with the competition part of the reference point hypothesis. Figure I provides an illustration of 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.[15]

Insert Figure I here

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). 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.[16] 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)).[17]

Interestingly, however, the higher prices in flexible contracts are not always sufficient to prohibit sellers from shading. Sellers provide low quality in 25 percent of the flexible contracts in which the good state is realized (see Table II). Despite the substantially lower prices shading is much less frequent in rigid contracts. Low quality is provided only in 6 percent of rigid contracts. The difference in the frequency of shading between the two

[15] 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.

[16] The "auction outcome" in Table II is the average of the fixed prices in all rigid contracts concluded at date 0. 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.

[17] 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.

contract types is statistically significant (non-parametric signed rank test, p-value = 0.031 (one-sided)) and very stable over time (see Figure I).[18]

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 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.[19] 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

[18] 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%.
[19] In the bad state of nature the session level frequencies of high quality in flexible contracts are 70%, 46%, 83%, 58% and 83%.

signed rank, p-value = 0.031 (one-sided)).[20] 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)).[21] 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 total, 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 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.[22] While the presence of rigid contracts contradicts the standard prediction and supports the reference point hypothesis and the positive time trend indicates that learning over time may play an important role, it is still somewhat surprising that the large difference in average profits has not motivated more borrowers to select rigid contracts. We will discuss possible reasons for this finding in more detail later on.

The discussion of Table II and Figure I has shown that in the aggregate our findings are supportive for the reference point hypotheses. Next, we analyze whether the underlying behavioral patterns are also in line with the assumption the contracts provide reference points for trading relationships. We start with seller's 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.

[20] 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.
[21] 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.
[22] 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).

There is no noteworthy correlation between prices and the frequency of normal quality.[23] 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 then 60 percent of the contracts. The frequency of normal quality is clearly increasing in prices, 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. 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). 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

[23] 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).

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

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 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 do 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]

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

Insert Table IV here

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 report of the regression results in Footnote 22), it is rather astonishing that not more buyers have chosen the rigid contract. One explanation for this observation may be that the advantages of rigid contracts over flexible contracts are subtle and hard to understand. The participants' instructions explain in detail that flexible contracts allow for trade in both states, while trade in rigid contracts is confined to the good state of nature. However, the participants have to understand or learn the fact that flexible contracts may also lead to adverse behavior. Within the experiment buyers have two important sources of information: their own experience and the market information provided at the end of the each period. Learning from own experience alone is rather difficult. The reason is that there are many possible constellations (i.e., combinations of contract type, state and price) and only few periods to explore them. That is why we provided buyers with the market information (average profits in both contract types in all previous periods, market share of each contract type in current period). However, since average profits are determined by realized states of nature, prices and quality choices, it is still not trivial to understand the behavioral mechanisms underlying the price difference. In fact, in our post-experimental questionnaires only about 30 percent of the buyers mention the price-dependence of quality in flexible contracts as a reason for not choosing flexible contracts (in response to the open question: "Have you chosen contracts with a price range? If so, why? If not, why not?"). A second reason may be found in buyers' risk aversion. Even if buyers understand the behavioral forces at work, they may still 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.

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 prices 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 only chosen 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.

IV.B. Discussion

Our results support the hypothesis that contracts provide reference points for trading relationships. When the buyer chooses a rigid contract and a competitive auction determines the price ex-ante, sellers do not engage in shading even if the realized price is very low. When the buyer implements a flexible contract, which allows him to adjust the price ex-post, many sellers cooperate only if the buyer is willing to pay a high price, otherwise they engage in costly sabotage. To the best of our knowledge this effect has not been documented empirically before. However, part of our interpretation of the data relies on the well known fact that many people are willing to punish their trading partners for inappropriate behavior, even if this is costly for them and yields no material gain. Since such behavior has been extensively studied in theoretical and experimental 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 consideration, especially fairness concerns, into account. The simplest versions of these models are theories of inequity aversion (Fehr and Schmidt 1999, Bolton and Ockenfels 2000), which 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 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 very 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. However, while inequity aversion is in line with sellers' behavior in flexible contracts, it cannot explain why sellers do not shade in rigid contracts. Since models of inequity aversion are based solely on the allocation of final payoffs they would predict the same price dependency of quality in both types of contracts.[25]

At date 1 the main difference between the two contract types in our experiment is that rigid contracts allow for only 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 can offer only an unequal division of the pie and a setup where the proposer can also choose more balanced allocations of payoffs. Falk et al. (2003) implement

[25] Using the models of inequity aversion to derive full-fletched 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.

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, which gives 5 points to each player and an unequal allocation in which he gets 8 points while the responder is left with 2 points. In another treatment the proposer can offer only the unequal allocation. The evidence shows that the frequency of rejections if the proposer chooses the unequal allocation strongly depends on whether the equal split is available to the proposer or not. When the equal split could have been chosen, the unequal offer is rejected in roughly 45 percent of the cases. When the unequal allocation was the only alternative available to the proposer, the rejection frequency was only 18 percent. While these findings are not in line with models of inequity aversion, they can be explained by more sophisticated theories of reciprocity which take into account that the evaluation of an outcome may depend not only on the realized allocation of payoffs but also on people's belief about the intentions of their trading partner (see Falk and Fischbacher 2006).[26] Intuitively, intention based models explain the different rejection rates of the unequal allocation in the two treatments of Falk et al. (2003) as follows: If the proposer chooses the unequal allocation when the equal split would have been available, it is clear that the proposer intentionally created an "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.

While the results of Falk et al. (2003) appear to be in line with our finding that there is less shading in rigid contracts than in flexible contracts, there are fundamental differences between the two experiments. In the ultimatum games of Falk et al. (2003) the proposer's strategy space is exogenously restricted by the experimenter. 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. To understand this point, imagine a version of the ultimatum games of Falk et al. (2003) in which it is common knowledge that the proposer can choose the treatment he wants to play. In this case the reciprocity theory of Falk and Fischbacher (2006) predicts that there is no longer a difference between the rejection rates for the unequal allocation across treatments. The reason is that the proposer would always have had the possibility to implement the equal

[26] Falk and Fischbacher (2006) develop a model, in which payoffs and intentions matter for the fairness evaluation of an outcome. For theories, in which reciprocity is modeled as entirely intention driven, see also Rabin (1993) and Dufwenberg and Kirchsteiger (2004).

split. If he chooses a path that ends up in the unequal allocation, this would be interpreted as intentional unfairness and lead to many rejections. Thus, if we take the view that sellers and buyers anticipate that competition leads to low auction outcomes, intention based reciprocity models cannot explain why there is less shading in rigid contracts than in flexible contracts.

A crucial difference between the fairness models discussed above and Hart and Moore (2008) is that the latter explicitly assume that the presence of competition at date 0 is the reason why the contract concluded at date 0 turns into a reference point for the non-competitive transaction at date 1. The idea is that competition adds an objective dimension to the bargaining process, such that the trading parties accept the outcome of a competitive market as "acceptable" or "fair". As a consequence, sellers do 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]

V. 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 implements 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 D 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.

In order to investigate the existence of reference-dependent shading activities in trading relationships in a clean and controlled way, we implemented an experimental setup, which intentionally abstracts from interesting aspects of Hart and Moore (2008). While contract choice is endogenous in our experiment, the structure of the contracts and the degree of flexibility are predetermined. In future work it would be very interesting to set up experiments which allow the investigation of truly endogenous contract structures. This would allow testing interesting predictions of Hart and Moore (2008) regarding optimal contracts and could thereby shed new light on the behavioral foundations of long-term trading agreements and the employment relationship.

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.

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., 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.

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. 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	
R2	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

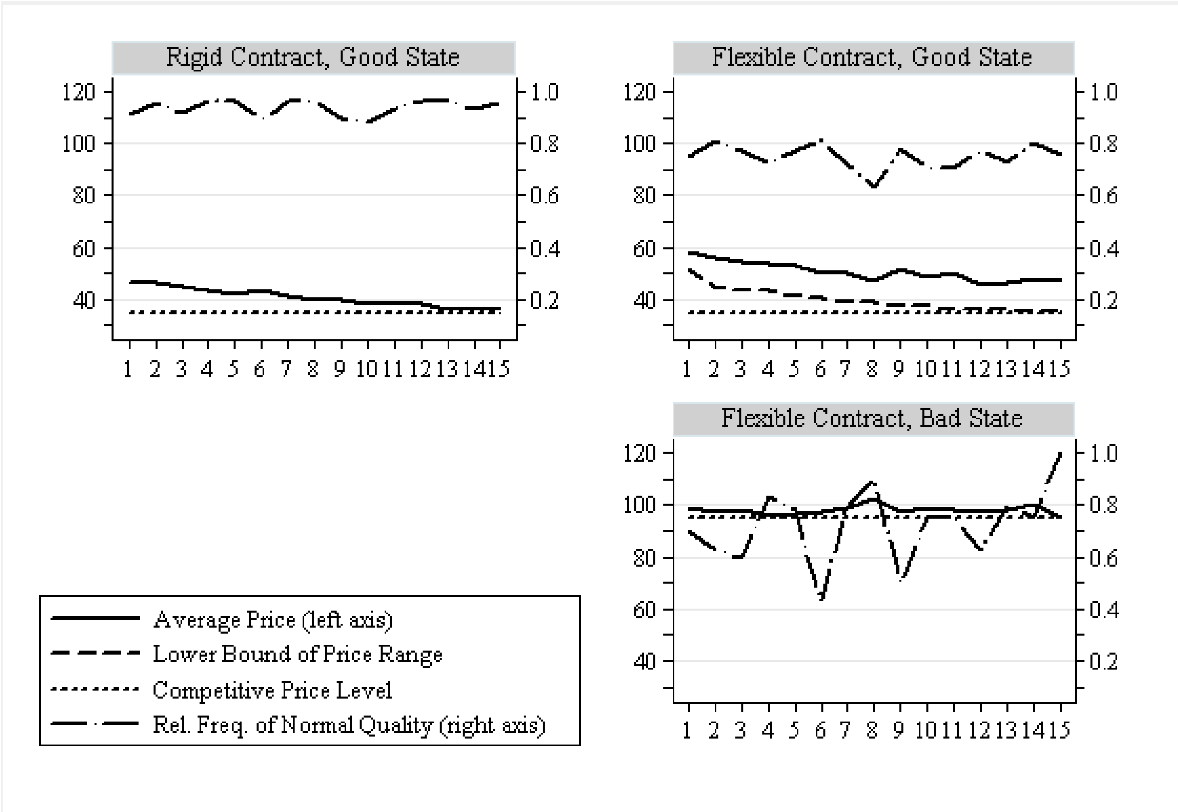

Figure II: Price Distribution and Frequency of Normal Quality Conditional on Price

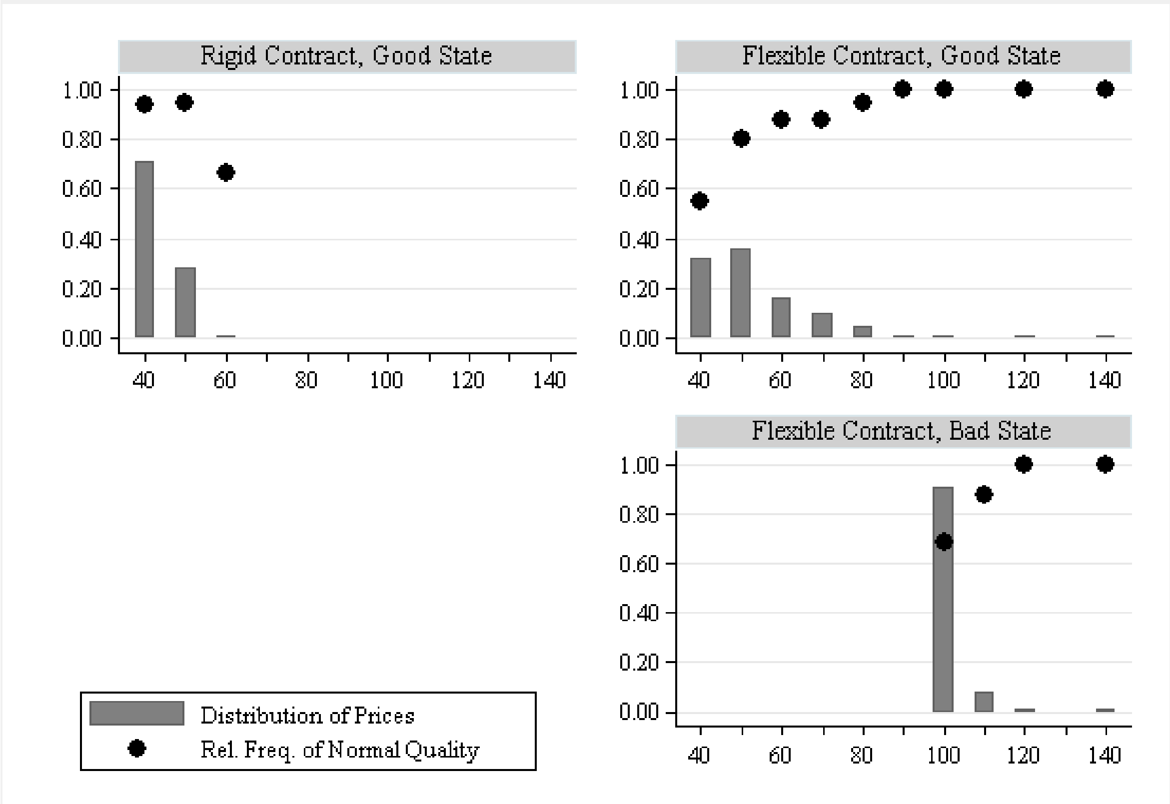

Figure III: Accepted Number of Rigid and Flexible Contracts per Individual

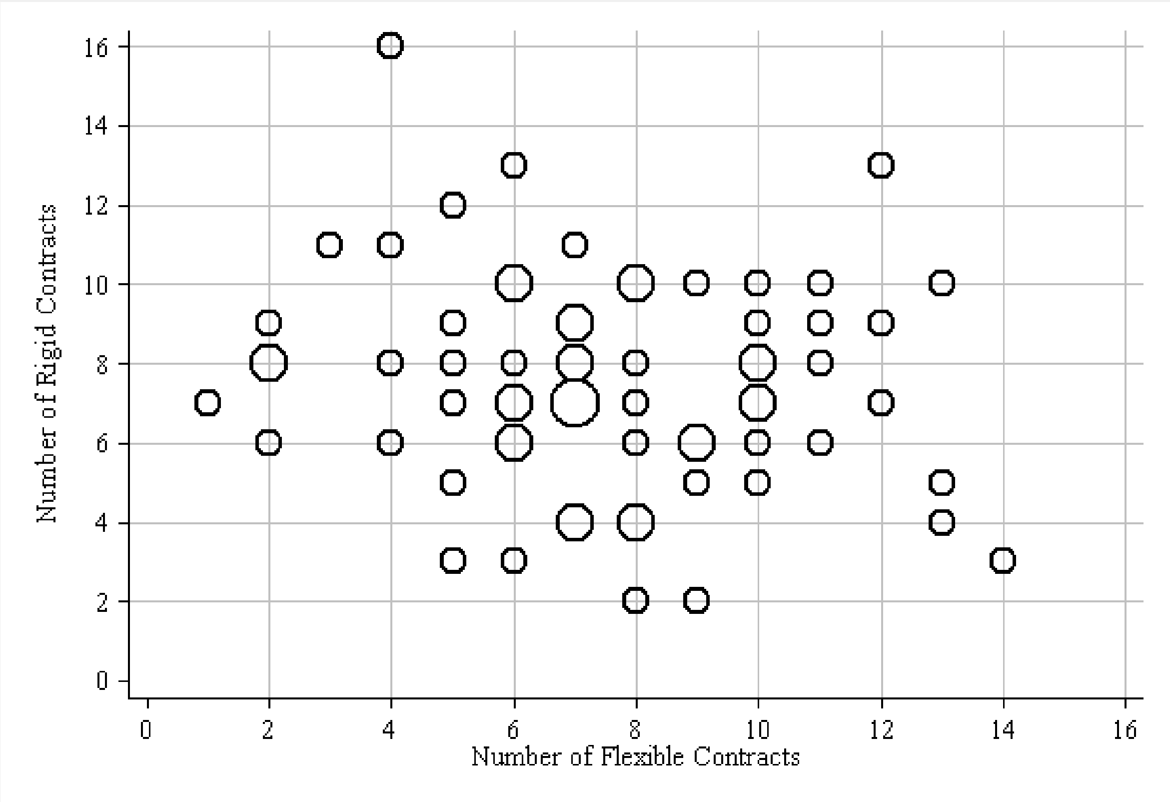

Figure IV: Price Distribution and Average Profit of Buyer conditional on Price