

Contracts as Reference Points—Experimental Evidence

By ERNST FEHR, OLIVER HART, AND CHRISTIAN ZEHNDER*

Hart and John Moore (2008) introduce new behavioral assumptions that can explain long-term contracts and the employment relation. We examine experimentally their idea that contracts serve as reference points. The evidence confirms the prediction that there is a trade-off between rigidity and flexibility. Flexible contracts—which would dominate rigid contracts under standard assumptions—cause significant shading in ex post performance, while under rigid contracts much less shading occurs. The experiment appears to reveal a new behavioral force: ex ante competition legitimizes the terms of a contract, and grievement and shading occur mainly about outcomes within the contract. (JEL D44, D86, J41)

In a recent paper Hart and Moore (2008) develop a theory which provides a basis for long-term contracts in the absence of noncontractible investments. The theory is also capable of rationalizing the employment contract, which fixes wages in advance and leaves discretion to the employer. However, the theory rests on strong behavioral assumptions that lack direct empirical support and deviate in important ways from the assumptions made in standard contract theory. For this reason, and because the theory has the potential to cast new light on the theory of the firm¹, it seems important to test it. In this paper, we provide such a test by carrying out a controlled laboratory experiment. In doing so, we identify new behavioral forces that cannot be explained either by traditional contract theory or by existing behavioral models. These forces are, however, predicted by the Hart-Moore notion that competitively determined contracts serve as reference points.

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

*Fehr: Institute for Empirical Research in Economics, University of Zurich, Blümlisalpstrasse 10, CH-8006 Zürich (e-mail: efehr@iew.uzh.ch); Hart: Department of Economics, Harvard University, Littauer Center 220, 1805 Cambridge Street, Cambridge, MA 02138 (e-mail: ohart@harvard.edu); Zehnder: Faculty of Business and Economics, University of Lausanne, Quartier UNIL-Dorigny, Internef 612, CH-1015 Lausanne (e-mail: christian.zehnder@unil.ch). We would like to acknowledge financial support from the US National Science Foundation through the National Bureau for Economic Research and the Research Priority Program of the University of Zurich on the “Foundations of Human Social Behavior.” We would like to thank Philippe Aghion, Björn Bartling, Gary Charness, Tore Ellingsen, Ray Fisman, Bob Gibbons, Steve Leider, Bentley MacLeod, Al Roth, Klaus Schmidt, Birger Wernerfelt, and three anonymous referees for helpful comments.

¹ See, e.g., Hart (2009) and Hart and Bengt R. Holmstrom (2010).

While this approach has been useful for studying asset ownership (see Sanford J. Grossman and Hart 1986, Hart and Moore 1990), it has been less useful for studying the employment relationship and the internal organization of large firms. In order to broaden the approach, Hart and Moore (2008) develop a theory based on the idea that an *ex ante* contract, negotiated under competitive conditions, shapes parties' entitlements regarding *ex post* outcomes. The basic idea is that parties do not feel entitled to outcomes *outside* the contract but may feel entitled to different outcomes *within* the contract. If a party does not get what he feels entitled to, he is aggrieved and shades by providing perfunctory rather than consummate performance, causing deadweight losses. This approach yields a trade-off between contractual rigidity and flexibility. A flexible contract is good in that parties can adjust to the state of nature, but bad in that there can be a lot of aggravation and shading. A rigid contract is good in that there is little shading, but bad in that parties cannot adjust to the state of nature.

Hart and Moore (2008) argue that it is the combination of *ex ante* competition and *ex post* lock-in—what Oliver E. Williamson (1985) calls the “fundamental transformation”—that makes an initial contract a useful and salient reference point. Competition provides objectivity to the contract terms, since market forces define what each party brings to the relationship, and market participants will therefore perceive the initial contract to be “fair.” Of course, it is possible that parties will perceive a contract to be “fair” even if it is not negotiated under competitive conditions. To determine the range of applicability of the Hart-Moore model it is important to know the answer to this question. We will therefore investigate the role of *ex ante* competition in our experiment.

Although some of the assumptions underlying the Hart-Moore model are broadly consistent with well-established behavioral concepts such as reference-dependent preferences (e.g., Daniel Kahnemann and Amos N. Tversky 1979; Botond Köszegi and Matthew Rabin 2006), self-serving biases (e.g., Linda Babcock and George Loewenstein 1997), and social preferences (e.g., Rabin 1993, Fehr and Klaus M. Schmidt 1999), there is not yet any empirical evidence that directly supports the idea that contracts are reference points for trading relationships.²

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

² In the Hart-Moore model it is supposed that each party feels entitled to the most favorable outcome permitted by the contract. This assumption implies that the trading parties exhibit a very strong self-serving bias. Other theories of reference points rely on less extreme assumptions. Köszegi and Rabin (2006), for example, suppose that the reference point is determined by parties' (rational) expectations about what they get in equilibrium. Hart and Moore (2008) illustrate that their theory is robust to such alternative specifications (see page 33 in their paper). Hart and Moore's assumptions about shading behavior are strongly related to the work on fairness in economics (see Fehr and Schmidt 2003 for a review). However, while existing fairness models explain why people may engage in dysfunctional behavior when they feel treated unfairly, these theories do not take into account that the terms of a competitively determined contract may shape peoples' fairness perceptions in an important way. We discuss the implications of the differences between existing fairness models and the theory of Hart and Moore (2008) in much more detail in Section IV.

may be required in the *ex ante* contract. However, under the assumptions of the model, this leads to *ex post* aggrievement and shading.

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

In the experiment buyers and sellers contract and then trade. Each transaction involves two dates. At date 0 the trading parties interact in a competitive market and sign a contract. Supply exceeds demand, so that sellers compete for contracts. After signing a contract, a buyer and a seller form a bilateral relationship. They trade at date 1 if the date 0 contract allows for a mutually profitable exchange. At date 0 there is uncertainty about the state of nature, *i.e.*, the trading parties do not yet know the seller's cost, which can be low (the good state of nature) or high (the bad state of nature). The uncertainty is resolved at date 1. However, while the seller's cost is observable, it is not verifiable.

Date 0 contracts are determined as follows. Each buyer decides whether to offer a rigid or a flexible contract. A rigid contract determines a single (fixed) price, while a flexible contract allows for a range of prices. Flexibility can be helpful, because it allows the price to adjust to the seller's cost at date 1. After a buyer has chosen a contract type, a competitive auction determines which seller gets the contract, and the contract terms. In the case of a rigid contract the auction determines the single (fixed) price; in the case of a flexible contract the auction determines the lower bound of the price range (the upper bound is exogenous). At date 1 the parties observe the seller's cost. Trade is possible only if the date 0 contract includes a price that covers cost. Competition ensures that the price in the rigid contract is sufficiently low that trade is possible only in the good state, while trade is possible in both states in the flexible contract. If trade is possible the buyer chooses a price from those allowed by the contract and the seller chooses quality, that is, whether to shade. Shading has a small cost for the seller but greatly reduces the buyer's value. If trade is impossible, the parties realize their outside options.

Under the assumptions of the standard economic model (rationality, selfishness and subgame perfection), the prediction for this experiment is straightforward. Since shading is costly, sellers should never shade, irrespective of the contract type and the price. Buyers should anticipate the sellers' behavior and therefore always choose the lowest price above seller cost permitted by their contract. The competitive auction used to assign contracts to sellers should ensure that the seller's profit is zero and that the entire surplus from the transaction goes to the buyer. Since only the flexible contract allows for trade in the bad state, the flexible contract yields more surplus than the rigid contract, and buyers should always choose the flexible contract.

However, if the behavioral assumptions of Hart and Moore (2008) apply, the predictions are different. The assumption that contracts are reference points does not affect the prediction concerning the competitive auction outcomes. But if competitively determined contract terms define reference points, the contract type may affect the sellers' quality choice. Since rigid contracts pin down outcomes, sellers get what they expect and should not be aggrieved. Accordingly, shading should not occur in rigid contracts. In flexible contracts, in contrast, sellers may be aggrieved if they get

a lower price than they had hoped for. This may trigger shading. In response, buyers may either offer a higher price or accept the possibility of getting low quality. Either way, the reference dependent behavior of sellers has a negative impact on the buyers' profit in flexible contracts. Thus, if the willingness to engage in shading is strong enough, buyers may find that rigid contracts are more profitable.

The results of the experiment are largely in line with Hart and Moore (2008). The auction process indeed induces strong competition for contracts. Both the fixed price in rigid contracts and the lower bound of the price range in flexible contracts converge to the competitive level over time (the level at which the sellers break even in the good state). However, despite the fact that, in principle, buyers have the possibility to pay the same prices in both types of contracts when the good state is realized, we observe that buyers pay significantly higher prices in flexible contracts. Moreover, depending on the price paid, there is considerable seller shading in flexible contracts in the good state. In contrast, there is almost no shading in rigid contracts. Under the parameter values of the experiment, the rigid contract is more profitable than the flexible contract even though it precludes trade in the bad state. Furthermore, a substantial fraction of buyers choose the rigid contract.

We also carry out two robustness checks. In one we reduce the range of prices in the flexible contract and find that shading declines, as predicted by Hart and Moore (2008). In the other we eliminate *ex ante* competition. We find that contracts no longer serve as reference points; in particular, there is significant shading in rigid contracts. In other words, as hypothesized in the Hart and Moore theory, the fundamental transformation from *ex ante* competition to *ex post* bilateral monopoly is associated with significant behavioral effects that influence the relative attractiveness of rigid and flexible contracts.

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

Our evidence becomes even more puzzling—when viewed through the lens of traditional theories of fairness—if we compare the low frequency of shading in rigid contracts with what happens under flexible contracts. In the latter we observe a lot of shading even though the sellers receive a higher share of the gains from trade than under rigid contracts. This pattern of shading across contract types makes perfect sense, however, if competitively determined contracts provide reference points

which possess special normative status. Our finding that the elimination of ex ante competition significantly increases shading in rigid contracts further reinforces this interpretation. In other words, in the presence of competitively determined contract terms the buyers can hide their unfairness behind the veil of competition because they are no longer blamed for unfair outcomes under rigid contracts, although—by offering a rigid contract—their actions contribute to the unfair outcome. We thus believe that our experiment reveals a new behavioral force: ex ante competition legitimizes the terms of the contract, and aggrievement occurs mainly about outcomes within the contract and not about the contract itself.

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

I. Experimental Design

We present our experimental design in Section IA, discuss the characteristics of our experimental robustness checks in Section IB, and describe the laboratory procedures in Section IC.

A. Details of the Baseline 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 σ . There are two states of nature: a good state ($\sigma = g$), in which the seller's production costs are low, and a bad state ($\sigma = b$), 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: $\pi_S = p - c(q, \sigma)$.

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

TABLE 1—EXPERIMENTAL PARAMETERS

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

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

higher than the production costs for normal quality: $c(q^l, \sigma) > c(q^n, \sigma)$. The idea is that it is most convenient for sellers if they simply do their job. They can, however, sabotage output (at a small cost) if they want to.³ For each unit of the product which a seller cannot sell—either because he lost the contract to the other seller in his trading group at date 0 or because his contract does not allow for a mutually beneficial trade at date 1—he realizes an outside option of $x_S = 10$. When a buyer is unable to trade a unit of the product at date 1, he also realizes an outside option of $x_B = 10$. Table 1 summarizes the cost and value parameters of the experiment:

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.

Step 2: The buyer's contract choice

Before buyers' contracts are auctioned off to sellers, each buyer has to decide which contract type t he wants to offer in this period. It is important to note that the buyer can choose only the type of contract; the actual terms of the contract are determined in the competitive auction later on. Specifically, the buyer can choose a rigid contract ($t = r$) or a flexible contract ($t = f$). A rigid contract fixes the price at date 0. The level of the fixed price p^r is determined in an auction in such a way that p^r lies in the interval $[c(q^l, g) + x_S, 75] = [35, 75]$.⁴ A flexible contract, in contrast, specifies a price range $[p^l, p^h]$ at date 0 out of which the buyer chooses the actual price at date 1. The upper bound of the price range is exogenously fixed and equal to the buyer's valuation of the product when the seller provides normal quality: $p^h = v(q^h) = 140$.

³ This means that our set-up differs from the typical assumptions made in the gift exchange literature (e.g., Fehr, Kirchsteiger, and Arno Riedl 1993; R. Lynn Hannan, John H. Kagel, and Donald V. Moser 2002; Charness, Guillaume R. Frechette, and Kagel 2004; Jordi Brandts and Charness 2004). In these papers the pecuniary incentive for workers (i.e., sellers) is to provide the minimal effort (i.e., quality) level, whereas in our paper the normal quality level maximizes the sellers' earnings.

⁴ The minimum of 35 for the fixed price ensures that the seller cannot make losses relative to his outside option in the good state even if he provides low quality. This feature guarantees that sellers do not refrain from choosing low quality just because they want to avoid losses (loss aversion). The maximum of 75 for the fixed price ensures that the price is always below the seller's cost in the bad state of nature. This guarantees that trade cannot occur if the bad state is realized. However, in the experiment the maximum was never binding.

The lower bound of the price range is determined in an auction in such a way that $p^l \in [35, 75]$, the same interval as for the fixed price in rigid contracts.

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 both cases the auction starts off at 35 and then increases by one unit every half second. Each of the two sellers has a button that allows him to accept the contract at any time during the auction. Thus, the first seller who is willing to accept the displayed fixed price or the displayed lower price bound respectively gets the contract. The seller who loses the auction and does not get the contract realizes the outside option x_S .

Date 1: Trade

Step 4: Determination of the state of nature

After the contracts have been auctioned off to the sellers, a computerized random device determines the state of nature for each contract independently. Both sellers and buyers observe the realized state for their contracts and are informed whether a mutually beneficial transaction can take place or not. Trade can always take place when the buyer 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, and a mutually beneficial transaction is 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 determines the actual trading price. In a rigid contract the buyer does not have a choice, since the price has been fixed at date 0. In a flexible contract, the buyer can choose the price. In the good state the buyer can choose any price $p \in [p^l, 140]$. 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 $p \in [c(q^l, b) + x_S, v(q^h)] = [95, 140]$.⁵

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 that 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 1).

Step 7: Profit calculations

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

⁵ Again we do not allow prices to be such that the seller can make losses by choosing low quality, since we want to avoid the possibility that people refrain from shading because of loss aversion (see also footnote 4).

Step 8: Market information for the buyers

Subsequent to viewing the profit screen buyers also get some aggregated information about the market outcome. Specifically, they are informed about profits of buyers in both contract types averaged over all past periods. Furthermore, they learn how many buyers have chosen the rigid contract and the flexible contract in the current period.⁶

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.

Remark: We have ruled out lump sum transfers and restricted flexible contracts to those in which the buyer chooses the price ex post. We discuss this further in IIIB.

B. Experimental Robustness Checks

We provide two experimental robustness checks.

Reduced Flexibility.—Our baseline treatment compares completely rigid contracts with maximally flexible contracts: the upper bound in the price range is exogenously fixed at the buyers' maximal willingness to pay, $p^h = v(q^h) = 140$. Studying these extreme cases gives us the best chance of illustrating the trade-off between rigidity and flexibility. However, the buyers would prefer a contract with less flexibility: the price never needs to go above $p = c(q^l, b) + x_s = 95$ in order to permit trade in the bad state, and, according to Hart and Moore (2008), a lower upper bound should reduce aggrievement and shading. In fact with an upper bound of 95, buyers with a flexible contract have only one price available in the bad state of nature, $p = 95$, and there should be no aggrievement at all. In our reduced flexibility treatment we replace the upper bound of 140 in the flexible contract with 95.

No Competition.—As noted in the introduction, Hart and Moore (2008) argue that competition will provide objectivity to the contract terms, and parties will therefore perceive the initial contract to be fair. Of course, it is possible that parties will perceive a contract to be fair even if it is not negotiated under competitive conditions. In the no competition treatment we see whether this is the case.

The interaction between buyers and sellers takes place in the same way as in the baseline treatment except for the determination of the contract terms at date 0 (i.e., the fixed price in rigid contracts and the lower bound in flexible contracts). As in the baseline treatment the first step is that the buyer chooses his contract type. However, while the baseline treatment subsequently relies on competitive auctions, the contract terms in the no competition condition are determined by an exogenous random device. The random device draws the terms out of the distribution of auction outcomes generated by the baseline treatment. This guarantees that any behavioral changes between the baseline and the no competition treatments are driven by the difference in the determination of contract terms and not by differences in the contract terms themselves. After

⁶ The aim of the provision of this information was to make learning easier for buyers. Since our set-up allows for many possible constellations (two contract types, two states of nature, two quality levels, many prices), learning from individual experience is rather difficult.

the contract terms have been set for both buyers in an interaction group, each contract is randomly (and independently) assigned to one of the two sellers (i.e., as in the baseline treatment each seller can end up with zero, one, or two contracts).

C. Subjects, Payments, and Procedures

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

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

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

The experiment was programmed and conducted with *z-Tree* (Fischbacher 2007). We conducted five sessions of the baseline treatment, two sessions of the reduced flexibility treatment, and five sessions of the no competition treatment. We had 28 subjects (14 buyers and 14 sellers) in ten of our 12 sessions and—owing to no-shows—24 subjects (12 buyers and 12 sellers) in the remaining two sessions. This yields a total number of 328 participants in the experiment. A session lasted approximately two hours, and subjects earned on average 59 Swiss Francs (including a show up fee of 10 Swiss Francs).

II. Behavioral Predictions

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

A. Predictions under Pure Self-Interest

If we assume common knowledge of rationality and money-maximizing behavior, the predictions for the baseline treatment of our experiment are straightforward.⁷

⁷ We do not believe that these assumptions provide an accurate description of our participants' behavior. However, since the standard incomplete contracts literature is based on the self-interest hypothesis, we feel that

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

The predictions for our two robustness checks are equally simple. The reduced flexibility condition and the no competition treatment should lead to the same outcome as the baseline condition. As long as the upper bound is high enough to allow for trade in the bad state ($p^h \geq 95$), changing the upper bound does not affect the arguments given above. Since the contract terms will be identical to the terms realized in the baseline condition, eliminating competition at date 0 should also not change anything.

We summarize the prediction of the standard economic model as the

Standard Hypothesis:

- (i) 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., $p^r = p^l = 35$.
- (ii) Sellers never choose low quality irrespective of the contract type and price level. Buyers always choose the lowest price available in flexible contracts.
- (iii) Buyers' profits are higher in flexible contracts than in rigid contracts. Therefore, buyers prefer flexible contracts.
- (iv) Lowering the upper bound in flexible contracts does not change outcomes.
- (v) Eliminating ex ante competition does not change outcomes.

B. Predictions if Contracts are Reference Points

According to Hart and Moore (2008), an ex ante contract, negotiated under competitive conditions, shapes parties' entitlements regarding ex post outcomes. In Hart and Moore (2008), a party compares the ex post outcome to the most favorable outcome permitted by the contract, and if he does not get what he feels entitled to he is aggrieved and shades on noncontractible aspects of performance. In the Appendix

these predictions provide an important benchmark.

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.

we extend Hart and Moore (2008) to allow for the case where parties may feel entitled to an outcome other than the most favorable outcome. The model's predictions are broadly similar to Hart and Moore (2008). Rigid contracts pin down outcomes, sellers get what they expect, and so sellers are not aggrieved. Accordingly, shading should not occur in rigid contracts. However, in a flexible contract, the seller may be aggrieved and shade if he gets a lower price than he had hoped for. We show that the heterogeneity in seller entitlements implies that the frequency of shading is decreasing in price. Given this the buyer will either increase the price in flexible contracts to avoid shading or accept the possibility of getting low quality. Thus, although the flexible contract guarantees trade in both states, the reference dependent behavior of sellers has a negative impact on the buyers' profit. Hence flexible contracts may be less profitable than rigid contracts. Reference dependent behavior does not, however, change the auction outcome: rivalry between sellers still ensures that the lower bound of the price range in flexible contracts, and the fixed price in rigid contracts, is 35. (Recall that we do not allow the price to fall below 35; see footnote 4.)

The assumption that contracts act as reference points also suggests that our experimental robustness checks should have an impact on trading parties' behavior. With the upper bound of the price range in flexible contracts reduced from 140 to 95, buyers have only one price available in the bad state of nature ($p = 95$), and so sellers should not be aggrieved, and there should be no shading. In addition, decreasing the upper bound may also lower (some) sellers' reference prices and decrease shading in the good state of nature in flexible contracts.⁹ As a consequence, flexible contracts may be more profitable, and the buyers may therefore choose the flexible contract more often than in the baseline treatment.

Eliminating ex ante competition is also predicted to have an effect. In the absence of ex ante competition low prices in rigid contracts are no longer justified by competitive market forces at date 0, and so we expect to see more shading in rigid contracts.

These considerations lead to the

Reference Point Hypothesis:

- (i) 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., $p^r = p^l = 35$.
- (ii) In rigid contracts sellers never choose low quality irrespective of the price level. In flexible contracts sellers' quality provision is price dependent. Heterogeneity in seller entitlements implies that the frequency of shading is decreasing in the price. Given the price dependence of quality, buyers may not choose the lowest price available in flexible contracts.

⁹ Lowering the upper bound may affect reservation prices in the good state in two ways: (i) When the upper bound is high ($p^h = 140$), some sellers may have reference prices above 95. Lowering the upper bound reduces these reference prices to 95. We assume that sellers never hope for a price outside the contract. (ii) Lowering the upper bound may also affect reference prices of sellers who hope for a price of less than 95, when the upper bound is high. This is the case if reference prices are a function of the bounds of the price range. Our formal analysis in the Appendix allows for this possibility, since we assume that the reference price is a function of the contract type.

- (iii) Buyers' 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.
- (iv) Lowering the upper bound of the price range leads to less shading in flexible contracts, in particular in the bad state of nature.
- (v) Eliminating ex ante competition increases shading in rigid contracts.

C. Discussion of Design Features

It is important to emphasize that the aim of this paper is not to determine whether people succeed in choosing an optimal contract. We are interested rather in the more fundamental question of whether the underlying behavioral assumptions of the Hart-Moore model are empirically relevant. To study this question in a clean and controlled way, we have intentionally abstracted from some theoretical features of the model. In this section we discuss to what extent the simplifications in our experimental set-up affect our predictions.

We have supposed that only the seller can shade, and that the buyer's value is certain. In the Hart-Moore model either of these simplifications implies that the first-best can be achieved.¹⁰ In other words, the trade-off between contractual flexibility and rigidity is destroyed. 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 lump sum transfers and consider only contracts in which the buyer chooses the ex post price.

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

Hart and Moore (2008) assume that trading parties 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

¹⁰ On the one hand, if the buyer's value is certain, a buyer can offer a rigid contract with the fixed price equal to $v(q^n)$. Trade will occur in both states, there will be no aggrievement since there is only one price, and sellers will provide normal quality. Redistribution of surplus from sellers to buyers (because of competition) can be achieved through lump sum payments conditional on winning the contract auction. On the other hand, if the buyer's value is uncertain but only sellers can shade, a buyer can offer a flexible contract in which the seller has the right to choose the price ex post. Sellers will choose a price equal to the buyer's realized value of $v(q^n)$, and since the sellers get their most favorable outcome they will not shade; also by assumption buyers cannot shade. Redistribution of surplus from sellers to buyers (because of competition) is again achieved through lump sum payments conditional on winning the auction.

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 be problematic. The reason is that indifference would not rule out equilibrium shading under standard economic assumptions in our set-up. In order to make sure that shading cannot be explained if people are motivated by pure self-interest, we implemented costly shading. However, since the increase in the sellers' costs is low ($c(q^l, \sigma) - c(q^n, \sigma) = 5$) relative to the damage which shading imposes on the buyer ($v(q^n) - v(q^l) = 40$), our set-up is still in line with the spirit of the model.¹¹

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 attractive only if the disadvantage due to the nonexistence 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 ($w^g = 0.8$).¹²

III. Results

In this section we present and discuss our results. The analysis of our data at the aggregate level in Section IIIA reveals that the outcomes largely confirm the reference point hypothesis and contradict the predictions of standard economic theory. In Section IIIB we demonstrate how our findings can be explained in light of the differences in the price dependence of sellers' performance across the two contract types. Sections IIIC and IIID illustrate that the predictions of the reference point theory are also relevant for individual behavior of sellers and buyers. In Section IIIE we discuss the results of our robustness checks.

A. Aggregate Findings

Table 2 and Figure 1 summarize the main results for our baseline treatment. Table 1 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 1 displays the development of prices and quality choices over time.

Auction Outcomes in Rigid and Flexible Contracts.—Figure 1 illustrates the power of competition in the auction phase of our experiment. The fixed price in rigid

¹¹ Alternatively we could have chosen to make the provision of low quality slightly less costly than the provision of normal quality. However, the case of costly shading probably leads to stronger effects. It seems more likely that aggrievement triggers costly shading than that the absence of aggrievement causes people to engage in costly voluntary cooperation. The reason is that aggrievement certainly causes a negative sentiment, while the absence of aggrievement may be completely neutral and may not imply the positive sentiment necessary to induce costly cooperation. Of course, this remains an empirical question that should be addressed in future work.

¹² Another way to make sure that we have a sufficient number of observations in both contract types would have been to assign contract types exogenously. However, this would have changed the spirit of the experiment in a fundamental way. From the perspective of the seller, it certainly makes a big difference whether the buyer himself chooses to limit his ability to adjust the price ex post or whether this is imposed by the experimenter (see the discussion in Section IV).

TABLE 2—SUMMARY OF OUTCOMES IN RIGID AND FLEXIBLE CONTRACTS (*Baseline*)

Contract type	Rigid contract		Flexible contract	
	Good	Bad	Good	Bad
State of nature				
Average price	40.7	—	51.1	98.4
Relative frequency 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
Relative frequency of contract		0.50		0.50

Notes: The table summarizes the outcomes for rigid and flexible contracts in both states of nature. All numbers are based on the data of all 5 sessions. *Average price* is the average of the trading price and *Relative frequency of normal quality* measures how often the seller has chosen the normal quality. For rigid contracts this information is available only 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 (seller) (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.

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 2). A nonparametric signed rank test using session averages as observations confirms that the auction outcomes for rigid contracts and flexible contracts are not significantly different.¹³

Prices and Quality in Rigid and Flexible Contracts.—The fact that auction outcomes do not differ across contract types implies that, in principle, the buyers 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 the incentive to increase their prices above the lower price bound, because low prices may cause sellers to be aggrieved 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 2). 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 1). 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

¹³ The session averages for the fixed price in rigid contracts are: 37.3, 40.7, 40.5, 41.2, and 43.4. The session averages for the lower bound of the price range in flexible contracts are: 37.5, 41.6, 40.5, 38.8, and 43.4.

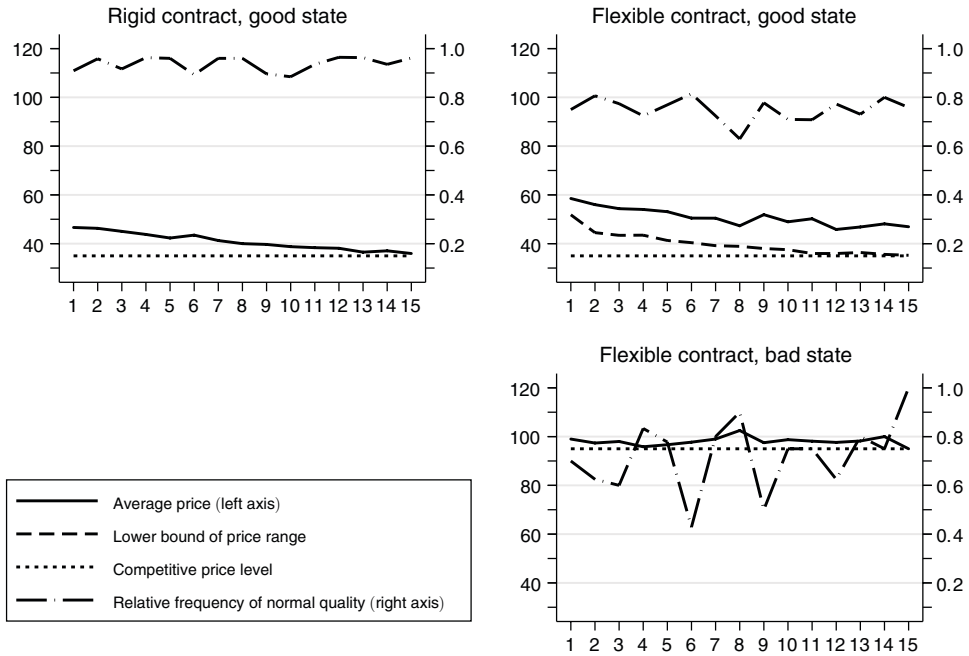


FIGURE 1. DEVELOPMENT OF QUALITY AND PRICES OVER TIME (BASELINE)

prices are around 40 on average and converge to the competitive level of 35 over time.¹⁴ This implies that in the good state buyers pay on average substantially higher prices in flexible contracts than in rigid contracts. A nonparametric signed rank test confirms that the price difference between rigid and flexible contracts is statistically significant (p -value = 0.031 (one-sided)).¹⁵

Although prices are close to the competitive level, shading is almost absent in rigid contracts. Sellers provide normal quality in 94 percent of the cases in which the good state is realized. In flexible contracts, however, the higher prices are not always sufficient to prohibit sellers from shading. In the good state sellers provide normal quality in only 75 percent of the cases (see Table 2). The difference in the frequency of shading between the two contract types is statistically significant (nonparametric signed rank test, p -value = 0.031 (one-sided)) and very stable over time (see Figure 1).¹⁶

Payoffs and Contract Choice.—The differences in price and quality levels have important implications for payoffs of buyers and sellers in the good state of nature. Since prices are higher and quality is lower in flexible contracts, buyers earn, on average, considerably lower payoffs in flexible contracts (78.9) than in rigid contracts

¹⁴ The “auction outcome” in Table 2 is the average of the fixed prices in all rigid contracts. The “price” in Table 2 and Figure 1, 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, there is no systematic difference between the two.

¹⁵ 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.

¹⁶ In the good state the session level frequencies of high quality in rigid contracts are (in percent): 89, 97, 95, 91, and 96. The corresponding numbers for flexible contracts are (in percent): 78, 76, 79, 67, and 75.

(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 2). Both payoff differences are highly significant according to nonparametric signed rank tests (Sellers: p -value = 0.031 (one-sided), Buyers: p -value = 0.031 (one-sided)).¹⁷

In the bad state of nature rigid contracts do not allow for trade. Accordingly, buyers and sellers realize their outside options. In flexible contracts trade takes place and buyers must offer a price of at least 95. We observe that buyers pay on average a price of 98.4 (see Table 2). 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 nonparametric signed rank test shows that this difference is not statistically significant.¹⁸ We will later investigate whether the price setting strategies of buyers in flexible contracts reflect profit maximizing behavior. Since outside options generate a payoff of only 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 (see Table 2).

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 (nonparametric signed rank, p -value = 0.031 (one-sided)).¹⁹ Sellers, in contrast, are better off in flexible contracts. Average seller payoffs in rigid contracts are 18.1, compared to 27.9 in flexible contracts. Also this difference is statistically significant (nonparametric signed rank, p -value = 0.031 (one-sided)).²⁰ The finding that rigid contracts yield higher profits for buyers than flexible contracts is, of course, highly dependent on the choice of parameters. It is certainly easy to find other parameter constellations which yield the opposite results (e.g., higher probability for bad state, weaker impact of shading on buyer's value, etc.). However, our findings illustrate not only 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.

¹⁷ The session averages for buyer profits in the good state are: 98.3, 97.9, 97.4, 95.3, 95.0 (rigid contracts) and 79.3, 80.6, 82.8, 76.6, 76.0 (flexible contracts). The session averages for seller profits in the good state are: 16.8, 20.6, 20.5, 20.6, 23.2 (rigid contracts) and 30.6, 28.5, 28.0, 28.6, 32.8 (flexible contracts).

¹⁸ In the bad state of nature the session level frequencies of high quality in flexible contracts are (in percent): 70, 46, 83, 58, and 83.

¹⁹ The session averages for buyer payoffs in rigid contracts are: 80.1, 80.1, 78.8, 75.0, and 75.3. The session averages for buyer payoffs in flexible contracts are: 71.0, 69.9, 71.4, 65.3, and 67.2.

²⁰ The session averages for seller 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.

In contrast to the standard prediction, buyers choose rigid contracts in 50 percent of the cases (see Table 2). 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 56 percent in period 15. An OLS regression of the fraction of rigid contracts on periods indicates that this positive time trend is statistically significant.²¹

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

The discussion of Table 2 and Figure 1 has shown that in the aggregate our findings support the reference point hypothesis. Next, we analyze whether the underlying behavioral patterns are also in line with the assumption that contracts provide reference points for trading relationships. We start with sellers' performance choices. Figure 2 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 shows the relative frequency with which each price level is chosen by buyers. Notice that prices on the horizontal axis are rounded to the nearest multiple of ten. The figure provides strong support for the seller behavior predicted by the reference point hypothesis. In rigid contracts sellers almost always choose normal quality even if prices are very close to the competitive level. There is no noteworthy correlation between prices and the frequency of normal quality.²² For flexible contracts, in contrast, the figure suggests a strong positive correlation between prices and the willingness to provide normal quality in both states of the world. If prices are close to the competitive level in the good state of nature, normal quality is chosen in less than 60 percent of the contracts. The frequency of normal quality is clearly increasing in price, but in order to reach the same average quality as observed in rigid contracts, buyers must raise their price to a level of at least 75. In the bad state of nature prices close to the lowest possible level also trigger a lot of shading. At prices between 95 and 104 sellers provide normal quality in less than 70 percent of the contracts. Also in the bad state substantial price increases are necessary to reach a high quality level on average.

We provide statistical backup for our observations on price dependence of quality with a regression analysis in Table 3. 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 offer 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

²¹ The regression uses one observation per period and session. The dependent variable is the fraction of rigid contracts; the explanatory variable is period. The estimated coefficients are as follows: Constant = 0.43, p -value < 0.001; Period = 0.010, p -value = 0.002 (p -values are based on robust standard errors).

²² The low frequency of normal quality when prices are around 60 should be ignored, because it is based on a very small number of outliers (see price distribution in Figure 2).

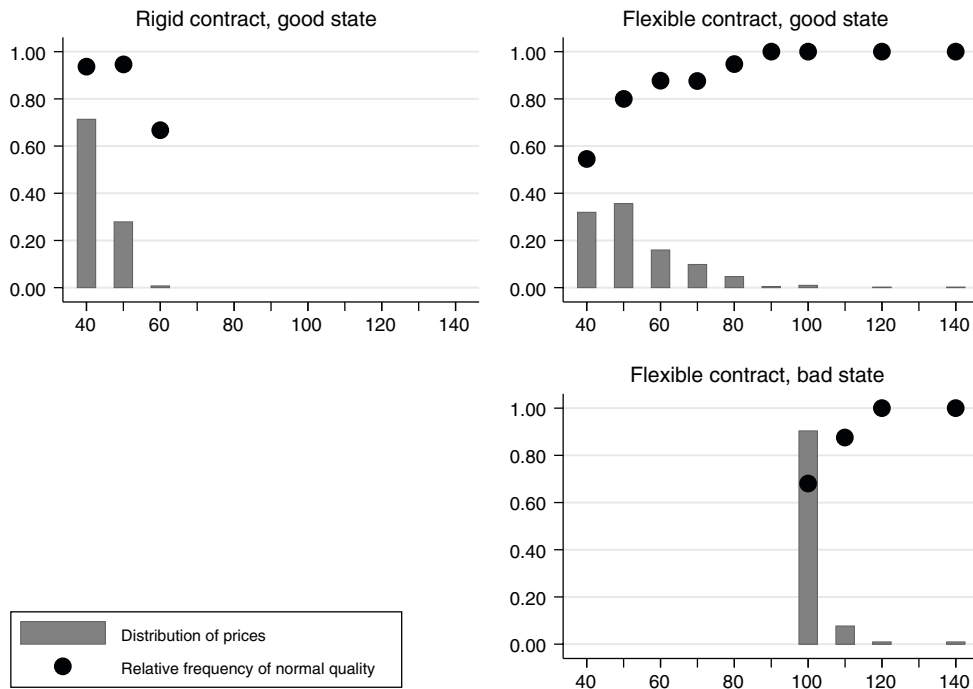


FIGURE 2. PRICE DISTRIBUTION AND FREQUENCY OF NORMAL QUALITY CONDITIONAL ON PRICE (Baseline)

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 ($-0.34, p\text{-value} < 0.01$). The regression also confirms the statistical significance of the positive impact of higher price increments on sellers’ quality choices in flexible contracts ($F\text{-test: Price increment} + \text{Price incr.} \times \text{flex. contr.} = 0, p\text{-value} < 0.01$).²³ In column 2 we show that a probit estimation (marginal effects reported) using the same set of variables yields results similar to the ones of the linear probability model used in column 1. Column 3 investigates the bad state of nature. We regress the indicator variable for choosing normal quality on price increments (now defined as the difference between price and the lowest possible price of 95). The constant indicates that the frequency of normal quality is only 0.66 when buyers pay the lowest possible price in the bad state of nature. In addition, the significant coefficient confirms that there is also a significant impact of price increments on quality in the bad state of nature. Column 4 documents that a probit estimation yields similar results (marginal effects reported).

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

TABLE 3—PRICE DEPENDENCE OF QUALITY ACROSS CONTRACT TYPES (*Baseline*)

Dependent variable	Quality [$\sigma = g$]		Quality [$\sigma = b$]	
	OLS (1)	Probit [ME] (2)	OLS (3)	Probit [ME] (4)
Price increment	0.000 [0.002]	0.000 [0.004]	0.013* [0.005]	0.023*** [0.009]
Flexible contract	-0.335*** [0.060]	-0.298*** [0.060]		
Price increment \times flex	0.009* [0.004]	0.009* [0.005]		
Constant	0.936*** [0.025]		0.657 [0.075]	
Observations	805	805	104	104
R^2	0.13		0.03	

Notes: *Price increment* is defined as price minus 35 in columns 1 and 2 and as price minus 95 in columns 3 and 4. *Flexible contract* is an indicator variable which is unity if the contract is of the flexible type and zero otherwise. *Price increment \times 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.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

C. Individual Seller Behavior: Quality Choice within and across Contracts

The analysis of sellers' behavior in Figure 2 and Table 3 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 3 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, and so we can compare sellers' performance choices across contract types.

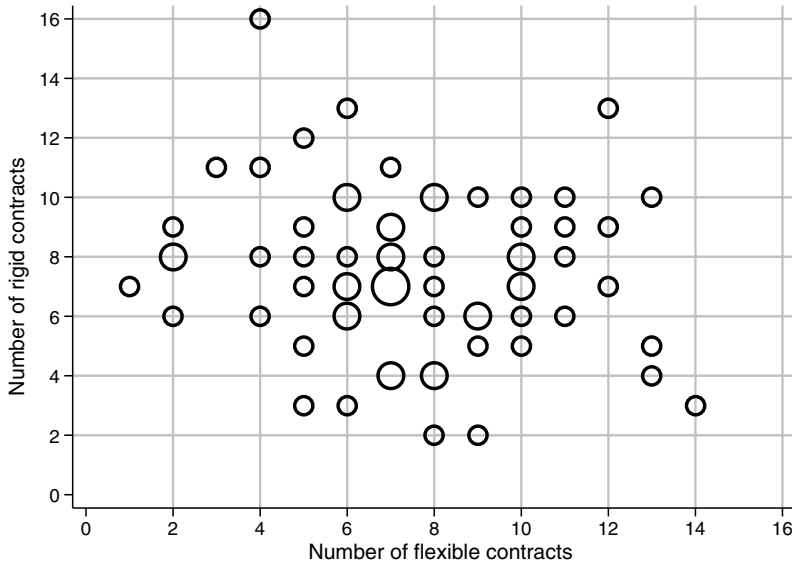


FIGURE 3. ACCEPTED NUMBER OF RIGID AND FLEXIBLE CONTRACTS PER INDIVIDUAL SELLER (*Baseline*)

In Table 4 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. Twenty-seven 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 IIA), 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 4 shows that sellers who never provide low quality have concluded a lower number of flexible contracts and receive, on average, higher price offers in these contracts (especially in the good state of nature). These two factors make it less likely that a seller with a given feeling of entitlement 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 4). A closer look reveals that seven of these sellers show behavioral patterns which are “almost in line” with the prediction of the reference point hypothesis: they provide low quality exactly once in a rigid contract, and they shade more often in flexible contracts than in rigid contracts, in both absolute and relative terms. Only ten of our 68 sellers make decisions which are clearly not in line with the reference point hypothesis.

In addition, the reference point hypothesis also suggests a positive (or zero) correlation of prices and quality in flexible contracts for each individual seller. However,

TABLE 4—SELLERS' QUALITY CHOICES AT THE INDIVIDUAL LEVEL (*Baseline*)

	Consistent with RPH		Inconsistent with RPH		Total
	No	Flexible only	Mostly flexible	Flexible and rigid	
Shading					
Number of sellers	27	24	7	10	68
Population fraction	0.40	0.35	0.10	0.15	1
<i>Rigid contracts</i>					
Average price [$\sigma = g$]	40.6	41.3	41.6	39.8	40.8
Relative frequency of q^n [$\sigma = g$]	1.00	1.00	0.83	0.67	0.93
Average number of contracts	8.2	6.5	8.1	7.3	7.5
<i>Flexible contracts</i>					
Average price [$\sigma = g$]	54.2	49.4	50	53.4	51.9
Relative frequency of q^n [$\sigma = g$]	1.00	0.67	0.47	0.70	0.78
Average price [$\sigma = b$]	98.7	97.7	98.5	99.6	98.4
Relative frequency of q^n [$\sigma = b$]	1.00	0.61	0.58	0.56	0.75
Average number of contracts	6.3	8.5	9.3	7.1	7.5

Notes: RPH stands for reference point hypothesis. The shading categories are defined as follows: “No” stands for sellers who always provide normal quality irrespective of the contract type. “Flexible only” are sellers who never provide low quality in rigid contracts but provide low quality in flexible contracts at least once. “Mostly flexible” are sellers who provide low quality exactly once in rigid contracts and more than once in flexible contracts. “Flexible and 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 2, because numbers reported in this table are averages over individual averages, while Table 2 directly averages over all contracts.

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 4). In seven of these 18 cases the correlation is statistically significant at the 10 percent level. The remaining six sellers have either never provided low quality in the good state (two cases) or the correlation is insignificantly negative. If we do the same analysis for the seven sellers whose quality choices are “almost in line” with the reference point hypothesis (column 3 of Table 4), we find a positive correlation of prices and quality for each individual. In five cases the correlation is significant at the 10 percent level.²⁴

D. Individual Buyer Behavior: Price Setting and Contract Choice

Our analysis has shown that seller behavior is by and large in line with the reference point hypothesis. It is interesting to examine how buyers react to the behavioral patterns of sellers when they choose a contract type and set prices in flexible contracts. From Table 2 we know that overall buyers choose the rigid contract in 50 percent of the cases. Although we have shown that there is a positive time trend (see the regression results in footnote 21), it may appear rather astonishing that the significantly higher average payoffs have not motivated more buyers to choose the rigid contract. However, since the higher average payoff in rigid contracts comes with a

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

higher variance (see the payoffs per state in Table 2), buyers' risk aversion (or loss aversion) may play an important role in this finding. The buyers may perceive the possibility of ending up with the no-trade outcome of rigid contracts as too risky. Our post-experimental questionnaires confirm this presumption. About 50 percent of buyers indicate that they didn't want to take the chance of getting only the outside option of 10 in case of the realization of the bad state and therefore went for flexible contracts.

A second important issue regarding the buyers' behavior is their response to the price dependence of quality in flexible contracts. Are buyers able to figure out the profit-maximizing price, or do they lose money in flexible contracts, because they fail to adopt the proper price? Figure 4 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 4 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 8 percent of contracts buyers pick prices in the range from 105 to 115. These prices yield only slightly lower average payoffs of 28. Higher prices, which yield much lower payoffs, are chosen in only 2 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.

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 2). This illustrates that in our set-up the impact of sellers' reference dependence is strong enough to change the optimal structure of the contract relative to the standard economic model.

E. Experimental Robustness Checks

In this section we report the results of our experimental robustness checks. Recall that in the reduced flexibility treatment we replace the upper bound of 140 in the flexible contract by 95, and in the no competition treatment we replace the auction by an exogenous randomization device.

Reducing Flexibility.—Lowering the upper bound of the price range in flexible contracts does not significantly affect the outcomes of rigid contracts. In both

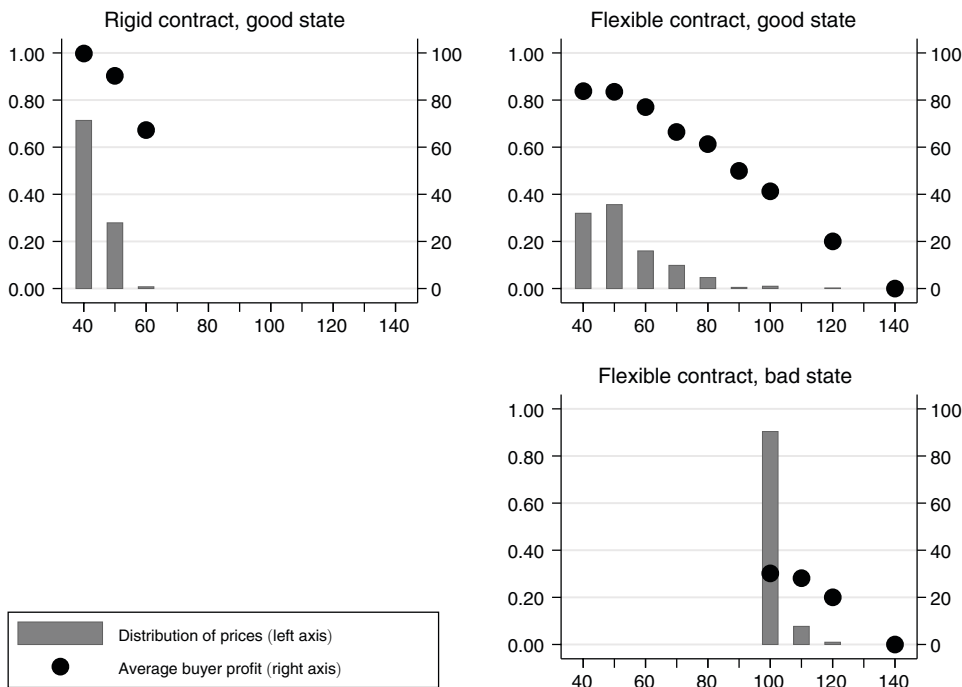


FIGURE 4. PRICE DISTRIBUTION AND AVERAGE PROFIT OF BUYERS CONDITIONAL ON PRICE (*Baseline*)

the reduced flexibility and baseline conditions fixed prices are relatively close to the competitive level (reduced flexibility: 41.3/baseline: 40.7) and sellers choose normal quality most of the time (reduced flexibility: 95 percent/baseline: 94 percent). There is no statistically significant difference in either dimension.²⁵ This reinforces the validity of the hypothesis that sellers accept competitively determined fixed prices in rigid contracts as salient reference points and rarely engage in shading.

As predicted, however, lowering the upper bound of the price range makes quite a difference in flexible contracts. In the bad state of the baseline condition where buyers can pick any price between 95 and 140, sellers are likely to be disappointed because most buyers choose prices close to the minimum (the average price paid is 98.4). As a consequence we observe a lot of shading (the relative frequency of normal quality is only 70 percent). In the reduced flexibility condition the upper bound of the price level and the minimally required price in the bad state are identical, i.e., the only price available to buyers in the bad state is 95. This reduction in flexibility has a very positive impact on quality in the bad state. Although prices are now lower

²⁵ The session averages for fixed prices in the baseline condition are: 37.3, 40.8, 40.8, 41.0, and 43.4. The corresponding values in the reduced flexibility condition are: 42.5 and 39.7. The session level frequencies of high quality in the baseline condition are (in percent): 89, 97, 95, 91, and 96. The corresponding numbers for the reduced flexibility condition are (in percent): 95 and 96. Regressions (OLS) of price and quality choice on a treatment dummy for the reduced flexibility condition (robust standard errors adjusted for clustering at the session level) reveal that there is no difference (price: p -value = 0.653/quality: p -value = 0.326).

(95 instead 98.4), the relative frequency of normal quality increases to 95 percent. This increase in quality is statistically significant.²⁶ It is interesting to note that the observed quality level in the bad state of flexible contracts is now identical to the quality level in rigid contracts.

The lower level of the upper bound also has an impact on quality in flexible contracts when the good state of nature is realized. In the baseline condition buyers pay on average a price of 51.1, and the relative frequency of normal quality is 0.75. In the reduced flexibility treatment the average price is 46.8, and the relative frequency of normal quality is 0.81. Both the reduction in prices and the increase in quality are statistically significant.²⁷ This suggests that reducing the upper bound also affects reference points in the good state. It seems that sellers no longer feel entitled to the same high prices as in the baseline case, and accordingly, buyers have the chance to offer lower prices and still get higher quality.

Since lowering the upper bound of the price range reduces prices and shading in flexible contracts and does not have any impact on rigid contracts, it also affects the relative attractiveness of the two contract types. While trading in rigid contracts is equally profitable for buyers in both treatments (buyer profit in good state: 96.8 (in both conditions)), trading in flexible contracts becomes more attractive in both states (buyer profit in good state: 78.9 (baseline) versus 85.5 (reduced flexibility)/buyer profit in bad state: 29.7 (baseline) versus 42.8 (reduced flexibility)).²⁸ In response, we observe that buyers choose the rigid contract significantly less often when the upper bound of the price range is low (share of rigid contracts: 50 percent (baseline) versus 37 percent (reduced flexibility)).²⁹

Eliminating Ex Ante Competition.—In this treatment we eliminate ex ante competition in order to see whether contracts serve as reference points in the absence of the fundamental transformation.

As expected (see the reference point hypothesis in Section IIB), removing the fundamental transformation increases shading in rigid contracts. Although prices are identical to those in the baseline condition (by design), sellers are more likely to provide low quality when the fixed prices are determined in a noncompetitive way. While sellers provide normal quality in 94 percent of the cases in the baseline condition, the relative frequency of normal quality drops to 84 percent in the no

²⁶ The session level frequencies of high quality in the baseline condition are (in percent): 70, 46, 83, 58, and 83. The corresponding numbers for the reduced flexibility condition are (in percent): 96 and 93. A regression of quality choice on a treatment dummy for the reduced flexibility condition (robust standard errors adjusted for clustering at the session level) shows the significance of the difference (p -value = 0.009).

²⁷ The session averages for prices in the baseline condition are: 51.7, 49.7, 49.0, 50.2, and 54.0. The corresponding values in the reduced flexibility condition are: 47.3 and 46.3. The session level frequencies of high quality in the baseline condition are (in percent): 78, 76, 79, 67, and 75. The corresponding numbers for the reduced flexibility condition are (in percent): 80 and 81. Regressions (OLS) of price and quality choice on a treatment dummy for the reduced flexibility condition (robust standard errors adjusted for clustering at the session level) show the significance of the differences (price: p -value = 0.004/quality: p -value = 0.038).

²⁸ The increase in buyer profits in flexible contracts is statistically significant in both states according to regressions (OLS) of buyer profit on a treatment dummy for the reduced flexibility condition (robust standard errors adjusted for clustering at the session level, good state: p -value = 0.002, bad state: p -value = 0.001).

²⁹ The decrease in the frequency of rigid contracts is significant according to a probit regression of an indicator variable for rigid contracts on a treatment dummy for the reduced flexibility condition (robust standard errors adjusted for clustering at the session level, p -value = 0.017).

competition treatment. This difference is highly significant (nonparametric rank-sum test, p -value = 0.004 (one-sided)).³⁰

In flexible contracts the elimination of ex ante competition does not change the fact that many sellers are willing to provide high quality only if the buyer offers a sufficiently high price. This is evident from Table 5 (which provides the same analysis as Table 3 for the baseline condition). Column 1 shows that in the good state sellers provide normal quality in only 71 percent of the cases when the price is 35 (see the sum of the constant and the coefficient for *Flexible Contract*), but their propensity to choose normal quality significantly increases if buyers pay higher prices (F -test of $Price\ increment + Flex.\ contr. \times Price\ inc. = 0$, p -value = 0.072).³¹ Column 3 reveals that in the bad state sellers provide normal quality in 81 percent of the cases when the price is 95, and higher prices have again a positive impact on quality (see the significantly positive coefficient for *Price increment*, p -value = 0.018).³² Comparing quality and prices with the corresponding values in the baseline condition shows that sellers in the no competition treatment are somewhat less inclined to shade on performance in flexible contracts, and buyers offer lower prices, although most of these effects are not significant.³³

Comparing rigid and flexible contracts, we see from the regression analysis in Table 5 that the elimination of ex ante competition substantially reduces the differences in shading activities across contract types. In the baseline condition sellers in flexible contracts are much more likely to engage in shading than sellers in rigid contracts when prices are close to the competitive level (see the large and highly significant coefficients for *flexible contract* in columns 1 and 2 of Table 3). In the no competition treatment the difference in the intensity of shading at low prices across contract types is much smaller and not significant or barely significant depending on the estimation procedure (p -value = 0.141 (OLS), p -value = 0.083 (probit)). This implies that the elimination of ex ante competition eliminates much of the trade-off between contractual flexibility and rigidity. Rigid contracts are no longer more profitable than flexible contracts (nonparametric signed-rank test, p -value = 0.156 (one-sided)),³⁴ and as a consequence, buyers are much less likely to choose the rigid contract in the no competition condition (28 percent) than in the baseline treatment (50 percent).³⁵

It is also interesting to compare the results of our no competition treatment to those of Fehr, Zehnder, and Hart (2009). That paper studies a slightly different way of eliminating ex ante competition in our experimental set-up. Instead of letting the

³⁰ The session level frequencies of high quality in the baseline condition are (in percent): 89, 97, 95, 91, and 96. The corresponding numbers for the no competition condition are (in percent): 88, 85, 88, 85, and 73.

³¹ The probit estimation in column 2 confirms the finding of the linear probability model.

³² The probit estimation in column 2 confirms the finding of the linear probability model.

³³ Average prices are 51.1 (good state) and 98.4 (bad state) in the baseline condition and 46.3 (good state) and 97.4 (bad state) in the no competition treatment. Only the price difference in the good state is significant (nonparametric rank-sum tests (one-sided), p -value = 0.004 (good state), p -value = 0.111 (bad state)). While the relative frequencies of normal quality are 75 percent (good state) and 70 percent (bad state) in the baseline condition, they turn out to be 78 percent (good state) and 84 percent (bad state) in the no competition treatment. Neither of the differences in the quality dimension is significant (nonparametric rank-sum test (one-sided), p -value = 0.155 (good state), p -value = 0.111 (bad state)).

³⁴ The session averages for buyer profits in rigid contracts are 79.4, 80.9, 80.4, 79.1, and 69.3. The corresponding values in flexible contracts are: 80.0, 76.0, 73.6, 71.7, and 73.3.

³⁵ The decrease in the frequency of rigid contracts is significant according to a probit regression of an indicator variable for rigid contracts on a treatment dummy for the no competition condition (robust standard errors adjusted for clustering at the session level, p -value < 0.001).

TABLE 5—PRICE DEPENDENCE OF QUALITY ACROSS CONTRACT TYPES (*No Competition*)

Dependent variable	Quality [$\sigma = g$]		Quality [$\sigma = b$]	
	OLS (1)	Probit [ME] (2)	OLS (3)	Probit [ME] (4)
Price increment	0.002 [0.005]	0.002 [0.006]	0.012** [0.003]	0.016*** [0.006]
Flexible contract	-0.110 [0.060]	-0.102* [0.051]		
Price increment \times flex	0.004 [0.006]	0.004 [0.007]		
Constant	0.823*** [0.041]		0.811 [0.038]	
Observation	820	820	149	149
R^2	0.02		0.01	

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 increment \times 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.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

random device determine the contract terms after the buyer has decided about his contract type, Fehr, Zehnder, and Hart (2009) provide the buyer with an exogenous basis price before he chooses between the rigid and the flexible contract. If the buyer chooses a rigid contract the basis price becomes the fixed price; otherwise the basis price determines the lower bound of the price range. While this treatment is a simple and natural way to eliminate competition in our set-up, there is one important caveat: on top of eliminating competition, it also changes the information available to the buyer when he decides about the contract type. When the buyer picks his contract in the baseline condition, he does not yet know the exact outcome of the subsequent auction. In the treatment of Fehr, Zehnder, and Hart (2009) he is already informed about the basis price. The no competition treatment in this paper is designed to resolve this potential problem.

The findings of Fehr, Zehnder, and Hart (2009) are broadly similar to the results of our no competition treatment. Fehr, Zehnder, and Hart (2009) find that the elimination of competition does not have any effects on the outcomes of flexible contracts but strongly increases shading in rigid contracts (the frequency of normal quality drops from 94 percent (baseline) to 67 percent). The finding that our no competition treatment has a less pronounced negative impact on quality in rigid contracts indicates that the buyer's information state when he decides about the contract plays a role: a low price in a rigid contract is less annoying to sellers if the buyer faces some uncertainty about the actual level of the fixed price when he chooses his contract. However, even with this uncertainty, the elimination of ex ante competition significantly increases shading in rigid contracts. This reinforces the interpretation that ex ante competition is important for the emergence of contractual reference points because it enables the parties to hide their unfairness behind the veil of the market.

IV. Discussion

We interpret the emergence of shading activities in our experiment as a specific manifestation of the more general observation in the behavioral literature that many people are willing to engage in costly punishment of inappropriate or unfair behavior. Since costly punishment has been extensively studied in theoretical and empirical work on social preferences, it is important to understand how our study relates to this literature.

Theories of social preferences assume that people are not solely motivated by their material self-interest but also take social considerations, especially fairness concerns, into account. There are two main classes of fairness theories: theories based on inequity aversion and reciprocity models. Models of inequity aversion (Fehr and Schmidt 1999, Bolton and Ockenfels 2000) suppose that people dislike inequitable outcomes and are therefore willing to forgo material payoff in order to prevent these outcomes from occurring. Despite their simplicity, models of inequity aversion can explain many experimental results which seem puzzling from the perspective of the standard self-interest model (see, e.g., Fehr and Schmidt 2003). Most important for our purpose is the evidence from the ultimatum game. The typical result in this bargaining game is that proposers offer a considerable share of the pie because they rightly fear that responders reject low amounts (see, e.g., Werner Güth, Rolf Schmittberger, and Bernd Schwarze 1982; Colin F. Camerer and Richard H. Thaler 1995).

The most obvious parallel between our work and the ultimatum game literature is the sellers' observed tendency to engage in shading when buyers offer low prices in flexible contracts (observed in all our treatments). Since paying low prices leads to payoff inequality to the advantage of the buyer, inequity aversion implies that sellers should shade on performance in order to reduce this payoff inequality. Moreover, inequity aversion is in line with the observation that in the absence of competition many sellers shade in rigid contracts (no competition treatment and Fehr, Zehnder, and Hart 2009). However, inequity aversion also predicts a high frequency of shading in rigid contracts in the presence of competition, since prices are typically close to the competitive level in such contracts. Thus inequity aversion cannot explain our central finding that sellers almost never provide low quality in rigid contracts when contract terms are determined in an *ex ante* competitive market.³⁶

The reciprocity models take into account the possibility that the same outcome may trigger different fairness perceptions depending on how the outcome came about. Specifically, these models assume that the evaluation of an outcome depends not only on the realized allocation of payoffs but also on people's beliefs about the intentions of their trading partner (Rabin 1993; Dufwenberg and Kirchsteiger 2004; Charness and Rabin 2002; Falk and Fischbacher 2006). The idea is that people perceive an unfavorable outcome as more unfair if another player intentionally

³⁶ Using the models of inequity aversion to derive fully fledged predictions for our experiment is not trivial. One problem is that the models do not specify precisely how the reference group relative to which individuals evaluate their payoffs is determined (e.g., sellers could compare their payoff to the (expected) payoff of both buyers and the other seller in their group or only to the buyer and the other seller involved in the trade or even only to their direct trading partner). However, while different assumptions with respect to the reference group may influence the predicted behavior in auctions and contracts, they do not affect the prediction that the price-quality relation should be independent of the contract type.

chooses the outcome than if the payoff distribution is beyond the other player's control. The main difference between the two contract types in our experiment is that rigid contracts restrict the buyer's choice set to a predetermined outcome which is typically not 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 set-up where the proposer is forced to offer a small fraction of the pie and a set-up where the proposer can choose between offering small and large fractions of the pie. Falk, Fehr, and Fischbacher (2003) implement two treatments in an ultimatum game experiment which make such a comparison possible. In one of their treatments the proposer, who is endowed with ten points, can choose between the equal split (five points to each player) and an unequal allocation in which the proposer gets eight points while the responder is left with two points. In another treatment the proposer has no choice but to offer the unequal allocation (eight for himself and two for the responder). The evidence shows that the rejection rate for the unequal allocation strongly depends on whether the equal split is available or not. When the proposer could have chosen the equal split, the unequal offer is rejected in roughly 44 percent of the cases. When the unequal allocation was the only alternative available to the proposer, the rejection frequency was only 18 percent.³⁷

One important difference between Falk, Fehr, and Fischbacher (2003)'s set-up and ours is that the proposer's strategy space in Falk et al. is chosen by the experimenter, i.e., each proposer is assigned either to a game with only one option or to a game with two options, whereas in our experiment, the contract types are endogenously determined, i.e., the buyers themselves choose whether they want to interact within a rigid or a flexible contract. If the proposers in Falk, Fehr, and Fischbacher (2003) had had the choice between the two games, reciprocity theories would no longer predict a difference in rejection rates for the unequal allocation across games. Choosing the game with the restricted choice set would be perceived as an intentionally unfair action which calls for punishment. In the context of our set-up a similar argument can be made with respect to the choice of a rigid contract. The rigid contract is not attractive for the seller: in the good state trade will take place at a low price, and in the bad state the seller is forced to take his unattractive outside option. Thus, a buyer who wants to signal generosity should avoid rigid contracts and rather choose flexible contracts and pay high prices. Accordingly, reciprocity theories would predict that sellers punish buyers not only for paying low prices in flexible contracts, but also for choosing unfair rigid contracts.

In fact, this is what we observe in our no competition treatment and in Fehr, Zehnder, and Hart (2009). However, we rarely observe such punishments in the baseline treatment and the reduced flexibility condition.

A crucial difference between the fairness and reciprocity models discussed above and Hart and Moore (2008) is that the latter explicitly assumes that people perceive outcomes of a competitive bargaining process to be acceptable or fair. This

³⁷ Charness and Rabin (2002) report evidence from similar games (see Berk23 and Berk27, page 844). Their results show the same pattern, although both rejection rates are much smaller (9 percent if the equal split was available and 0 percent otherwise).

assumption is crucial for the interpretation of the differences in shading behavior across our treatments. While the no competition treatment and the experiment in Fehr, Zehnder, and Hart (2009) rely on exogenously determined contract terms, the baseline condition and the reduced flexibility treatment let competitive auctions determine the terms. Thus, the low frequency of shading in rigid contracts suggests that sellers do *not* blame the buyer for the realization of unequal payoffs in a rigid contract, but rather view this realization 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.³⁸

The result that sellers earn higher average payoffs in flexible contracts than in rigid contracts is to a certain extent an artifact of our simplified design. In the more general theoretical setting of Hart and Moore (2008) the availability of ex ante lump sum transfers would imply that the differences in seller payoffs across contracts should disappear. However, the fact that in the presence of ex ante competition sellers shade significantly less in rigid contracts, even though they earn substantially lower profits than in flexible contracts, makes our results stronger. People's fairness considerations work against the reference point hypothesis. Reciprocity suggests that sellers may want to punish buyers for choosing the contract which gives them lower payoffs. Since the introduction of lump sum transfers would equalize sellers' payoffs across contracts, such fairness concerns would become less relevant, so that shading should decrease even further in rigid contracts.

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 implement a contract with very rigid terms such that outcomes are pinned down from the outset and sellers know exactly what to expect, the same unfavorable outcomes do not trigger shading. Given uncertainty about the state of nature, these behavioral regularities imply a trade-off between contractual rigidity and flexibility. While flexible contracts are desirable because they allow the buyer to adjust the outcome to the state of nature, rigid contracts have the advantage that they avoid inefficient shading activities.

³⁸ Recent experimental work shows that people can also avoid punishment for unequal outcomes by shifting the relevant decision to another person (see Björn Bartling and Fischbacher 2008, and Lucas C. Coffman 2009). In Coffman (2009) 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, he decides about the price and player B is forced to keep at least an amount equal to the price when he plays the game with player C. In both cases player C observes the outcome and has the possibility to attribute punishment points to player A. It turns out that player A is punished less 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 investigate experimentally how these different ways of shifting blame are related.

In the absence of ex ante competition, we find that contracts no longer serve as reference points; in particular, there is significant shading in rigid contracts. In other words, the fundamental transformation does seem to be important for the Hart and Moore (2008) theory.

One important limitation of our paper is that it has ignored renegotiation. Suppose that a fixed price contract is written at date 0 and the bad state occurs. Assume that the parties have the chance to renegotiate to enjoy the gains from trade. Will they do so? Will such a mutually beneficial renegotiation create bad feeling, given that the price was already fixed, and hence lead to shading? Does the possibility of renegotiation open the door to hold-up; that is, will the seller try to force a renegotiation in order to grab more of the surplus, even if trade would occur in the absence of renegotiation?³⁹ Experiments can clarify the answers to these important questions and thereby provide new understanding of the nature of contractual arrangements.

APPENDIX

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

$$(A1) \quad u_S = \pi_S - \theta \max[(p^R(t, \sigma) - p), 0] I(q),$$

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

Rigid contracts never lead to aggrievement. The buyer has no other possibility than to pay the fixed price which has been determined in the competitive auction. Since the fixed price is also the seller's reference price, the seller gets exactly what he feels entitled to. Thus, whenever the good state is realized, there is trade and the seller delivers the normal quality. If the bad state is realized, trade does not take

³⁹ As in Hart (2009).

place and the buyer realizes his outside option. Accordingly the buyer's expected payoff from choosing the rigid contract can be expressed as follows:

$$(A2) \quad E\pi_B^r = w^s[v(q^n) - p^r] + (1 - w^s)x^B.$$

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

$$(A3) \quad p^T(\sigma) = p^R(f, \sigma) - [c(q^l, \sigma) - c(q^n, \sigma)]/\theta.$$

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 $F_\sigma(\cdot)$ be the distribution function of threshold prices in the state σ . After observing the realization of the state at date 1 a buyer who has concluded a flexible contract chooses his price as follows:

$$(A4) \quad p_\sigma^f = \arg \max v(q^n)F_\sigma(p) + v(q^l)[1 - F_\sigma(p)] - p.$$

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

$$(A5) \quad E\pi_B^f = w^s[v(q^n)F_\sigma(p_g^f) + v(q^l)(1 - F_\sigma(p_g^f)) - p_g^f] \\ + (1 - w^s)[v(q^n)F_\sigma(p_b^f) + v(q^l)(1 - F_\sigma(p_b^f)) - p_b^f].$$

A profit maximizing buyer chooses a rigid contract if $E\pi_B^r > E\pi_B^f$; otherwise he chooses the flexible contract. This establishes the trade-off between contractual flexibility and rigidity described in the reference point hypothesis in Section IIB: i) Our assumption that sellers can completely offset aggrievement by shading implies that they are willing to accept a contract at any available price. Accordingly, the competitive auctions will drive the fixed price in rigid contracts and the lower bound of the price range in flexible contracts down to the competitive level ($p^r = p^l = 35$). ii) Rigid contracts have the advantage that buyers can pay low prices and not suffer from shading in the good state. The downside of rigid contracts is that there is no

trade in the bad state. iii) Flexible contracts guarantee trade in both states of nature. However, buyers have to increase the price above the competitive level and/or bear the consequences of shading in both states of nature. This reduces buyers' payoffs.

REFERENCES

- Babcock, Linda, and George Loewenstein.** 1997. "Explaining Bargaining Impasse: The Role of Self-Serving Biases." *Journal of Economic Perspectives*, 11(1): 109–26.
- Bartling, Björn, and Urs Fischbacher.** 2008. "Shifting the Blame: On Delegation and Responsibility." University of Zurich Institute for Empirical Research in Economics Working Paper 380.
- Bolton, Gary E., and Axel Ockenfels.** 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review*, 90(1): 166–93.
- Brandts, Jordi, and Gary Charness.** 2004. "Do Labour Market Conditions Affect Gift Exchange? Some Experimental Evidence." *Economic Journal*, 114(497): 684–708.
- Camerer, Colin F., and Richard H. Thaler.** 1995. "Ultimatums, Dictators and Manners." *Journal of Economic Perspectives*, 9(2): 209–19.
- Charness, Gary, and Matthew Rabin.** 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics*, 117(3): 817–69.
- Charness, Gary, Guillaume R. Frechette, and John H. Kagel.** 2004. "How Robust Is Laboratory Gift Exchange?" *Experimental Economics*, 7(2): 189–205.
- Coffman, Lucas C.** 2009. "Intermediation Reduces Punishment (and Reward)." Unpublished.
- Dufwenberg, Martin, and Georg Kirchsteiger.** 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior*, 47(2): 268–98.
- Falk, Armin, and Urs Fischbacher.** 2006. "A Theory of Reciprocity." *Games and Economic Behavior*, 54(2): 293–315.
- Falk, Armin, Ernst Fehr, and Urs Fischbacher.** 2003. "On the Nature of Fair Behavior." *Economic Inquiry*, 41(1): 20–26.
- Fehr, Ernst, and Klaus M. Schmidt.** 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 114(3): 817–68.
- Fehr, Ernst, and Klaus M. Schmidt.** 2003. "Theories of Fairness and Reciprocity: Evidence and Economic Applications." In *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, Vol. 1, ed. Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky, 208–57. New York: Cambridge University Press.
- Fehr, Ernst, George Kirchsteiger, and Arno Riedl.** 1993. "Does Fairness Prevent Market Clearing? An Experimental Investigation." *Quarterly Journal of Economics*, 108(2): 437–59.
- Fehr, Ernst, Christian Zehnder, and Oliver Hart.** 2009. "Contracts, Reference Points, and Competition: Behavioral Effects of the Fundamental Transformation." *Journal of the European Economic Association*, 7(2–3): 561–72.
- Fischbacher, Urs.** 2007. "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics*, 10(2): 171–78.
- Greiner, Ben.** 2004. "An Online Recruitment System for Economic Experiments." In *Forschung und Wissenschaftliches Rechnen*. Gesellschaft für wissenschaftliche Datenverarbeitung Bericht, Vol. 63, ed. Kurt Kremer and Volker Macho, 79–93. Göttingen, Germany: Gesellschaft für wissenschaftliche Datenverarbeitung.
- Grossman, Sanford J., and Oliver D. Hart.** 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy*, 94(4): 691–719.
- Güth, Werner, Rolf Schmittberger, and Bernd Schwarze.** 1982. "An Experimental Analysis of Ultimatum Bargaining." *Journal of Economic Behavior and Organization*, 3(4): 367–88.
- Hannan, R. Lynn, John H. Kagel, and Donald V. Moser.** 2002. "Partial Gift Exchange in an Experimental Labor Market, Impact of Subject Population Differences, and Effort Requests on Behavior." *Journal of Labor Economics*, 20(4): 923–51.
- Hart, Oliver.** 2009. "Hold-up, Asset Ownership, and Reference Points." *Quarterly Journal of Economics*, 124(1): 267–300.
- Hart, Oliver, and Bengt Holmström.** 2010. "A Theory of Firm Scope." *Quarterly Journal of Economics*, 125(2): 483–513.
- Hart, Oliver, and John Moore.** 1990. "Property Rights and the Nature of the Firm." *Journal of Political Economy*, 98(6): 1119–58.
- Hart, Oliver, and John Moore.** 2008. "Contracts as Reference Points." *Quarterly Journal of Economics*, 123(1): 1–48.

- Kahneman, Daniel, and Amos Tversky.** 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263–91.
- Kőszegi, Botond, and Matthew Rabin.** 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 121(4): 1133–65.
- Rabin, Matthew.** 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 83(5): 1281–1302.
- Williamson, Oliver E.** 1985. *The Economic Institutions of Capitalism*. New York: Free Press.