Trust and Trustworthiness in Imbalanced Markets

WALKER, MATTHEW,JAMES

Trust and Trustworthiness in Imbalanced Markets

Matthew James Walker

Department of Economics and Finance

Durham University

June 2021

This thesis is submitted for the degree of

Doctor of Philosophy

To my family, friends, and fiancée.

Trust and Trustworthiness in Imbalanced Markets

Matthew J. Walker

Abstract

In this thesis, we use the methodology of experimental economics to investigate issues of trust and trustworthiness in procurement and the supply chain.

First, we develop a procurement model in which both seller-side and buyer-side decisions are endogenous to the trading relationship. We investigate how a real-world contractual incentive mechanism, retainage, can be used to overcome the seller-side moral hazard problem. We find that retainage can improve trade efficiency but increases market prices and may deter participation. We offer managerial insights on how to design the retainage mechanism, conditional on levels of trust and trustworthiness.

Second, we extend the procurement model to incorporate a contingent contract and show analytically that this contract can mitigate the seller-side moral hazard problem. We observe in a lab experiment that suppliers strategically adjust their bids with a contingent contract and that the contingent contract has unintended behavioural consequences, with buyers rewarding sellers less for the quality of works delivered. The results have managerial implications for the use of hierarchical elements in contracts.

Third, we analyse theoretically and experimentally the horizontal effects of supply chain late payments. We show that if firms discount payment received after the standard term, then late payments feed into higher prices and reduced competition. Reneging on a standard payment term entails a penalty for the buyer set by a third-party. If this penalty is not set carefully, a welfare loss arises due to price externalities. We demonstrate how free-riding payment behaviours may emerge among financially weaker buyers in the firm population.

Declaration

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, this has been cited accordingly. No part of the thesis has previously been submitted for a degree in this or any other institution.

Chapter 2 is based on a joint research paper with Prof. Elena Katok and Prof. Jason Shachat. An earlier draft of this paper entitled “Trust and Trustworthiness in Procurement Contracts with Retainage” was published by Chapman University as ESI Working Paper 20-34. All three authors contributed significantly to this project.

Chapter 3 is based on a sole-authored research paper. An earlier draft of this paper entitled “Trade Contingencies in Procurement Interactions” was circulated as my Job Market Paper during the European Job Market 2020 – 2021.

Chapter 4 is based on a joint research paper with Dr. Kyle Hyndman. An earlier draft of this paper entitled “A Theoretical and Experimental Investigation into the Welfare Consequences of Late Payments” is available at SSRN: http://dx.doi.org/10.2139/ssrn.3752360. Both authors contributed significantly to this project.

Acknowledgements

I am indebted to my advisor, Prof. Jason Shachat, for taking me on as his student four years ago and for providing me with an exceptional source of knowledge, guidance, and opportunities throughout the PhD process. Jason, thank you for believing in my research and always challenging me to go further. I would also like to express my sincere gratitude to Prof. Daniel Zizzo, Dr. Kenju Kamei, and Dr. Kyle Hyndman, from whom I have learnt so much and who have each given me much time and effort to provide advice, support, and feedback. I could not have written this thesis without them.

I wish to thank Prof. Elena Katok for generously hosting me at the Center and Laboratory for Behavioral Operations and Economics of The University of Texas at Dallas during the Spring of 2019, and for invaluable inputs into the production of this thesis. The research and training visit to Dallas proved a highly formative experience from which I benefited greatly. I would like to thank all those involved for making my time there welcoming and productive, in particular Prof. Gary Bolton, Prof. Ernan Haruvy, Dr. Simon Siegenthaler, Dr. Blair Flicker, Dr. Bojun Gu and Ms. Joan Besch. Ms. Jennifer Burson provided excellent research assistance in the conduct of the experiments for Chapter 2.

I am also very grateful to the following people for helpful reviews and feedback during my doctoral studies: Prof. Timo Heinrich, Prof. Hong Il Yoo, Dr. Jinrui Pan, Prof. Morten Lau, Prof. Nejat Anbarci, Prof. Nicolas Fugger, Dr. Ruth Beer, Ms. Sabrina Bonzelet, Dr. Matthew Embrey, Prof. Pierre Dubois, participants of The International Foundation for Research in Experimental Economics Graduate Student Workshop in Experimental Economics 2019, academic audiences at Durham University, Newcastle University, University of Cologne, The University of Texas at Dallas, The Society for the Advancement of Behavioral Economics Annual Conference 2020, The Economic Science Association Global Meeting 2020, the Utah Experimental Economics Conference 2020, the Behavioral Operations Conference 2019, and several anonymous referees. I also thank Dr. Alexandros Karakostas and Dr. Morgan Beeson for their help and support at the Experimental and Behavioural Economics Lab of Newcastle University.

I gratefully acknowledge financial assistance from the Economic and Social Research Council (Award No. ES/R500963/1), Durham University Business School and The University of Texas at Dallas. I am also thankful to Mr. Stuart Gladwell, Prof. Deborah Riby and everyone at the NINE Doctoral Training Partnership for their support during the PhD and especially during this past year of uncertainty.

Next, I would like to thank my PhD colleagues and friends at Durham University, Jiunn Wang, Helena Brennan, Lioba Wendling, Artem Nesterov, Xiaoxiao Ma, Sijia Wei, Qinjuan Wan, Johannes Schmalisch, Katy Tabero, Ajaz Ahmed and David Woroniuk for interesting discussions and helpful conversations. I greatly enjoyed being a part of the Newcastle Economics Research Development community, and in particular the friendship of Scott Kirkman and Eduardo Gonzalo Almorox.

Finally, my very deepest appreciation goes to my family and friends, who have made these (many) years of study possible. Thanks above all to my parents, Bob and Debbie, to my grandparents, John and Shirley, and to my brother, Dan, for their enduring support. And to Katie, who has shared in the ups and downs, been a constant source of new ideas, and kept me sane in the process.

Contents

List of Tables

List of Figures

1 Motivation and Related Literature

A tension between competition and cooperation characterises many long-term relationships between firms and suppliers. Recent decades have seen auctions, and competitive tender processes more generally, emerge as important mechanisms for the selection of suppliers and the determination of contractual terms in the presence of scarcity and market imbalances. A distinguishing feature of procurement auctions (or reverse auctions), as opposed to sales auctions, is that the bidding process represents the beginning rather than the end of the transaction (Fugger, Gretschko and Pollrich 2019). Thus, there is scope for moral hazard to become a major source of friction during the trading relationship.

Overcoming this tension between competition and cooperation requires the careful design of incentives to foster trust and trustworthiness among the agents involved. Mutual trust and trustworthiness underpin the majority of successful economic relationships. The oft-cited quotation from Nobel Laureate Kenneth J. Arrow bears repeating: “Virtually every commercial transaction has within itself an element of trust, certainly any transaction conducted over a period of time” (Arrow 1972, p. 357).

This thesis uses the methodology of experimental economics to investigate issues of trust and trustworthiness in procurement and the supply chain. We take inspiration from Özer and Zheng (2019), who argue that “developing structurally new games and experimental design … are necessary to better understand the role of trust and trustworthiness in decision-making because they are *not* abstract issues” (Section 14.4.3). This thesis consists of three independent experimental studies, each addressing a distinct economic issue in which trust and trustworthiness influence the outcome of the trading relationship.

In Chapter 2, we analyse a procurement setting in which it is difficult to write enforceable contracts that condition price upon quality. If higher quality is also costly to deliver, supplier non-performance becomes an acute risk, particularly when there is intense competition for the contract. An established incentive mechanism used to mitigate the problem of supplier non-performance is retainage, in which the buyer sets aside a portion of the purchase price. After project completion, the buyer determines the amount of retainage that is released to the seller, considering any defects that arise. While generally a feasible contract form to implement, the practical difficulties in assessing completion introduce a moral hazard for the buyer. We develop a structurally new game and experimental design to offer managerial insights on how retainage principles mediate trust and trustworthiness in competitive procurement settings with moral hazard. The results suggest that retainage can deliver a significant quality improvement but inflates tender prices. In high retainage structures, there is a trade-off between trade efficiency and supplier participation in request for bids. We further develop a model of fair payment norms and offer managerial insights on how to design the retainage mechanism, conditional on prevailing levels of trust and beliefs about fairness.

In Chapter 3, we extend the environment considered in Chapter 2 to incorporate a contingent contract. A characteristic feature of procurement relationships is the supplier's ability to influence the buyer's valuation after allocation of the contract. Since it is difficult to write a contract that accounts for all possible contingencies, parties to the contract may understand their performance obligations differently. We use laboratory experiments to investigate the behavioural consequences of contingent contracting in procurement relationships characterised by two-sided moral hazard. The contract is allocated at a sealed-bid reverse auction (tender), contains a minimum payment obligation on the buyer, and specifies a deferred payment that is tied to the buyer's and supplier's performance. Verification of supplier performance is imperfect, and so contractual disputes may arise. We develop a game-theoretic procurement model and find qualitative support in the experiment for the model's prediction that suppliers strategically adjust their bids with a contingent contract. In contrast with the theory, the contingent contract crowds out buyer reciprocity, who reward suppliers less for quality delivered. Thus, the contingent contract does not significantly increase the efficiency of trade relative to a non-contingent contract. The findings have managerial implications for the establishment of trust in procurement relationships and for the unintended effects of hierarchical elements in contracts.

In Chapter 4, we model issues of trust and trustworthiness in relation to the payment term. Specifically, we analyse how the expectation of a late payment affects market entry and price competition. Buyers first send a signal to potential suppliers about their intended payment term, which may be standard (e.g., 30 days) or extended (e.g., 60 to 120 days). Suppliers then decide whether to incur a fixed and irreversible cost to enter into price competition. After the seller and winning bid is determined, the buyer chooses the *ex-post* payment date, which may or may not coincide with the *ex-ante* payment term. We show theoretically that if firms discount payment received after the standard term, then payment delays feed into higher consumer prices and reduced competition. Reneging on a standard payment term entails a penalty for the buyer set by a third-party. If this penalty is not set carefully, a welfare loss may arise due to price spillover effects in the market. We provide experimental evidence that the probability of an on-time payment responds to the penalty for reneging and demonstrate that free-riding behaviour may emerge among financially weaker buyers. We also find that seller over-entry may distort payment behaviours in the market. The findings have economic implications for the horizontal effects of supply chain payment practices and for the design of regulatory interventions to deter late payments.

1.1. Reverse Auctions with Moral Hazard

Related studies in the economics and operations management literatures have emphasised the role of the auction institution in determining procurement efficiency. When quantifiable attributes such as reliability and expertise vary across sellers, *ex-ante* auction scoring rules that commit the buyer to an allocation may be used to achieve efficient outcomes (e.g. Che 1993, Santamaria 2015). Often, however, this is not the case and buyers must rely on price-based selection mechanisms. In (first-price) reverse auctions, buyers must select the bidder offering the lowest price as the auction winner. In buyer-determined auctions, buyers are free to choose a price from among competing suppliers.

Engelbrecht-Wiggans et al. (2007) observe experimentally that buyer-determined auctions are welfare-improving only if enough suppliers compete and the correlation between the supplier's cost and value generated is high. Dynamic-bid price-based auctions with bidding credits can also generate greater aggregate and individual surplus for buyers and sellers than sealed-bid buyer-determined mechanisms (Shachat and

Swarthout 2010). High levels of price transparency in open-bid relative to sealed-bid auctions generate consistently lower levels of buyer surplus, and this differential effect intensifies when the quality of individual bidders is common knowledge (Haruvy and Katok 2013). Price transparency can also damage inter-firm relationships and deter relationship-specific investments by heightening supplier suspicions of buyer opportunism (Jap 2007).

The first experimental study of procurement contracts in reverse auctions with moral hazard was conducted by Cox et al. (1996). They compare the efficiency of fixed-price and cost-sharing contracts. In a fixed-price contract, the buyer pays an amount equal to the lowest or second-lowest bid depending on the auction format. In a cost-sharing contract, the buyer pays this amount plus (or minus) an agreed proportion of the cost overrun (or saving). The seller's hazard of opportunism is increasing in the cost-sharing rate. Cox et al.'s main finding is of a trade-off between budgetary expense and efficiency. Although contracts with a greater cost-sharing element result in a lower procurement expense, they are also less efficient due to the heightened seller-side moral hazard problem.[1]

Reverse auctions with seller-side moral hazard have also been studied in the behavioural operations management literature (for a survey of this literature, see Elmaghraby and Katok 2018). Closely related to the experiments reported in Chapters 2 and 3 are two behavioural studies that examine the efficiency of fixed-price (i.e., zero retainage) contracts in reverse auctions with moral hazard when quality is endogenous to the trading relationship. Fugger, Katok and Wambach (2019) show that simply giving buyers the option to select a seller who did not place the lowest bid significantly raises prices and quality levels relative to a binding price-based auction. Interactions are one-shot, no reputation information is available, and buyers must accept vulnerability to loss so as to incentivise high quality. Fugger et al. employ a multi-level cost and quality design, across two different buyer valuation schedules. Their data reveal that buyer-determined

[1] Cost-sharing arrangements tend to be more appropriate for complex projects, accompanied by low degrees of design completeness (Bajari and Tadelis 2001). In these situations, a buyer commitment to negotiate with one seller can outperform an auction when there is adverse selection, scope for product improvements and/or costly renegotiation (Herweg and Schmidt 2017).

auctions yield a robust improvement in cooperation and efficiency, which is explained using a model of inequity-averse preferences.

An earlier experiment was conducted by Brosig-Koch and Heinrich (2014), who consider a buyer-determined auction in which buyers can condition procurement acceptance decisions on past seller performance. Providing reputation information in this way significantly increases buyer profits and procurement quality, relative to the price-based auction format.

Buyer-determined auctions clearly have an important trust-building role in procurement. They are not, however, always palatable from a regulatory standpoint, not least because they may facilitate bidder collusion if there is uncertainty as to how non-price seller attributes affect a buyer's decision (Fugger et al. 2016), or if there is discrimination in the award decision (Verdeaux 2003).

1.2. Social Norms, Competition, and Contract Types

Chapter 2 is further related to the well-established experimental economics literature examining the influence of social norms on competition. Separating these motives is rarely possible using field data; the experimental method enables us to control preferences and the institutions involved. Our baseline model integrates the gift-exchange game of Fehr et al. (1993) into an auction setting. In a typical gift-exchange game, participants are assigned to the role of either buyer or seller and participate in a two stage exchange. First, the buyer sets a price. Second, the seller produces a costly quality level that maps to the buyer's value. In the absence of reputational information, and if there is a preannounced and finite number of repetitions, then the standard perfect equilibrium has the seller incurring the minimum production cost and receiving the lowest available price.

In contrast to the equilibrium prediction with rational profit maximisers, experimental labour market studies of the gift exchange game observed a positive relationship between price and quality which is robust to the market environment (Anderhub et al. 2002, Brandts and Charness 2004, Fehr et al. 1997, Fehr et al. 1998, Keser and Willinger 2000). This data supports the fair wage–effort hypothesis (Akerlof 1982, Akerlof and

Yellen 1990). That norms exert a significant behavioural influence in competitive markets is particularly suggestive of their economic importance: in market experiments, convergence to competitive equilibria usually obtains within a few rounds (see Smith 1982 for a seminal discussion).

Reciprocal behaviour can be rationalised with theories of social preferences (e.g., Bolton and Ockenfels 2000, Charness and Rabin 2002, Fehr and Schmidt 1999). It is less pronounced in one-shot interactions than under relational contracting (Brown et al. 2004) and when outcomes are a stochastic rather than deterministic function of agent effort (Rubin and Sheremeta 2016). Previous gift-exchange experiments have also observed differences in the relative effectiveness of contract types. Fehr et al. (2007) find that contracts in which the principal promises a voluntary bonus payment to the agent can incentivise greater effort than monitoring contracts that impose a probabilistic fine. This result is explained theoretically by the interaction of fair and self-interested types in the population. Karakostas et al. (2017) challenge the generality of this finding. If agents perceive an output-contingent fine as a hostile act, then they may be predisposed to shirking in the monitoring contract. The authors support this assertion with experimental evidence that a revenue-sharing contract is preferred by most principals to a bonus contract.

1.3. Incomplete Contracting

Chapter 3 is further related to the economics literature on incomplete contracting in procurement. In a seminal contribution, Tirole (1986) discusses the implications of contract design for non-contractible relation-specific investments in a bilateral buyer-seller relationship. Relatively few papers address the issue of a buyer's inability to commit not to renegotiate in competitive procurement.[2] Waehrer (1995) permits sellers to renege on their auction bids after cost uncertainty is resolved. Anticipating this, bidding competition is more intense. Wang (2000) and Shachat and Tan (2015) also consider forms of renegotiation after a price-based reverse auction. A feature shared by these studies and

[2] Renegotiation may also arise for informational purposes (Onderstal and Yang 2020).

the model developed here is that the auction institution allocates the contract but the initial price is non-binding.

Herweg and Schwarz (2018) consider the effects of renegotiation due to a contingency not specified in the initial procurement contract (cost overruns). Sellers are endowed with a private cost type, which is observed by the buyer after the auction and before contract renegotiation. The buyer can specify one of two designs during the initial contracting stage, which differ according to their complexity. The cost advantage of the efficient seller type is increasing in the design complexity and the generalised Nash bargaining solution implies that the buyer should minimise this cost differential by specifying the less complex design up front. The final price is higher if either the cost of the more complex design is higher for the winning seller, or if the winning seller is endowed with enough bargaining power. Since the low cost type always wins the auction, the outcome is efficient when there are no bargaining frictions.

Herweg and Schmidt (2020) analyse inefficiencies associated with *ex-ante* information asymmetries in procurement project design. If a price-based auction allocates the procurement contract, sellers have no incentive to reveal their private information about potential design flaws. This can produce inefficient renegotiation, inefficient production and/or inefficient design. The authors show that a two-stage auction process relying on an independent arbitrator to *ex-post* verify the payoff consequences of unforeseen events can implement an efficient procurement outcome, even when the set of all possible events is unknown. This stands in contrast to the usual assumption in the mechanism design literature that all model parameters are common knowledge *ex-ante* when implementing an allocation. *Ex-post* verifiability enables the arbitrator to complete the mechanism by separating the problems of inducing truthful reporting of design flaws and truthful reporting of production costs. This is motivated by the established role of arbitrators in real-world dispute resolution of complex goods.

Gretschko and Pollrich (2021) abstract from frictions in renegotiation and identify an alternative, information revelation, channel through which incomplete procurement contracts may be costly for the buyer. They consider a dynamic model in which production takes place over two periods. In each period, sellers submit bids based on their private cost information. An exogenous shock is realised at the beginning of the second period, which affects the value – but not ranking – of the sellers' production costs. Thus, the optimal mechanism selects the same seller in each period. In the absence of contractual incompleteness, an option contract can implement the surplus-maximizing allocation. When contracts cannot be conditioned on cost realizations, the buyer must negotiate prices in the second period. Implementing the efficient allocation now requires restricting the buyer's information about the seller's cost type. Gretschko and Pollrich show that this can be achieved in equilibrium by using an English auction with optimal reserve price, followed by a take-it-or-leave-it offer, if the exogenous production cost shock is publicly observable and there is no supplier switching during the procurement relationship.

McAfee and McMillan (1986) show that, with seller-side moral hazard, the best linear contract under risk-neutrality trades off the expected marginal benefit to the buyer of sharing in more of the seller's cost uncertainty (lower payment due to greater bidding-competition) against the expected marginal cost (higher payment due to a distortion in effort incentives).[3] The seller's cost in this context consists of a base cost drawn from a known distribution, a random cost shock and a discretionary cost reduction. While the total cost is observable by everyone after procurement occurs, the discretionary cost reduction is decided by the seller after submitting his bid and has an effort cost that is unobservable to the buyer. If sellers are risk-averse, then there is an additional benefit to the buyer of sharing in more of the seller's cost uncertainty, as in the classic principal-agent analysis (e.g., Weitzman 1980).

1.4. Bertrand Entry Games

Chapter 4 is further related to industrial organization models with non-zero entry costs. Lang and Rosenthal (1991) show that with symmetric randomization over entry and bids, the expected price is non-decreasing in the number of potential entrants. This

[3] Linear contracts are commonly used in practice due to their ease of implementation. McAfee and McMillan (1987b) show that there are many cases in which the optimal general mechanism is a linear contract.

counter-intuitive result arises from a trade-off: while the larger number of potential entrants pushes down the lowest-order bidding statistic, this effect is overshadowed by the combination of each supplier being less likely to submit a bid, and any bid that is submitted being at a higher level. Sharkey and Sibley (1993) set out a similar model, permitting non-linear pricing in the form of two-part tariffs. An alternative approach allows for observed entry and yields similar qualitative predictions for the case of symmetric firms (Elberfeld and Wolfstetter 1999).

The Bertrand competition model with entry is related to the theoretical literature on auctions with an endogenous number of bidders. Herein, equilibrium is also formulated in terms of symmetric mixed strategies (Levin and Smith 1994, Samuelson 1985). These models arose as a response to dissatisfaction with deterministic entry and asymmetric equilibria in pure strategies (Engelbrecht-Wiggans 1987, McAfee and McMillan 1987a). Pevnitskaya (2004) extends the endogenous auction entry model to account for risk preferences. Lu (2010) and Moreno and Wooders (2011) incorporate heterogeneous entry costs. The main result of this literature is that it is not always optimal for a reserve price to exceed the auctioneer's valuation and screen potential bidders. This contrasts to the case of a fixed number of bidders.

A series of experiments test the performance of the stochastic entry model in independent private value auctions with endogenous entry. Smith and Levin (2001) find support for the mixed strategy equilibrium, using a one-stage entry game in which payoffs are linked to the number of market participants via the risk-neutral Nash bidding function. This design ensures that incentives are not influenced by differences in subjects' bidding abilities. Cox et al. (2001) test the behavioural validity of the model in a common value first-price auction where private signals are learnt after the entry cost is sunk. Subjects are informed about the market size before bidding and the opportunity cost of submitting a bid is varied between subjects. Observed behaviour is consistent with the comparative statics of the model. Subjects tend not to enter in periods where the entry cost exceeds expected equilibrium profits.[4]

[4] In a follow-up experiment, Casari and Cason (2016) incorporate qualified entry into a common value auction experiment with voluntary entry. A separate experimental literature examines

Palfrey and Pevnitskaya (2008) test the stochastic entry model in a first-price private value auction setting. They find over-entry across a range of outside option levels and value distributions. There is strong evidence of a self-selection effect, with less risk-averse bidders reaching the second stage.

Reverse auctions with wealth-constrained bidders have been studied theoretically and empirically by Chang et al. (2016).[5] Their environment is based on a private plus common cost structure. Sellers are endowed with the same degree of bargaining power but differ in their ability to absorb losses. Strong sellers can accommodate any loss without defaulting, whereas wealth-constrained sellers always default. Both buyers and sellers incur a cost in the event of seller bankruptcy. In a standard first-price auction, strong sellers can afford to bid relatively more aggressively. The addition of a post-auction opportunity to renegotiate the price changes the predictions when there are wealth-constrained sellers in the population. These sellers can use the credible threat of default to renegotiate higher prices, thereby distorting their strategic incentives at the auction. A trade-off emerges between the probability of default and the higher prices paid by buyers. Chang (2019) observes that renegotiation opportunities may even induce sellers to increase their credibility of default, in anticipation of greater bargaining power.

Finally, Chapter 4 also connects with research at the intersection between operations management and finance, in particular on trade credit (for a survey, see Seifert et al. 2013). Operations models typically specify that buyers are willing to pay a premium for the vertical operational role of trade credit, such as to deter moral hazard (Babich and Tang 2012) or share demand risk (Kouvelis and Zhao 2012). Suppliers may in turn make financing agreements available as a signal of quality (Long et al. 1993). The horizontal effects of trade credit are less well developed. Peura et al. (2017) use a Bertrand model with liquidity shocks to show how suppliers might benefit from the use of trade credit

---

subjects' selection into and valuation of different auction formats (Engelbrecht-Wiggans and Katok 2005, Ertaç et al. 2011, Ivanova-Stenzel and Salmon 2008, 2011). Aycinena et al. (2018) compare revenue outcomes across independent private value auction formats and information structures in which potential bidders know their valuation before making the entry decision.

[5] Engelmann et al. (2020) study the effects of buyer default at online consumer auctions. In this setting, strategic behaviour comes from the bidders rather than the auctioneer, while there is typically a reserve bidder to ensure that trade occurs.

contracts because of a softening in horizontal price competition. Firms in a stronger financial position may be able to use trade credit to exclude their weaker competitors. These studies, however, assume timely payment and no supplier reaction to payment beyond the agreed term, which differs from the model developed here.

2 Procurement Contracts with Retainage

"Many transactions will potentially be too costly to undertake if the participants cannot rely on efficient and equitable adaptation to those unforeseen contingencies"

- David M. Kreps (1990, p. 92).

2.1. Introduction

It is common practice for public and private sector entities to rely on competitive procurement to obtain goods and deliver projects.[6] For standardised goods, price competition promotes productive efficiency. For non-standardised goods, such as complex construction or infrastructure projects, the benefits of competition are less straightforward. Procurers often rely on *ex-post* incentives to mitigate the risk of supplier non-performance.[7] It can be difficult, however, for the procurer to write complete and enforceable agreements *ex-ante* that condition price on the quality of works delivered (Chakravarty and MacLeod 2009, Gretschko and Pollrich 2021). Supplier cost-cutting is an ever-present issue (Lo et al. 2007, Midler 2007). This may manifest itself in reduced quality materials or unethical/unsustainable production processes (Chen and Lee 2016, Guo et al. 2015). Intense competition for the contract may further increase suppliers' incentives to cut corners later on (Chaturvedi 2021), or adversely affect relationship-building (Emiliani and Stec 2005).

An understudied incentive mechanism used in procurement to mitigate supplier moral hazard is retainage. A retainage provision, or retention as it is known outside the United States, is a pre-agreed percentage of the contractual price withheld from a seller by the buyer. The buyer in this context might be a client, main contractor or sub-contractor withholding money from a lower tier. Retainage has its origins in nineteenth-

[6] The World Bank estimates that, on average, public procurement constitutes 14.5% of gross domestic product globally (Djankov et al. 2017).

[7] See, e.g., Bajari et al. (2014) in highway procurement. An alternative mechanism to overcome the tension between competition and cooperation in procurement is proposed by Chakraborty et al. (2021). In a mixed adverse selection and moral hazard model, they show that incentives to mitigate shirking by sellers can be provided either by limiting the number of bidders, or by using an inefficient auction allocation rule.

century British railway construction, when it was set at 20% of contractual value (Bausman 2004). Today, typical provisions range from 3% to 10% and provisions are found across most standard construction contract types (Cox et al. 2011, Nabi Mohamad et al. 2021).[8] On substantial completion, retainage is released back to the seller, minus any deductions for defects. Measures to safeguard retainage vary by country and locale.[9]

In this chapter, we develop a structurally new game and experimental design to offer managerial insights on how retainage principles mediate trust and trustworthiness in procurement settings with moral hazard. Specifically, we investigate how and when contractual retainage can be used to mitigate the seller moral hazard problem in a two-tier supply chain. If operating as intended, retainage circumvents the difficulties in writing a complete contract and effectively aligns project incentives (Raina and Tookey 2013). An alternative view cited among practitioners is that retainage negatively impacts contractors' cashflow, thereby acting as a financial constraint and generating a counterproductive increase in procurement costs. Recognition of a potential hidden cost of retainage has driven a downward trend in the maximum retainage provision permitted by several US states (ASA 2018). The efficiency of using retainage to mitigate moral hazard is understudied and is challenging to measure using empirical data due to the nuances and complexity of each construction project. The lab enables us to isolate the causal effect of retainage on bids, quality and profits, without the confounds of project-specific factors or alternative mechanisms (e.g., repeated interactions) observed in the real world.

An important consideration in the implementation of retainage is what, in practice, constitutes *substantial completion*. Legal scholars have long recognised the difficulties

[8] In the UK, standard building contracts are produced by the Joint Contracts Tribunal, which encourages the holding of retention monies until practical completion. In the US, standard Design-Build agreements include those produced by the American Institute of Architects, ConsensusDocs and the Engineers Joint Contract Documents Committee. These contracts often specify that half of the retainage money be released immediately, while the remainder is released after the expiration of a defects liability period.

[9] The European Commission (2009, clause 41) prescribes that retainage monies "are not paid until the satisfaction of conditions specified in the contract for the payment of such amounts or until defects have been rectified". In New Zealand, the Construction Contracts Amendment Act 2015 provides additional protection for the payment of retainage monies to sub-contractors. In China, retainage applies to pre-specified defects liability periods and at the time of writing enjoys additional financial guarantees from the Agricultural Bank of China.

inherent in determining such a doctrine (Thomas et al. 1995).[10] Failure of trade parties to understand their contractual obligations is one of the leading causes of construction disputes (Arcadis 2020). Litigation is often lengthy to pursue. A costly dispute between the Californian construction contractor, FTR, and its client Rio School District, over the latter's failure to release more than half a million dollars of retainage persisted for many years before being resolved in FTR's favour in 2015.[11] Recently, a construction sector consultation commissioned by the UK Department of Business, Energy & Industrial Strategy (Pye Tait Consulting 2017) found that late and non-payment of retainage monies from clients to contractors is commonplace, especially among lower tier suppliers. Whether or not the withholding of retainage monies is justified is often unclear, precisely because of the difficulties in verifying substantial performance. What is clear from the report is that a substantial fraction of clients believes their overall project costs are higher because of retainage, that retainage induces the possibility of opportunistic payment behaviour and that tender prices reflect this countervailing buyer moral hazard.

The buyer moral hazard problem is accentuated when suppliers make relation-specific investments before a contract is confirmed. In 2016, German automaker Volkswagen (VW) cancelled orders worth 500 million euros with two component suppliers, in the wake of the emissions scandal that forced VW to cut approximately one billion euros in costs. The cancellation came too late for their suppliers, however, who had already spent 58 million euros making factory alterations in preparation to receive the order (Rauwald 2016).

Similar situations can arise in the construction industry. One example is when a project proceeds based on a *letter of intent* (LOI). There are various types of LOI, from a mere handshake agreement stating the intention of parties to trade, to an interim contract that is replaced by a binding contract on expiration, to a non-binding trade agreement in its own right. The peril of supplying under an LOI, without a concrete

[10] Corbin (1919) captured the essence over a century ago: "What constitutes substantial performance must be determined with reference to the particular facts in each case. The question is always one of degree and its solution must be doubtful in many cases" (p. 761).
[11] FTR International, Inc. v Rio School District. California Court of Appeal, 2015.

payment schedule, is demonstrated by a notable English contract law case.[12] In 2005, RTS Systems won a competitive tender to supply improved food packaging for the German dairy manufacturer Müller. Work began based on an LOI and Müller paid RTS only 30% of the agreed price up front and a further 40% later on. After expiration of the LOI and repeated deferral in the execution of a binding contract, Müller alleged product defects and refused to pay RTS the remaining 30% of the tender price. A protracted and costly legal battle ensued, centred around the basis for which a contractual agreement existed. The Supreme Court Justice pronounced on judgement day that "the moral of the story is to agree first and to start work later". Nevertheless, today "the use of LOIs remains widespread in the construction industry" (Wevill 2015, p. 29). The buyer, while not explicitly designating withheld monies as retainage, withholds a high percentage of the purchase price up front. It is not difficult to specify cost structures which expose suppliers to losses and deter them from participating in requests for bids.

The previous anecdotal evidence suggests that, by varying the percentage of the contract price withheld by the buyer until after delivery of the project, we shift the relative burdens of trust between buyer and seller. To that end, we compare the performance of procurement contracts in which there is either (i) zero retainage – a fixed-price contract, (ii) a retainage provision set such that suppliers can adjust their bid upwards to compensate for the increased risk of non-payment of monies, or (iii) high retainage in which the buyer pays only a small percentage of the contract price up front, and consequently suppliers cannot fully compensate for future production costs incurred by bidding higher. These three arrangements approximate the market conditions (rather than specific parameters) discussed above. That is, there is either a one-sided seller moral hazard problem, a two-sided buyer and seller moral hazard problem, or a one-sided buyer moral hazard problem.

We develop an analytical model that consists of a sealed-bid reverse auction (tender) followed by a bilateral trade interaction. Specifically, a single buyer seeks to procure one unit of an indivisible good (e.g., construction of a new school) from a pool of multiple

[12] RTS Flexible Systems Limited v Molkerei Alois Müller GmbH [2010] 14 UKSC.

pre-qualified suppliers. There is a commitment to procure at the lowest price (if at all) and the contract contains a fixed retainage provision. After allocation of the contract, the seller takes costly action to deliver the project and the buyer realizes the project value as a function of the seller's action. The buyer then has discretion over the retainage return decision. In other words, the buyer decides how to allocate retainage monies to compensate the seller for performance delivered.

The retainage-related part of the transaction is predicated on trust, and so game-theoretic arguments based on standard preferences predict that the retainage mechanism generates no quality improvement. Trust between agents (i.e., managers) is an important driver of supply chain success (Cerić 2016). Thus, we develop a model of fair payment norms, in which the buyer may be trustworthy or untrustworthy. We adopt Özer and Zheng's (2019, p. 497) definition of trustworthiness as a voluntary behaviour "in a way not to take advantage of the trustor's vulnerable position when faced with a self-serving decision that conflicts with the trustor's objective". Whereas an untrustworthy buyer always withholds retainage from the seller, a trustworthy buyer distributes retainage according to some known and exogenous fairness norm. Trust is defined as the seller's belief about interacting with a trustworthy buyer (Herold 2010).

Standard preferences imply that trade efficiency will be low in procurement contracts (i) and (ii) as a consequence of the seller moral hazard problem, and nil in contract (iii) as the market unravels. By contrast, our model of fair payment norms demonstrates that if there exists sufficient trust in the market then high quality delivery emerges as an equilibrium outcome in anticipation of a reciprocal retainage payment. Designed appropriately and considering what constitutes a fair outcome in a particular procurement relationship, retainage can mitigate the seller moral hazard problem. We fit our behavioural model to the data and provide a characterization of the optimal retainage level, given prevailing levels of trust and beliefs about fairness. By doing so, we contribute to a growing behavioural operations literature addressing how social preferences influence supply chain contracting (Beer et al. 2018, Cui et al. 2007, Davis and Hyndman 2018, Hu et al. 2017, Katok and Pavlov 2013, Loch and Wu 2008).

Our experimental data offers several insights for practitioners in construction and related procurement settings: (1) retainage mitigates the seller moral hazard problem and increases trade efficiency; (2) retainage provisions inflate tender prices; (3) high retainage structures deter supplier participation in the contracting process and undermine overall procurement efficiency (which we will refer to here as global efficiency); and (4) suppliers may inadequately adjust their bids in anticipation of uncertain retainage returns, leaving them vulnerable to losses for production costs incurred.

2.2. Model and Theory

Consider a one-shot interaction in which a single buyer seeks to procure one unit of an indivisible good from a group of n pre-qualified suppliers, indexed by i. A first-price sealed-bid reverse auction determines selection of the winning supplier (henceforth the seller) and the contract price. Auction participation is voluntary for suppliers, and the buyer can choose to not purchase after observing the contract price.[13] The seller can produce either a high- or low-quality unit, but the setting prohibits quality contingent contracts.

In our setting, the contract price is less binding than usual as purchases are made with retainage provisions. Such a provision includes a retainage proportion ρ, which is a fraction of the contract price withheld from the seller until after unit production. A general interpretation of ρ is as the degree of price flexibility. After production, the buyer observes her valuation and (indirectly) the seller's action. The amount of retainage released to the seller is then at the buyer's discretion. The retainage proportion thus regulates each party's trust burden: at low levels, the buyer possesses limited insurance against low quality production; at high levels, the seller is vulnerable to financially damaging retainage return decisions. The sequence of events is displayed in Figure 2.1.

In the Bidding Stage, each supplier simultaneously submits his bid, b_i, or chooses not to participate in the auction process. If at least one seller submits a bid, the one submitting the lowest bid wins the auction and the contract price is the winner's bid,

[13] Discretion clauses are common in construction tenders.

$p = min\{b_1, \cdots, b_n\}$. Ties are broken randomly. If no bid is submitted, all parties earn zero and we call this outcome "market unravelling". When an auction succeeds, the profile of bidding-stage actions is announced before the next stage.

In the Procurement Stage, the buyer either accepts the winning bid and pays $(1-\rho)p$ to the winner, or she accepts it resulting in all parties earning zero. The buyer's Procurement Stage action is $a \in \{a^0, a^1\}$, where a^0 is a rejection and a^1 is an acceptance. A buyer's agreement to purchase initiates a fundamental transformation (Williamson 1985), which describes the transition from an *ex-ante* competitive market in which multiple suppliers have the opportunity to tender their bid, to an *ex-post* bilateral trade relationship between the buyer and seller.

In the Production Stage, the seller chooses to produce either a high- or low-quality unit, $q_i \in \{q^L, q^H\}$. The seller incurs a sunk production cost c^j for quality level j. Production cost schedules are the same across suppliers and this is common knowledge. A seller's cost of producing high quality is greater than his cost of producing low quality, $c^H > c^L > 0$. Likewise, the buyer's valuation of the unit is increasing in quality and given by v^j. Trade is preferred to no trade and surplus is increasing in quality, i.e., $v^H - c^H > v^L - c^L \geq 0$.

Finally, in the Payment Stage, the buyer observes the seller's chosen quality level and then selects a fraction, r, of the retainage money to return to the seller. The interaction ends and profits are realised.

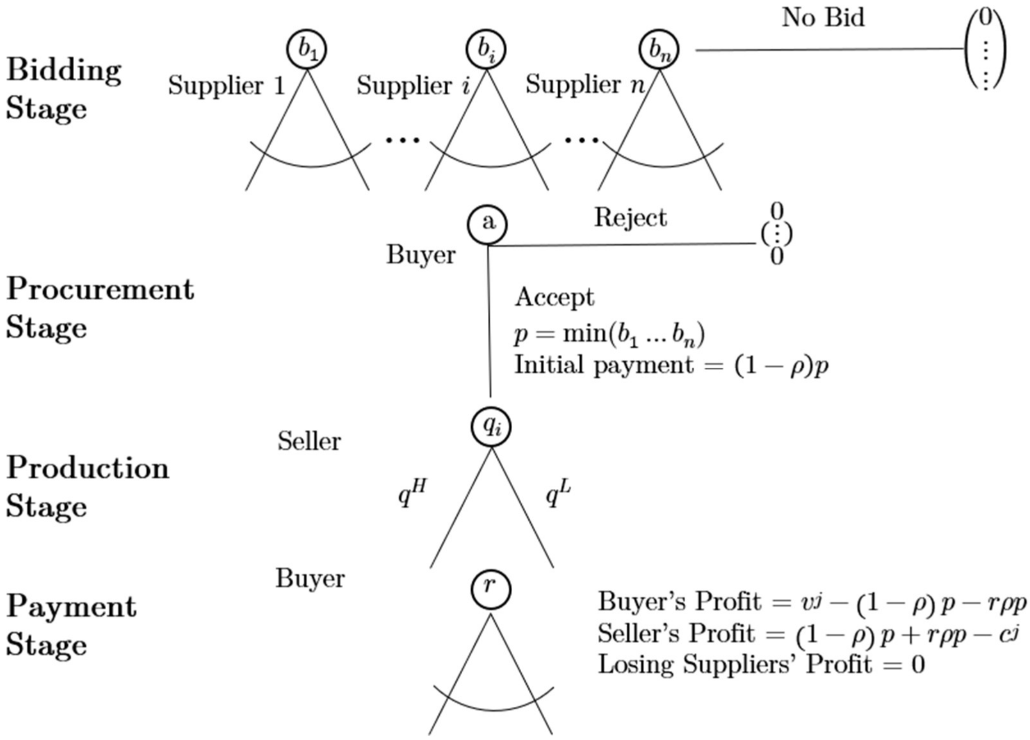
Formally, a supplier i's strategy has two components. These are a Bidding Stage action $b_i \in \{[c^L, v^H], b^0\}$ and a quality choice function $q_i(b_i | b_i = p, a = a^1)$. Bids can be submitted from a continuous interval between the seller's minimum production cost and a reserve price, which without loss we set to equal the buyer's maximum valuation for the unit, and b^0 is non-participation. A buyer's strategy also has two components: a procurement decision function, $a(p)$; and a retainage return function, $r(p, q_i |\ a = a^1)$.

The profits of the transacting buyer and seller are,

$$\text{Buyer's profit :} \qquad \Pi_B = v^j - (1-\rho)p - r\rho p$$

$$\text{Seller's profit :} \qquad \Pi_S = (1-\rho)p + r\rho p - c^j$$

Figure 2.1. The sequence of events in our procurement model.



Notes: This is an extensive game tree representation of the strategic interaction. Suppliers move first and either submit a bid b_i at auction or choose not to participate. If a market forms, the buyer can either accept to trade with the lowest bidder and make a guaranteed payment equal to $(1-\rho)p$, or refuse the transaction. The winning supplier (seller) then selects to produce a high quality, q^H, or low quality, q^L, product and incurs the production cost . The buyer is informed about the product value, v^j, and decides on a discretionary proportion r of the retainage money ρp to return to the seller. The seller earns a profit equal to the difference between total payment received and the cost incurred. The buyer earns a profit equal to the difference between value received and the total payment made. Non-trading parties earn zero profit.

Our primary interest is in whether and under what conditions there exists an equilibrium solution that implements high quality. High quality is implementable if it yields a non-negative expected payoff to the transacting parties (buyer and seller participate voluntarily) and it is incentive compatible (preferred to low quality by the seller).

2.2.1. Standard theory

In a first approach, we assume that the buyer and sellers act as strict expected profit-maximisers. We proceed to solve for the subgame perfect equilibria (SPNE) of the game using backward induction.[14]

Proposition 2.1.

A. *For* $\rho \leq 1 - \frac{c^L}{v^H}$*, there is a unique SPNE in which* $b_i^* = \frac{c^L}{(1-\rho)} \forall\ i$, $a^* = a^1$, $q_i^* = q^L$ *and* $r^* = 0$.

B. *For* $\rho > 1 - \frac{c^L}{v^H}$*, there is a unique SPNE in which* $b_i^* = b^0\ \forall\ i$ *, and the market unravels; off the equilibrium path,* $a^* = a^1$, $q_i^* = q^L$ *and* $r^* = 0$.

Proof. A. The trade quality and contract price are known in the Payment Stage (for the special case of zero retainage, the Payment Stage is obsolete). Since the buyer's profit is decreasing in the retainage return fraction r, the buyer will always choose $r = 0$. In the Production Stage, with certain zero retainage payment, low quality always yields higher profit than high quality. In the Procurement Stage, the buyer anticipates this and accepts to procure for all $p \leq v^L/(1-\rho)$. In the Bidding Stage, each seller's expected profit maximizing bid is the cost of low quality marked up in proportion to the retainage money vulnerable to loss, generating an expected profit of zero. Deviation to a higher price would also yield zero profit, due to the binding auction selection rule. Since $\rho \leq 1 - c^L/v^H$, withdrawing from the market would not improve their position. For all bids strictly above $c^L/(1-\rho)$, a supplier has an incentive to unilaterally undercut his competitors in the continuous bid interval: by doing so, he can increase the probability of trade from $1/n$ to one. The full trade surplus accrues to the buyer. B. For higher retainage proportions, $\rho > 1 - c^L/v^H$, there exists no price in the available bid interval that would induce a seller to participate. Thus, the only rationalizable action in the Bidding Stage is b^0. Participation would enable the buyer not only to appropriate

[14] Other Nash equilibria exist at higher prices and which are associated with non-participation, but since none of these outcomes are compatible with the production of a high quality unit, we ignore them here.

the transaction surplus, but to appropriate part of the seller's cost outlay. Off the equilibrium path, the rationale for payment, production and procurement decisions are the same as in part A. ■

Proposition 2.1 describes how the retainage proportion regulates the burden of trust between buyer and seller. Since the buyer always returns zero retainage, the proportion only influences seller participation and bids. Below a retainage threshold determined by the ratio of a seller's minimum production cost to the buyer's maximum unit valuation, suppliers will participate in the reverse auction, submit a bid that covers their marginal cost and is acceptable to the buyer, and produce low quality; the equilibrium bid in this interval is strictly increasing in the retainage proportion. Above said threshold, no supplier wishes to participate in the auction process – no available bid at least breaks even – and so the market unravels.

Corollary 2.1. *Under standard preferences, (i) high quality is not an implementable outcome, (ii) equilibrium bids and the probability of market unravelling are non-decreasing in ρ.*

Proof. Follows directly from Proposition 2.1. ■

2.2.2. A model of fair payment norms

Let us consider an alternative approach in which we depart from standard assumptions on the buyer's preferences. The buyer is one of two types: with probability α, the buyer is trustworthy (T); with probability $(1-\alpha)$, the buyer is untrustworthy (U). The prior probability α is common knowledge to all agents. An untrustworthy buyer returns zero retainage, as before. A trustworthy buyer can be relied upon to make a fair retainage return decision in exchange for *high* quality, constrained by the terms of the contract. That is, whereas low quality unambiguously warrants zero retainage return, the buyer moral hazard permits subjectivity in what constitutes fair compensation for good performance. This formulation permits us to incorporate both positive and negative reciprocity in the model.[15]

Models of distributional fairness were originally proposed to explain individual behaviour (e.g., Bolton and Ockenfels 2000, Fehr and Schmidt 1999) and typically assume that the fair reference point in a bilateral relationship is the 50-50 surplus split. While this assumption may be appropriate to describe some firm relationships, not all supply chain relationships are created equal. Thus, following Cui et al. (2007), we specify that the fair reference point for a trustworthy buyer is γ multiplied by the seller's profit. Production costs are included in profit comparisons, based on prior evidence from hold-up experiments that buyers consider sunk costs when making decisions on surplus divisions (Ellingsen and Johannesson 2004). The parameter γ is exogenous and represents a fairness norm, which in practice may be specific to the industry, locale or transaction.

There are some notable special cases in this model of fair payment norms. The standard theory is captured by $\alpha = 0$. When $\rho = 0$, the model collapses to an extensive-form game of complete information and Proposition 2.1 applies. In fixed-price contracts, high quality is never an implementable outcome regardless of the degree of trust. From now on we assume price flexibility ($\rho > 0$) and some probability that the buyer is trustworthy ($\alpha > 0$). The model is analysed as an extensive-form game of incomplete information. We apply a perfect Bayesian equilibrium (PBE) solution concept and restrict attention to pure strategy equilibria. This requires that beliefs correspond to objective probabilities for all equilibrium actions.

Suppose again that suppliers compete anonymously in the Bidding Stage and that in the Procurement Stage, the buyer accepts any bid from which she expects to profit. After the fundamental transformation takes place, the seller is uncertain as to the probability with which the buyer is trustworthy. Trust is defined in our setup as the seller's belief α about interacting with a trustworthy buyer. There is a direct mapping between trust beliefs and the seller's choice to produce high quality. Our approach enables the

[15] The model generalises straightforwardly to the case in which low quality is also rewarded by trustworthy types. We find little evidence for this in our experiment. The alternate specification introduces a non-monotonicity into the relationship between trust and the incentive compatibility of high quality, which we return to below.

characterization of threshold trust levels at which high quality may emerge at the population level. Bolton and Ockenfels (2000) provide a similar application of their ERC theory ("the α model") to data from gift-exchange experiments.[16]

In the Payment Stage, the buyer returns zero retainage with probability at least $(1 - \alpha)$. An untrustworthy buyer will always return zero ($r_U = 0$). A trustworthy buyer also returns zero on receipt of low quality. If the seller delivers high quality, then a trustworthy buyer returns the retainage fraction necessary to achieve a profit distribution as close as possible to the fair reference point. In this situation, the best-response retainage return function for a trustworthy buyer is as follows:

$$r_T(p, q^H; \gamma) = \frac{v^j + \gamma c^j + (\rho - 1)p(1 + \gamma)}{\rho p(1 + \gamma)}, \tag{2.1}$$

where the return in (2.1) is bounded in the unit interval. Thus, conditional on the winning bid, the optimal return for a trustworthy buyer produces either an interior or corner solution. For ease of notation, we define $r_T^j = r_T(p, q^j; \gamma)$. The main insights of this section would be unchanged by assuming that a trustworthy buyer offers a large enough flat reward to the seller for producing high quality. By introducing a fairness norm into the environment, we consider a variable reward and provide some foundation as to how this might be determined in practice.

Corollary 2.2. *The trustworthy buyer's retainage return fraction is (i) non-decreasing in quality, and (ii) non-increasing in* γ*.*

Proof. Follows directly from $v^H - c^H > v^L - c^L$ and inspection of the trustworthy buyer's best-response retainage return function, $\frac{\partial r_T}{\partial \gamma} = (c^j - v^j)/\rho p(1 + \gamma)^2 < 0$. ■

In the Production Stage, the seller anticipates the retainage return decision of the buyer and chooses a quality level j to maximise his expected monetary payoff as follows:

[16] Bolton and Ockenfels succinctly explain the rationale for such an approach as follows: "much of what we need to know has to do with the thresholds at which behaviour deviates from the 'more money is preferred to less' assumption" (p. 167).

$$E[\Pi_S] = (1-\rho)p + E\big(r_T^j\big)\rho p - c^j \quad ,$$

where a seller's expectation of the retainage return fraction given his quality and bid is driven by his posterior belief about the buyer's trustworthiness, with $E(r_T^L) = 0$ and $E(r_T^H) = \alpha r_T^H$.

Definition 2.1. *The breakeven bid associated with quality level j is $\underline{b}^j = \frac{c^j}{1-\rho(1-E(r_T^j))}$.*

Intuitively, the higher the seller's trust and/or the trustworthy buyer's retainage return, the lower the bid that a seller can profitably submit. For correct beliefs on α, any bid accompanying q^j below the breakeven level is weakly dominated by a bid equal to $\underline{b}^j$.

Corollary 2.3. *Equilibrium bids and the probability of market unravelling are (i) non-decreasing in ρ, (ii) non-increasing in α, and (iii) non-increasing in γ.*

Proof. If $\underline{b} = min\{\underline{b}^L, \underline{b}^H\}$ always forms part of an equilibrium supplier bidding strategy, then this result follows directly from Definition 2.1, Corollary 2.2 and the upper bound restriction on the bid interval. Suppose instead that there is an equilibrium supplier strategy that contains a bid other than $\underline{b}$. Holding constant the quality level, any bid below $\underline{b}$ would be strictly dominated by a bid equal to $\underline{b}$. Given the binding auction selection rule, it would also be a weakly dominant strategy for a competitor to undercut any bid above $\underline{b}$ and deliver the quality level associated with $\underline{b}$. Thus, any alternative bid cannot be part of an equilibrium supplier strategy. ■

We can now characterise the conditions under which there exists a PBE associated with the production of a high quality unit in our environment.

Proposition 2.2.

If $\frac{A}{1-C(1-A)} \leq \rho \leq \frac{B}{C}$, where $A = 1 - \frac{c^L}{c^H}$, $B = 1 - \frac{c^H}{v^H}$, $C = 1 - \alpha r_T^H$, then there exists a PBE in pure strategies in which $b_i^* = \underline{b}^H \ \forall\ i$, $a^* = a^1$, $q_i^* = q^H$, $r_T^* = r_T^H$ and $r_U^* = 0$.

Proof. Given Corollary 2.3, the proof amounts to showing that there exist feasible conditions under which $\underline{b}^H \leq \underline{b}^L$. This constitutes the condition on ρ. The left-hand inequality is obtained by equating the quality-specific breakeven bids per Definition 2.1. An additional condition necessary for this result to hold, $\underline{b}^L > c^H$, is never binding because $C(1-A) > 0$. The right-hand inequality is obtained by setting $\underline{b}^H \leq v^H$. Suppose the condition on ρ is satisfied and so the seller prefers high quality during the Production Stage in anticipation of the expected retainage return equal to αr_T^H. In the Procurement Stage, the buyer will accept a price equal to $\underline{b}^H$ independently of her type, because it guarantees her a non-negative profit. The procurement decision is uninformative, and the seller does not update his prior belief about α. In the Bidding Stage, there is no incentive for a supplier to deviate to a bid other than $\underline{b}^H$ (Corollary 2.3). Thus, there exists a PBE in pure strategies characterised as follows: $b_i^* = \underline{b}^H \ \forall \ i$, $a^* = a^1$, $q_i^* = q^H$, $r_T^* = r_T^H$ and $r_U^* = 0$. ■

Proposition 2.2 states that conditional on the prevailing levels of trust and beliefs about fairness, the retainage proportion can be set appropriately to incentivise high quality as an equilibrium strategy and mitigate the seller moral hazard problem. We call this the *implementable retainage interval*. The equilibrium bid is $\underline{b}^H$ and so the buyer appropriates all the gains from trade. The buyer will always find such an outcome profitable and the buyer types will separate in their final retainage return decision per the discussion above. The two inequalities that define this interval ensure that $\underline{b}^H \in [c^L, v^H]$ and $\underline{b}^H < \underline{b}^L$. In other words, the breakeven bid associated with high quality is feasible and below that of low quality.

To understand the bounds on the implementable retainage interval intuitively, note that A is a measure of seller moral hazard, B is a measure of the trade surplus generated by high quality, and C is a measure of the expected retainage return fraction lost if the seller produces a high quality unit. As A approaches one from below, the seller has greater cost incentive to produce low quality. As B approaches one from below, there are greater potential gains from producing a high quality unit. Conditional on the winning bid, high quality is more difficult to implement using retainage (i.e., the interval narrows) the larger the seller moral hazard and the smaller the trade surplus from high quality. Similarly, high quality is easier to implement using retainage the smaller the associated expected loss of retainage.

Corollary 2.4. *There is a direct and positive correspondence between trust and the size of the implementable retainage interval.*

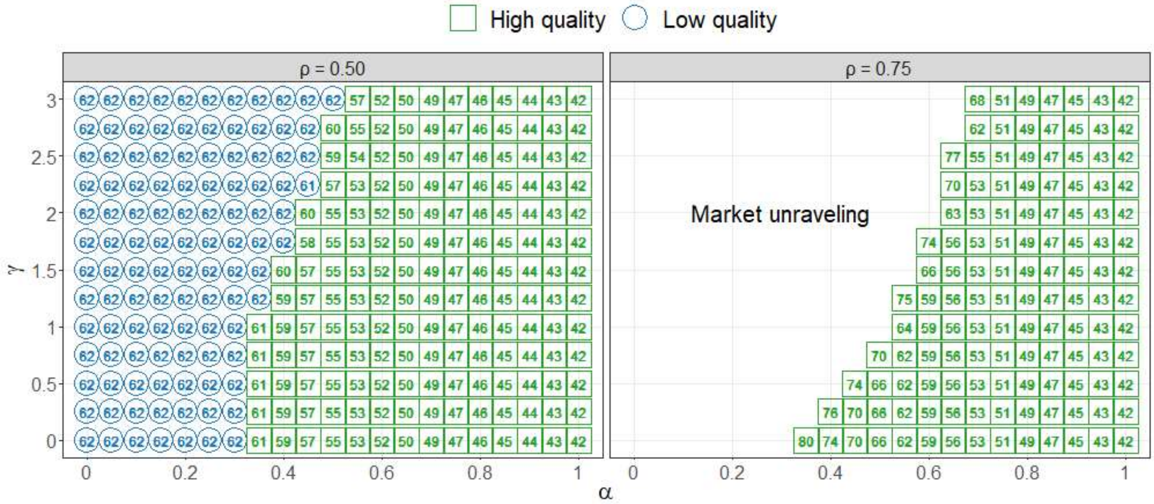
Proof. Follows directly from inspection of the implementable retainage interval. ■

To demonstrate the trade-offs between bids, participation and quality in our environment, consider an example. Suppose that $n = 2$ suppliers compete to win a procurement contract, with $c^L = 0.30$, $c^H = 0.40$, $v^L = 0.35$ and $v^H = 0.80$ (units in tens of thousands).[17] The contract contains a retainage provision, with $\rho = 0.50$ or $\rho = 0.75$. We will test these parameter values in the experiment. Figure 2.2 presents the spectrum of PBE outcomes in the (α, γ) space that are supported by the model of fair payment norms for each retainage arrangement. Blue circles (green squares) in the figure indicate beliefs for which the seller delivers low quality (high quality) as part of the equilibrium strategy. The number inside the shape indicates the equilibrium bid. The empty region indicates market unravelling.

In the implementable regions, the equilibrium bid is (weakly) decreasing in α and increasing in γ. Fixing γ and moving horizontally from left to right, high quality outcome is implementable above a certain trust threshold. Similarly, fixing α and moving vertically downwards, high quality is implementable above a certain fairness threshold, provided α is high enough. The right panel demonstrates that in a high retainage arrangement, the behavioural model can support high quality in addition to market unravelling as an equilibrium outcome.

[17] By setting $c^H > v^L$, we ensure that with zero retainage provision, the buyer accepts vulnerability due to uncertain seller production. This is in-keeping with Özer and Zheng's (2019) definition of trust.

Figure 2.2. Example: equilibrium outcomes supported by the behavioural model.



Notes: The figure displays equilibrium bidding strategies in (α, γ) space for $\rho = 0.50$ and $\rho = 0.75$. All computations are based on the following cost and valuation parameter values: $c^L = 0.30$, $c^H = 0.40$, $v^L = 0.35$ and $v^H = 0.80$. The numbers inside the shapes are the equilibrium bid amounts associated with the indicated quality level. The empty regions indicate market unravelling. Due to space constraints, we only consider $\gamma \leq 3$. Bids are discrete with minimum increment 0.01, which reflects the experimental implementation.

We remark that if trustworthy buyers were also to reward *low* quality according to some – same or different – fairness norm γ, uncertainty about the buyer's type (i.e., $0 < \alpha < 1$) would become a necessary condition to incentivise sellers to deliver high quality. That is, a non-monotonicity would be introduced into the relationship between trust and the implementable retainage interval. Above a certain threshold, the high probability of encountering a trustworthy buyer would allow the seller to submit a low bid and still be insured against loss in case of producing a low quality unit. In this situation, the lower bound on the implementable retainage interval becomes $A/(D - C(1 - A))$, where $D = 1 - \alpha r_T^L$ is the expected retainage return fraction lost if the seller produces a low quality unit. From Corollary 2.2, we have $D \geq C$, and so this new lower bound is non-decreasing in α. As a consequence, the seller would have an incentive to undercut his competitor in the Bidding stage and produce low quality ($\underline{b}^L < \underline{b}^H$). This is always the case for $\alpha = 1$, i.e., when the seller knows that the buyer is trustworthy.

2.3. Experiment and Hypotheses

2.3.1. Experimental design and procedures

To isolate the effect of contractual retainage on trade outcomes in the absence of confounds typically observed in the field, we conduct an experiment. In our lab setting, we consider the case of two suppliers.

We employed a between-subjects design with three treatments that varied the retainage provision, $\rho \in \{0, 0.50, 0.75\}$ – see Table 2.1. The $\rho = 0$ treatment benchmarks previous experiments for which payment of the winning bid in full is binding on the buyer, i.e., a fixed-price contract (cf. *Auction* in Fugger, Katok and Wambach 2019). The $\rho = 0.50$ treatment simulates a setting in which suppliers can fully offset the risk of partial or non-receipt of retainage monies by increasing their bids at auction. By contrast, in the $\rho = 0.75$ treatment, there is no bid available at which a supplier can ensure to avoid a loss during trade. Thus, in this treatment, profitable trade can proceed based on the buyer's intent to compensate the seller for costs incurred. We selected values $\rho = 0.50$ and $\rho = 0.75$ for the non-zero retainage treatments because they are the easiest retainage provisions for subjects to comprehend within the appropriate intervals from Proposition 2.1. The valuation and cost parameters associated with high and low quality are also summarised in Table 2.1. These values were displayed on the computer screens of all subjects. We restrict bids to be integers.

Subjects were randomly allocated to one of the three treatments and no subject participated in more than one session. Each treatment included six independent cohorts. There were three cohorts of the same treatment in every session. Each cohort consisted of three buyers and six sellers, who were randomly matched across 30 procurement interactions.[18] A bespoke algorithm guaranteed that no participant in a cohort played with the same pair of individuals in any two consecutive interactions. A total of 162 human subjects participated in our sessions, which were conducted at the lab of a large public university in the United States. Participants were students recruited using web-

[18] We did not inform subjects of the cohort size, to mitigate the possibility of tacit collusion in what might be considered a small cohort (see Katok 2011 for a discussion).

based recruitment software. Students and professionals appear to demonstrate qualitatively similar behaviours in the lab (Fréchette 2015). Of particular relevance for the generalisability of the experiments reported in this chapter and the next, procurement managers and construction industry professionals have also been observed to display the same decision-making biases as students (Dyer et al. 1989, Bolton et al. 2012).

Table 2.1 – Experimental treatments and parameter values.

Treatment	Retainage Level	Buyer's valuation (q^L, q^H)	Seller's cost (q^L, q^H)
1	$\rho = 0.00$		
2	$\rho = 0.50$	$(35, 80)$	$(30, 40)$
3	$\rho = 0.75$		

All sessions followed the same protocol. Upon arrival, participants were seated at computer terminals and handed a written copy of the instructions to read in private. Terminals had physical dividers to prevent subjects from seeing the screens of other participants. The instructions were played from an audio recording at the front to ensure the description of the game was common knowledge and delivery consistent across sessions. The task was explained to subjects using a cover story related to the application of interest and the instructions included concrete terms such as "Buyer" and "Seller". This was a deliberate choice to improve subject understanding (see Cooper and Kagel 2003) and increase external validity (see Katok 2017).[19] Participants completed a computerised test of understanding before being assigned to their role as a buyer or seller and matched into their first interaction group. Roles remained fixed throughout. Communication was prohibited and all interactions were anonymous. Own-group feedback was provided between periods. This information remained available in a history table to reinforce the game-theoretic assumption of "perfect recall". At the end of a session, participants answered a non-incentivised questionnaire to elicit demographic

[19] The instructions were framed in neutral language (see Zizzo 2010). We avoided the term "retainage". Audio recordings are available on request.

information, and attitudes to trust and risk. The experimental interface was programmed using oTree software (Chen et al. 2016).

Subjects received monetary incentives. Each subject was paid his or her summed experiment earnings privately and in cash at the end of a session, in addition to a $5 show-up fee. Payment was made sequentially, with sufficient time intervals between any two subjects to mitigate against the possibility of side-payments. We used a symmetric exchange rate of 20 experiment currency units (ECU) to $1. Average subject earnings were $17.70 in the $\rho = 0$ treatment, $22.70 in the $\rho = 0.50$ treatment and $25.00 in the $\rho = 0.75$ treatment. Sessions lasted 60 to 75 minutes. Each subject received a non-refundable endowment of 7 ECU per period to cover potential losses. Subjects were informed that they would not leave the session with less than the show-up fee.[20] To reinforce the one-shot nature of interactions, we did not inform subjects about cumulative earnings until after the final period.

2.3.2. Hypotheses

Below we outline the key hypotheses of interest in our experiment, based on the standard theory (*ST*) and the behavioural model of fair payment norms (*BM*). Our first two hypotheses relate to the *ex-ante* competitive market.

Hypothesis 2.1. Participation. *ST and BM both predict that the probability of market unravelling will be highest in* $\rho = 0.75$*; unravelling is an absolute prediction of ST only.*

Hypothesis 2.1 captures the potential anti-competitive effect of a procurement arrangement in which the seller cannot guarantee to cover future production costs. Market unravelling is the unique equilibrium prediction under *ST*. Thus, observing substantial participation rates in $\rho = 0.75$ would offer support for the relevance of *BM*. No set of beliefs under *BM* can sustain non-participation in $\rho = 0$ or $\rho = 0.50$ and so market unravelling in these treatments is expected to be negligible.

[20] In $\rho = 0.75$, limited liability was imposed for two sellers. All results reported below hold if we exclude these two subjects from our analysis. In a pilot experiment, we tested the most extreme seller trust arrangement of 100% retainage. Seller losses, however, became a problem. Summary statistics for this variant are available on request.

Hypothesis 2.2. Prices. *ST predicts prices will be rank ordered* $p_{\rho=0} < p_{\rho=0.5} < p_{\rho=0.75}$*; BM predicts bids will be lowest in* $\rho = 0$*.*

Hypothesis 2.2 captures the potential inflationary effect of retainage, as reflected in tender prices. *ST* predicts that, conditional on trade, as the retainage proportion increases bidders will submit higher prices. *BM* offers no belief-independent comparative static between the non-zero retainage treatments but does predict that sellers will bid lowest in $\rho = 0$.

The next two hypotheses incorporate the *ex-post* bilateral trade relationship between the buyer and seller. At this point, we define a measure of *trade efficiency* as the proportion of trade surplus realized out of the total available given the sellers' bidding decisions. For our parameter values, low (high) quality corresponds to trade efficiency of 0.125 (1.00).

Hypothesis 2.3. Trade efficiency. *ST predicts no difference in trade efficiency among treatments; BM predicts that trade efficiency is weakly higher in* $\rho > 0$ *than in* $\rho = 0$*.*

Hypothesis 2.3 represents the main empirical test of interest: can retainage mitigate the moral hazard problem and incentivise high quality? Our theoretical analysis suggests that in $\rho = 0.50$ and $\rho = 0.75$, high quality is implementable under *BM* but not under *ST*. In $\rho = 0$, neither model can rationalise high quality as an equilibrium outcome. In $\rho = 0.75$, if we observe trade then *BM* tells us that it is most likely to be efficient. The full set of equilibrium seller strategies under *BM* are provided in Figure 2.2. Our measure of trade efficiency does not consider the surplus-reducing effect of market unravelling. To this end, we construct an additional *global efficiency* measure defined as the proportion of trade surplus realised out of the total attainable, i.e., if the market always attracts at least one bid. To what extent the higher likelihood of market unravelling in $\rho = 0.75$ undermines global efficiency, relative to the other retainage arrangements, is an empirical question and so we do not place a formal hypothesis on the trade-off between market unravelling and trade efficiency.

Hypothesis 2.4. Reciprocity. *ST predicts that the retainage return will be independent of quality; BM predicts that the retainage return will be positively correlated with the seller's quality choice.*

Hypothesis 2.4 tests for the presence of buyer reciprocity. Prior economic experiments suggest that the existence of positive and negative reciprocity is robust in the lab to the imposition of demanding market institutions (e.g., Fehr and Falk 1999). If a positive correlation is observed between quality and the retainage return in our experiment, this type of behaviour would be consistent only with *BM*.

2.4. Experimental Results

In this section, we first outline the main aggregate results in relation to our four experimental hypotheses. Second, we analyse outcomes at the cohort-level, to gain greater insight into the evolution of market dynamics. Finally, we examine market choices at the individual-level, to better understand how retainage influences buyer and seller decision-making and conduct a robustness analysis to check for possible experiment confounds.

2.4.1. Aggregate findings

Table 2.2 summarises average seller participation rates, prices and quality levels, along with buyer acceptance rates and the retainage returns. The table also presents summary statistics on market unravelling, our efficiency measures and profits.[21] Consistent with Hypothesis 2.1, sellers nearly always bid in $\rho = 0$ and $\rho = 0.50$. By pairwise comparisons, the percentage of participating bidders is significantly lower (64%) in $\rho = 0.75$ (both p-values = 0.015). This variable seller participation is reflected in different

[21] For each treatment, we have data from 540 matching groups. Since there is no interaction between subjects playing in different cohorts, each cohort is considered a statistically independent observation. We employ two-tailed Signed-Rank tests for one-sample comparisons and Wilcoxon-Mann-Whitney tests for two-sample comparisons, correcting for multiple testing using Holm's (1979) p-value adjustment method. We acknowledge the potential *caveat* of arbitrary static correlations within sessions (Fréchette 2012).

rates of market unravelling. In $\rho = 0.75$, 22% of auctions fail to attract a single bidder. By contrast, no market unravels in the lower retainage arrangements.

Result 2.1. *High retainage structures deter supplier participation in the contracting process.*

Table 2.2 – Cohort means and standard deviations.

	$\rho = 0$	$\rho = 0.50$	$\rho = 0.75$
Panel A: Decision variables			
Seller participation	0.99	0.98	0.64
	(0.01)	(0.03)	(0.26)
Price	35.1	57.8	57.2
	(1.72)	(5.42)	(8.08)
Buyer acceptance	0.60	0.91	0.98
	(0.20)	(0.08)	(0.01)
High quality	0.06	0.30	0.50
	(0.02)	(0.20)	(0.19)
Retainage return (low)		0.06	0.16
		(0.07)	(0.08)
Retainage return (high)		0.32	0.40
		(0.16)	(0.16)
Panel B: Market outcomes			
Market unravelling	0.00	0.00	0.22
	(0.00)	(0.00)	(0.26)
Trade efficiency	0.11	0.36	0.56
	(0.03)	(0.17)	(0.17)
Global efficiency	0.11	0.36	0.46
	(0.03)	(0.17)	(0.25)
Buyer profit	2.83	15.8	31.4
	(1.07)	(7.65)	(2.49)
Seller profit	4.43	-0.25	-8.75
	(1.54)	(1.60)	(5.71)

Notes: Mean (SD) values for the key parameters in our experiment based on cohort averages. Profit data are per round and exclude the endowment.

That sellers in $\rho = 0.75$ still choose to submit a bid in the majority of auctions, despite the high vulnerability to loss due to uncertain retainage return behaviour of the buyer, is preliminary evidence to suggest a role for trust in the decision-making process. *BM* thus provides a more satisfactory description of sellers' participation behaviours than *ST*. Further evidence that the retainage level shifts the burden of trust in trade relationships away from the buyer can be inferred from buyers' acceptance behaviours. The acceptance rates are increasing in the retainage level, from 60% in $\rho = 0$ to 91% in $\rho = 0.50$ and 98% in $\rho = 0.75$.

Average contract prices are higher in $\rho = 0.50$ (57.8) and in $\rho = 0.75$ (57.2) than in $\rho = 0$ (35.1). These pairwise differences versus the zero retainage treatment are highly significant (both p-values < 0.01). We also reject the point prediction of *ST* that prices equal 30 in $\rho = 0$ (p-value $= 0.031$); we fail to reject an average price of 60 in $\rho = 0.50$ (p-value $= 0.563$). There is no significant difference in prices between the two non-zero retainage treatments, although prices are more variable when $\rho = 0.75$. Thus, the aggregate price data supports the price differences of *BM* outlined in Hypothesis 2.2, but rejects the rank ordering of *ST*.

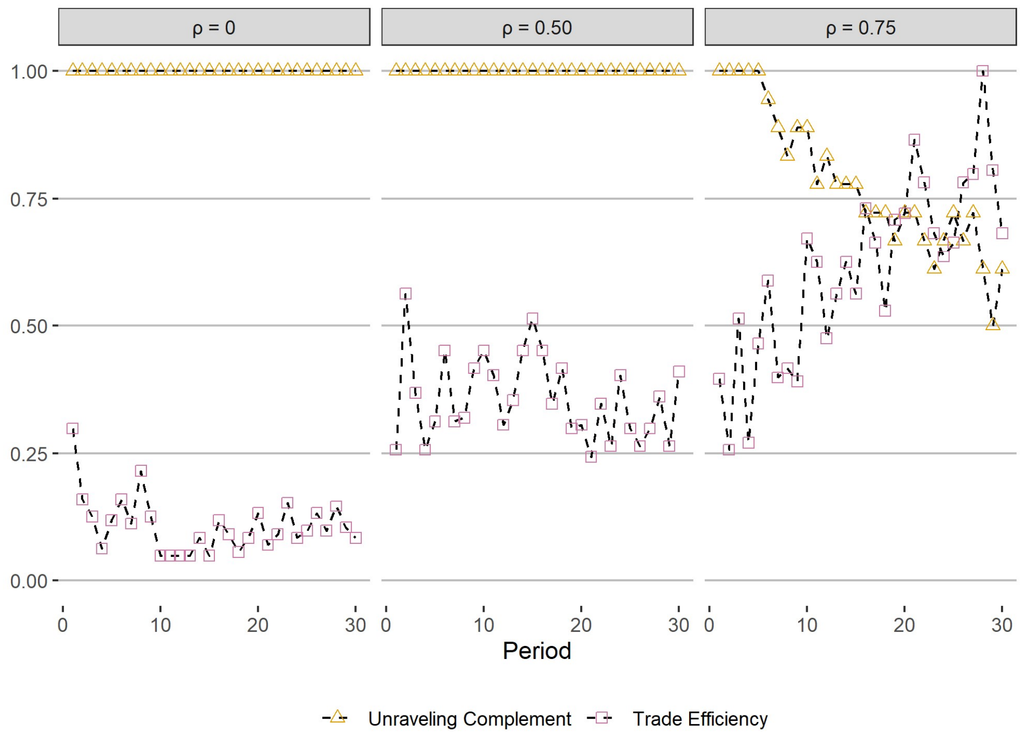
Result 2.2. *Retainage provisions inflate tender prices.*

Results 2.1 and 2.2 underscore the potential for retainage to have an anti-competitive effect on the procurement process. The flip side is that retainage significantly improves average quality levels. In $\rho = 0$, low quality is chosen by the seller in 93.9% of transactions. We fail to reject the null that trade efficiency attained its Nash level of 12.5% (p-value $= 0.31$). In $\rho = 0.50$, the proportion of transactions associated with high quality is 30.2%, and in $\rho = 0.75$ it is 50.4%. The difference in quality levels between the non-zero retainage treatments is significant at the 10% level (p-value $= 0.093$) and is consistent with the prediction of *BM* that, conditional on attracting seller bids, trade in the high retainage arrangement is more likely to be of high quality. The relative quality choice frequencies in these treatments generate significant trade efficiency gains when compared to $\rho = 0$ (both p-values < 0.01). As stated in Hypothesis 2.3, such an improvement can be rationalised only by *BM*.

Result 2.3a. *Retainage mitigates the seller moral hazard problem and increases trade efficiency.*

Owing to the observed market unravelling, global efficiency in $\rho = 0.75$ is 10 percentage points lower than trade efficiency and is not significantly higher than in $\rho = 0.50$ (p-value = 0.589). To further investigate the trade-off between the seller's quality and participation decisions, in Figure 2.3 we plot a time series of the market unravelling complement and trade efficiency across the 30 periods in our experiment. The trend differences in trade efficiency are pronounced. In $\rho = 0$, trade efficiency fluctuates about its Nash equilibrium level within the 0-25% interval. In $\rho = 0.50$, trade efficiency fluctuates within the 25-50% interval. In both of these treatments, the variability of trade efficiency declines over time, while the market never unravels.

Figure 2.3. Decomposition of market unravelling and trade efficiency over time.



Notes: Based on 18 matching groups per treatment in a period. Unravelling Complement is one minus the proportion of auctions which failed to attract a single bidder. Trade efficiency is a measure of surplus realized divided by surplus made available.

Most interesting is $\rho = 0.75$, in which trade efficiency begins in the 25-50% interval then trends upwards over time. By the end of the session and notwithstanding an end-game effect, nearly all transactions that take place are of high quality, which as we observed in Figure 2.2 (right panel) is a direct prediction of *BM.* Meanwhile, after period five, there is a marked fall in the number of auctions attracting at least one bidder. This trend continues into the final period, at which point around half of markets unravelled. Notably, in the second half of the experiment, the difference in global efficiency between $\rho = 0$ and $\rho = 0.75$ is not significant at conventional levels (p-value = 0.13). On the other hand, global efficiency remains significantly higher in $\rho = 0.50$ than in $\rho = 0$ after period 15 (p-value = 0.015).

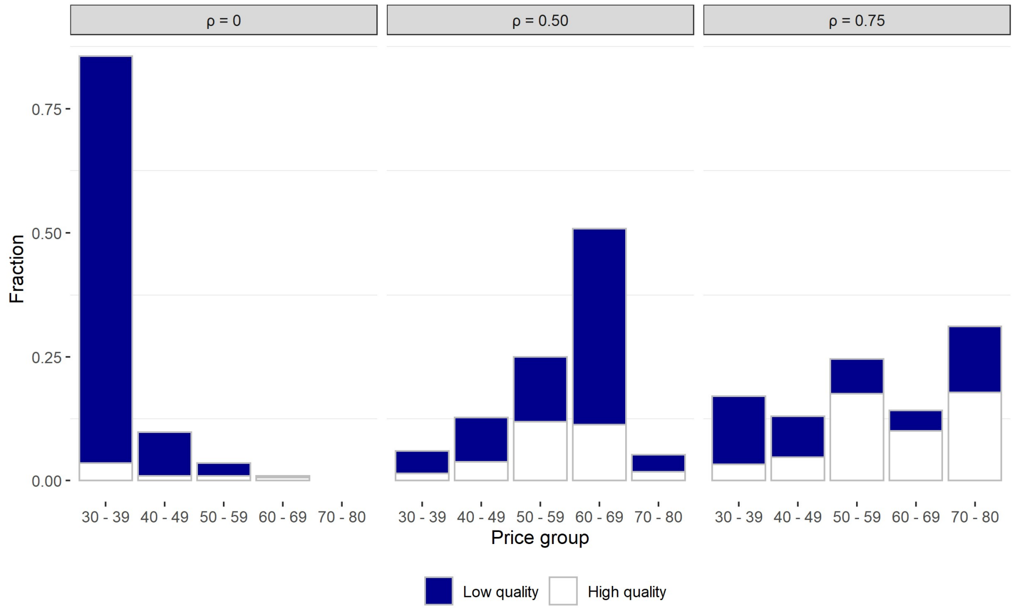
Result 2.3b. *High retainage structures undermine global efficiency.*

In Figure 2.4 we present relative frequencies of trade by price and quality, conditional on at least one bid submitted and on buyer acceptance. Prices are in intervals of 10. More than 85% of transactions in $\rho = 0$ are in the 30-39 interval and the quality of these transactions is near-uniformly low. The majority of transactions in $\rho = 0.50$ fall into the interval 50-69, with low quality most commonly observed at prices above 60, and high quality more likely at prices below 60. In $\rho = 0.75$, high quality is the majority choice for sellers at prices above 50. These patterns offer indirect evidence to support *BM.*

Turning to Hypothesis 2.4, we reject the null of independence between the quality level and the buyer's retainage return predicted by *ST,* in favour of the reciprocal relationship predicted by *BM.* In $\rho = 0.50$, the seller's probability of receiving a non-zero retainage return is 31% after choosing low quality and 61.3% after choosing high quality. Buyers in this treatment return just 6.1% in exchange for low quality and 31.9% of retainage monies in exchange for high quality. In $\rho = 0.75$, where the retainage represents a larger share of the price, return rates are 15.6% and 40.4%, respectively. The premiums paid for high quality in each treatment are significant (both p-values = 0.031). Positive reciprocity is evident in the comparative distributions of non-zero retainage returns by quality level (see Figure 2.5).

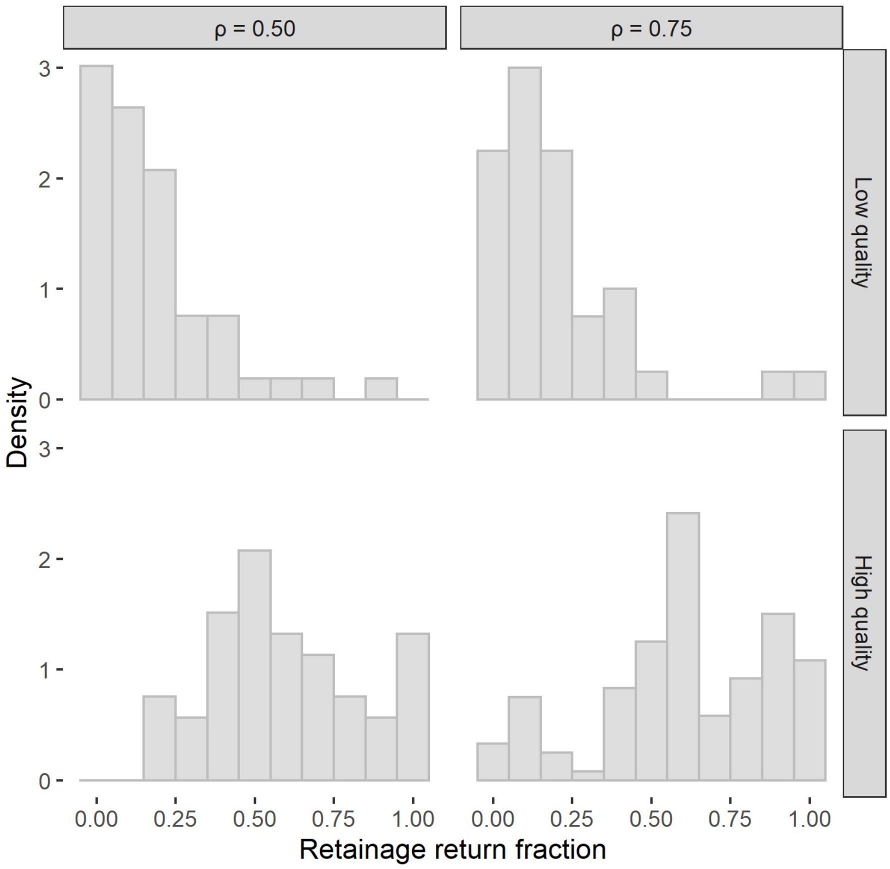
Result 2.4. *Buyers reciprocate high quality with a more generous retainage return.*

Figure 2.4. Relative frequencies of trade by price group and quality level.



Notes: Relative frequencies are conditional on trade.

Figure 2.5. Distributions of non-zero retainage returns by quality level.



Notes: Histograms of strictly positive retainage returns.

Retainage also has implications for the distribution of profits. In $\rho = 0$, sellers are able to maintain a positive profit and even earn slightly more (4.43) than buyers (2.83) on average, although this difference is not significant (p-value = 0.156). Buyers gain substantially from the introduction of retainage. In $\rho = 0.50$, sellers earn approximately zero (which would be expected given the market imbalance) and buyers earn 15.8, a significant profit differential (p-value = 0.031). Sellers fare significantly worse in $\rho = 0.75$, incurring an average loss of 8.75, suggesting that their trust in the buyer's willingness to reciprocate is often misplaced.

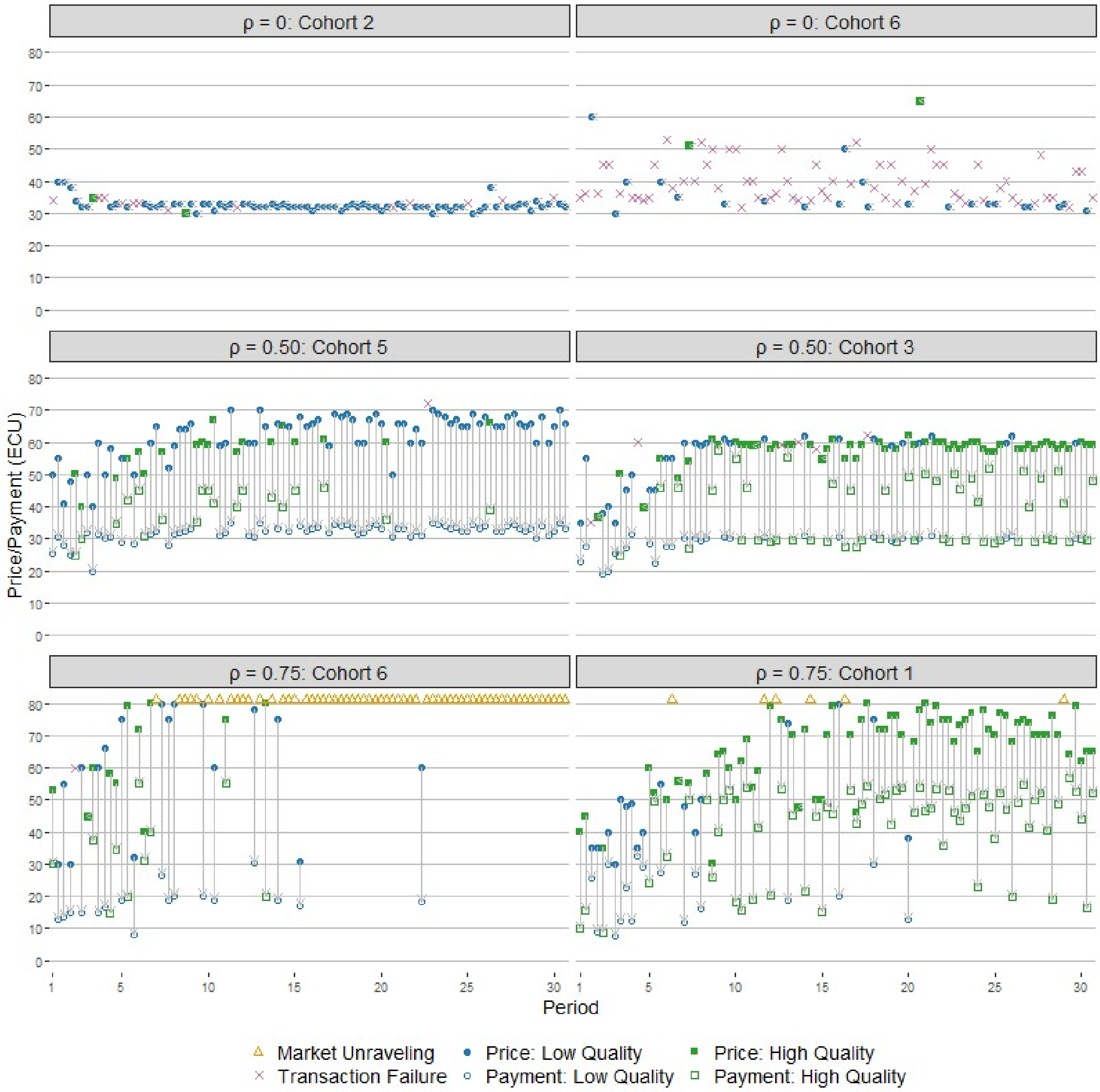
Our theoretical analysis is equilibrium-based. To check whether subjects' learning in the experiment is an important behavioural factor, we split the dataset into three blocks of 10 periods and conduct a formal analysis of the differences in decision-making over time. The full results of this analysis are provided in the appendix. There is some evidence of learning early on, but that behaviour appears to converge after period 10, with no significant differences in buyer or seller decisions between the second and third blocks.

2.4.2. Cohort dynamics

To obtain insight into the market dynamics observed over the course of the experiment, we consider each cohort separately and plot the outcomes per interaction group and period, in relation to the winning bid (Figure 2.6). This approach yields 90 observations per cohort. The figure displays outcomes for two representative cohorts per treatment, which capture the main trends that emerge. An open triangle is an instance of market unravelling and a cross is transaction failure due to buyer rejection of the winning bid. A solid circle (square) is an accepted winning bid at which the seller produced low (high) quality, and an open circle (square) is the corresponding total payment. The arrows capture the differential between agreed price and total payment.

The top two panels in Figure 2.6 correspond to $\rho = 0$. The top-left cohort exemplifies the disciplining power of competition with a fixed-price contract: after some early adjustments, sellers submit bids in the 30-35 interval and choose low quality. The top-right cohort reveals a different dynamic: sellers attempt to elicit buyer acceptances at prices above the buyer's value of low quality. Buyers reject most such attempts.

Figure 2.6. Representative cohort outcomes over time.



Notes: Panels display procurement outcomes over time in the specified treatment and cohort. An open triangle is an instance of market unravelling, in which neither seller submitted a bid. A cross is a transaction failure, in which a buyer rejected the winning bid. A solid circle is an accepted winning bid at which the seller produced low quality. A solid square is an accepted winning bid at which the seller produced high quality. An open circle (square) are the corresponding total payments in instances where these differ from the winning bid. In such instances, the vertical arrows represent the price-payment differential.

The middle two panels in the figure reveal a reverse bidding trend in $\rho = 0.50$, as sellers learn to adjust their bids upwards to account for retainage monies lost over time. In the middle-left cohort, sellers produce low quality, marking up their bids proportional to the associated cost. In the middle-right cohort, sellers produce high quality and buyers reward this with a positive retainage return, compensating sellers for the increased delivery cost.

The bottom two panels in the figure highlight variable participation in $\rho = 0.75$. In the bottom-left cohort, sellers gradually choose not to participate in request for bids. Two-thirds of markets unravel in this cohort. Sellers in the bottom-right cohort are willing to participate and produce high quality. Where a buyer fails to reciprocate, the seller's downside is larger than in $\rho = 0.50$, resulting in losses on average.

2.4.3. Individual decision-making

To analyse individual bidding decisions in the *ex-ante* competitive market, we code each seller non-participation decision as a bid at one increment above the highest available. We then use a Tobit estimator that censors the bid from above (Table 2.3). We include the once-lagged competitor's bid and once-lagged trade as predictors, along with a time trend and control variables from the post-experiment questionnaire. Standard errors are clustered at the subject-level.[22]

In all three treatments, we infer that bids exhibit a positive dependency on the most recently matched competitor's bid and that subjects who traded in the most recent period learn to submit lower bids. The smaller sample size in column (3) reflects the elevated seller non-participation in $\rho = 0.75$. In an (unreported) regression, we further estimate determinants of participation in $\rho = 0.75$. Only the (negative) time trend is a significant predictor of entry decisions (p-value = 0.011).

[22] All results reported in this section are qualitatively unchanged if we account for non-independence of observations by clustering standard errors at the cohort-level or using a random effects estimator. A few subjects did not provide a complete set of responses in the post-experiment questionnaire. The results are robust to dropping the control variables and using the full dataset. The once-lagged competitor's bid and once-lagged trade variables are correlated and so problematic for causal inference; the sign and significance of the partial effects are unchanged by running separate regressions in which we include only one of the two variables as a predictor.

Table 2.3 – Censored regression of sellers' bids in the *ex-ante* competitive market.

	Bid		
	$\rho = 0$	$\rho = 0.5$	$\rho = 0.75$
	(1)	(2)	(3)
Competitor's bid $_{t-1}$	0.24***	0.51***	0.38***
	(0.05)	(0.08)	(0.06)
Trade $_{t-1}$	-4.68***	-6.35***	-5.76*
	(0.86)	(1.04)	(2.48)
Period	-0.16*	0.16***	0.55***
	(0.06)	(0.03)	(0.12)
Constant	26.73***	29.98***	48.99***
	(5.90)	(8.76)	(12.78)
Control variables[1]	Yes	Yes	Yes
Log likelihood	-3390	-3448	-2033
Observations	942	995	603

Notes: The models were estimated using a Tobit regression; seller non-participation decisions are coded as bids equal to 81. Robust standard errors in parentheses, clustered at the subject level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

[1] The controls include self-reported risk and trust indices, dummies for female and economics or business major, age and income rank (coefficient estimates are available on request).

In Table 2.4 we conduct a series of Logit regressions to investigate the determinants of buyer and seller decisions in the *ex-post* trade relationship with a retainage contract.[23] In columns (1) and (2), the dependent variable is a dummy variable for high quality seller production. In columns (3) and (4), we consider specifications in which the dependent variable is the buyer's retainage return fraction and fit a fractional response model.

There is a significant quadratic relationship between price and quality in both treatments. For $\rho = 0.50$, this is consistent with the observation from *BM* that above a certain price, low quality is the unique equilibrium choice (left panel in Figure 2.2). The positive and significant coefficient estimate on the once-lagged retainage return in $\rho = 0.50$ suggests that retainage provisions are most effective at mitigating seller-side moral hazard when the seller has had a recent positive trade experience with this type of contract. In $\rho = 0.75$, we also observe a positive effect of the once-lagged return on the

[23] We also conducted regression analyses for $\rho = 0$ with the buyer's acceptance and seller's quality decision as dependent variables. Acceptances are significantly decreasing in winning bid (*p*-value = 0.001) and high quality is more likely at higher prices (*p*-value = 0.029).

probability of high quality (p-value = 0.083). There is also a highly significant positive time trend in this treatment. This appears to capture the withdrawal of less trusting sellers from the market (see next section) and suggests that high retainage structures are most effective at inducing high quality once trust is established.

There is strong evidence to support the positive reciprocity prediction of *BM* in Hypothesis 2.4 – high quality has a significant positive effect on the retainage return in both treatments. The significant negative relationship between price and the retainage return suggests that buyers are concerned about distributional outcomes in their return decision.

Table 2.4 – Logit regression analysis of buyer and seller decisions in the *ex-post* bilateral trade relationships with the retainage contract.

	High quality		Retainage return[2]	
	$\rho = 0.50$	$\rho = 0.75$	$\rho = 0.50$	$\rho = 0.75$
	(1)	(2)	(3)	(4)
Price	0.95*	0.35**	-0.07***	-0.02**
	(0.40)	(0.11)	(0.01)	(0.01)
Price^2	-0.01*	-0.003**		
	(0.004)	(0.001)		
Retainage return $_{t-1}$	1.72**	0.68		
	(0.58)	(0.39)		
High quality			3.16***	1.84***
			(0.48)	(0.33)
Period	0.01	0.16***	-0.03	-0.03
	(0.02)	(0.04)	(0.02)	(0.01)
Constant	-27.61*	-4.75	12.02***	-0.74
	(11.66)	(3.29)	(2.29)	(1.77)
Control Variables[1]	Yes	Yes	Yes	Yes
Log Likelihood	-120.04	-73.22	-121.5	-200.3
Observations	236	178	492	415

Notes: Coefficient estimates are on the logit scale with robust standard errors in parentheses, clustered at the subject level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

[1] The controls include self-reported risk and trust indices, dummies for female and economics or business major, age and income rank (coefficient estimates are available on request).

[2] Fit a fractional response model.

Finally, we conduct a robustness analysis to test for the presence of two possible experiment confounds in the $\rho = 0.50$ and $\rho = 0.75$ treatments. We use the distinctiveness, relevance and plausibility (DRP) method of Zizzo (2013).

(i) Wealth effects. Subjects' behaviour might change as they accumulate earnings over the course of the experiment at variable rates. Wealth effects are distinctive from the determinants of behaviour considered in the behavioural model. They are same-direction relevant for the seller's choice of high quality. If subjects exhibit decreasing absolute risk aversion, then the accumulation of earnings over the experiment may lead them to choose the riskier choice (high quality) more often in the knowledge that they have a buffer against possible losses.

(ii) Limited liability. Limited liability is distinctive and relevant in the opposing direction to wealth effects: as cumulative earnings fall towards zero, sellers may exhibit more risk-seeking behaviour in the knowledge that their downside is limited to zero. It is not "magnifying glass relevant" to the extent that firms entering into procurement transactions are vulnerable to going out of business if they make losses on multiple transactions.

To investigate the plausibility of these possible confounds, we re-run the regression analyses in columns (1) and (2) of Table 2.4, with the additional inclusion of a seller's unlimited cumulative earnings as a regressor. The results are presented in Table 2.5. The determinants of high quality are qualitatively unchanged in both treatments. Cumulative earnings display no statistical relationship with the seller's choice of high quality in the $\rho = 0.50$ treatment. In the $\rho = 0.75$ treatment, there is a positive effect of cumulative earnings on the probability of high quality, although the effect size is small. There remains a significant upward trend in market quality over time in the $p = 0.75$ treatment which is independent of earnings.

We will provide evidence in the next section to suggest that this trend captures the trust-building between buyers and sellers that remain in the market.

Table 2.5 – Cumulative earnings and sellers' quality decisions.

	High quality			
	$\rho = 0.50$		$\rho = 0.75$	
	$\hat{\beta}$	dy/dx	$\hat{\beta}$	dy/dx
Price	0.91*	0.15**	0.25*	0.03*
	(0.39)	(0.06)	(0.12)	(0.01)
Price ^2	-0.01*	-0.001**	-0.002	-0.0003*
	(0.003)	(0.001)	(0.001)	(0.0001)
Retainage return $_{t-1}$	1.86**	0.31***	0.48	0.06
	(0.62)	(0.09)	(0.39)	(0.05)
Period	0.09	0.02	0.17***	0.02***
	(0.07)	(0.01)	(0.04)	(0.004)
Cumulative earnings $_{<t}$	-0.01	-0.002	0.01**	0.001**
	(0.01)	(0.002)	(0.003)	(0.0004)
Constant	-26.66*		-2.25	
	(11.24)		(3.54)	
Control Variables[1]	Yes		Yes	
Log Likelihood	-118.8		-70.12	
Observations	236		178	

Notes: Coefficient estimates ($\hat{\beta}$) are on the logit scale, with average marginal effects (dy/dx) reported alongside. Robust standard errors are in parentheses, clustered at the subject level.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.
[1] The controls include self-reported risk and trust indices, dummies for female and economics or business major, age and income rank (coefficient estimates are available on request).

2.5. Behavioural Model Estimation

The findings outlined in the previous section suggest that our model of fair payment norms can organise several behavioural patterns observed in the data. In the theoretical analysis, we reduced the buyer population to two types: trustworthy and untrustworthy. A crude look at the individual-level buyer return data after period 10 supports the existence of different buyer types in the experiment. Across the pooled retainage data, 18 out of 36 buyers either returned zero or made a partial retainage return in exchange for high quality which resulted in the seller incurring a loss. This number increases to 28 in exchange for low quality, which suggests that buyers exhibit both positive and negative reciprocity in the experiment and provides some empirical justification for our assumption that trustworthy buyers reward only high quality. Moreover, 15 buyers always returned some positive fraction to the seller in exchange for high quality, with mean returns among these subjects often exceeding 50% of the retainage amount.

To examine seller trust and beliefs about fairness, we conduct a maximum likelihood estimation of our behavioural model's parameters in a logit choice framework. The probability that seller i chooses high quality in period t conditional on price p and retainage ρ is:

$$Pr_{it}(q^H) = \frac{e^{\lambda \bullet \Delta U}}{1 + e^{\lambda \bullet \Delta U}}, \tag{2.2}$$

where $\Delta U = \alpha r_T(p_{it}, q^H; \gamma)\rho p_{it} - (c^H - c^L)$, and λ is the rationality parameter which is inversely related to the level of decision error (McKelvey and Palfrey 1995). For the trust parameter, we specify $\alpha = \beta_0 + \beta_1 I_{it,0.75} + \beta_2 t + \beta_3 t \bullet I_{it,0.75}$, where $I_{it,0.75}$ is a dummy variable for $\rho = 0.75$, which captures level and trend differences between our retainage treatments. We restrict $\gamma > 0$ by transforming $\tilde{\gamma} = \ln(\gamma)$ to keep the optimization problem unconstrained. The resulting likelihood function is given by:

$$L(\alpha, \gamma, \lambda) = \prod_{i=1}^{N} \prod_{t=1}^{T} [Pr_{it}(q^H)]^{q_{it}} [1 - Pr_{it}(q^H)]^{1-q_{it}}, \tag{2.3}$$

where the total number of sellers is N and the total number of periods is T. We cluster robust standard errors at the subject level. Based on our learning analysis in the previous section and to mitigate issues of serial correlation (Davis 2015, p. 334), we drop observations from the first block of 10 periods. This leaves us with a total of 573 transactions across the two treatments, of which 322 are from $\rho = 0.50$ and 251 are from $\rho = 0.75$. The results of this estimation are presented in Table 2.6. We consider three specifications.

First, we consider the standard theory (*Baseline*) in which we restrict the behavioural model parameters to equal zero. The standard theory does not explain the data well, with low seller rationality and an estimated λ not significantly different from zero (i.e., all errors).

Second, we estimate the constant parameters of our behavioural model (*BM-1*). The parameter estimates suggest that on average, sellers believe that there is a one in three chance of encountering a trustworthy buyer and that the fair reference point yields the buyer approximately 1.3× the seller's profit. The 95% confidence interval for our γ estimate includes equal profit-sharing. The α and γ parameter estimates are significantly different from zero (both p-values < 0.001) and rationality increases. The behavioural model also overwhelmingly outperforms the baseline when comparing the log-likelihood, AIC and BIC values and based on the results from a nested likelihood ratio test ($\chi^2 = 133.86$, p-value < 0.001).

Table 2.6 – Results of the structural estimation.

	Baseline	BM-1	BM-2
	(1)	(2)	(3)
α			
β_0		0.312***	0.392***
		(0.019)	(0.024)
β_1			-0.170***
			(0.038)
β_2			-0.002
			(0.001)
β_3			0.005*
			(0.002)
γ		1.297***	1.489***
		(0.291)	(0.195)
λ	0.019	0.452***	0.820***
	(0.020)	(0.076)	(0.204)
AIC	791.44	661.58	638.42
BIC	795.79	674.63	664.53
Log Likelihood	-394.72	-327.79	-313.21
Observations	573	573	573
(Clusters)	(67)	(67)	(67)

Notes: Robust standard errors in parentheses, clustered at the subject level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. The estimation is based on the Logit choice framework in (2). We specify $\alpha = \beta_0 + \beta_1 I_{it,0.75} + \beta_2 t + \beta_3 t \bullet I_{it,0.75}$, where $I_{it,0.75}$ is a dummy variable for $\rho = 0.75$.

Third, we re-estimate the behavioural model and allow α to vary as a function of the retainage level and time (*BM-2*). In this specification, trust starts off higher in $\rho = 0.50$, with sellers assigning 40% probability to the probability of buyer trustworthiness. The results reinforce our earlier observation that trust increases over time in $\rho = 0.75$, with a significant positive interaction between the treatment dummy and the time trend (p-value < 0.01). By contrast, there is some evidence of a fall in trust over time in $\rho = 0.50$, although this effect is statistically weak (p-value $= 0.064$). The fair reference point is estimated to be slightly higher at around 1.5× the seller's profit, although the confidence intervals between *BM-1* and *BM-2* overlap. Seller rationality in this specification is not significantly below one. We prefer the full model based on all available statistical comparisons ($\chi^2 = 29.16$, p-value < 0.001).

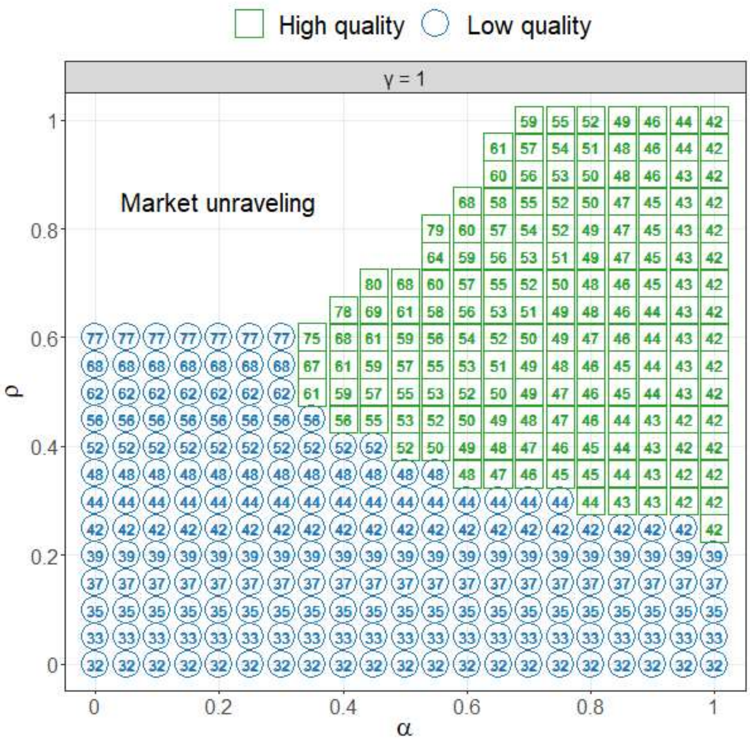
The behavioural estimation implies that seller trust levels in our experiment were between $\alpha = 0.3$ and 0.4, and the perceived fair reference point γ at slightly above one. We benchmark our α estimate against previous studies. Using a least absolute deviation approach, Bolton and Ockenfels (2000) estimate $\alpha = 0.50$ in the gift-exchange experiments of Fehr et al. (1993) and $\alpha = 0.42$ in the trust game experiments of Berg et al. (1995).[24] The structure of our *ex-post* bilateral trade relationship is closer to the trust game and our α estimate reflects this.

Payoff equality is normatively appealing (Konow 2003) and commonly found at the firm level (Andreoni and Bernheim 2009). Given this observation, it is interesting to infer what the behavioural estimates imply for the optimal retainage level, as formulated in the equilibrium analysis. In Figure 2.7, we depict equilibrium outcomes in the (α, ρ) space, given our estimate for the fair reference point $\gamma \approx 1$. For $\rho = 0.50$, we can infer from the figure that our upper bound estimate of $\alpha = 0.4$ lies on the threshold at which high quality is an implementable seller strategy. The equilibrium bid at this coordinate (59) is also very close to the average observed in the data (57.8). For $\rho = 0.75$, the level of trust required to implement high quality is 0.55, greater than our estimate from the data. This suggests substantial heterogeneity in sellers' beliefs in this treatment, which is not captured by our model but implied by the unravelling in certain cohorts of this treatment. It also suggests that some sellers were overly optimistic in their beliefs about others' trustworthiness (for a related result, see Li et al. 2019).

---

[24] We obtain similar estimates for α using a least absolute deviation approach on the observed profit data (details are contained in the appendix).

Figure 2.7. Implementable retainage levels supported by the behavioural model estimation.



Notes: The figure displays equilibrium seller strategies in (α, ρ) space for $\gamma = 1$.

2.6. Discussion and Conclusion

The managerial literature has devoted substantial attention to how auction institution design can improve procurement efficiency. In competitive procurement, standard principal-agent models show that high product quality equilibria do not exist due to moral hazard. Thus, we see preferences among procurers for repeat purchases from a smaller group of certified suppliers and the establishment of reputational mechanisms. The identification of alternative incentive mechanisms that can reap the benefits of price competition while sustaining cooperation is of considerable managerial value.

In our experiments, we find that retainage significantly increases the probability of high quality project delivery relative to the benchmark zero retainage case. There is a trade-off between trade efficiency and transaction fairness, due to the difficulty in sustaining trustworthy buyer payment behaviour over time. When sellers are engaged informally, with no payment guarantee to cover production costs, we observe a significant unravelling in markets over time. This reduces overall procurement efficiency, despite the higher quality of transactions remaining in the market. We develop a model of fair payment norms and show that, by explicitly incorporating trust and trustworthiness into the analysis, we can improve our understanding of decision-making at the individual-level.

The experimental results are consistent with the finding of Fugger, Katok and Wambach (2019) that reverse auctions perform poorly from an efficiency perspective with rigid fixed-price contracts when there exists seller-side moral hazard. The benefits of mechanisms that permit buyers to incorporate aspects other than price in their procurement decisions have been discussed in detail (see, e.g., Englebrecht-Wiggans et al. 2007). Here we demonstrate that attention should also be given to the flexibility of the contractual price, even when price cannot be conditioned upon quality.

The main message for procurement managers and construction industry practitioners is as follows. Retainage mechanisms, and deferred payments more generally, have a useful role in incentivizing sellers to provide high quality. This recommendation comes with two *caveats*. First, procurement arrangements based on intent may have an anti-competitive effect by discouraging suppliers from participating in the market. Second, if retainage monies are not properly administered, sellers are vulnerable to the payment malpractice of an untrustworthy buyer. This can undermine seller profits relative to fixed-price contracts and the deductibility of retainage throughout the supply chain would likely amplify these negative effects. Practitioners should remain cognizant of industry norms when designing contractual incentives.

Our behavioural model offers further insights into how to design retainage incentives. If high trust levels have been built up over time in the industry, then engaging suppliers based partially on intent may be an effective strategy to keep projects on schedule. If trust levels are low, agreeing a formal contract with retainage provision before initiating works is preferable. Factors determining trust attitudes and behaviours should be considered here. Exogenous market uncertainty, for example in the buyer's value, can affect beliefs about trustworthiness (Özer et al. 2011). Competition among buyers may reduce trustworthy behaviours (Spiliotopoulou et al. 2016). Özer and Zheng (2019) also emphasise that the target and context of trust matter. Perceptions of trust in a specific managerial network may substitute for generalised trust in supply chain interactions (Choi et al. 2020). Trust and trustworthiness may be easier to facilitate in small pools of suppliers engaged in repeated interactions, in which case our estimates for the impact of retainage on transaction efficiency are conservative. Higher prevailing trust has been observed in collectivist societies (e.g., in China) when there is the prospect of long-term supply chain relationships, but it may be reduced in cross-border transactions if in-group bias emerges (Özer et al. 2014). Investigating cross-cultural differences in transaction norms would be an interesting avenue for further work.

An implication of the analysis is that, with retainage, the first-order efficiency problem becomes a second-order distributional problem. Future research might therefore also examine the effects of incorporating dispute resolution into the setup. A limitation of our setup is that sellers are endowed with no bargaining power after the auction and cannot dispute the buyer's retainage return decision. This assumption is clearly restrictive. In a consumer setting, Andreoni (2018) observes that some legal enforcement is required to induce buyer trust when goods are sold with a "satisfaction guarantee". Nevertheless, formal litigation is often costly and alternative dispute mechanisms underdeveloped. Consideration could also be given to the buyer's payment record in the procurement process. Reputation has been observed to improve seller trustworthiness in reverse auctions with moral hazard (Brosig-Koch and Heinrich 2014). Sellers may trust more those buyers who have previously proved themselves to be trustworthy payers.

The type of incentives examined in this study are a feature of performance bonuses in principal-agent relationships more generally. Employment contracts often contain a fixed base payment, with the promise of discretionary rewards in the future (see also Fehr et al. 2007, Lee et al. 2018). In certain industries, such as finance or professional sports, these voluntary bonus payments may be the most lucrative part of the compensation package. A similar mechanism is used in the rental housing market by landlords, who withhold deposits as a percentage of the total rental price – typically amounting to one or two month's rent – to incentivise tenants to take good care of their property.[25] When there is an excess supply of agents, the motivating effects of performance bonuses are not obvious (MacLeod and Malcomson 1998). We provide evidence from a competitive bidding environment to suggest that implicit incentives can help mitigate the moral hazard problem, under such conditions of market imbalance.

Finally, an open question is the extent to which models of fairness, originally developed to explain personal exchanges, apply to the impersonal setting of firm decision-making. We find that deriving common thresholds as to the likelihood of a buyer or seller in the market acting in good faith is a useful one. We acknowledge, however, that in repeated interactions the reference point may be path-dependent. This possibility is not captured by our static equilibrium model and is a productive avenue for further theoretical and empirical work.

[25] Arrangements to pay "half now, half later" are also popular with criminal and gangster organizations, at least as portrayed in the movies.

3 The Behavioural Consequences of Contingent Contracting in Procurement Relationships

"Doveryai, no proveryai; Trust, but verify"

- A famous Russian proverb.

3.1. Introduction

As noted in the previous chapter, a characteristic feature of procurement relationships is the supplier's ability to influence the buyer's valuation after allocation of the contract. Since it can be difficult to verify supplier non-performance, contracts are often incomplete and parties to the contract may understand their performance obligations differently, leading to liability risks and moral hazard issues. When these differences are important, disputes over supplier and buyer performance arise, particularly when suppliers are incentivised to cut costs and reduce prices.

Confronted with the administrative reality that complex contracts are incomplete (e.g., the bounded rationality argument of Simon 1947), contracts governing long-term procurement relationships often contain "hierarchical" elements, such as dispute resolution provisions, to mitigate *ex-post* opportunism.[26] The theoretical underpinnings of hierarchy come from transaction cost economics (see Williamson 1985). Buyer and supplier opportunism necessitates contractual safeguards, particularly under conditions of asset-specificity that typify the procurement of critical products, which may involve specific production equipment, knowledge, and qualifications. For example, a recent survey of procurement professionals found that 65% of procurement projects last more than three years and just 18% of buyers switch suppliers during a long-term procurement project (Gretschko and Pollrich 2021). The governance of contractual relations is "greatly complicated" by this kind of economic arrangement, in which parties are bilaterally dependent (Williamson 1988, p. 71).

[26] Other hierarchical elements are structures of authority, risk-sharing, operating procedures and change orders (Stinchcombe and Heimer 1985). These elements may not be mutually exclusive.

A high-profile case arose in 2016, when the German automaker Volkswagen announced a halt to production at several assembly plants due to a contractual dispute with two major component suppliers, ES-Automobil Guss GmbH and Car Trim GmbH (Smith 2016). The dispute stemmed from Volkswagen cancelling payments worth 500 million euros, as it pressured critical suppliers to cut costs following the diesel emissions scandal that broke earlier in the year. These non-payment events came after the suppliers had invested tens of millions of euros into building relations-specific capacity for the contracts. The result was a costly and inefficient disagreement in which both parties reneged on performance obligations.

Another common source of disputes are *ex-post* payment adjustments in construction projects which are not written into the initial contract. For example, the buyer may deduct liquidated damages in response to suspected seller negligence and/or a failure to meet specification standards. Alternatively, an unforeseen contingency may arise during trade and the parties will try to negotiate a variation order that alters the final payment amount. Price ranges are a common response to contractual indefiniteness, such as when parties write an initial agreement and leave specific terms up for renegotiation within pre-defined limits (Ben-Shahar 2004, p. 424). If agreement is not possible, parties may end up in arbitration or the courts, depending on the provisions in the contract. The value of construction disputes is often large: a global industry report estimated the *average* value in 2019 at US $30.7 million (Arcadis 2020). The shadow of potential dispute may lead contractors to "strategically manipulate their bids in anticipation of changes to the pay" (Bajari et al. 2014, p. 1292).

In this chapter, we investigate how hierarchical contractual elements may exacerbate – rather than mitigate – hazards of opportunism and influence strategic buyer and supplier behaviours in procurement relationships. To do so, we construct a stylised game-theoretic procurement model and examine, relative to this normative benchmark, the behavioural consequences of two contract types. The contracts under study either provide, or do not provide, for the resolution of trade contingencies via costly dispute.

To fix ideas, consider the construction of the Channel Tunnel between Britain and mainland Europe, which opened in May 1994 (see Genus 1997, for a detailed account).

The Channel Tunnel was a complex large-scale construction project that required various administrative provisions and contract flexibility to mitigate non-performance risks. A lump sum contract to design and build the tunnel was awarded to TransManche Link (TML) in 1986, and a Disputes Panel created to resolve disagreements between TML and the client, Eurotunnel. This project involved bilateral dependence, *ex-post* opportunism and an incomplete contract; clearly, fertile grounds for dispute. Perhaps it is unsurprising then that the lump sum contract became a major source of conflict during the relationship. Claims running to hundreds of millions of pounds were made by both parties for compensation due to non-performance of the other. As time went on, both parties became reliant on the hierarchical elements of the contract, in particular the use or arbitration to resolve disputes. This eroded trust in the TML-Eurotunnel relationship. The Channel Tunnel effectively became "a zero-sum 'game' where the parties compete antagonistically rather more than they seem to cooperate" (Genus 1997, p. 434).

In our setting, a single buyer seeks to procure one unit of an indivisible good from a pool of pre-qualified suppliers. The contract is allocated via a first-price sealed-bid reverse auction (tender).[27] Reverse auctions are less likely to be used for the procurement of complex private-sector construction projects, but they remain a requirement for most public-sector projects (Bajari et al. 2009, 2014). The contract contains a minimum payment obligation on the buyer. Additionally, the contract contains provision for a contingent payment to be made from buyer to supplier in the event of dispute over performance (see Skrzypacz 2013 for a general characterisation of auctions with contingent payments). Contingent contracts reduce risk by sharing it between the trade parties (Bazerman and Gillespie 1999). Thus, they can align incentives for responsible contractual behaviour when the contract price cannot be conditioned perfectly on all outcomes.[28] Supplier performance is defined here by the quality of works provided, e.g., the quality

[27] First-price auction formats are typically used in procurement (Engelbrecht-Wiggans et al. 2007).

[28] The importance of responsible contractual behaviour has gained renewed attention in the wake of the Covid-19 pandemic. Indeed, the UK government released specific guidance on this issue: "Responsible and fair behaviour in contracts now – in particular in dealing with potential disputes – will result in better long-term outcomes for jobs and our economy" (Cabinet Office 2020).

of materials used or the ethicality of production practices. Buyer performance is defined by payment of monies requested in the initial contract.

We model the contingent payment as an arbitrator, although it finds more general interpretation as the enforceability of deferred value. As alluded to above, arbitration clauses are common in construction and engineering projects as a means of dispute resolution.[29] Standard building contracts in Britain are produced by the Joint Contracts Tribunal, which provides model rules for the resolution of payment disputes by arbitration. Until 2007, arbitration was the preferred dispute resolution method in standard United States construction contracts by default, while construction case filings at the American Arbitration Association continue to increase, up six percent in 2019, with complex new cases rising at twice this rate. [30]

The role of an arbitrator is "to remember who sunk costs in the past, and to ensure that future compensation is paid in a way that creates the *correct* incentives to make such investments" (Crawford 1985, p. 375).[31] Traditionally, the economics literature examined the distributional implications of arbitration, given a fixed surplus to divide (Ashenfelter and Bloom 1984, Crawford 1981, Farber 1980, Farber and Katz 1979). Gabuthy and Muthoo (2019) drop the fixed surplus assumption and show that the mere presence of an arbitrator can improve incentives of bilateral trade parties to make mutual relationship-specific investments.

We consider the efficiency consequences of arbitration in a quite different environment. After conclusion of the tender process, the winning supplier (seller) receives the guaranteed payment component of the contract and takes costly action to determine the project quality, high or low. Thus, while the pre-qualified suppliers are homogeneous *ex-*

---

[29] Unlike litigation, there is generally no public record of arbitration proceedings.

[30] Since 2007, US construction contracts require trade parties to check a box specifying the preferred dispute resolution method (e.g., Document A101-2017 at https://www.ai-acontracts.org/). The US case filings statistic is obtained from American Arbitration Association Annual Report at https://www.adr.org/annual-reports. UK building contracts are obtained from: https://www.jctltd.co.uk/.

[31] Arbitration finds varied applications for resolving disputes. The most well-known is for the settlement of wage disputes in labour bargaining. Mandatory consumer arbitration clauses are also prevalent. For example, as of 2020, the eBay user agreement contains a requirement to submit dispute claims, e.g., due to an unpaid auction bid, to binding and final arbitration.

ante, the homogeneity assumption is not necessarily satisfied *ex-post*. After observing her value from trade, the buyer has discretion to propose a final price constrained by the bounds of the contract. This combination of buyer- and seller-side moral hazard and implicit non-homogeneity of suppliers after moving from competitive market to bilateral monopoly lays the foundations for an inefficient outcome if the contract is not designed carefully and may lead parties to end up in dispute (see Williamson 1973, p. 318). In such circumstances, an arbitrator is invoked to provide binding resolution. The arbitrator must verify the claims and adjudicate on how to compensate the trade parties according to some pre-defined reference point. Knowledge of the reference point may come from experience or be explicitly written into the contract. No restriction is placed on what the arbitrator deems a fair settlement; but the arbitrator is not always able to intervene. Specifically, there is an exogenous probability with which the arbitrator can verify quality. This imperfection reflects the reality that quality verification is not feasible in every trading scenario.

The contingent contract is defined by the combination of price flexibility and the probability of quality verification. For high quality to emerge as an equilibrium outcome, there is a "goldilocks region" in which the probability that quality can be verified is large enough to reward high quality, but not so large as to allow suppliers to reduce their bids too far during the tender process. The outcome is first-best in the absence of verification costs and benefits the buyer, but not seller, owing to the strong nature of *ex-ante* competition.

In the second part of the chapter, we test the predictions of the model in a lab experiment. The lab has found long-standing use as a test-bed for more general arbitration mechanisms because it enables control over factors such as the arbitrator's preferences, costs, and valuations.[32] No prior experiment has tested the relative effectiveness of contingent versus non-contingent contracts in reverse auctions with moral hazard.[33] Behavioural considerations may distort the effectiveness of supply chain contracts (e.g.,

[32] For a survey of this literature, see Kuhn (2009).

[33] Chen et al. (2009) conduct a theoretical analysis of contingent contracting in reverse auctions with adverse selection.

Ho and Zhang 2008, Loch and Wu 2008); it is therefore interesting to consider the effects of contingent contracting on actual decision-making. Hart and Moore (2008) argue that competitively determined contractual terms may serve as a reference point for later actions in relationships that influence the gains from trade.[34] As a consequence, trade may be inefficient not as a result of allocation to the inefficient "type" *ex-ante*, but as a result of reneging on contractual performance *ex-post*. This contrasts with the traditional property rights approach (Hart and Moore 1990) in which trade parties can always renegotiate to an efficient outcome.

Our experimental results suggest that with a contingent contract, bidding is more aggressive and high quality is the more profitable seller strategy. However, a contingent contract does not increase trade efficiency relative to the equivalent non-contingent contract. Instead, we observe that the contingent contract crowds out the reciprocity of buyers, a significant fraction of whom exhibit reference-dependent fairness preferences. This crowding out effect erodes suppliers' relation-specific trust, as inferred from their willingness to deliver high quality. The study thus relates to a literature in experimental economics recognising the potential for explicit incentives to be counterproductive.

3.2. Model and Theory

We first recap features of the environment that are shared with Chapter 2. A single buyer seeks to procure one unit of an indivisible good from a pool of $n \geq 2$ pre-qualified suppliers, indexed by i.

At *date 0*, the pre-qualified suppliers compete on price (only) to deliver the unit. That is, pre-qualified suppliers are *ex-ante* symmetric in their delivery capability. Each supplier submits a sealed bid b_i at a first-price reverse auction (tender), with a commitment to procure at the lowest price. The selected supplier is the one submitting the lowest priced bid b (henceforth, the seller). If two or more suppliers submit the same bid, the tie is broken randomly. A bilateral trade relationship forms between buyer and

[34] Related are experiments investigating whether renegotiable contracts can solve the hold-up problem (Bartling and Schmidt 2015, Brandts et al. 2015, Davis and Leider 2018, Fehr et al. 2011, Fehr et al. 2015, Hoppe and Schmitz 2011).

seller. Any non-selected supplier is no longer considered in the interaction and earns zero payoff.

At *date 1*, the seller determines the quality of the unit. We again discretise the quality space such that the seller either delivers a low-quality ($q = q^L$) or high-quality ($q = q^H$) unit. The seller's quality choice maps directly to his own unit cost c^j, and the buyer's unit valuation v^j. The cost and valuation schedules are common knowledge. We impose the same assumptions on these schedules as before.

At *date 2*, the buyer observes her valuation for the unit and proposes total payment y, which must at least fulfil her minimum obligation in the contract (to be defined).[35]

The final payment p from buyer to seller is then determined and payoffs from the interaction are realised. The payoffs of the buyer and seller are summarised as follows,

$$\pi_B = v^j - p. \tag{3.1}$$

$$\pi_S = p - c^j. \tag{3.2}$$

The focus of this chapter is on how the final payment p is determined, how these payment incentives influence equilibrium buyer and seller behaviours and the associated efficiency consequences. Consider the following contract:

$$P = (1 - \alpha)b + \beta Z, \tag{3.3}$$

for some constants α, $\beta \in [0,1]$ and some (possibly non-linear) contingent payment function Z. The standard auction theory setup with a fixed-price contract equal to the winning bid is captured by $\alpha = \beta = 0$. More generally, (3.3) represents a cash bid plus fixed contract. For example, McAfee and McMillan (1986) replace Z with the seller's exogenously assigned cost type, in which case $\alpha = \beta = 1$ is a cost-plus contract, and $\beta > 0$ is a linear incentive contract for which the seller has a share in cost overruns.[36]

[35] The results in this section would be unchanged by switching the order of date 1 and 2 actions.
[36] McAfee and McMillan (1986) also include a fixed fee. A fixed fee has no influence on the equilibrium analysis because an increase in this fee would be offset by a decrease in the bid amount.

We will interpret the first term in (3.3) as the buyer's minimum payment obligation at date 2. Thus, $y \geq (1-\alpha)b$. We limit attention to $0 < \alpha < 1$ and so this parameter is the maximum bid discount; without loss, we set $\beta = 1$. The function Z considered here captures a particular hierarchical element of contracts for large-scale procurement projects: dispute resolution. We assume that quality can only be imperfectly verified by third parties, such as an arbitrator. This is an incomplete contracting assumption (Hart 2017). It is motivated by the observation that in non-commoditised procurement, the seller's costly effort underlying the quality decision may be unobserved and so it is difficult to write a complete contract that ties payment to outcomes.

We incorporate this uncertainty via an exogenous probability σ with which delivered quality can be verified. A specific interpretation of σ is the probability with which an arbitrator can *ex-post* adjudicate on a dispute (see e.g., Herweg and Schmidt 2020). Since we use this interpretation in the experiment, we will adopt it here. A more general interpretation of σ is as the enforceability of deferred value.

The dispute resolution function is defined as follows:

$$Z := \begin{cases} z(b, q^j), & w.p.\, \sigma; \\ 0, & w.p.\, (1-\sigma). \end{cases} \tag{3.4}$$

Which specifies a deferred payment from buyer to seller, where $z(b, q^j) \in [0, \alpha b]$ is a function that maps the seller's bidding strategy to a deferred payment in the non-negative quadrant, bounded by αb. This preference may be inferred from the contract or be an unwritten but agreed upon trading norm. Either way, it represents what an (unbiased) arbitrator would deem a fair distribution of profits after completion of the fact-finding process.

Award of the preferred deferred payment is contingent on the verifiability of quality and – similar to a debt or insurance security – caps total payment at the winning bid. Note that setting a higher cap to incorporate cost overruns would not change the model's insights; by setting the cap equal to the seller's bid, we interpret this action as the seller's claim on the surplus.

We make three assumptions on the dispute resolution function in (3.4) for any combination of winning bid b and quality level j: (i) $z(b, q^j) \geq c^j - (1-\alpha)b$; (ii) $z(b, q^H) \geq$

$z(b, q^L)$; (iii) for any $b' < b$, $z(b', q^j) \geq z(b, q^j)$ and $b - b' \geq z(b', q^j) - z(b, q^j)$. The first assumption ensures that (where feasible) a fair payment should cover the seller's costs known *ex-ante*. The second assumption simply implies that the deferred payment rewards high over low quality. The third assumption requires that the deferred payment offsets (in part or in full) a reduction in the buyer's minimum payment obligation. Thus, all else equal, a contract containing a lower minimum obligation is associated with a (weakly) higher deferred payment.

It remains to define what constitutes a dispute in our model.

Definition 3.1 *(Dispute). A dispute is recorded if and only if $y < (1 - \alpha)b + Z$.*

That is, a dispute is recorded if the buyer's claim on the surplus at date 2 is strictly below the fair settlement and – by our assumptions on Z – below the seller's claim on the surplus. If a dispute is recorded and quality can be verified, then the buyer pays a fixed verification cost $k > 0$.[37] This cost accrues outside the model, such as to a third-party arbitrator. To avoid uninteresting cases, we assume that k does not exceed the gains from trade. If no dispute is recorded or quality cannot be verified, then the seller receives the buyer's proposed payment (i.e., $p = y$).

The game can be analysed as an extensive-form game of complete information. Formally, a pre-qualified supplier i's strategy has two components: a bid $b_i \geq c^L$, and a quality choice function $q_i(b_i)$. Bids below marginal cost are never profitable for the seller. We place no restriction on the maximum permissible bid except that it be high enough to guarantee a seller minimum payment to cover his cost schedule (i.e., at least $c(q^H)/(1 - \alpha)$). A buyer's strategy consists of a payment function $y(b, q^j) \geq (1 - \alpha)b$, i.e., a payment that satisfies the minimum contractual obligation.

[37] In practice, arbitration costs may be borne by the unsuccessful party or shared between the dispute parties. To simplify the exposition, we employ the former rule here, but the main insights of the model are unchanged by using a shared cost allocation. Common knowledge about the arbitration cost has real-world application. The German Arbitration Institute, for example, makes its cost schedule publicly available and provides a cost calculator tool on its website (http://www.disarb.org/en/).

The solution concept is SPNE in pure strategies and we solve the game by using backward induction. As in standard Bertrand competition models with bounded demand, complete information, fixed marginal costs and zero fixed costs, there is no mixed strategy equilibrium (see e.g., Baye and Morgan 1999). We are interested in whether there exists a procurement outcome in which high quality is implementable. A necessary condition for implementability is incentive compatibility, and so the definitions proceed in order.

Definition 3.2 *(Incentive compatibility). A quality level is incentive compatible if it is strictly preferred by the seller to the other quality level.*

Definition 3.3 *(Implementability). A quality level is implementable if it is an incentive compatible component of an equilibrium seller strategy.*

3.2.1. Non-contingent contract

If quality cannot be verified (i.e., $\sigma = 0$), then we have a non-contingent contract.

Proposition 3.1 *. There is a unique SPNE in which* $b_i^* = \frac{c^L}{(1-\alpha)}\ \forall\ i,\ q_i^* = q^L$ and $y^* = (1-\alpha)b$.

Proof. See Proposition 2.1A. ■

Thus, with a non-contingent contract, high quality is not implementable. As we showed in the previous chapter, other-regarding preferences can sustain high quality in equilibrium with a non-contingent contract when there is flexibility in the price. This is the appropriate benchmark for assessing the relative effectiveness of a contingent contract in the experiment.

3.2.2. Contingent contract

If quality can be verified with some positive probability (i.e., $\sigma > 0$), then we have a contingent contract. There is now the possibility of dispute because $Z \geq 0$. The dispute payoffs of buyer and seller are as follows,

$$d_B = v^j - (1-\sigma)y - \sigma[(1-\alpha)b + z(b, q^j) + k]. \quad (3.5)$$

$$d_S = (1-\sigma)y + \sigma[(1-\alpha)b + z(b, q^j)] - c^j. \quad (3.6)$$

The results derived below are qualitatively unchanged by specifying a utility function with risk aversion so long as the dispute utilities remain monotonic over the price grid. In the appendix, we show that the model predictions still hold under risk aversion for the dispute resolution function used in our experiment. For the remainder of this section, we assume that the buyer and suppliers act as risk neutral profit-maximisers.

At date 2, the buyer observes her valuation for the good and proposes to compensate the seller as follows for quality delivered,

$$y = \begin{cases} (1-\alpha)b, & if \ \frac{\sigma k}{(1-\sigma)} < z(b, q^j); \\ (1-\alpha)b + z(b, q^j), & if \ \frac{\sigma k}{(1-\sigma)} > z(b, q^j). \end{cases} \quad (3.7)$$

If the buyer's expected cost of verification is less than her expected benefit from dispute, then the buyer fulfils her minimum payment obligation in the contract (only) and the two parties end up in dispute. If the expected cost of verification is greater than her expected benefit from dispute, then the buyer makes final payment in line with the arbitrator's preference and no dispute is recorded. Since the buyer's dispute payoff is non-increasing in y, it is not in her interest to propose an interior payment.

At date 1, anticipating the buyer's payment function in (3.7) and given his bid, the seller chooses the quality level. The seller's incentive compatibility constraint for high quality is the union of the three cases below. The left-hand inequality defines the buyer's expected cost of verification, and the right-hand inequality is the seller's required marginal return to high quality. To economise on notation, we define $c^{HL} = c^H - c^L$ and $z^{HL} = z(b, q^H) - z(b, q^L)$.

$$z(b, q^H) \leq \tfrac{\sigma k}{(1-\sigma)}: \quad z^{HL} > c^{HL}. \quad (3.8a)$$

$$z(b, q^L) < \tfrac{\sigma k}{(1-\sigma)} < z(b, q^H): \quad \sigma z^{HL} - (1-\sigma) z(b, q^L) > c^{HL}. \quad (3.8b)$$

$$z(b, q^L) \geq \tfrac{\sigma k}{(1-\sigma)}: \quad \sigma z^{HL} > c^{HL}. \tag{3.8c}$$

At high expected verification cost (3.8a), the seller knows his return to high quality with certainty and incentive compatibility depends on the deferred payment premium for high quality. As σ approaches one, the buyer prefers to make a fair payment and avoid dispute.

Corollary 3.1 . *At high expected verification cost, high quality is not implementable with a contingent contract.*

Proof. If cost case (3.8a) applies, then from (3.7) the payment is $p = (1 - \alpha)b + z(b, q^j)$ and no dispute is recorded. By our assumption on the cost schedule, $\pi_S(b_i = c^L, q_i = q^H) < \pi_S(b_i = c^L, q_i = q^L) = 0$ and so low quality is incentive compatible for any supplier i at a bid equal to c^L. For any higher bid, each supplier has an incentive to undercut his competitor and deliver low quality due to the binding auction selection rule and the zero-payoff for a non-selected supplier. Any bid below c^L is not profitable. Thus, high quality cannot be part of an equilibrium seller strategy. ■

An interesting implication of Corollary 3.1 is that, as σ approaches one, the buyer's claim on the surplus is irrelevant and high quality is not incentive compatible because potential suppliers compete away their rents at date 0. We obtain a second corollary.

Corollary 3.2 . *Perfect verification (i.e., $\sigma = 1$) cannot implement high quality.*

From now on, suppose that the expected verification cost $k < z(b, q^H)$ and so the buyer reciprocates high quality with the minimum required payment. In these cases (3.8b, 3.8c), high quality is incentive compatible only if the probability of quality verification is high enough. Clearly, there is a trade-off on σ to simultaneously satisfying the buyer's cost condition and the seller's incentive compatibility constraint. Thus, for high quality to be implementable, there is a "goldilocks region" in which σ is large enough to reward high quality but not so large as to allow suppliers to reduce their bids too far in the date 0 auction. A third corollary follows directly from Definition 3.2.

Corollary 3.3 . *High quality is implementable if and only if there exists at least one seller strategy for which* $d_S(b, q^L) < 0 \le d_S(b, q^H)$.

Note that the inequality in Corollary 3.3 is never satisfied with a non-contingent contract because, if $\sigma = 0$, then $d_S(b, q^L) = \pi_S(b, q^L) > \pi_S(b, q^H) = d_S(b, q^H)$ for all b. Given our assumptions on the dispute resolution function in (3.4), the seller's dispute payoff in (3.6) is non-decreasing as the winning bid increases. Thus, there always exists at least one bid which in combination with low quality yields the seller an expected loss. For high quality to be implementable, at least one of these bids must also yield the seller a non-negative profit when combined with high quality. We will see now that when such a bid exists, high quality is the unique equilibrium outcome.

Proposition 3.2 . *There exists a contingent contract for which the unique SPNE outcome is* $b_i^* = \frac{c^H - \sigma z(b, q^H)}{(1-\alpha)}$ $\forall\ i$, $q_i^* = q^H$ and $y^* = (1 - \alpha)b$.

Proof. By our earlier assumption on the expected verification cost, the buyer chooses the minimum payment at date 2 in exchange for high quality. A supplier's breakeven bid associated with high quality, obtained by setting the dispute payoff in (3.6) equal to zero, is $\underline{b}^H = (c^H - \sigma z(b, q^H))/(1 - \alpha)$. Suppose that there exists at least one seller strategy for which $d_S(b, q^L) < 0 \le d_S(b, q^H)$. Owing to monotonicity of the dispute payoff function in the winning bid, $\underline{b}^H$ must be the smallest bid associated with this strategy. At date 0, if two or more pre-qualified suppliers submit a bid equal to $\underline{b}^H$ then every supplier earns zero profit in expectation. Any upward deviation would yield zero profit, due to the binding auction selection rule. Undercutting this bid would yield the seller a loss in expectation independent of the chosen quality level. Thus, this is the unique equilibrium outcome. ■

From inspection of the equilibrium bids in Propositions 3.1 and 3.2 and the incentive compatibility constraint in (3.8), we infer a final corollary.[38]

[38] Corollary 3.4 requires $c^{HL} < \sigma z(b, q^H)$, which is true if high quality is incentive compatible.

Corollary 3.4 . *Any contingent contract that implements high quality is associated with lower equilibrium bids than the equivalent contract with* $\sigma = 0$.

We will now put more structure on the dispute resolution function in (3.4) and specify a parametric form for the arbitrator's preferences, which will then be used in the experiment.

3.2.3. A specific functional form

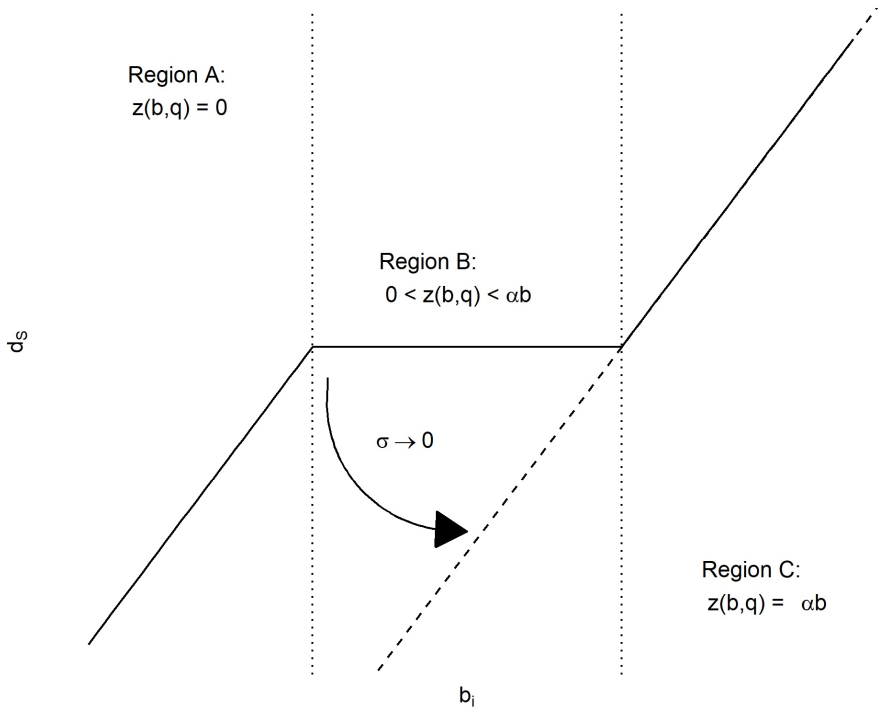
Consider the following mapping from actions to the arbitrator's preferred payment gross of the minimum contractual obligation:

$$z(b, q^j) + (1 - \alpha)b = \mu v^j + (1 - \mu)c^j \tag{3.9}$$

Where $0 < \mu < 1$ is the arbitrator's preference parameter and $z(b, q^j) \in [0, \alpha b]$. This specifies a deferred payment from buyer to seller as a convex combination of the trade surplus, determined by the seller's quality decision at date 1. The payment is again constrained to lie in the interval between the buyer's minimum payment obligation and the seller's bid. It is straightforward to verify that (3.9) satisfies our earlier three assumptions on the dispute resolution function.

In Figure 3.1, we sketch the seller's dispute payoff associated with (3.9) as a function of the bidding strategy. Without loss, we set $k = 0$. The solid line in the figure corresponds to the payoff function with perfect quality verification (i.e., $\sigma = 1$), and the dashed line with no quality verification (i.e., $\sigma = 0$). The payoff function is quality-specific and has two pivot points at the vertical dotted lines in the figure. Incentive compatibility corresponds to regions in which the seller dispute payoff to high quality is above the payoff to low quality. At bids below the lower pivot point, the arbitrator's preference cannot be implemented and so the arbitrator awards the seller his bid in full (region A). As bids rise above the upper pivot point, the seller received more than his fair share of the surplus and so the arbitrator awards the minimum payment obligation (region C). At bids in-between, the arbitrator awards the preferred amount exactly (region B).

Figure 3.1. Seller i's dispute payoff as a function of bid b_i and quality level $q = q^j$.

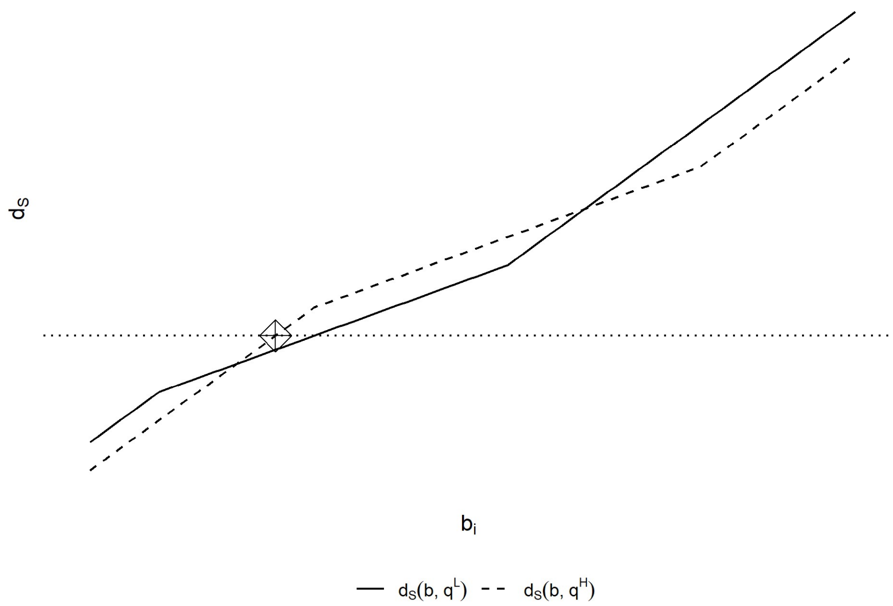


Notes: The solid line is seller i's dispute payoff d_S as a function of bid b_i and quality level $q = q^j$ with perfect quality verification (i.e., $\sigma = 1$). The dashed line is the corresponding payoff function with no quality verification (i.e., $\sigma = 0$).

As the arbitrator's preference μ increases, the pivot points are found over higher bids, the arbitrator is more favourable to the seller, and the payoff line in region B shifts up. As the maximum bid discount α increases, the payoff function in Figure 3.1 shifts down because of the greater potential for seller losses when quality cannot be verified. Similarly, region B is found over a wider subset of bids owing to the greater contractual flexibility. As the probability of quality verification σ falls below one, the solid line payoff function in region A shifts down and the payoff function in region B pivots towards the dashed line. Thus, σ determines the slope of the payoff function in region B. When $\sigma > 0$, the payoff functions to high and low quality either never intersect, or they intersect at two points in the price grid. At the limit, $\sigma = 0$ and the two segments form a straight line as in standard Bertrand competition with homogenous goods; region B collapses to a single price and high quality is not incentive compatible.

An example contingent contract for $0 < \sigma < 1$ that implements high quality is presented in Figure 3.2. The equilibrium bid, marked on the figure by a diamond plus, is at the point where the dispute payoff function to high quality crosses the zero-profit line. This bid is below (above) the corresponding equilibrium bid when $\sigma = 0$ $(= 1)$. The first intersection between the payoffs to low and high quality occurs below the zero-profit line. For zero verification cost, the first-best outcome is achieved. For any non-zero verification cost, a Pareto improving second-best outcome can be implemented. We will now test the effectiveness of the contingent contract in the lab.

Figure 3.2. An example contract that implements high quality.



Notes: The solid (dashed) line is seller i's dispute payoff to low (high) quality. The dotted line is the zero-profit line. The equilibrium bid is marked by a diamond plus.

3.3. Experimental Setup

3.3.1. Treatments and parameters

In the experimental setting, one buyer faces two potential sellers. The maximum bid discount α is set at 0.75, which permits substantial flexibility in the contract and makes the payment trade-offs salient. There are three experimental conditions (Table 3.1), which vary according to the probability of quality verification – using a between-subjects design – and the arbitrator's preference from (3.9) – using a within-subjects design. The first condition simulates the non-contingent contract, and so the probability of quality verification $\sigma = 0$ (the "*Voluntary*" treatment). In this treatment, any payment to the seller above one-quarter of the winning bid is made by the buyer voluntarily. In the second and third conditions, we introduce a contingent contract, where the probability of quality verification $\sigma = 0.5$ (collectively, the "*Arbitrator*" treatment). This treatment is split into two sub-treatments, in which the arbitrator either prefers to award $\mu = 1/3$ or 2/3 of the trade surplus to the seller. By varying the arbitrator's preference parameter within-subjects, we conduct a less restrictive test of the model and consider relationships in which the contingency favours the buyer and seller. We deliberately avoid invoking norms associated with an equal surplus split, which have been observed in prior bargaining studies (Andreoni and Bernheim 2009).[39] The fixed verification cost $k = 2$ throughout (more details below).

Table 3.1 – Treatment matrix.

Treatment	*Voluntary*	*Arbitrator*	
Sub-treatment		$\mu = 1/3$	$\mu = 2/3$
Probability of quality verification	$\sigma = 0$	$\sigma = 0.5$	
Maximum bid discount	$\alpha = 0.75$		

The cost and valuation schedules are common knowledge in the experiment (Table 3.2). Low quality generates a trading surplus of 20 and high quality generates a surplus of 60. Low quality costs the seller 30 to deliver and is valued by the buyer at 50. High quality costs the seller 40 to deliver and is valued by the buyer at 100. Consistent with the investment game literature in experimental economics, the surplus multiplier is three. The bid increment is set at one, the minimum permissible bid at 30 and the maximum permissible bid at 200. The maximum bid is high enough to ensure that potential sellers can always submit a profitable bid associated with either low or high

[39] Aside from the normative appeal of the equal split, there is no immediate reason why subjects should behave differently for values of μ other than 1/3 or 2/3 (where Proposition 3.2 applies).

quality. It is also low enough to ensure that there is no bid at which a buyer could be forced into a purchase that might yield a loss.

Table 3.2 – Cost and valuation schedules.

Quality level q^j	$q = q^L$	$q = q^H$
Buyer's valuation: v^j	50	100
Seller's cost: c^j	30	40

3.3.2. Session protocol

All subjects participated in a sequence of 30 trade interactions. Each interaction was divided into distinct phases. There were two phases of every interaction in the *Voluntary* treatment and three phases of every interaction in the *Arbitrator* treatment.

The first two phases were the same in both treatments. In phase 1, the potential sellers in a group each submitted a sealed bid at a first-price reverse auction. At the same time as choosing a bid, each potential seller also chose a quality level, high or low, to be delivered conditional on winning the auction. This variant on the strategy method enabled twice as many quality observations to be collected, without changing the strategic nature of the game.[40] The seller who submitted the lower bid in the first phase won the auction, with ties broken randomly. The winning and losing bids and the minimum payment obligation were then revealed within the group. This information was presented to subjects as a contract price range, with the lower bound price equal to one-quarter of the winning bid and the upper bound price equal to the winning bid. The seller that submitted the losing bid in phase 1 earned zero profit for the interaction.

In phase 2, the buyer observed the winning seller's chosen quality level and proposed a payment from within the contract price range. In *Voluntary*, the buyer's proposal was the final payment and the interaction ended there. In *Arbitrator*, there was a third (non-

[40] Further benefits were that it helped to preserve anonymity of the winning seller and minimise wait times. A complete strategy method in *Arbitrator* would have required sellers to specify a quality choice for every possible bid, which is not practical.

decision) phase in which appeal to an arbitrator could be triggered if the buyer's proposal was less than a reference arbitrator price. The arbitrator's price was the amount from between the buyer's proposal and the seller's bid that minimised the distance between the final payment and the mapping from arbitrator's preferred surplus division to the gross deferred payment in (3.9). This mapping was common knowledge in the experiment (see implementation details below). If the buyer's proposed price was below the arbitrator's price, then an appeal to the arbitrator was triggered. The arbitrator was available to set a final price on one-half of appeals, determined at random at the onset of phase 3, in which case the buyer paid the verification cost.[41] If no appeal was triggered, or the arbitrator unavailable, then the seller received the buyer's proposed price.

To illustrate, suppose the winning bid is 128 and the seller selects high quality. The contract price range is [32, 128]. The buyer proposes a price of 65. In *Voluntary*, this is the final payment; the buyer earns 35 and the seller earns 25. In *Arbitrator*, the arbitrator's price would be 60 when $\mu = 1/3$ or 80 when $\mu = 2/3$. In the first situation, no appeal is triggered, and profits are the same as in *Voluntary*. In the second situation, an appeal to the arbitrator is triggered. If the arbitrator is available, then a two-point verification fee is levied on the buyer and the arbitrator's price is used for payment; the buyer earns 18 and the seller earns 40. If the arbitrator is unavailable, then profits are the same as in *Voluntary*.

3.3.3. Hypotheses

Proposition 3.1 states that trade will be inefficient (i.e., low quality) in *Voluntary*. We chose the contract parameters in *Arbitrator* such that high quality is the unique implementable outcome for all levels of risk aversion and so Proposition 3.2 applies. In the appendix, we demonstrate the implementable quality levels for the full set of contingent contracts.[42] The cost of quality verification is low enough to ensure that the buyer should always propose her minimum payment obligation in the experiment, which

---

[41] This cost was not factored into calculation of the arbitrator's price.

[42] The subset of contracts for which high quality is implementable is decreasing in the seller's degree of risk aversion.

gives us scope to examine the behavioural effects of the hierarchical contract element. In Table 3.3, we present the point predictions for bid, quality, and profits in each treatment. To within an increment, the equilibrium bid in *Voluntary* is 120, in *Arbitrator* with $\mu = 1/3$ is 80, and with $\mu = 2/3$ is 64. Due to the strong nature of supplier competition, buyers are expected to appropriate the full surplus.

Table 3.3 – Equilibrium predictions for bids, quality, and profits in the experiment.

Treatment	*Voluntary*	*Arbitrator*	
Sub-treatment		$\mu = 1/3$	$\mu = 2/3$
Measure			
b_i	120	80	64
q	Low quality	High quality	High quality
π_B	20	60	60
π_S	0	0	0

Notes: A potential seller's bid is b_i; the quality level chosen by the auction winner is q; the buyer's price proposal is y. The buyer's expected profit is π_B and the seller's expected profit is π_S . Numbers are to within one increment due to discreteness of the experimental price grid.

The comparative hypotheses to be tested are summarised as follows:

Hypothesis 3.1. Bidding. *(i) Winning bids are lower in Arbitrator than in Voluntary; and (ii) winning bids are higher when* $\mu = 1/3$ *than when* $\mu = 2/3$.

Hypothesis 3.2. Efficiency. *The probability of high quality is greater in Arbitrator than in Voluntary.*

Hypothesis 3.3. Reciprocity. *There are no differences in the buyer's payment behaviour between Arbitrator and Voluntary.*

Hypothesis 3.4. Profits. *The buyer's profit is higher in Arbitrator than in Voluntary; sellers' profits are unchanged between Arbitrator and Voluntary.*

3.3.4. Implementation

The total number of subjects recruited for the experiment was 108. Subjects were recruited using the web-based software hroot (Bock et al. 2014) and allocated at random to one of the two treatments, *Voluntary* or *Arbitrator*. Each treatment included six independent cohorts and no subject participated in more than one treatment. Every cohort had nine human subjects, three buyers and six sellers, who were matched into groups of one buyer and two sellers using a stranger matching protocol between interactions. An algorithm ensured no subject played with the same two players in consecutive interactions, and subjects were informed of this. In the procurement of critical products or long-term projects, firms typically compete for contracts within a fixed pool of potential sellers. While buyers and potential sellers vary between interactions, they often meet again at some future date. Thus, our matching protocol is arguably more appropriate than a perfect stranger matching protocol for the target setting. There were two cohorts (18 subjects) of the same treatment in every session.[43] To minimise the possibility of tacit collusion, subjects were not informed about the cohort size.

To facilitate understanding, the task was presented using a cover story. A shipping company sought to procure an engine for its cargo ship. The choice of a shipping company was deliberate. Subjects are less likely to have experience from daily life with a shipping company than, perhaps, an airline (see Alekseev et al. 2017 on the merits and pitfalls of using a meaningful instruction frame). One-third of subjects assumed the role of a shipping company (buyer); and two-thirds the role of a supplier.

In *Arbitrator*, the sequence of trade interactions was broken into two blocks of 15 interactions, which constituted the two sub-treatments. To control for order effects, the block sequence in this treatment followed a crossover design, with one-half of the subjects in a session assigned to the sequence $\mu = \{1/3, 2/3\}$ and the other half assigned to $\mu = \{2/3, 1/3\}$. The probabilistic nature of the arbitrator was explained using a die

[43] Data from one cohort of the *Voluntary* treatment was discarded after discovery of a comprehension issue. We subsequently conducted an additional session of this treatment, with nine (new) subjects, to fill the missing cohort. Summary statistics for the discarded cohort are available on request.

roll. Subjects were informed that if an appeal to the arbitrator was triggered, the computer would roll a fair six-sided die. If the die came up one to three, the arbitrator would be unavailable, and the seller would receive the buyer's proposed price as payment. If the die came up four to six, the arbitrator would be available, and the seller would receive the arbitrator's price as payment.[44] The mapping from arbitrator's preference to price described in (3.9) was explained to subjects in terms of a two-to-one buyer (seller) to seller (buyer) profit ratio. Sellers were informed about the arbitrator's price associated with each combination of bid and quality level. Buyers were informed about the arbitrator's price associated with each proposal and whether an appeal would be triggered.

At the end of each interaction, feedback was provided about the outcomes of a subject's own interaction group (only) on bids, quality, payment, and profits. This feedback was provided to all groups simultaneously, to prevent subjects inferring the identities of others in their interaction group. Private feedback remained available in a history table to facilitate learning. The losing seller only observed the auction outcome.

The experiments were computerised and programmed in oTree (Chen et al. 2016). To aid replicability and ensure common knowledge, video instructions were created and played at the start of an experimental session.[45] Subjects had to answer a set of comprehension questions correctly before proceeding to two trial rounds, in which they were guided through the decision screens specific to their role. Roles were assigned randomly before the trial rounds and remained fixed. Each seller had their own auction in the trial rounds and these rounds were non-incentivised. The main experiment followed immediately. At the end of a session, subjects completed a survey to elicit demographic information, risk and trust attitudes.[46]

Subjects received money for their participation, which was paid in private and in cash at the end of a session. As is often the case in auction experiments, there was the

[44] The randomization was successful: the arbitrator was available on 49.3 percent of appeals in the experiment, which is not significantly different from 50 percent (p-value = 0.690, two-tailed binomial test).

[45] A compendium of instructions and links to the video recordings are included in the appendix.

[46] Since the predictions do not depend on risk preferences, we do not use an incentivized elicitation. Summary statistics from the post-experiment survey are presented in the appendix.

possibility of losses. Subjects assigned a buyer (seller) role therefore began each interaction with an endowment of five (ten) points.[47] Any profit or loss was added to or subtracted from the endowment. Sessions lasted sixty to seventy-five minutes. The total points from all rounds were multiplied by a pre-determined exchange rate of forty points per one British pound. Subjects were informed that the minimum amount they could leave the session with was £4.[48] Average earnings for buyers were £28 and for sellers £10. Payoff inequality does not influence the theoretical predictions, because we assume standard preferences. Empirically, however, a concern for fairness may influence behaviour (e.g., Fehr and Schmidt 1999, Bolton and Ockenfels 2000). We return to this point when analysing the experimental results.

3.4. Experimental Results

This section presents findings of the lab experiment. First, we conduct an aggregate analysis. The cohort is the independent level of observation and is used for all statistical comparison tests in Section 3.4.1. Second, we analyse the individual bidding strategies pursued by sellers, in relation to the theory. Third, we estimate a mixture model and find that a significant fraction of buyers exhibit reference-dependent preferences.

3.4.1. Aggregate findings

The main results of the experiment are summarised in Table 3.4. Cohort averages are provided for winning bids, the frequency of high quality, the buyer's proposal for the seller's share of the trade surplus (as inferred from the buyer's proposed price), and profits. An efficiency measure is provided, which captures the realised percentage of trade surplus net of verification cost incurred by the buyer. The point predictions from Table 3.3 are included in square brackets.

---

[47] The relative endowment sizes compensate for the expectation that, on average, each seller participates in half as many transactions as the buyer.

[48] One subject encountered a limited liability problem. The results are qualitatively unchanged by inclusion or exclusion of the cohort in which this subject participated. For completeness, we include this cohort.

Table 3.4 – Aggregate summary statistics from the experiment.

Treatment	*Voluntary*		*Arbitrator*			
Sub-treatment			$\mu = 1/3$		$\mu = 2/3$	
Measure						
Winning bid	105.64	[120]	65.29***	[80]	72.96***	[64]
Quality	0.45	[0]	0.43	[1]	0.53	[1]
Proposed seller share	0.13		-0.18***		-0.05**	
Low quality	0.13		-0.23***		-0.13***	
High quality	0.13		-0.11**		-0.05	
Buyer profit	33.85	[20]	33.8	[60]	32.32	[60]
Low quality	17.32		18.35		13.93	
High quality	52.41		52.83		48.21	
Seller profit	4.79	[0]	3.35	[0]	8.01	[0]
Low quality	2.68		0.78		5.24	
High quality	7.59		6.51		11.11	
Efficiency	63%	[33%]	61%	[96.7%]	67%	[96.7%]

Notes: The table summarises the main outcomes of the lab experiment. Values are median cohort averages, based on six independent cohorts per treatment. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All statistical tests are based on two-tailed non-parametric Wilcoxon-Mann-Whitney tests versus the *Voluntary* treatment. The point prediction of the theory is provided in square brackets. Variable definitions: *Winning bid* is the lowest bid submitted by potential sellers in the auction phase; *Quality* is the relative frequency that high quality is chosen for delivery by the winning seller; *Proposed seller share* is the buyer's proposal for the seller's share of the trade surplus (as inferred from the proposed price); *Buyer profit* is the trade profit per round for the buyer; *Seller profit* is the trade profit per round for the winning seller; *Efficiency* is the realised percentage of attainable trade surplus net of any verification cost incurred.

Bidding. The theory predicts that winning bids are lower in the *Arbitrator* treatment and that the equilibrium bid is a function of the arbitrator's preference. The data strongly supports the first prediction, but not the second.

Both bids and winning bids (the lower-order bidding statistic) differ between the two treatments. In *Voluntary*, the average bid submitted in the auction is 115.77 and in the pooled *Arbitrator* data, 78.55. The average winning bids are 105.64 and 69.24, respectively. Both differences are statistically significant in the direction predicted by

Hypothesis 3.1 (p-value < 0.01 for both comparisons).[49] Pairwise comparisons between the *Voluntary* treatment and the *Arbitrator* sub-treatments yield similar statistical differences.[50] The distribution of bids in *Voluntary* is stochastically larger than in *Arbitrator* (p-value < 0.001, Kolmogorov-Smirnov test, one-tailed). In *Voluntary*, bids are in line with the point prediction of 120 (p-value $= 0.563$).

Result 3.1a. *We find support for Hypothesis 3.1(i). A contingent contract results in significantly lower winning bids than a non-contingent contract.*

The theory is less successful at explaining bidding behaviour within-subjects in *Arbitrator*. When the arbitrator favours the buyer ($\mu = 1/3$), rather than the seller ($\mu = 2/3$), we expect sellers to adjust their bids upwards to compensate for the lower expected price in the event of winning the auction and the arbitrator being available. This is not observed in the data.

Average bids are 74.71 when $\mu = 1/3$ and 81.06 when $\mu = 2/3$. Winning bids are 65.29 and 72.96, respectively. Neither pairwise difference is significant (p-value $= 0.563$ and p-value $= 0.313$). The failure to adjust bids in response to a change in μ is not concealed by learning. If we restrict attention to experienced sellers only – periods 11 to 15 and 26 to 30 – neither bids nor winning bids are significantly different (p-value $= 1.00$ and p-value $= 0.563$). The null finding is also not a consequence of behavioural spillover effects. Spillover effects are possible in *Arbitrator* because μ is varied within subjects between the first and second half of the experiment (see Bednar et al. 2012, for a discussion of the importance of accounting for spillovers in economic experiments). The results in this section are qualitatively unchanged if we only consider data from periods one to fifteen and so only consider between-subjects variation.[51] This suggests that any spillover effects, if present, did not significantly alter subjects' behaviour.

---

[49] Unless otherwise stated, all p-values in this section are based on Wilcoxon-Mann-Whitney tests for two-sample comparisons and Wilcoxon Signed-Rank tests for one-sample comparisons. We report one-sided p-values if the theory predicts a direction, else two-sided.

[50] *Voluntary* vs. $\mu = 1/3$, p-value < 0.01 for both; *Voluntary* vs. $\mu = 2/3$, p-value $= 0.021$ and p-value < 0.01.

[51] Summary statistics for the spillover analyses are included in the appendix.

Result 3.1b. *We reject Hypothesis 3.1(ii). Potential sellers fail to adjust their bids based on expected values in response to a change in the arbitrator's preference.*

Efficiency. The theory predicts that high quality trading relationships are more likely to be observed with a contingent contract independently of the arbitrator's preferences. There is no systematic evidence for this in the data.

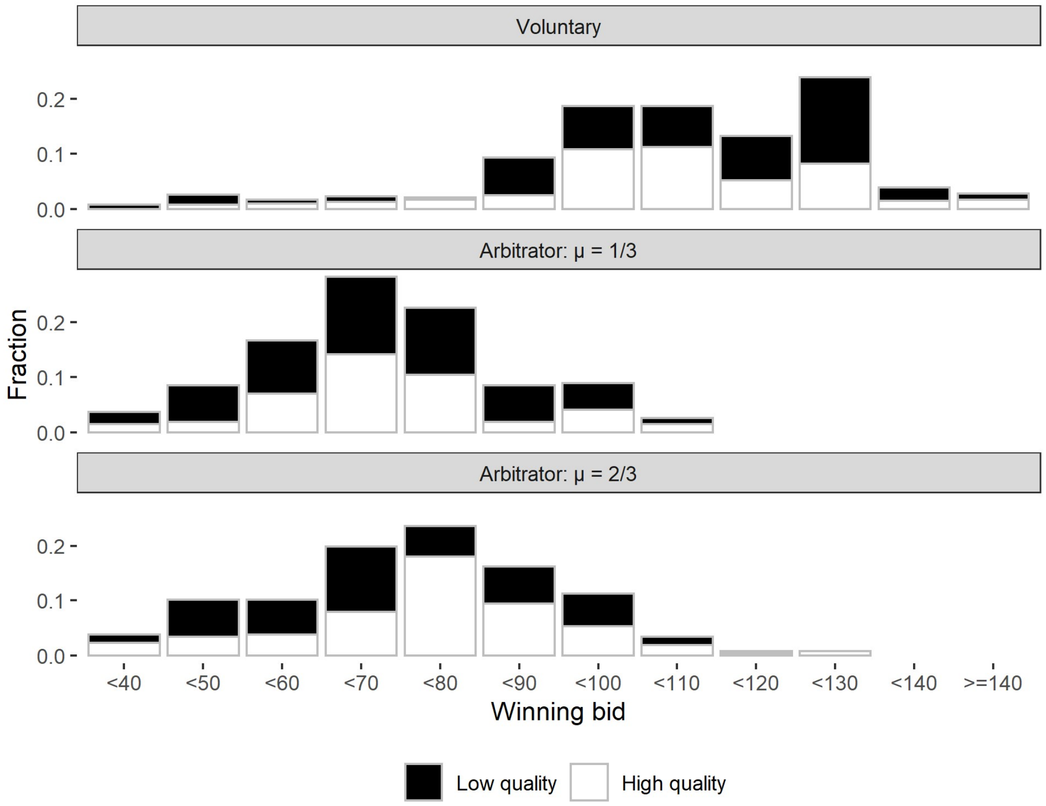
High quality trades are observed across the winning bid distributions in both treatments (see Figure 3.3). The relative frequencies of high quality delivered by the winning seller in the *Voluntary* and *Arbitrator* treatments are forty-five and forty-eight percent, respectively. This difference is not significant (p-value = 0.350). There is also no significant difference in quality between the *Arbitrator* sub-treatments (p-value = 0.313).

Consistent with the model, the relative frequency of dispute is high at close to eighty percent in both *Arbitrator* sub-treatments. After taking the verification cost into account, average trade efficiency is sixty-three percent in *Voluntary* and sixty-four percent in *Arbitrator*. This difference is not statistically significant (p-value = 0.531). There is no significant pairwise difference between *Voluntary* and either of the *Arbitrator* sub-treatments (*Voluntary* vs. $\mu = 1/3$, p-value = 0.803 and vs. $\mu = 2/3$, p-value = 0.294).

Result 3.2. *We reject Hypothesis 3.2. A contingent contract does not significantly increase trade efficiency relative to a non-contingent contract.*

Reciprocity. The model assumes standard preferences and so predicts that buyers satisfy their minimum payment obligation only. Given the vast literature in experimental economics documenting social preferences, it is unsurprising that this assumption is falsified behaviourally. More interesting is that buyer proposals (in profit-sharing terms) are significantly less generous with a contingent contract.

Figure 3.3. Distribution of winning bids and the relative frequency of high quality.



In *Voluntary*, the buyer on average proposes a price equal to 37.6 percent of the winning bid. In *Arbitrator*, the average proposal is forty-two percent, but of a lower average price. Both rates are significantly above the minimum required one-quarter (p-value = 0.016 for both comparisons) and this is consistent with the existence of social preferences among buyers.[52] There is a sustained gap between lower bound contract price and the average buyer proposal in the experiment over time. The average final price in *Arbitrator* is 40.78, which coincides with the equilibrium prediction of compensation equal to the seller's delivery cost of high quality. By contrast, in *Voluntary* the average final price is 38.93, significantly above the cost of low quality (p-value = 0.031).

The seller surplus shares implied by these proposals differ markedly between treatments and we interpret this as a measure of buyer reciprocity. Average price proposals in *Voluntary* imply a seller profit equal to 13.26 percent of the transaction surplus. The picture is very different in *Arbitrator*, where proposals imply a seller loss equal to 14.65

[52] Since proposals cannot be below one-quarter of the winning bid, we use a one-tailed test.

percent of the transaction surplus. The difference is highly significant (p-value < 0.01). The arbitration clause was triggered on four out of five transactions. Revealingly, while buyers in *Voluntary* offer the seller a similar surplus share independently of the quality level, buyers in *Arbitrator* offer significantly less in exchange for low quality than for high quality (p-value $= 0.031$). This is not driven by differences in bids accompanying the two quality levels (see next sub-section) and implies that buyers are less forgiving of a seller's reluctance to deliver high quality with a contingent contract.

Result 3.3. *We reject Hypothesis 3.3. A contingent contract crowds out buyer reciprocity and the arbitrator acts as a partial substitute for reciprocity in the determination of final prices.*

Profits. The standard theory predicts that sellers compete away their rents in the auction and so the buyer appropriates the full trade surplus, yielding higher profits in *Arbitrator*. This prediction is only partially borne out in the data.

In both the *Voluntary* and *Arbitrator* treatments, buyers earn significantly higher profits than sellers (p-value $= 0.016$ for both comparisons). The gains from high quality trade are also apparent. Buyers and sellers each earn roughly three times as much from high quality trades as low quality trades, in line with the surplus multiplier. Yet buyers do not benefit from the presence of an arbitrator. In fact, buyer profits are remarkably stable across experimental conditions, at 33.85 in *Voluntary*, 33.80 in *Arbitrator* when $\mu = 1/3$ and 32.32 when $\mu = 2/3$. Neither pairwise difference is significant.[53]

The average seller profit is 4.79 in *Voluntary* and 6.42 in *Arbitrator*. This increase is not significant (p-value $= 0.818$). It does, however, mask some underlying differences between the *Arbitrator* sub-treatments. Seller profits are 3.35 when $\mu = 1/3$ and 8.01 when $\mu = 2/3$. While profits in the $\mu = 2/3$ sub-treatment of *Arbitrator* are not significantly different from the *Voluntary* treatment (p-value $= 0.310$), we cannot rule out that this is due to a lack of statistical power – see the robustness check below. More data is required to make concrete inferences on whether, and the conditions under which,

[53] *Voluntary* vs. $\mu = 1/3$, p-value $= 0.758$ and vs. $\mu = 2/3$, p-value $= 0.758$; $\mu = 1/3$ vs. $2/3$, p-value $= 1.00$.

an arbitration mechanism can improve the seller's position in competitive procurement interactions.

Result 3.4. *We find partial support for Hypothesis 3.4. Buyers earn significantly more than sellers, but neither trade party significantly benefits from a contingent contract.*

Robustness check: Since the number of independent observations per treatment is small, low statistical power might be a concern for any null result in the hypothesised direction. There is disaggregated data from 540 matching groups in each treatment, which yields greater statistical power to detect effects than the more conservative mean comparison tests used above. Results 3.1 to 3.4 are qualitatively unchanged if we conduct a regression analysis on the matching group level data, accounting for intra-cohort dependencies and small-sample considerations using the wild cluster bootstrap method (Cameron et al. 2008). The results of this exercise are contained in the appendix.

3.4.2. Seller bidding strategies

In this sub-section, we analyse seller bidding strategies formulated during the auction phase of the experiment. There is data from 1,070 strategies submitted in *Voluntary* and 1,059 strategies submitted in *Arbitrator*.[54] High quality is a component of forty-two percent of bidding strategies in *Voluntary*. In *Arbitrator*, this is true of forty-six (fifty-five) percent of strategies when $\mu = 1/3$ ($\mu = 2/3$), which is a weakly significant increase over *Voluntary* (p-value = 0.066). Thus, whereas a contingent contract does not significantly improve the winning quality delivered, there is some evidence that it promotes higher quality among the pool of potential sellers. This implies a negative selection effect of the reverse auction.

As a first step, it is instructive to examine average bids that accompany the respective quality levels at the cohort level. In *Voluntary*, the median bid for low quality strategies

[54] The difference is due to the use of a hard time-out protocol in the experimental sessions. In the event of time-out, the default bid was 200, the quality level was decided at random and the price proposal was set equal to the winning bid. The rate of data loss is less than one percent, which seems a reasonable trade-off given the benefits of the protocol for minimizing delays during a session.

is 118.35, very close to the equilibrium prediction [120]. The median bid for high quality strategies is significantly lower at 109.71 (p-value = 0.031). This suggests that, with a non-contingent contract, deviation to a lower bid may have been driven by a belief that high quality is a "gift" that buyers will reciprocate. By contrast, in *Arbitrator* the median bid for high quality strategies at 81.86 is significantly higher than for low quality strategies at 74.46 (p-value = 0.063). One explanation is that potential sellers anticipate lower proposed surplus shares with a contingent contract and adjust their bids to compensate; i.e., sellers must insure themselves against relatively lower price offers.

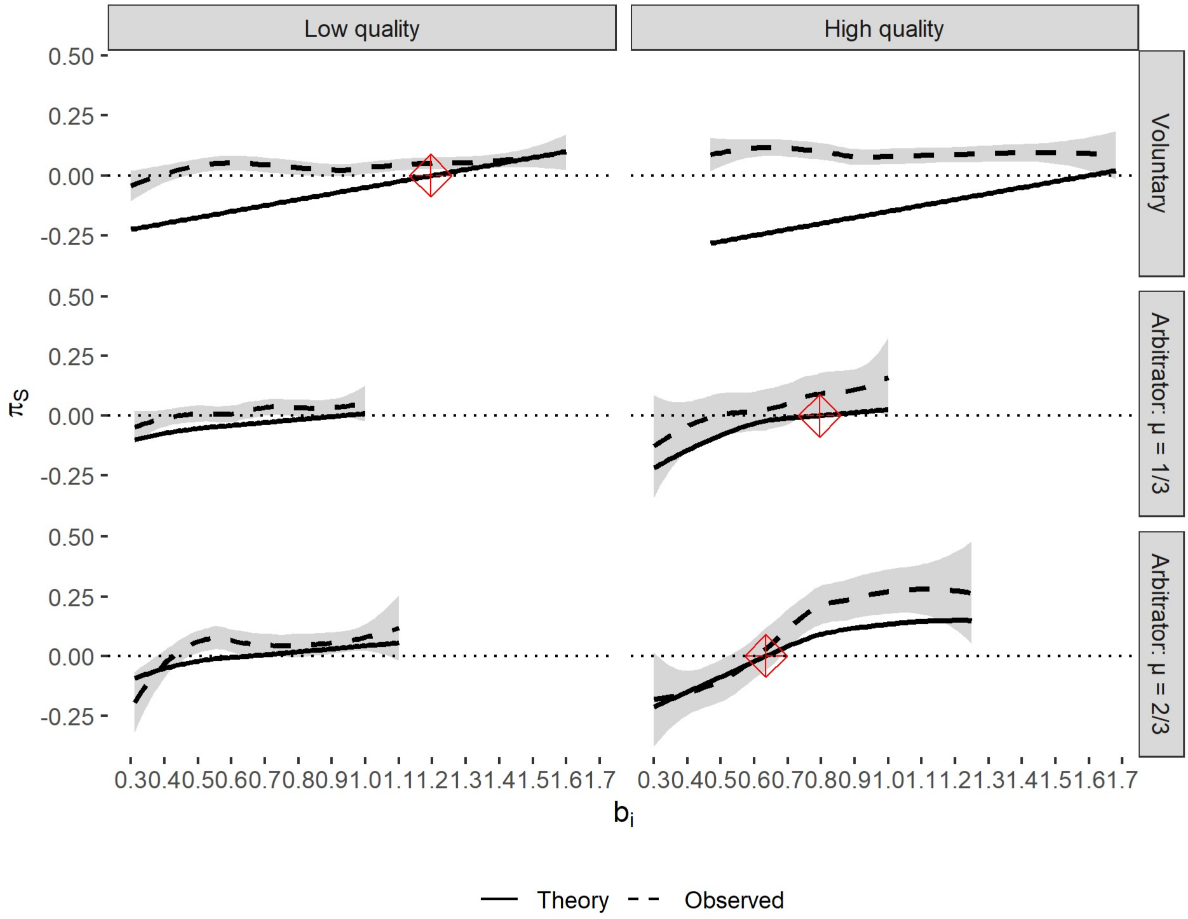
In Figure 3.4 we compare *actual* versus *predicted* seller profits per quality level, as a function of the winning bid. The SPNE bid in each experimental condition is marked with a red diamond plus. Consistent with the aggregate results, seller profit levels are stable across bids in *Voluntary*. On average, observed seller profits remain flat above the break-even level, independent of the bid and quality choice. In *Arbitrator*, however, seller profits are sensitive to the winning bid and more closely track the prediction based on the assumption that buyers act as strict profit maximisers. Note that at the median winning bid in each *Arbitrator* sub-treatment, high quality is incentive compatible based on actual profits.

Table 3.5 presents the results of regression analyses to investigate the determinants of suppliers' bidding strategies. First, to examine factors affecting bids, we use random effects regressions with three levels of dependencies and controlling for a range of covariates (models 1 to 4, adapted from Moffatt 2015, Chapter 4.7). The dependent variable is the bid of subject i in cohort j in period t. To account for intra-session correlation, the variance is estimated at subject and cohort levels.[55] We observe a positive dependency of current bids on the once-lagged competitor's bid, which captures the strategic complementarities of Bertrand competition. Bids trend downwards over time in both *Voluntary* and *Arbitrator* treatments. In models 2 and 4, we add the buyer's proposed

[55] The specification is $y_{ijt} = \beta_0 + \beta_1 x_{ijt} + \beta_2 z_i + \beta_3 o_j + \beta_4 t + u_i + \upsilon_j + \varepsilon_{ijt}$, variances $V(u_i) = \sigma_u^2$, $V(\upsilon_j) = \sigma_\upsilon^2$ and $V(\varepsilon_{ijt}) = \sigma_\varepsilon^2$. The three levels are $i = 1, \dots, 36$ (sellers), $j = 1, \dots, 6$ (cohorts), $t = 1, \dots, 30$ (periods). Independent variables are specific to the subject-cohort-period (x_{ijt}), subject only (z_i) or cohort only (o_j).

bid discount from the previous period as a regressor to measure the influence of buyer reciprocity. This measure drives significantly higher bids in both treatments. The effect size is much larger in *Voluntary*, where a 10 percentage point increase in the most recently proposed bid discount leads suppliers to adjust their bid upward by nearly five points on average. In *Arbitrator*, the adjustment is just over one point.

Figure 3.4. Seller profits as a function of the winning bid: Actual versus predicted.



Notes: The lines are loess smoothers and the shaded regions for the observed data are 95% confidence intervals. The red diamond plus is the equilibrium bid level.

Table 3.5 – Determinants of bid strategies: the most recent bid discount drives higher bids and lower probability of high quality, but the effect weakens in Arbitrator.

Dependent variable	Bid				Prob. High Quality	
Treatment	*Voluntary*		*Arbitrator*		*Voluntary*	*Arbitrator*
Model	(1)	(2)	(3)	(4)	(5)	(6)
Competitor's bid $_{t\text{-}1}$	0.222***	0.218***	0.075***	0.061**		
	(0.024)	(0.029)	(0.024)	(0.027)		
Bid (IV = Bid $_{t\text{-}1}$*)*					0.011	-0.005
					(0.010)	(0.013)
Proposed bid discount		0.471***		0.119***	-0.026***	-0.009***
(ppt) $_{t\text{-}1}$		(0.048)		(0.034)	(0.007)	(0.003)
Period	-0.302***	-0.132*	-0.510***	-0.679***	0.005	-0.011
	(0.063)	(0.077)	(0.077)	(0.101)	(0.011)	(0.013)
$\mu = 2/3$			1.44	2.225		0.096
			(1.28)	(1.643)		(0.182)
OrderHL			8.37	11.791		0.313
			(9.38)	(8.911)		(0.529)
Constant	87.62***	61.375***	85.38***	53.062***	0.453	-1.185
	(13.58)	(9.692)	(19.95)	(19.639)	(1.266)	(1.840)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Random effects (subject and cohort)	Yes	Yes	Yes	Yes	No	No
Cohort fixed effect	No	No	No	No	Yes	Yes
Observations	1,037	512	990	488	512	488
Subjects	36	36	36	36	36	36
Cohorts	6	6	6	6	6	6
Wald χ^2	149.8	266.9	85.08	82.46	56	29.98
χ^2 *p*-value	0.000	0.000	0.000	0.000	0.000	0.012
σ_v	15.01	8.400	11.08	10.46		
σ_u	9.601	5.511	5.437	4.639		
σ_ε	16.68	13.56	20.05	17.89		

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The models in columns 1 to 4 are estimated using multilevel mixed effects linear regression and include two random effects intercepts that capture intra-session correlation at the subject and cohort levels, respectively. Coefficient estimates are presented, with standard errors in parentheses. The models in columns 5 and 6 are estimated using instrumental variable (IV) probit regression and include cohort fixed effects. Coefficient estimates are on the z-score scale, with robust standard errors clustered at the subject level. *Bid* is the continuous endogenous regressor and the excluded instrument is the once-lagged bid. The IV first stage is reported in the appendix. *Proposed bid discount (ppt)* $_{t\text{-}1}$ is one minus the ratio of the buyer's proposed price to the winning bid in the previous period, expressed in percentage points. The following control variables (not shown) are included: dummy for being female; dummy for being an economics and finance major; two Likert scales for self-reported willingness to take risks in general and in financial matters; dummy for reporting trust in strangers; and a generalised trust index. *OrderHL* is a dummy variable to indicate that the cohort followed the sequence $\mu = \{2/3, 1/3\}$ in the *Arbitrator* treatment.

To examine the factors affecting sellers' quality choices, we estimate instrumental variable probit regressions in which quality is the dependent variable and bid is the endogenous regressor, which is instrumented for by the once-lagged bid (models 5 and 6). The instrumental variable estimator is used to mitigate the issue of simultaneity between bids and quality choices inherent in the strategy method application. The instrument passes recognised strength tests – see the appendix for diagnostic statistics. Standard errors are adjusted using the Huber-White sandwich estimator of variance at the subject level. Cohort fixed effects are included in both regressions, based on a joint test of their significance (p-value < 0.001, two-tailed Wald test).

In both treatments, the only significant predictor of the seller's willingness to deliver high quality is the buyer's most recently proposed bid discount.[56] In *Voluntary*, a 10 percentage point reduction in the most recent bid discount increases the probability of high quality by 9.4 percentage points. Again, the effect size is smaller in *Arbitrator*, where a 10 percentage point reduction in the most recent bid discount increases the probability of high quality by just 3.8 percentage points. The effect is robust to different values of the arbitrator's preference. This suggests that seller reciprocity is also attenuated with a contingent contract, which hampers the effectiveness of the contingent contract for improving trade efficiency. This offers further support for Results 3.2 and 3.3 above.

3.4.3. Reference-dependent buyers

In this sub-section, we analyse buyer behaviour during the price proposal phase of the experiment. There are eighteen buyers in each treatment and every buyer is observed over thirty trading periods. We have data from 535 price proposal decisions in each treatment. In *Voluntary*, only 23.7 percent of buyer choices are of the lower bound price (i.e., the maximum bid discount). This varies by the seller's chosen quality level, equal to 31.4 percent for low quality and less than 15 percent for high quality. There are

[56] Reassuringly, the order in which the values of μ are presented to sellers in *Arbitrator* has no significant effect on either bid or quality decisions. No significant order effects are found in determining buyer proposals either (see the appendix for details).

masses at the salient fifty-fifty surplus split, but the most common allocations are found at three-to-one and five-to-one buyer to seller ratios.

In *Arbitrator*, proposals of the lower bound price are observed more frequently: 46.8 percent of choices when $\mu = 1/3$, and 50.8 percent of choices when $\mu = 2/3$. These frequencies are stable across quality levels. Thus, a seller in this treatment could expect to encounter a strictly profit-maximising buyer on roughly one-half of transactions. Around one-in-five buyer choices in *Arbitrator* are "mimicking". That is, the buyer proposes the reference arbitrator price, given the winning bid and quality level delivered. Note that the arbitrator's price does not necessarily coincide with the final price even when the arbitrator is available, because of the constraint to award at least the buyer's proposal. The frequency of mimicking choices is stable across quality levels. Most of the remaining proposals are in the region between the lower bound price and the reference arbitrator price. Less than four percent of proposals award the seller more than what the arbitrator deems fair.

In both treatments, buyers tend towards strictly profit-maximising choices as the experiment progresses. Non-parametric bootstrap tests of a positive time trend in the number of lower bound proposals show strong significance across conditions.[57] The standard theory cannot explain the persistence of mimicking choices in *Arbitrator*. When $\mu = 1/3$, no time trend is observed in the average number of mimicking choices (p-value $= 0.493$); when $\mu = 2/3$, there is a significant *positive* trend (p-value $= 0.012$).

One explanation for the persistence of such choices is that some buyers have reference-dependent fairness preferences over outcomes. In the experiment, subjects are informed in advance about the arbitrator's preference and when this would change. Other-regarding buyers may then have anchored their proposals on the arbitrator's preference and perceived this as the fair reference point. If the reference point is determined by recently held expectations (e.g., Kőszegi and Rabin 2006), then we would expect the proportion of mimicking choices to be stable across the *Arbitrator* sub-treatments. The

---

[57] Based on 999 replications, *Voluntary*: Mean = 0.365, S.E. = 0.191, p-value = 0.025; *Arbitrator*, $\mu = 1/3$: Mean = 0.763, S.E. = 0.269, p-value < 0.001; *Arbitrator*, $\mu = 2/3$: Mean = 0.553, S.E. = 0.270, p-value = 0.01.

plausibility of the reference-dependent argument is suggested by the raw choice data: 22.3% (20.7%) of choices are mimicking when $\mu = 1/3$ $(= 2/3)$, respectively.

To test the reference-dependent argument, we separate buyers in *Arbitrator* into one of two behavioural types: self-interested (*Self*) or fair (*Fair*). We then estimate the parameters underlying each type's price proposal function using a finite mixture model. This structural approach will help us to identify the determinants of buyer behaviour in the experiment. The mixture model estimation procedure is adapted from Moffatt (2015) and full details are contained in the appendix. Self-interested buyers are associated with maximal bid discounts. For fair buyers, we specify a latent model in which price proposals anchor on the arbitrator's preferred price with a normally distributed error term. An implicit assumption is that subjects do not switch between types.[58] A two-limit tobit model is appropriate, where the limits correspond to the lower/upper bounds of the contract price range and are specific to each buyer and period. The type frequencies are given by the mixing fractions γ_{self} and γ_{fair}.

The maximum likelihood estimates from this procedure are summarised in Table 3.6. Bootstrapped standard errors are computed, with resampling clusters at the cohort level. The estimated fraction of self-interested (fair) buyers in *Arbitrator* is 0.611 (0.389) (p-value < 0.001). Consistent with the reference-dependent argument, the coefficient estimate on the arbitrator's preferred price is not significantly different from one (*p*-value $= 0.853$, two-sided Wald test). In Figure 3.5, panel (a), we display posterior type probabilities as a function of the number of lower bound price proposals. The posteriors separate buyers into one of the two types. A cluster of buyers in *Arbitrator* never choose the lower bound price. Self-interested types choose the lower bound price on more than ten transactions. Panel (b) of the figure presents the histogram of proposals by type, as a share of trade surplus. We exclude any proposal equal to the winning bid (35 observations) because these do not necessarily reflect a buyer's true preference. For fair types, the largest densities can be found at 1/3 and 2/3 of the surplus. If we include proposals equal to the winning bid, the mass at 1/3 of the surplus increases further.

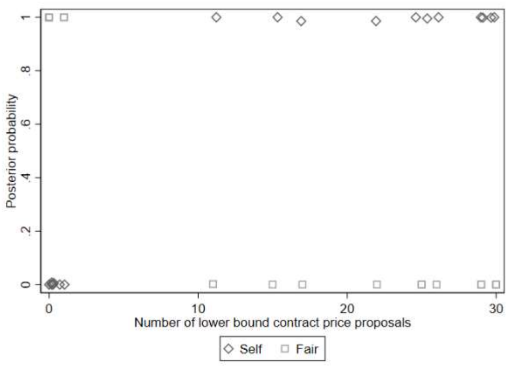
[58] Buyers appear to gravitate towards one strategy during the experiment, rather than mixing between different strategies (see the appendix).

Table 3.6 – Maximum likelihood estimates from the finite mixture two-limit Tobit model of buyer behaviour in the Arbitrator treatment.

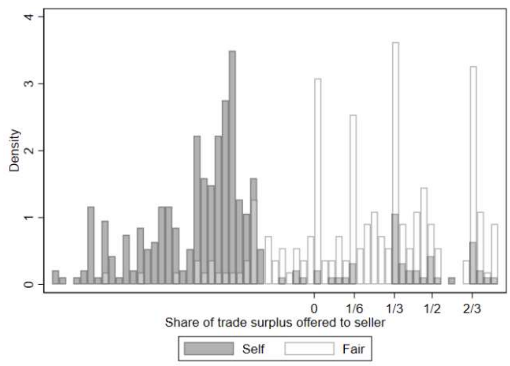
Fair buyers	
Arbitrator's preferred price	0.971 (0.158)***
Constant	-3.64 (5.76)
Error term sd.	8.51 (6.54)
Mixing fractions	
γ_{self}	0.611 (0.084)***
γ_{fair}	0.389 (0.084)***
Observations	535
Wald χ^2	37.97
χ^2 p-value	0.000
(Subjects, Periods)	(18, 30)

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors in parentheses, clustered at the cohort level, based on 999 bootstrap replications.

Figure 3.5. Finite mixture model validation.



(a) Posterior type probabilities.



(b) Histogram of buyer proposals.

Notes: Panel (a) contains a jittered scatter plot of posterior type probabilities from the finite mixture two-limit Tobit model. Panel (b) contains a histogram of buyer price proposals as a share of the trade surplus, excluding upper censored price proposals and one outlying observation in which a buyer offered more than the full trade surplus to the seller.

3.5. Discussion and Conclusion

Hierarchical elements, such as dispute resolution provisions, are common elements of procurement contracts when there are moral hazard and liability issues. Contingent contracts are intended to avoid the prohibitive costs of lengthy court proceedings and align incentives to maximise gains from trade. We use a combination of theory and experiments to examine the behavioural consequences of contingent contracting in a procurement relationship characterised by two-sided moral hazard. The contingency is modelled as an arbitration clause; the seller is determined by a reverse auction; and the buyer has discretion to deduct monies from the seller in exchange for quality delivered. In theory, a non-contingent contract yields low quality trade, while a contingent contract can improve trade efficiency and generate more competitive bids.

We observe in a lab experiment that, while the contingent contract is robust in the sense that it relies only on the regard of each party to their own interest, it also crowds out reciprocity. Buyers reward suppliers less (in profit-sharing terms) for quality delivered. Sellers are in turn less willing to deliver high quality and so the contingent contract does not significantly increase the efficiency of trade relative to a non-contingent contract. This crowding out effect is consistent with evidence from labour market experiments on the demotivating effects of explicit incentives (Fehr and Gächter 2000, Fehr et al. 2007), and with principal-agent literature on the hidden costs of incentive devices (Falk and Kosfeld 2006).

The model developed here can rationalise why powerful corporations may favour the use of arbitration agreements, from simple profit-maximizing behaviour. In theory, the buyer appropriates the increased surplus from trading with a contingent contract. More data is required to inform on whether a contingent contract improves the seller's position empirically. An interesting implication of the analysis arises from the prediction that uncertainty is a necessary condition for the contingent contract to be successful. That is, when sellers compete on price, guaranteeing quality verification may be inefficient. Since the total verification cost increases in the relative frequency with which an arbitrator is employed, any industry scheme guaranteeing the availability of an arbitrator during trade may incur an unnecessary and counter-productive expense. Potential sellers may use such information to lower their bids and increase their probability of winning the contract. This observation is consistent with anecdotal evidence that sellers lower their bids in expectation of later recovering costs through a dispute mechanism.[59]

From a trade efficiency perspective, both flexible price contract types tested here perform favourably when compared to the allocation of a rigid, fixed-price contract in reverse auctions with moral hazard (Brosig-Koch and Heinrich 2014, Fugger, Katok and Wambach 2019). There are, of course, alternative mechanisms to mitigate the seller-side moral hazard problem, such as buyer-determined auctions and systems of reputation. And since there is a cost involved in organising an arbitration scheme, these mechanisms may well be complementary and/or preferable in certain settings. That said, mechanisms that can preserve price competition have transparency benefits beyond those captured in the model. For this reason, moving away from a commitment to procure at the lowest price is not always palatable for the regulator.

There are several limitations of the model that might be addressed in future work. First, an optimal dispute resolution mechanism was employed in the experiment, as implemented by a computer program. Whether or not trade parties make optimal use of dispute resolution is an interesting question for future work. An alternative design would be to assess how payment norms emerge endogenously if a human arbitrator(s) were introduced into the experiment. Third-party punishment behaviours have been observed to serve a useful function in enforcing distributional norms (Fehr and Fischbacher 2004).

Second, the implementation of the arbitrator's preference is a known and deterministic function of the transaction price and quality. In practice, the final deferred value is likely to involve noise, especially when actions are difficult to verify. This *caveat* should be kept in mind if drawing policy implications from the analytical results, which may

[59] For example, see the following extract taken from a survey of industry practitioners: "Some interviewees noted that they had observed project participants bidding a lower price upon the expectation of recovering sums through variation orders which, if disputed, could be arbitrated" (Pinsent Masons 2019, p 7).

overstate the effects of a contingent contract in real-world dispute resolution. One avenue for further research would be to add a stochastic component into the preference set.

Third, the nature of supplier competition considered here is strong. Small supplier profit margins are characteristic of many procurement settings. For example, publicly traded highway construction firms in the sample of Bajari et al. (2014) report profit margins of less than three percent. However, strong supplier competition is not representative of all procurement settings and it would be useful to extend the analysis to a balanced market environment, a market in which the seller holds the balance of power, and/or alternative auction selection rules.

Finally, the assumption that buyers are endowed with full bargaining power is often used in the contract-theoretic literature to create the most severe hold-up problem (Hart and Moore 1999, Hoppe and Schmitz 2011). The model here could be generalised to allow for an investigation of a seller's relative bargaining power in determining the final trading price. Future work might also consider the effects of contingent contracting in procurement with multi-sourcing of suppliers. As Volkswagen found to its detriment in 2016, over-dependence on any one supplier can lead to costly dispute.

4 The Horizontal Effects of Late Payments

"Imagine walking into a shop, taking what you want up to the counter, and then, when the time comes to pay, saying 'thanks – but I think I'll just take these now and pay later', and casually walking out with your items. You obviously wouldn't dream of it – so why is this kind of behaviour acceptable when it comes to paying small business suppliers?"

- Stephen P. Kelly, CEO, Sage (2017).[60]

4.1. Introduction

Existing models of price competition assume that suppliers receive payment on demand. Yet often, suppliers must account for the possibility of late payment when formulating their bidding strategies. As a consequence, they may value prospective revenue below the posted price. A multi-national industry report uncovered that more than one in ten invoices issued by small and medium-sized enterprises are paid with delay, ranging from 8% in Brazil, to 12.5% in the US, and 18% in the UK and Singapore. This translates to over 1 trillion US dollars in payments made outside of the agreed term globally (Miller and Wongsaroj 2017). Late payments on large invoices are equally as likely as those on small invoices, and so from a managerial perspective pose a significant business risk. Research released by the British Federation of Small Businesses revealed that the failure of clients to pay on-time was perceived as the single greatest risk faced by suppliers (FSB 2018).[61] A separate representative survey of UK businesses from 2020 found that 89% of respondents pay their suppliers late (Bottomline 2020).

Payment delays have a negative effect on supplier cash flows. There is a disproportionate impact on smaller companies, which are typically more liquidity constrained.

[60] https://www.linkedin.com/pulse/lets-call-time-late-payments-stephen-kelly/.

[61] In the FSB survey, more than 80% of businesses reported experience of receiving payments beyond the due date. This practice spanned multiple business sectors, from construction to retail, with one in five suppliers to the public sector also affected. As of January 2021, £23.4 billion worth of late invoices are owed to businesses in Britain alone (Small Business Commissioner 2021).

Indeed, a culture of poor payment practice extends beyond missed invoice dates. For example, Carillion - the collapsed facilities management and construction services group – were “notorious late payers” and even requested a discount from suppliers in exchange for use of an early repayment facility. Evidence submitted to a parliamentary select committee in 2018 suggests that Carillion had doubled its payment term and used suppliers as a line of credit. Such practices are made possible by the power imbalance between smaller contractors and their larger clients.[62]

The aggregate impact of these payment practices is stark. Late payments were estimated to be responsible for the closure of 50,000 companies in the UK during a single year, undermining investment and hiring (FSB 2016). Market commentators routinely claim that they inflate consumer prices.[63] In EU member states, a significant statistical relationship has been documented between the timeliness of business-to-business payments and firm survival rates (Connell 2014, Conti et al. 2020). Late payments have also been identified with a withholding of supplies and a counter-productive increase in procurement costs for buyers (Howorth and Reber 2003). An obvious challenge in inferring causality from empirical data on the impact of late payments is endogeneity. Markets in which late payments abound may be characterised by financially weak firms, who are more likely to mark up their prices to achieve positive profits or otherwise exit the market. Therefore it is extremely difficult to separate the horizontal effects of late payments from broader financial constraints which are themselves inherent in field data.

In this chapter, we use theory and experiments to circumvent the endogeneity problem and isolate the channel through which payment delays influence supplier bidding strategies and competition. The analytical results build from earlier models of Bertrand competition in the industrial organization literature. We assume simultaneous entry and pricing with non-zero entry costs. The number and identity of rival entrants is unobserved before bidding and the Nash equilibrium is characterised in symmetric mixed strategies (Lang and Rosenthal 1991, Levin and Smith 1994). To focus on the role of

[62] See “Committees publish correspondence” (2018).

[63] The following quote is indicative of this type of reasoning, “the additional financing costs that suppliers incur because they aren’t being paid promptly work their way back into higher prices for consumers” (Strom 2015).

late payments in entry and pricing decisions, we assume away other factors, such as repeated interactions and renegotiation, which might confound the underlying channel of interest. While we acknowledge that this may be restrictive from a practical perspective, we consider the absence of these factors a strength of the experimental methodology in enabling us to address the research question of interest, that is, to identify the specific horizontal (rather than vertical) effects of late payments.

We also consider aspects of the regulatory environment that might influence a buyer's choice of payment strategy. Specifically, we analyse the impact of an exogenous penalty (e.g., a regulatory fine) imposed on the buyer for reneging on an intention to pay within the standard payment term. This type of penalty is an active area of policy debate. In the UK, legislation enacted in 2016 established the Office of the Small Business Commissioner with the remit to crack down on the endemic culture of late payment. New powers proposed by the government include the ability to levy fines on firms that fail to meet their payment obligations (Small Business Commissioner 2021). The EU directive 2011/7/EU on combating late payment in commercial transactions (2011) sets minimum rules specifying that public and private companies must pay their invoices within 30 and 60 days, respectively, or else incur a financial penalty.

One strategy used by buyers to avoid paying beyond the due date is to extend long payment terms. Absent deterrent penalties and/or in situations where buyers are endowed with high bargaining power, extending long payment terms makes commercial sense to maximise capital usage.[64] Even where deterrent regulation exists, such as 2011/7/EU, there are exceptions for trade parties who "expressly agree" to longer payment periods (p. L 48/2).[65] There is growing empirical evidence that many large companies agree long payment terms with their suppliers (IACCM 2015). That is, it is not only financially weak buyers who pay late. The following extract from The New York Times (NYT) is a case in point: "Exhibit A is Cisco Systems, one of the largest

[64] And has even propagated a new business model. See: https://www.bloomberg.com/news/articles/2019-08-16/how-late-payments-to-vendors-spawned-a-new-business-quicktake.

[65] Moreover, where exceptions should not apply, enforcement of prompt payment terms is variable. See, for example, the recent judgment in Commission v Italy (Directive combating late payment) (C-122/18).

technology companies in the world, which announced last year that it would wait a full 60 days to pay its small-business suppliers - mostly because it had found that that was what other big companies were doing" (Seligson 2011). Another NYT commentary pointed to the popularity of this tactic among the world's largest food processing companies: "Diageo, the European spirits company, now asks for 90 days to pay its bills. Mondelez, Mars and Kellogg seek 120 days. The list of companies doing the same reads like a grocery store version of Who's Who" (Strom 2015).

Given the observed payment strategies, we model the interaction as a signalling game. In our setting, firms possess a discount factor by which they mark down late payments relative to the nominal payment value. We interpret this discount factor as a firm's type. For ease of analysis, we confine our theoretical environment to two types of buyer: Strong and Weak. Weak buyers place greater value on late payment than Strong buyers. The timeline of events in our model is as follows. First, the buyer observes her type and signals an *ex-ante* payment term. We consider two payment terms: standard (e.g., 30 days) or extended (e.g., 60 to 120 days). Second, potential suppliers observe the buyer's signal, but not her underlying type. They must then decide whether or not to incur a fixed entry cost and enter into price competition. The supplier that submits the lowest bid wins the right to contract with the buyer. Third, the buyer observes the winning bid and selects the final payment date *ex-post*. The buyer is free to pay within or beyond the signalled payment term but a penalty is imposed on the buyer for reneging on a standard term. The contribution of the signalling structure is to enable us to understand how the institutional environment (i.e., exogenous penalty) interacts with the buyer's financial capacity (i.e., discount factor) and the regulatory implications for setting an optimal penalty.

We show theoretically that payment delays feed into higher consumer prices and reduced supplier entry. Consistent with the anecdotal evidence described above, when the penalty imposed on buyers for paying suppliers late is low, even Strong buyers – who may benefit from large cash reserves, plentiful access to credit and by implication high discount factors – signal an extended payment term in equilibrium. Within deterrent institutional environments, whether buyers fulfil a standard payment term depends on the relative financial strength of buyers in the population. Financial strength *per se*, however, does not preclude the emergence of late payments in equilibrium. Our analysis can rationalise a late payment event if the cost of reneging on a standard payment term is not set optimally and there exist price spillover effects in the market. Moreover, we demonstrate that increasing the size of a penalty for late payment may not necessarily result in a welfare gain.

In the second part of the chapter, we design a lab experiment to test the predictions of the model. Experimentation has served as a useful tool for testing predictions of the classic Bertrand-Nash oligopoly model with constant and symmetric marginal costs of production and exogenous market size (Dufwenberg and Gneezy 2000, 2002, Dufwenberg et al. 2007, Fouraker and Siegel 1963). A private correspondence with Martin Dufwenberg in July 2019 confirmed that non-zero entry costs were not tested in the original set of Bertrand competition experiments, neither published nor unpublished. Several experiments have been conducted in the related Bertrand-Edgeworth environment, with capacity constraints (Fonseca and Normann 2008, 2013, Grether et al. 1988, Heymann et al. 2014, Kruse et al. 1994) and/or convex cost functions (Abbink and Brandts 2008, Argenton and Müller 2012). Recent lab studies have integrated asymmetric marginal costs into the Bertrand environment (Boone et al. 2012, Dugar and Mitra 2016). This line of inquiry complements theoretical work attempting to explain price dispersion, a real-world phenomenon at odds with marginal cost pricing and that supports the Bertrand paradox.[66] We employ a strategy method design to control for the potential confound of excessive seller entry observed in previous auction experiments with an endogenous number of bidders (e.g., Palfrey and Pevnitskaya 2008). This simplifies the game

[66] The Bertrand paradox disappears theoretically if we change assumptions underlying the model, including capacity constraints (Edgeworth 1925), convex costs (Dastidar 1995), product differentiation (Singh and Vives 1984), *ex-ante* demand commitments (Kreps and Scheinkman 1983), cost uncertainty (Wambach 1999), a probabilistic number of competitors (Janssen and Rasmusen 2002) or the introduction of unbounded monopoly profits (Baye and Morgan 1999). Behavioural explanations for price dispersion cover bounded rationality (Baye and Morgan 2004), step-level thinking (Gneezy 2005), or cognitive limitations due to perceived coarseness of the pricing grid (Fatas et al. 2014).

and permits us to focus on the horizontal effects of late payments for prices, which is not feasible with field data.

The experimental results largely support the model's predictions. The anticipation of late payments has anti-competitive effects among suppliers and results in higher transaction prices. When the penalty imposed for a late payment is low, the majority of buyers recognise that extending standard terms is not preferred and often signal extended payment terms up front. Nevertheless, within such an environment and in contrast to the theory, a significant minority of buyers - in particular Strong buyers - still offer standard payment terms in an attempt to induce greater supplier competition. We also observe that the propensity of Strong buyers to pay late is decreasing in the size of the penalty. Weak buyers are only deterred if the penalty is above a certain threshold; below this threshold, these buyers prefer to free ride on the price spillover effects generated by other firms in the market and a welfare loss arises. We also describe the results of a follow-up experiment that we conducted in which sellers can determine their own entry decisions.

4.2. Model and Theory

Consider a one-shot interaction in which a single buyer seeks to procure one unit of an indivisible good from a group of n pre-qualified suppliers, indexed by i. A first-price sealed-bid reverse auction determines selection of the winning supplier (henceforth the seller) and the contract price. Auction participation is voluntary for suppliers, and the buyer can choose to not purchase after observing the contract price.[67] The seller can produce either a high- or low-quality unit, but the setting prohibits quality contingent contracts.

4.2.1. Signalling game

We build from a one-shot Bertrand model with simultaneous entry and pricing. Suppose that there are two *ex-ante* homogenous potential suppliers, indexed by i. Each

[67] Discretion clauses are common in construction tenders.

supplier has the opportunity to submit a bid b_i to deliver a contract to a single buyer at date 0.[68] Conditional on at least one supplier entering into competition, the supplier that submits the lowest price bid wins the contract at a price p equal to the winning bid, with ties broken at random. Henceforth, we will denote the winning supplier as the seller. Bids can be any number in the interval $[0, R]$, where R is large and constitutes the reserve price at auction.

Before submitting a bid, each supplier must decide whether or not to incur a non-refundable entry cost $E > 0$. This is the cost of preparing a bid for tender, which is not imposed if a firm chooses not to enter into competition. Whether the rival supplier has entered is unobserved until after the entry and pricing decision is made. Production costs are normalised to zero. If no bid is submitted, the contract is not completed. In that case, all parties earn zero. Specifying that entry and pricing be simultaneous, rather than sequential, is a modelling decision. Sequential entry and pricing generate counter-intuitive predictions when potential suppliers are asymmetric, due to the nature of mixed strategy equilibrium (Elberfeld and Wolfstetter 1999).

We now embed the Bertrand entry and pricing model into a signalling game structure.

Stage 1. The buyer sends the potential suppliers a message, $m \in \{0, 1\}$, announcing her *ex-ante* intended payment term, where $m = 0$ is an offer of a standard payment term (e.g., 30 days) and $m = 1$ is an offer of an extended payment term (e.g., 60 to 120 days). This message is non-discriminatory, i.e., it does not depend on the identity of the seller.

Stage 2. The potential suppliers observe the buyer's message m and make their entry and bidding decisions. Conditional on at least one entering supplier, the seller and price are determined as described above.

Stage 3. The buyer observes the potential suppliers' entry and pricing decisions and selects the *ex-post* payment date $\tau \in \{0,1\}$ to pay the seller, where $\tau = 0$ denotes

[68] Restricting entry and price competition to just two suppliers is without loss of generality. See Lang and Rosenthal (1991) for the general case of M potential suppliers.

payment on or before the due date associated with a standard payment term (henceforth, "on-time"), and $\tau = 1$ denotes a payment made after this date (henceforth, "late").[69]

Just like how individuals prefer current income to delayed income of a similar magnitude, firms value a late payment *ex-post* at less than its nominal value. This insight is captured in the model by a discount function $D(\tau) = \delta_i^\tau$, with $D(1) < D(0) = 1$. The corresponding discount factor is given by $\delta_i \in (0, 1)$, which is firm-specific and might be determined by cash reserves, access to secured credit and/or the prevailing interest rate. We endow the buyer with private information about her discount factor, which we refer to as her type. This type is assigned by nature. We assume the existence of two buyer types, either Weak (δ_B^W) or Strong (δ_B^S), where $\delta_B^W < \delta_B^S$. The types are independently drawn from the same commonly known distribution. Potential suppliers hold the same prior belief concerning the buyer's type, described by a single number $x \in (0, 1)$. With probability x, the buyer is Strong; with probability $1 - x$, she is Weak. There is but one seller type δ_S, which is common knowledge.

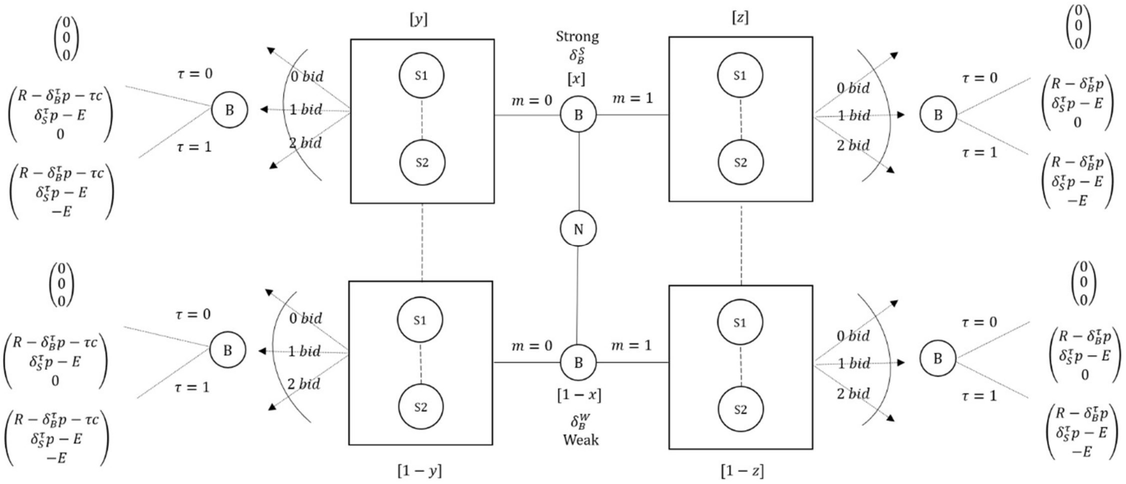
We incorporate a fixed penalty $c \geq 0$ to reneging on standard payment terms. We do not specify who accrues this cost, which remains outside the model. Potential suppliers find themselves at one of two information sets in Stage 2, each containing two nodes. After receiving the buyer's message, they update their belief about the buyer's type using Bayes' Rule. If the suppliers observe an offer of standard payment terms at Stage 1, then their posterior belief that the buyer is Strong is $y \in [0,1]$. If the potential suppliers observe an offer of extended payment terms, then they attach probability $(1 - z)$ to the event that the buyer is Weak. In Figure 4.1, we illustrate the extensive form game of incomplete information.[70]

---

[69] In this regard, we consider "Late" in a relative sense, i.e., relative to the reference date that would be implied by a standard payment term, which may be longer for a private entity than a public entity.

[70] Space constraints preclude proper representation of the entry and pricing continuation games.

Figure 4.1. Extensive-form representation of the model.



Notes: The first row in the payoff vectors corresponds to the buyer's profit, the second row corresponds to the seller's profit and the third row corresponds to the other supplier's profit. The buyer's valuation and reserve price at auction is R and the potential suppliers' entry cost is E. The buyer's discount factor is $\delta_B \in \{\delta_B^W, \delta_B^S\}$ and the seller's discount factor is δ_S. The buyer's *ex-ante* intended payment term is $m \in \{0,1\}$ and *ex-post* payment date is $\tau \in \{0,1\}$.

In the equilibrium analysis that follows, firms are risk-neutral and act to maximise expected profits. We discuss the qualitative implications of risk preferences when formulating the experimental hypotheses. Information about the distribution of types is common knowledge. The buyer's profit is $\pi_B = R - \delta_B^\tau p - max\ \{\tau - m, 0\} \bullet c$ and the seller's profit is $\pi_S = \delta_S^\tau p - E$.

4.2.2. Equilibrium analysis

The model can be analysed as an extensive-form game of incomplete information. The solution concept is weak PBE. Let q be the potential suppliers' belief at Stage 2 that payment is received late *ex-post*. If $q = 0$, potential suppliers expect to receive payment at date 0 with probability one; if $q = 1$, potential suppliers expect to receive payment at date 1 with probability one. We will find that there is a direct relationship between the beliefs about receipt of a late payment and the beliefs about a buyer's type. All proofs are relegated to the appendix.

A PBE of the signalling game is defined on the buyer's *ex-ante* intended payment term from Stage 1. If the two buyer types send the same message m, we have a pooling equilibrium on m (independent of the final *ex-post* payment date). If the two buyer types each send a different message m, we have a separating equilibrium. The probability q is equal to 0 or 1 in any equilibrium where the two buyer types choose the same *ex-post* payment date. There is an interior probability q in any pooling equilibrium where the two types choose different *ex-post* payment dates. We will find that no pooling equilibrium exists in which the Strong buyer pays late, and the Weak buyer on-time. Thus, the only interior probability of interest is $q = (1 - x)$.

To gain insight, we first examine the second stage as a standalone entry and pricing game. No Nash equilibrium exists in pure strategies at parameter values of interest. To understand this, note that for all p less than or equal to R, there exists an incentive for each supplier to undercut his competitor on price to increase own profit. Yet once the (discounted) price falls below the entry cost, each potential supplier would prefer not to bid for the contract. A unique symmetric equilibrium in mixed strategies does exist, however, consisting of an independent randomization by both players over their entry and bid decision for which there is no profitable unilateral deviation. This equilibrium is outlined in Lemma 4.1. The lemma uses Theorems 1 and 2 in Lang and Rosenthal (1991).

Lemma 4.1. *Suppose that the two potential suppliers have symmetric discount factors, with $\delta_S R > E$. There exists a unique symmetric Nash equilibrium of the entry and pricing stage characterised as follows: each supplier enters with probability $\mu(q)$ and submits a bid according to the cumulative distribution function (c.d.f.) $F(p,q)$, where*

$$\mu(q) = 1 - \frac{E}{(1-q+\delta_S q)R} \; ,$$

$$F(p,q) = \frac{1-\frac{E}{(1-q+\delta_S q)p}}{1-\frac{E}{(1-q+\delta_S q)R}} \quad \text{for} \quad p \in [\underline{p}, R], \underline{p} = \frac{E}{1-q+\delta_S q} \; .$$

The conditional price c.d.f. is

$$G(p,q) = \frac{\alpha_q(p)}{\beta_q};$$

Expected welfare is

$$W(p, q, m, \tau) = \beta_q \pi_B;$$

Where $\alpha_q(p) = 1 - \left(\frac{E}{(1-q+\delta_S q)p}\right)^2$ *and* $\beta_q = 1 - \left(\frac{E}{(1-q+\delta_S q)R}\right)^2$.

Lemma 4.1 states that in equilibrium, potential suppliers randomise second stage entry with the same fixed probability $\mu(q)$ and submit a bid according to the c.d.f. $F(p, q)$. Since equilibrium expected profits of potential suppliers net of the entry cost E are zero, expected total welfare $W(p, q, m, \tau)$ is determined by expected buyer surplus, given at least one entrant. The potential suppliers' symmetric equilibrium strategy depends on their expectation q about the buyer's *ex-post* payment date. Entry is decreasing in q. The c.d.f. of the winning bid conditional on at least one bid being submitted is $G(p, q)$, with density function $g(p, q)$. We define the expected price as $p_q = \int_{\underline{p}}^{R} pg(p, q)dp,$ which is increasing in q.[71]

We obtain the following corollary directly from Lemma 4.1.

Corollary 4.1. *Late payments are associated with higher prices and lower welfare.*

In an appendix, we generalise Lemma 4.1 to the asymmetric case in which potential suppliers possess different discount factors. The results are intuitive: the supplier with the highest discount factor (in absolute value terms) is most likely to enter into competition and earns economic rents in equilibrium. Since the comparative statics of entry and pricing are unchanged, we can restrict attention to the case of symmetric potential suppliers.

Now we consider the entry and pricing game at Stage 2 in context of the wider signalling structure. Each strategy profile of the three-stage game is an ordered list, $([A, B], [C, D]; [E, F], [G, H]; y, z)$, where A and B are the buyer's message and payment date if Strong, C and D are the buyer's message and payment date if Weak, E and F are the potential suppliers' entry and bidding decisions if $m = 0$ is observed, G and H are the potential suppliers' entry and bidding decisions if $m = 1$ is observed, and y and

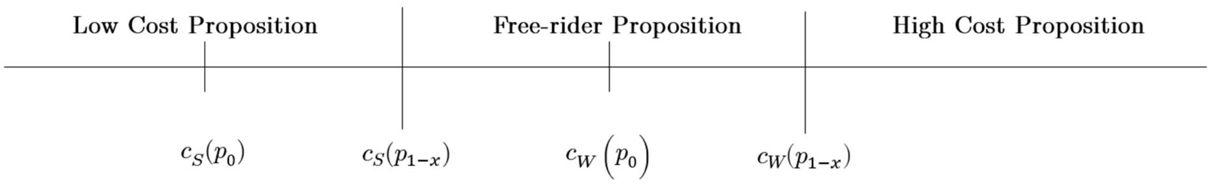
[71] The conditional c.d.f. is typically the observed distribution in naturally occurring markets.

z are the potential suppliers' posterior belief that the buyer is of Strong type at each information set.

In the continuation game associated with $m = 1$ (i.e., an extended payment term), the buyer has a dominant strategy to pay late in the final subgame and so $\tau = 1$. In the continuation game associated with $m = 0$ (i.e., a standard payment term), the buyer pays on-time in the final subgame if the cost from reneging on her intended payment term exceeds the benefit from a discounted late payment, that is if $c \geq (1 - \delta_B)p_q$. In that case, $\tau = 0$. We define the penalty at which this inequality binds in expectation for a Strong buyer by $c_S(p_q)$, which is an increasing function of the price. The threshold penalty required for a Weak buyer to pay on-time is $c_W(p_q)$, which is everywhere greater than for a Strong buyer.

To fully characterise equilibria in our institutional environment, we impose a *Sufficient Difference* restriction on the distribution of types (see the appendix for details). This restriction is without loss of generality. The propositions that we derive below pertain to three distinct penalty intervals, in relation to the buyer's type (Figure 4.2). The buyer's type determines which payment term she prefers to offer in Stage 1.

Figure 4.2. Penalty intervals and corresponding propositions.



Low Cost Proposition.

A. *For all* $c < c_S(p_0)$, *there exists a unique pooling PBE, independent of the type distribution, characterised as:* $([1,1], [1,1]; [\mu(1), F(p, 1)], [\mu(1), F(p, 1)]; y, x)$.

Expected welfare is $W^a = \beta_1 \bullet (R - (x\delta_B^S + (1 - x)\delta_B^W)p_1)$.

B. *For all* $c \in [c_S(p_0), c_S(p_{1-x}))$*:*

i. *If* $\delta_B^S > \frac{\beta_0 p_0 - R(\beta_0 - \beta_1)}{\beta_1 p_1}$*, then in addition to A. there exists a separating PBE characterized as:* $([0,0],[1,1];[\mu(0),F(p,0)],[\mu(1),F(p,1)];1,0)$.

Expected welfare is $W^b = x \bullet \beta_0 \bullet (R - p_0) + (1-x) \bullet \beta_1 \bullet (R - \delta_B^W p_1)$.

ii. *If* $\delta_B^S < \frac{\beta_0 p_0 - R(\beta_0 - \beta_1)}{\beta_1 p_1}$*, then we return to the unique pooling equilibrium in A.*

The Low Cost Proposition (*A.*) is intuitive: a buyer prefers to offer and pay within an extended payment term independent of her discount factor if the penalty is too low to (at least) deter the Strong type from reneging on an offer of a standard payment term at the lowest expected equilibrium price. No buyer type prefers to honour a standard payment term at any equilibrium price defined by q. Buyers therefore pool on extended payment terms in the first stage. At higher penalties, up to the threshold required to deter the Strong type from reneging on a standard payment term at an expected price of p_{1-x} (the expected price if only the Strong buyer pays on-time in the continuation game associated with a standard payment term), the equilibrium outcome depends on the distribution of types.

If the threshold condition on the Strong type in (*B.i.*) holds, then a separating equilibrium exists in which the Strong buyer offers and pays within a standard term and the Weak buyer offers and pays within an extended term. This case does not rule out the existence of a pooling equilibrium on an extended payment term for certain off-the-path posterior beliefs. For example, if we specify $y \in [0, x)$. Such a specification is objectionable, however, because it implies that a Weak buyer is more likely to deviate to a standard payment term than a Strong buyer despite having less to gain. In weaker buyer populations, the separating equilibrium cannot be reached at any cost below $c_S(p_{1-x})$ and we return to the unique pooling equilibrium in case A. Note that if the Strong buyer's expected benefit from paying late exceeds her expected gain from increased competition, then extended payment terms are optimal.

Corollary 4.2. $W^b > W^a$ *if and only if* $\delta_B^S > \frac{\beta_0 p_0 - R(\beta_0 - \beta_1)}{\beta_1 p_1}$.

We now consider the opposing end of the institutional spectrum.

High Cost Proposition. *For all* $c \geq c_W(p_{1-x})$*:*

i. *If* $\delta_B^W > \frac{\beta_0 p_0 - R(\beta_0-\beta_1)}{\beta_1 p_1}$*, then there exists a pooling PBE characterised as:* $([0,0],[0,0];[\mu(0),F(p,0)],[\mu(1),F(p,1)];x,z)$*. This is unique if* $c \geq c_W(p_1)$*.*

Expected welfare is $W^c = \beta_0 \bullet (R - p_0)$*.*

ii. *If* $\delta_B^W < \frac{\beta_0 p_0 - R(\beta_0-\beta_1)}{\beta_1 p_1} < \delta_B^S$*, then there exists a separating PBE characterised as:* $([0,0],[1,1];[\mu(0),F(p,0)],[\mu(1),F(p,1)];1,0)$*. This is unique if* $c \geq c_W(p_1)$*.*

Expected welfare is W^b*.*

iii. *If* $\delta_B^S < \frac{\beta_0 p_0 - R(\beta_0-\beta_1)}{\beta_1 p_1}$*, then there exists a unique pooling PBE characterised as:* $([1,1],[1,1];[\mu(1),F(p,1)],[\mu(1),F(p,1)];y,x)$*.*

Expected total welfare is W^a*.*

The **High Cost Proposition** (*i.*) implies that, if the distribution of buyer types is strong enough, the imposition of a penalty that can deter the Weak type from reneging on a standard payment term at an expected price of p_{1-x} is sufficient to fully deter late payments. In this situation, both types pay on-time and W^c is the first-best outcome. There is maximal equilibrium entry and price competition, without the welfare-damaging cost of reneging on a standard payment term being imposed. If the buyer types fall either side of the threshold discount factor in (*ii.*), a high penalty can only induce adherence to payment on-time by the Strong type. In that case, the separating equilibrium associated with W^b re-emerges, with an additional condition on the Weak type due to the deterrent effect of the higher penalty in this interval. This condition was redundant before. Even with a very high penalty, extended payment terms will obtain if the buyer population is weak enough (*iii.*).

Corollary 4.3. $W^c > W^b$ *if and only if* $\delta_B^W > \frac{\beta_0 p_0 - R(\beta_0-\beta_1)}{\beta_1 p_1}$*.*

In each equilibrium outcome considered so far, the *ex-post* payment date is consistent with the buyer's *ex-ante* payment term. Moreover, the buyer must be weak enough to prefer to make late payment at the associated (higher) equilibrium price. Our final proposition considers the intermediate penalty interval, $c \in [c_S(p_{1-x}), c_W(p_{1-x}))$. This proposition can help us to understand why, in the real-world, we observe financially strong companies making a late payment despite offering a standard payment term.

Free-rider Proposition. *For all* $c \in [c_S(p_{1-x}), c_W(p_{1-x}))$*:*

If $\delta_B^W > \frac{\beta_{1-x}p_{1-x} - R(\beta_{1-x}-\beta_1)}{\beta_1 p_1}$*, then there exists a pooling PBE characterized as:* $([0{,}0], [0{,}1]; [\mu(1-x), F(p, 1-x)], [\mu(1), F(p, 1)]; x, z)$.

Expected welfare is $W^d = \beta_{1-x} \bullet (R - (x + (1-x)\delta_B^W)p_{1-x} - (1-x)c)$.

The **Free-rider Proposition** states that, if the expected competition benefit from offering a standard payment term is large enough and the buyer population strong enough, then the Weak(er) type can free ride on the timely payment of the Strong(er) type and profitably renege on her own payment term, incurring the penalty in equilibrium. In such situations, financial strength per se does not preclude late payments. Instead, even those buyers who do not place a high value on making a late payment may exploit the competition externalities from other timely payers in the market. Note that as x approaches zero, the threshold condition for δ_B^W goes to one because there are fewer Strong buyers for the Weak type to free-ride on. If $c \geq c_W(p_0)$ and there is equilibrium potential for the Weak type to gain with maximal entry and price competition relative to an extended payment term, then a second pooling equilibrium exists on $m = 0$ associated with payment on-time by both types. In weaker buyer populations, **High Cost Proposition** (*ii.* and *iii.*) apply. For all $c < c_W(p_1)$, if we admit objectionable beliefs, then the *caveat* on multiplicity discussed with regard to the **Low Cost proposition** (*i.*) applies to the **Free-rider proposition** and to **High Cost Proposition** (*ii.* and *iii.*).

Corollary 4.4. *Welfare loss:*

i. $W^d \leq W^c$ *if and only if* $c \geq \frac{\beta_{1-x}(R-(x+(1-x)\delta_B^W)p_{1-x}) - \beta_0(R-p_0)}{\beta_{1-x}(1-x)}$;

ii. $W^d \leq W^b$ *if and only if* $c \geq \frac{\beta_{1-x}(R-(x(1-\delta_B^W)+\delta_B^W)p_{1-x}) - x\beta_0(R-p_0) - (1-x)\beta_1(R-\delta_B^W p_1)}{\beta_{1-x}(1-x)}$;

iii. $W^d \leq W^a$ *if and only if* $c \geq \frac{\beta_{1-x}(R-(x+(1-x)\delta_B^W)p_{1-x}) - \beta_1(R-(x\delta_B^S+(1-x)\delta_B^W)p_1)}{\beta_{1-x}(1-x)}$.

There is a welfare loss associated with W^d if the cost of reneging on a standard payment term is set too high. This loss can occur relative to the first-best pooling equilibrium in which both types offer a standard payment term and pay on-time, that is the outcome associated with W^c, the separating equilibrium associated with W^b, and/or the pooling equilibrium on an extended payment term associated with W^a. The conditions in Corollary 4.4 are ordered, i.e., the condition on the penalty in *(iii)* is sufficient for (*ii*) and (*i*) to hold, while the condition in (*ii*) is sufficient for (*i*) to hold. A welfare loss is more likely to arise if Weak buyers are relatively more frequent in the population (i.e., as x falls).[72] This corollary suggests that uniformly increasing the size of a fine imposed on buyers for a late payment may not necessarily lead to a welfare gain. Thus, care should be taken in setting the level of any such fine.

4.3. Experimental Design

The experimental environment considers the case of two potential suppliers and augments earlier Bertrand competition experiments with costless entry (e.g., Dufwenberg and Gneezy 2000). Subjects make decisions in either a buyer or seller role, which is determined randomly at the beginning of the experiment and remains fixed throughout. Then, subjects are matched into groups of 1 buyer and 2 sellers. In our experiment, seller eligibility to customise the unit substitutes for entry. The buyer may purchase a single unit of a customised good from one of the two sellers. The buyer's valuation for the unit is 100. The unit is to be sold at an auction. One or both of the sellers may not be eligible to customise the unit. Thus, a transaction is not guaranteed. If a transaction takes place, the auction determines which seller delivers the unit and the transaction price.

Subjects with a seller role are assigned the common discount factor $\delta_S = 0.50$ for the entire experiment. Subjects with a buyer role are each randomly assigned a discount factor with equal probability from the set $\delta_B \in \{0.75, 0.95\}$. The two elements in this set correspond to the Weak and Strong buyer types, respectively. We denote these as types 1 and 2 in the experiment. The buyer's type remains private information and is drawn anew at the beginning of each period. We specify a discount factor for the Strong

[72] For example, with our experiment parameters, $W^d \leq W^c$ for all x; $W^d < W^b$ for any $x \leq 0.70$; and $W^d < W^a$ for any $x \leq 0.10$.

type strictly below one to ensure that the buyer has strict preference in the final payment decision. The sellers' discount factor and the distribution of buyer discount factors is common knowledge. Thus, the sellers' prior belief that their matched Buyer is of Strong type is one-half.[73] Each period of the experiment consists of three stages: a *Pre-Auction stage*, an *Auction stage*, and a *Post-Auction stage*.

In the *Pre-Auction stage*, the buyer learns her type for the period and sends a message to the two sellers in her group. This message signals her intended payment method for the unit, either "upon delivery", or "after delivery". These payment methods substitute for an offer of payment terms (either standard or extended). When describing the experiment, we use the language of the laboratory.

In the *Auction stage*, the sellers observe the buyer's message. Each seller submits a bid for which they would be willing to deliver the unit. This bid can be any integer from 25 to 100 (the reserve price). To give the theory its best shot, we control for possible seller over-entry by using a variant on the strategy method. Subjects do not choose whether to enter the auction. Instead, the computer determines whether each seller is eligible to customise the unit, according to a pre-determined probability. This probability is set equal to the equilibrium mixed strategy entry randomization, given correct beliefs (more on this later). Sellers are informed about their eligibility probability at the time of submitting a bid. Sellers are not informed about their own nor their competitor's eligibility realization until after submitting a bid.

Any seller that is determined by the computer not to be eligible can return his non-customised unit to the experimenter in exchange for 25 points (the entry cost). After both bids in a group have been submitted, the eligibility randomization is performed by the computer, independently for each seller. If the computer determines that both sellers are eligible, then the buyer purchases the unit from the seller that submitted the lower bid in the auction. Ties are broken at random. If the computer determines that only one seller is eligible, then the buyer purchases the unit from the eligible seller. The transaction price is the selected seller's bid.

[73] S*ufficient Difference* is satisfied: $\delta_B^S \geq 0.4 + 0.6\delta_B^W$.

In the *Post-Auction stage*, the buyer observes the price (but not the identity) of the selected seller and selects a final payment method for the unit. If final payment is made upon delivery, then both the buyer and seller value payment at the full transaction price. If final payment is made after delivery, then the seller values payment at half of the transaction price; a Weak buyer at 75% of the transaction price; and a Strong buyer at 95% of the transaction price.[74] If the buyer reneges on an intention to make payment upon delivery, then a penalty is levied on the buyer. Subjects are informed that the penalty accrues to an independent third party outside of the experiment.

Based on our model, we design three treatments which vary the size of this penalty using a between-subjects design. In the first treatment, which we call *Low-Cost* (*LC*), the penalty is set equal to 1. The penalty in this treatment is insufficient to deter either buyer type from reneging on an intention to make payment upon delivery and so the theory predicts a pooling equilibrium on intended payment after delivery followed by final payment after delivery by both types.

In the second treatment, which we call *High-Cost* (*HC*), the penalty is set equal to 16. Given the distribution of types, the penalty in this treatment is sufficient to fully deter reneging on an intention to pay upon delivery among buyers and so the theory predicts a pooling equilibrium on intended payment upon delivery followed by final payment upon delivery by both types.

In the third treatment, which we call *Free-Rider* (*FR*), the penalty is set equal to 8. This penalty is sufficient to deter only a Strong buyer from reneging on an intention to pay upon delivery and so the theory predicts the existence of a pooling equilibrium on intended payment upon delivery associated with divergence in payment methods during the final stage, with the Weak type free-riding on the Strong type.

We implement the strategy method entry randomization for our experiment parameters both on and off the equilibrium path (cf. Lemma 4.1). In *LC*, if the buyer signals

[74] To avoid any negative connotations, we do not refer to payment as "late" in the experiment. We make clear in the instructions that a buyer's intended payment method may not correspond with her final payment method: "*The final payment method selected can differ from the buyer's intended payment method*". Subjects must also answer a comprehension question which is related to this.

an intention to make payment upon delivery, then the computer is programmed to randomise seller eligibility with probability one-half. In *HC*, if the buyer signals an intention to make payment upon delivery, then the computer is programmed to randomise seller eligibility with probability three-quarters. In *FR*, if the buyer signals an intention to make payment upon delivery, then the computer is programmed to randomise seller eligibility with probability two-thirds. In all three treatments, if the buyer signals an intention to make payment after delivery, then the computer is programmed to randomise seller eligibility with probability one-half.

If a transaction occurs, then the buyer earns 100 minus her payment value and minus the penalty if applicable; the seller earns a profit equal to her payment value. Any seller who is determined by the computer not to be eligible earns a profit of 25. An eligible seller who submits the higher bid earns zero profit for that period. If neither seller is determined by the computer to be eligible, the buyer does not purchase a unit and so earns zero profit for the period.

4.3.1. Hypotheses

The preceding theory and design offer several testable hypotheses, which we summarise below. The point predictions are presented in Table 4.1.

Hypothesis 4.1 . *(i) Buyers will signal an intention to pay after delivery more often in LC than in HC or FR; and (ii) There will be no difference in intended payment decisions between Weak and Strong buyers in any treatment.*

The first part of Hypothesis 4.1 reflects that the buyer's cost of reneging on a standard payment term is too low in the *LC* treatment and so this signal is not credible. It is possible, however, that buyers will attempt to convince sellers of a (false) intention to pay upon delivery. The second part of Hypothesis 4.1 captures the nature of the predicted pooling equilibrium.

Table 4.1 – Theoretical predictions by treatment.

Treatment	*LC*	*HC*	*FR*
Intended payment	(AD, AD)	(UD, UD)	
Final payment	(AD, AD)	(UD, UD)	(UD, AD)
Minimum bid	50	25	33.33
Average price	66.67	40	50
Buyer profit	(27.5, 37.5)	(56.25, 56.25)	(44.44, 48.44)
Seller profit		25	

Notes: The first two rows indicate the buyer's intended and final payment decision for each type (Strong, Weak), where *UD = Upon Delivery* and *AD = After Delivery*. The third row indicates the minimum bid in the sellers' equilibrium support. The fourth row indicates the expected transaction price conditional on at least one seller being eligible to customise the unit. The fifth row indicates the expected buyer profit for each type. The final row indicates the expected seller profit gross of the entry cost.

Hypothesis 4.2. *(i) Buyers will make final payment after delivery more often in LC than in HC; and (ii) Weak buyers will make final payment after delivery more often than Strong buyers in FR.*

The first part of Hypothesis 4.2 states that, at low penalties for reneging on a standard payment term, the benefit of paying late dominates the associated cost for both buyer types, and vice-versa at high penalties. Of course, different behavioural norms could emerge. Prior studies suggest that certain individuals prefer to tell the truth (e.g., Charness and Dufwenberg 2006), which might manifest itself in buyers keeping to a standard payment term even when it is not in their self-interest. The second part of Hypothesis 4.2 states that if the penalty for reneging on a standard payment term is not set carefully, then Weak buyers will have an incentive to free ride on their Strong counterparts.

Hypothesis 4.3. *(i) Bids will be independent of the buyer's signal in LC, but higher after receiving an intention to pay after delivery in FR and HC; (ii) Transaction prices will be rank ordered such that $p_{LC} > p_{FR} > p_{HC}$; (iii) Buyer profit will be rank ordered such that $\pi_{B,HC} > \pi_{B,FR} > \pi_{B,L}$, there will be no difference in seller profit among treatments.*

This last hypothesis captures the negative indirect effect of late payments on consumer prices and welfare (i.e., expected buyer profit). In *LC*, the bid support is [50, 100] and the expected price is 66.67. Strong buyers earn 27.5 and Weak buyers earn 37.5. In *HC*, the bid support is [25, 100], sellers adjust their bids upwards in response to an offer of an extended payment term and the expected price is 40. Both buyer types earn 56.25. In *FR*, the bid support is [33, 100], sellers adjust their bids upwards in response to an offer of an extended payment term and the expected price is 50. Strong buyers earn 44.44 and Weak buyers earn 48.44. In all treatments, the expected seller profit gross of the entry cost is 25.

4.3.2. Procedure

The experiments were conducted online in the fall of 2020, using oTree and Zoom (Chen et al. 2016, Zhao et al. 2020). Unlike in a usual online experiment – and just as in a typical lab experiment – subjects attended sessions at a pre-specified date and time. All subjects were recruited from a lab database of predominantly graduate and undergraduate students at a large public university in the Southwestern United States and were randomly assigned to one of the three treatments. We recruited four independent cohorts per treatment. In each treatment, two cohorts repeated the interaction for 20 periods and the other two cohorts for 30 periods; in each cohort, we assigned 3 subjects to the role of a buyer and 6 subjects to the role of a seller.[75] Within a cohort, subjects were randomly matched into new groups each period. There were two cohorts of the same treatment in a session and the cohort size was not revealed to subjects. This protocol mitigated the possibility of tacit collusion (see Katok 2011).[76]

Between periods, feedback was provided on the buyer's pre-auction message, eligible bids, the buyer's final payment method and any penalty levied. All subjects also saw

[75] The experimental sessions with 20 rounds were conducted first; after confirming that these finished within an acceptable time in the online format, we extended the second sessions to 30 rounds. Two cohorts of *HC* had only 2 buyers and 4 sellers due to a lower than expected show-up rate on the day.

[76] A compendium of experimental instructions, together with a discussion of the challenges confronted in the online implementation and how we dealt with these, is contained in the appendix.

their own profit for the period. To minimise possible wealth effects, cumulative earnings were not displayed during the session. The profits of other players in a trio were also not displayed, to avoid making fairness concerns salient. Feedback was about a subject's own matching group in the period and not the wider cohort. This design further minimised the possibility of tacit collusion (see also Bruttel 2009, Dufwenberg and Gneezy 2002). As the game progressed, a summary of this information remained available to each subject privately in an on-screen history table, to reinforce the game-theoretic assumption of "perfect recall". Earnings from the experiment were converted to US Dollars at an exchange rate of 80 (50) points per dollar in sessions of 30 (20) periods and were paid in addition to a show-up fee of $5. Sessions lasted approximately 75 minutes. Average earnings were $20.15 for buyers and $15.17 for sellers.

Additionally, before moving the experiment online due to the Covid-19 shutdown, we collected data for two further cohorts of each treatment in a physical lab at a public university in the North East of England. To avoid mixing subject pools and experiment environments, we only present data collected from the online experiments below. A statistical analysis of the pooled data with six independent observations per treatment is contained in the appendix. Behaviour in the lab and online is statistically similar and the results reported below are qualitatively unchanged between the two samples.[77]

[77] For the pooled data, non-parametric tests support the theory in 22 out of 25 comparisons.

4.4. Experimental Results

In this section, we begin by examining summary statistics for the three experimental treatments. Since there is no interaction between subjects playing in different cohorts, each cohort is considered a statistically independent observation in the construction of treatment averages and standard deviations. We then conduct a formal test of our hypotheses using regression analysis.[78]

4.4.1. Summary Statistics

Table 4.2 contains aggregate data on observed payment decisions in the experiment and the relative frequencies with which the penalty for reneging on an intention to pay upon delivery was imposed on the buyer. We disaggregate this data by the buyer's type to examine how varying her value for making payment late influenced behaviour.

Table 4.2 – Payment decisions in the experiment (0 = UD; 1 = AD).

Treatment	*LC*		*HC*		*FR*		
			$\boldsymbol{H4.1(i)}$ (LC–HC–FR)				
Intended payment	0.62	(0.15)	0.39	(0.11)	0.32	(0.12)	
Weak Buyer	0.68	(0.15)	0.48	(0.31)	0.37	(0.18)	$\boldsymbol{H4.1(ii)}$
Strong Buyer	0.52	(0.24)	0.30	(0.13)	0.26	(0.06)	
			$\boldsymbol{H4.2(i)}$ (LC–HC)				
Final payment	0.89	(0.17)	0.36	(0.10)	0.50	(0.16)	
Weak Buyer	0.89	(0.22)	0.49	(0.34)	0.74	(0.27)	$\boldsymbol{H4.2(ii)}$
Strong Buyer	0.88	(0.12)	0.26	(0.14)	0.22	(0.07)	

Notes: Mean (SD) values based on 4 independent cohorts per treatment. Horizontal braces correspond to the two-sample comparisons of interest and associated hypotheses. Vertical braces correspond to the one-sample comparisons of interest and associated hypotheses. The results of non-parametric hypothesis tests for the pooled online and lab samples are contained in the appendix. *UD = Upon Delivery, AD = After Delivery.*

[78] All regression results reported in this section are robust to including the full set of demographic control variables elicited in the post-experiment questionnaire (see the appendix for details). The theory is not predicated on these factors and so we exclude them from the main analysis.

The theory predicts that both types signal an intention to pay after delivery in the *LC* treatment and upon delivery in the *HC* and *FR* treatments. In the experiment, the probability that a buyer signals an intention to pay after delivery is higher in *LC* (62%) than in *HC* (39%) or *FR* (32%), which supports Hypothesis 4.1(i). These differences are observed for both Weak and Strong buyers, which supports Hypothesis 4.1(ii). Nevertheless, we do observe a substantial fraction of buyers signalling an intention to pay upon delivery in *LC*. As a consequence, the number of eligible sellers is lower than expected in *HC* (actual 1.34 vs. expected 1.50) and *FR* (actual 1.21 vs. 1.33). We will discuss this in more detail below.

The aggregate data also offers qualitative support for Hypothesis 4.2. Buyers in *LC* are the most likely to make final payment after delivery (89%) and this does not vary with the buyer's type. Only 36% of buyers in *HC* and 50% of buyers in *FR* make final payment after delivery. Note that the latter probability would be expected across all buyers in *FR* if Weak buyers always free ride on their Strong counterparts. While this masks some variation between types, Weak buyers (74%) clearly pay after delivery more often than Strong buyers (22%). This is reflected in the penalty data: in *FR*, a penalty is imposed on 3% of transactions involving a Strong buyer and 42% of those involving a Weak buyer. The imposition of penalties is negligible in *HC*; 28% of transactions incur a penalty in *LC*.

In Table 4.3, we present aggregate data on bids, transaction prices and buyer and seller profits. To examine whether sellers anticipated that the signal would influence the buyer's final payment decision, bids are split by the buyer's intended payment decision. Consistent with Hypothesis 4.3(i), sellers appear only to respond to a credible signal. Whereas in *LC* sellers do not adjust their bids conditional on the intended payment decision, bids are higher in both *FR* and *HC* when the buyer signals an intention to pay after delivery. The symmetric mixed strategy equilibrium implies that bids will be randomised over a higher support in *LC* than in *FR,* and in *FR* than in *HC*, respectively. Similar qualitative differences are predicted for average transaction prices. We observe average bids of 67.54 in *LC*, 60.83 in *FR* and 54.94 in *HC*. The transaction prices conditional on entry are slightly lower (63.38, 50.19 and 56.83) but the ranking predicted by Hypothesis 4.3(ii) remains intact. Nevertheless, whereas in *LC* transaction prices are close to the point prediction, in both *HC* and *FR* they are higher than expected.

Table 4.3 – Bidding, prices and profits in the experiment.

Treatment	*LC*		*HC*		*FR*		
Bid							
Intention UD	66.76	(11.80)	52.46	(7.25)	57.55	(8.79)	}
Intention AD	67.82	(10.68)	61.42	(9.24)	68.23	(7.03)	} ***H*4.3(*i*)**
			***H*4.3(*ii*)**				
Price	63.38	(13.12)	50.19	(6.17)	56.83	(8.69)	
			***H*4.3(*iii*)**				
Buyer profit	34.70	(5.92)	47.52	(5.83)	39.30	(4.72)	
Weak Buyer	39.27	(6.20)	49.64	(6.96)	42.66	(6.39)	
Strong Buyer	28.78	(7.26)	45.55	(4.85)	36.08	(5.37)	
Seller profit	25.30	(2.12)	26.04	(2.73)	27.28	(2.83)	

Notes: Mean (SD) values based on 4 independent cohorts per treatment. Horizontal braces correspond to the two-sample comparisons of interest and associated hypotheses. Vertical braces correspond to the one-sample comparisons of interest and associated hypotheses. The results of non-parametric hypothesis tests for the pooled online and lab samples are contained in the appendix. *UD = Upon Delivery, AD = After Delivery.*

There is qualitative support for the ranking in buyer profits from Hypothesis 4.3(iii). On average, buyers earn the most in *HC* (47.52), the second-most in *FR* (39.30) and the least in *LC* (32.50). Seller profits are close to the entry cost of 25 in all three treatments. In *HC*, both buyer types pay on-time in equilibrium and so we expect no difference in profits. Weak buyers, however, earn a premium in this treatment of around four points (49.64 versus 45.55). Weak buyers also earn more than Strong buyers in *FR* (42.66 versus 36.08) which is similar to the differential predicted by the theory, albeit at a slightly lower absolute profit level. Weak buyers are predicted to earn 10 more points than Strong buyers in *LC* and this is broadly in line with the data (39.27 versus 28.78).

In the remainder of this section, we will conduct a formal test of our hypotheses using regression analysis and consider the observed buyer and seller strategies in more detail. Unless otherwise stated, reported *p*-values are based on two-sided Wald tests.

4.4.2. Buyer payment decisions

In Table 4.4, we examine determinants of the buyer's intended and final payment decisions in the experiment (0 = Upon Delivery, 1 = After Delivery). We also consider a specification in which the dependent variable is a dummy for whether the penalty was imposed in the transaction. The regressions are estimated using Logistic regression and include treatment dummies as regressors (the reference treatment is *LC*) and – in the second and third columns – their interaction with the buyer's type. We include cohort fixed effects and a linear time trend. Marginal effects are presented next to the coefficient estimate, with robust standard errors based on clustering at the subject level.

Result 4.1. *Buyers are most likely to signal an intention to pay after delivery in LC. Weak buyers are more likely to signal an intention to pay after delivery than Strong buyers in all treatments.*

Support for the first part of Result 4.1 is provided in column 1 of Table 4.4. Buyers in *FR* and *HC* are, respectively, 30% and 22% less likely to signal an intention to pay after delivery than buyers in *LC* (*p*-values = 0.014 and 0.071).[79] There is no significant difference in signals between buyers in *HC* and *FR* (*p*-value = 0.307). The second part of Result 4.1 applies to all treatments and is evident from the positive coefficient estimate effect on the Weak type dummy in specification 1.[80] Specifically, Weak buyers are 16% more likely to signal payment after delivery in the Pre-Auction stage than Strong buyers (*p*-value = 0.041). In summary, our data is in line with Hypothesis 4.1(i), but we reject Hypothesis 4.1(ii) in favour of an alternative hypothesis in which Weak buyers are more likely to signal extended payment terms up front.

---

[79] The effect versus *HC* is not robust to including individual-level control variables (see appendix).
[80] In an unreported regression, we find no significant effect on the buyer's intended payment decision of the interaction between Weak and *HC* (*p*-value = 0.933) or between Weak and *FR* (*p*-value = 0.843).

Table 4.4 – Logit regression analysis of payment decisions (0 = UD; 1 = AD).

Dependent variable	*Intended payment*		*Final payment*		*Penalty = 1*	
	$\hat{\beta}$	dy/dx	$\hat{\beta}$	dy/dx	$\hat{\beta}$	dy/dx
	(1)		(2)		(3)	
HC	-0.93*	-0.22*	-3.25***	-0.67***	-4.48***	-0.28***
	(0.52)	(0.12)	(0.95)	(0.13)	(1.07)	(0.08)
FR	-1.29**	-0.30**	-3.38***	-0.69***	-2.97***	-0.20***
	(0.53)	(0.12)	(0.89)	(0.12)	(1.08)	(0.07)
Weak	0.68**	0.16**	0.52	0.12	-0.52	-0.04
	(0.33)	(0.08)	(0.63)	(0.14)	(0.47)	(0.04)
*HC * Weak*			0.50	0.11	1.89	0.26
			(1.04)	(0.22)	(1.21)	(0.24)
*FR * Weak*			1.94**	0.35**	3.93***	0.65***
			(0.87)	(0.14)	(1.24)	(0.21)
Period	-0.02	-0.004	0.003	0.001	0.02	0.002
	(0.02)	(0.004)	(0.02)	(0.004)	(0.01)	(0.001)
Constant	0.17		2.31***		-0.28	
	(0.71)		(0.75)		(0.89)	
Observations	850		700		700	
Cohort fixed effects	Yes		Yes		Yes	
Log-likelihood	-533.40		-336.68		-251.27	
Wald test statistic	98.29***		274.96***		189.63***	

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates on the logit scale ($\hat{\beta}$) and average marginal effects (dy/dx), with robust standard errors in parentheses clustered at the subject level. The omitted experimental treatment is *LC*. The models are estimated using Logistic regression. The Wald test statistic is based on a Chi-squared test of joint regression significance. *UD = Upon Delivery, AD = After Delivery.*

Result 4.2. *Buyers are less likely to pay after delivery in HC than in LC regardless of the buyer's type. Weak buyers are more likely to pay after delivery in FR than Strong buyers and to incur the penalty for reneging on an intention to pay upon delivery.*

Evidence to support Result 4.2 is contained in columns 2 and 3 of Table 4.4. The probability of a late payment is 67% lower in *HC* than in *LC* (p-value < 0.01) and, as predicted by the theory, there is no significant effect of the buyer's type in this treatment on the final payment decision (whether main or interaction effect). By contrast in *FR*, Weak buyers are significantly more likely to pay late (p-value $= 0.026$) and incur the penalty (p-value < 0.01). Similar inferences are made for Weak buyers in *FR* versus *HC*

(p-values = 0.020 and < 0.01). Together, these findings are consistent with Hypothesis 4.2. There is no evidence to suggest the existence of a time trend in payment behaviours and the results are qualitatively unchanged if we split the regressions between the first and second halves of the experiment (see estimates in the appendix).

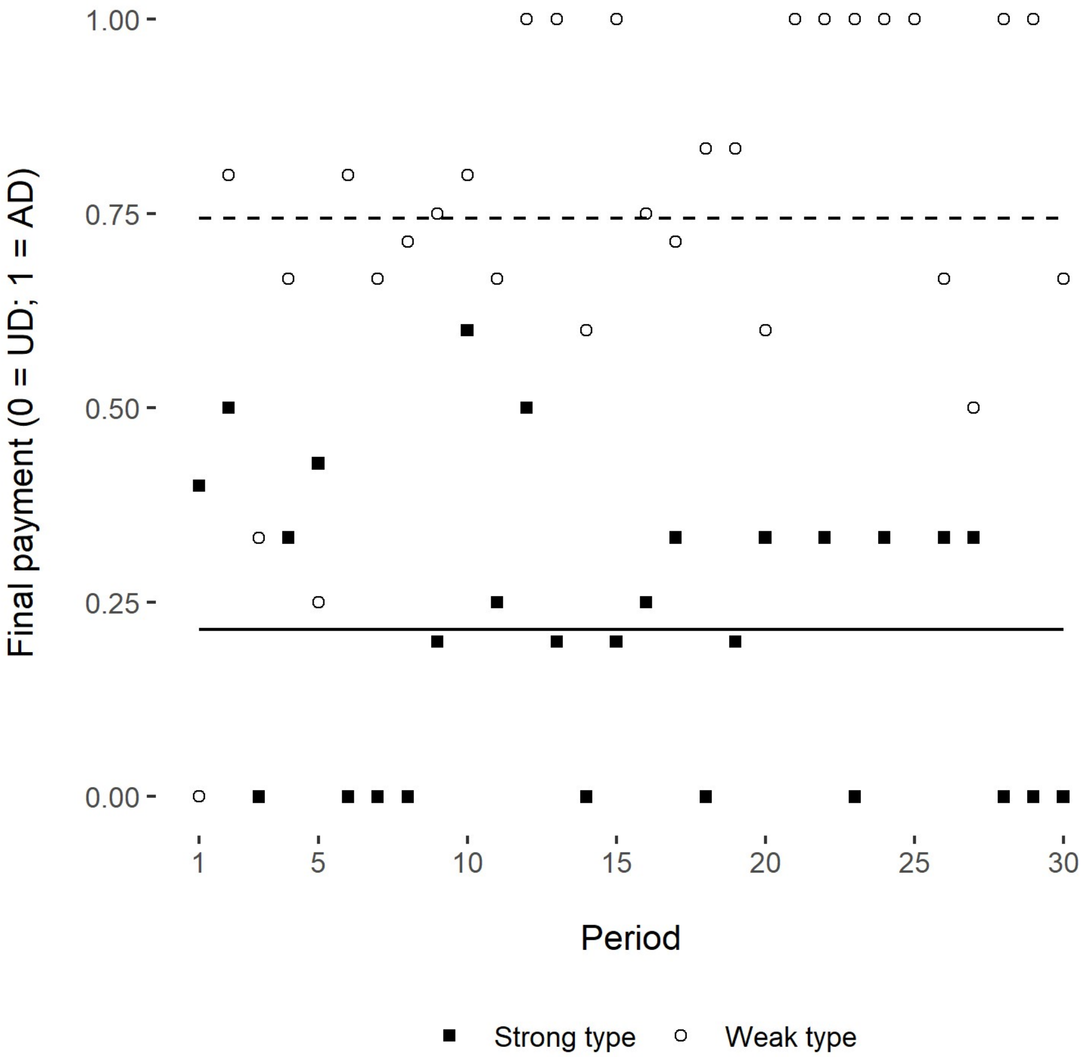
In Figure 4.3, we plot the relative frequency of final payments after delivery in *FR* over time and according to the buyer's type. The figure provides strong interocular support for the hypothesised free-riding by Weak buyers on price spillover effects in the market: there is a clear separation between Strong types who predominantly pay upon delivery and Weak types who predominantly pay after delivery. This separation persists over time. Given that both types more often signal an intention to pay upon delivery, this trend is associated with Weak types reneging on a significant fraction of their payment terms offered.

Results 1 and 2 are robust to issues of learning that might arise because the buyer's type is redrawn in each period (in the appendix, we analyse payment decisions for the first and second halves of the experiment separately). A robust behavioural deviation from the theory, however, is that buyers are more likely to incur the penalty in *LC* than in *HC* or *FR* (p-value < 0.01 for both comparisons). This observation applies in particular to Strong buyers, 48% of whom signal payment after delivery and 36% of whom incur the penalty in *LC*. To investigate off-equilibrium buyer strategies, in Table 4.5 we present relative frequencies for the vector of payment decisions observed in each treatment.[81] Frequencies for the Weak type are in square brackets and the equilibrium strategy for each type is in boldface. We include the percentage of markets that failed to produce at least one eligible seller, which we denote as an "unravel". As expected, the rate of unravelling decreases as the size of the penalty increases.[82]

[81] The off-equilibrium payment behaviours discussed here are not simply a facet of learning: the qualitative insights are unchanged if we restrict data to the last 15 periods only.

[82] In *LC*, the type randomization produced a slight imbalance between Weak and Strong types. Nevertheless, we fail to reject the null hypothesis that the probability of observing each type in our sample is 50% (p-value = 0.181, two-sided binomial test). For *LC*, "unravel" is a check of our eligibility randomization (expected 25%); for the other treatments, it depends on the mix of buyer's intended payment decisions (equilibrium-expected 6.25% in *HC* and 11.1% in *FR*).

Figure 4.3. Free-riding behaviour in the FR treatment.



Notes: *UD = Upon Delivery, AD = After Delivery*. The solid (dashed) line is the frequency of final payments after delivery in the FR treatment among Strong (Weak) buyer types.

Table 4.5 – Contingency table of buyer payment strategies for Strong [Weak] types by treatment. Equilibrium prediction in bold.

Treatment / Strategy	*LC*		*HC*		*FR*	
(UD, UD)	11	[10]	**90**	**[55]**	**89**	[30]
(UD, AD)	37	[36]	1	[3]	3	**[57]**
(AD, UD)	0	[0]	0	[0]	2	[1]
(AD, AD)	**49**	**[85]**	29	[45]	23	[44]
Unravel	24%		10.8%		17%	

Notes: The first (second) entry in the strategy indicates the intended (final) payment date. *UD = Upon Delivery, AD = After Delivery*; Unravel is the percentage of markets that failed to produce at least one eligible seller.

The equilibrium strategy is the modal strategy for each type in every treatment. In *LC*, the alternative strategy (UD, AD) is also popular among both Weak and in particular Strong types, suggesting that buyers recognised the possibility of influencing seller prices via their choice of payment terms *ex-ante* before the auction (recall that in this treatment, the buyer's choice of signal has no effect on the sellers' eligibility probability). The draw of circumventing a larger penalty in *FR* and *HC* via the signal is also evident: the second-most popular strategy for both types in these treatments is (AD, AD). The third-most popular strategy for Weak types in *FR* is (UD, UD). Conditional on the price, this strategy could be an act of self-interest or indicate a preference for truth-telling among certain subjects. We find some evidence to suggest the latter: 45 out of 52 instances in which buyers deviated from self-interest in their final payment decision were to honour an intention to pay on-time, which would be consistent with a notion of guilt aversion (e.g., Charness and Dufwenberg 2006).

A natural question is the extent to which the payment strategy of a buyer is consistent during the experiment, or whether there is evidence of mixing between strategies. In the appendix, we present within-buyer relative choice frequencies observed across periods, conditional on the assigned type. We also provide a weighted scatterplot of the strategy coordinates for each individual buyer in our experiment. These plots show largest concentrations at the equilibrium coordinates, but with some buyers displaying off-equilibrium behaviours in line with observations of the previous paragraph.

4.4.3. Seller bids, transaction prices and profits

We next analyse sellers' bidding behaviour in the Auction stage of the experiment. In Table 4.6 (columns 1 and 2), we regress seller bids and transaction prices on the treatment dummies (the reference treatment is *LC*) and their interaction with the buyer's intended payment decision (recall that the buyer's underlying type is unobserved by sellers). We again incorporate a linear time trend and cohort fixed effects. All regressions are estimated using OLS, with robust standard errors calculated based on clustering at the subject level.[83]

Result 4.3a. *Sellers mark up their bids in response to extended payment terms when a deterrent penalty is in place.*

Consistent with Hypothesis 4.3(i), sellers in *HC* and *FR* submit significantly higher bids after observing an intention to pay after than upon delivery (p-value < 0.01 for both interaction terms) and the intended payment decision has no significant impact on bids in *LC*. Thus, buyers' attempts in the *LC* treatment to influence bids through the signal are (on average) unsuccessful. This result provides empirical evidence that the use of extended payment terms may inflate consumer prices even when the prevailing institutional environment penalises buyers for reneging on a standard payment term.

Result 4.3b. *Bids and transaction prices are lower when the penalty for reneging on a standard payment term is higher.*

This result is line with Hypothesis 4.3(ii) and validates our conjecture that the anticipation of a late payment – due to the absence of deterrent regulation – feeds into higher consumer prices. Conditional on the equilibrium signal, we fail to reject our hypothesis that the average transaction price equals 40 in *HC*, 50 in *FR* and 66.67 in *LC* (p-values = 0.647, 0.828 and 0.132). Bids and transaction prices are significantly lower

[83] The results are robust to using a random effects panel estimator (see appendix).

in *HC* and in FR than in *LC* (p-value < 0.01 for all four comparisons). These measures are also significantly lower in *HC* than in *FR* (p-values = 0.039 and 0.032). There is a significant trend upwards in bids and prices over time in all three treatments. This is not captured by the static equilibrium theory and suggests some learning by sellers early on. Nevertheless, the comparative statics remain unchanged if we split the regressions between the first and second halves of the experiment (details in the appendix).

Table 4.6 – OLS regression analysis of bids, transaction prices and profits.

Dependent variable	*Bid*	*Price*	*Buyer profit*	*Seller profit*
	(1)	(2)	(3)	(4)
HC	-17.85***	-20.15***	20.42***	1.35
	(3.69)	(4.21)	(3.38)	(1.61)
FR	-10.42***	-12.41***	10.88***	3.13*
	(3.34)	(3.91)	(2.43)	(1.90)
Intention AD	-0.35	-2.18	-1.64	-3.92***
	(1.93)	(3.27)	(3.39)	(1.21)
*HC * Intention AD*	11.95***	14.68***	-13.37***	0.78
	(4.06)	(4.55)	(4.04)	(2.16)
*FR * Intention AD*	11.99***	17.71***	-14.19***	3.27*
	(4.25)	(4.19)	(3.58)	(1.68)
Weak			9.40***	0.51
			(2.48)	(1.12)
*HC * Weak*			1.35	-3.74**
			(1.79)	(1.82)
*FR * Weak*			-2.54	-6.81***
			(3.82)	(2.18)
Period	0.68***	0.57***	-0.35***	0.07
	(0.10)	(0.10)	(0.11)	(0.05)
Constant	62.81***	61.54***	30.61***	26.81***
	(2.80)	(3.41)	(2.36)	(1.38)
Observations	1,650	682	850	1650
Cohort fixed effects	Yes	Yes	Yes	Yes
R-squared	0.22	0.27	0.13	0.03
Wald test statistic	50.02***	27.24***	10.86***	4.03***

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates are presented, with robust standard errors in parentheses clustered at the subject level. The omitted experimental treatment is *LC*. The models are estimated using OLS regression. The Wald test statistic is based on an F-test of joint regression significance. *AD = After Delivery*.

To assess the predictive power of the unique symmetric equilibrium in mixed strategies of the Auction stage, in Figure 4.4 we compare the observed bid c.d.f. with the predicted c.d.f. for each treatment and conditioned on the buyer's signal. Each row in the figure corresponds to one of the three experimental treatments; the left-hand panels correspond to a signal of payment upon delivery, and the right-hand panels to a signal of payment after delivery. All tests for equality of probability distributions are based on two-sided Kolmogorov-Smirnov tests of the empirical bid c.d.f. versus that expected from Lemma 4.1.[84] We find the evidence in support of the mixed strategy equilibrium to be (itself) mixed.

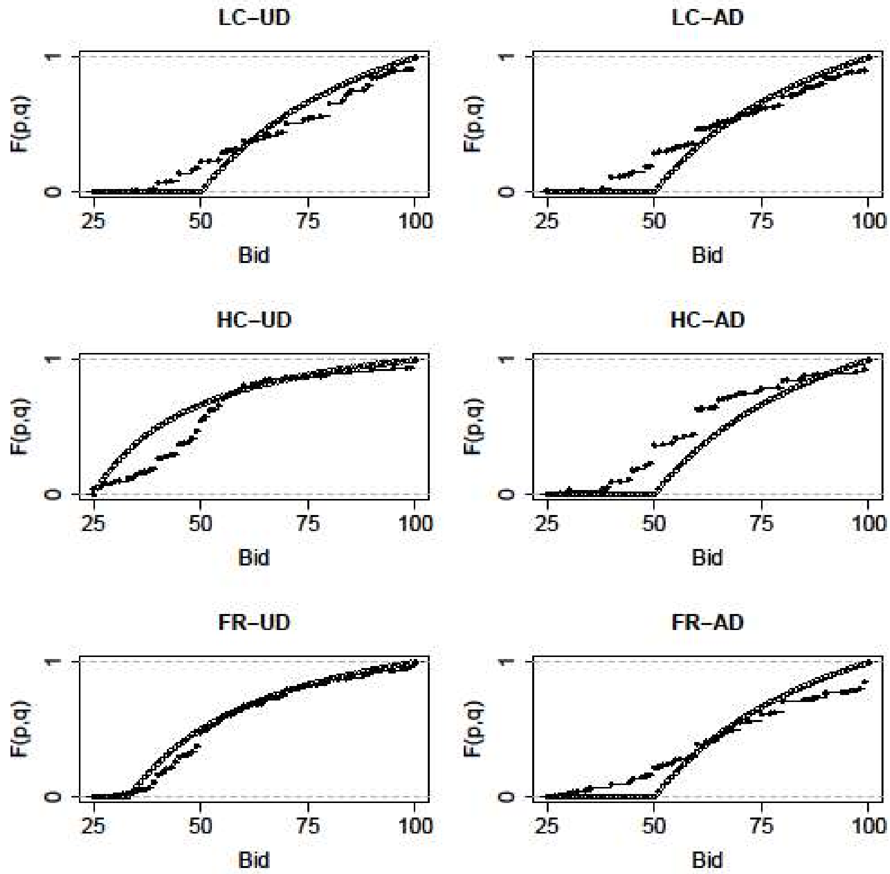
In *LC* (top row), observed bids are spread more evenly over the feasible bid range than predicted and a non-negligible fraction of bids are submitted below the lower bound of the equilibrium support. Bids in this treatment are similar regardless of the signal but we reject the theoretical distribution at each information set (p-value < 0.01 for both). In *HC* (middle row), we fail to reject the null of equality of bid distributions on-the-path (p-value $= 0.152$), although the empirical c.d.f. appears to lie to the right of the theoretical distribution at bids below 50; off-the-path, the distribution of bids is significantly lower than expected (p-value < 0.01) although, consistent with the theoretical bid support, sellers do not submit the lowest prices. In *FR* (bottom row), the observed distribution of bids is remarkably close to that predicted on-the-path (p-value $= 0.904$); off-the-path, the bid distribution diverges (p-value < 0.01).

Finally, we analyse the determinants of buyer and seller profits observed in the experiment (columns 3 and 4 of Table 4.6).

Result 4.3c. *The imposition of a deterrent penalty for reneging on a standard payment term increases consumer surplus.*

[84] All statistical inferences are unchanged by using the more powerful Epps and Singleton (1986) two-sample omnibus test, except that in *HC* we reject the null of equality of distributions on-the-path (p-value < 0.01).

Figure 4.4. Bid c.d.f. observed (•) versus theory (○), conditioned on the signal (UD = Upon Delivery, AD = After Delivery).



Buyers also earn significantly more in *HC* than in *FR* (p-value < 0.01), suggesting that the welfare-enhancing potential of a fully deterrent regulatory intervention is real. In its absence (i.e., in *LC*), Weak buyers end up significantly better off than Strong buyers (p-value < 0.01). The results in Table 4.3 reinforce the profitability of signalling a standard payment term in *HC* and *FR* relative to an alternative strategy of circumventing the penalty by signalling an extended payment term up front (p-value < 0.01 for both interaction terms).

In line with the nature of mixed strategy equilibrium, there is no statistical evidence that sellers' profits differ among treatments when interacting with a Strong buyer (p-value $= 0.265$). Sellers are worse off when confronted with a Weak buyer in *HC* (p-value $= 0.040$) and particularly in *FR* (p-value < 0.01), due to the propensity of these types to pay late. This finding underscores the potential for late payments to damage sellers' earnings in competitive market settings. Conditional on the equilibrium signal, sellers' profits are significantly above the expected level of 25 in *HC* and *FR* when interacting with a Strong buyer (p-value $= 0.037$ and p-value < 0.01), but not in *LC* (p-value $=$ 0.142). We fail to reject the hypothesis that sellers break even on average when interacting with a Weak buyer in any treatment.

4.5. Endogenous Entry

A common finding in entry game experiments is over-entry (Camerer and Lovallo 1999, Fischbacher and Thöni 2008, Goeree and Holt 2005). When the Nash equilibrium entry probabilities are less than 0.5, this can be explained theoretically by quantal response equilibrium (McKelvey and Palfrey 1995). In all three of our treatments, the Nash equilibrium probabilities are weakly above 0.5 and so over-entry would not be consistent with quantal response equilibrium. Previous auction experiments with endogenous entry, however, observe over-entry even at high equilibrium probabilities (Aycinena et al. 2018, Palfrey and Pevnitskaya 2008). This might be attributed to risk preferences, an additional utility from winning or relative payoff considerations. Alternatively, it may be an artefact of activity bias (Lei et al. 2001).

Since equilibrium payment strategies in our signalling model depend crucially on correct expectations about sellers' entry probabilities, we deliberately set these probabilities as exogenous in our experiment. This design increased experimental control and mitigated the potential confound of over-entry in the auction. To investigate whether over-entry would arise in our environment, we conducted an additional experiment which permitted sellers in the *HC* treatment to make both an entry and pricing decision at the Auction stage. We call this variant *High-Cost-Entry* (*HC-E*). The design and procedures for this variant were identical to the original *HC* treatment, with the following differences. At the Pre-Auction stage, buyers no longer knew the probability of a seller being eligible to customise the unit conditional on the signal. At the Auction stage, each seller had the option to return his or her non-customised unit to the experimenter without submitting a bid, i.e., entry was no longer randomised by the computer. We recruited four independent cohorts of 9 subjects for the *HC-E* treatment across two sessions.[85] On average, buyers earned more per round in *HC-E* than in *HC*, and vice versa for sellers (Buyers: 56.97 versus 46.60, p-value = 0.013; Sellers: 20.55 versus 26.36, p-value < 0.01; two-sided Wilcoxon-Mann-Whitney tests).[86] This is consistent with an over-entry story.

The novelty of *HC-E* is in the Auction stage and so we first examine entry and price statistics. In Table 4.7, we compare the new cohort frequencies of entry and conditional-on-entry winning bids, by the received signal, to those predicted in the symmetric mixed strategy equilibrium. We observe two diverging behavioural patterns among cohorts.

Table 4.7 – Aggregate entry, price and payment statistics for the HC-E cohorts.

Cohort	*1*		*2*		*3*		*4*	
Intention UD	*n = 136*		*n = 136*		*n = 18*		*n = 16*	
Entry [0.75]	0.79	(0.41)	0.75	(0.43)	0.78	(0.43)	0.81	(0.40)
Price	34.53	(15.22)	38.19	(12.45)	37.78	(9.05)	56.12	(11.03)
Final payment AD	0.04	(0.01)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Intention AD	*n = 44*		*n = 44*		*n = 162*		*n = 104*	
Entry [0.50]	0.66	(0.48)	0.57	(0.50)	0.80	(0.40)	0.77	(0.42)
Price	55.47	(20.03)	61.65	(14.08)	42.1	(18.08)	56.88	(16)
Final payment AD	1	(0.00)	1	(0.00)	1	(0.00)	1	(0.00)

Notes: Displayed are mean (SD) values at the subject level in each cohort; n = number of signals at a given information set (cohorts 1 to 3 had six sellers and so 180 auctions; cohort 4 had four sellers and so 120 auctions in total). The Nash equilibrium entry prediction is in square brackets. *UD = Upon Delivery, AD = After Delivery.*

In cohorts 1 and 2, approximately three in four signals indicate payment upon delivery and in response, sellers enter into 75 to 79% of auctions. This is close to the expected equilibrium entry frequency of 75%. On receipt of an intention to pay after delivery, sellers in these cohorts enter into 57 to 66% of auctions versus an expected frequency of 50%. A very different outcome is observed in cohorts 3 and 4: despite approximately

[85] One cohort had only 2 buyers and 4 sellers due to a lower-than-expected show-up rate.

[86] Unless otherwise stated, statistical tests in this section are based on linear regressions that control for cohort fixed effects and a linear time trend, with standard errors clustered at the subject level. The corresponding regression estimates are reported in the appendix. The entry results are qualitatively unchanged if we model the probability of entry as a logistic function.

nine in ten signals indicating payment after delivery, sellers enter into 80% of auctions. Thus, we find strong evidence of over-entry in *HC-E* off-the-path (p-value = 0.013).

Entry is lower in *HC-E* after receiving an intention to pay after delivery versus upon delivery (p-value = 0.052). We find no statistical evidence of a time trend in entry decisions. To test the mixed strategy equilibrium at the individual level, we follow Moffatt (2015, 16.2) and examine entry sequences for each subject separately (Table 4.8). Conditional on the received signal, we reject the Nash entry point prediction for 13 out of 22 sellers (p-value < 0.05, two-sided Binomial tests); we also reject randomness of the entry sequence for 11 subjects (p-value < 0.05, two-sided runs tests). There is no significant relationship between entry and self-reported risk preferences.

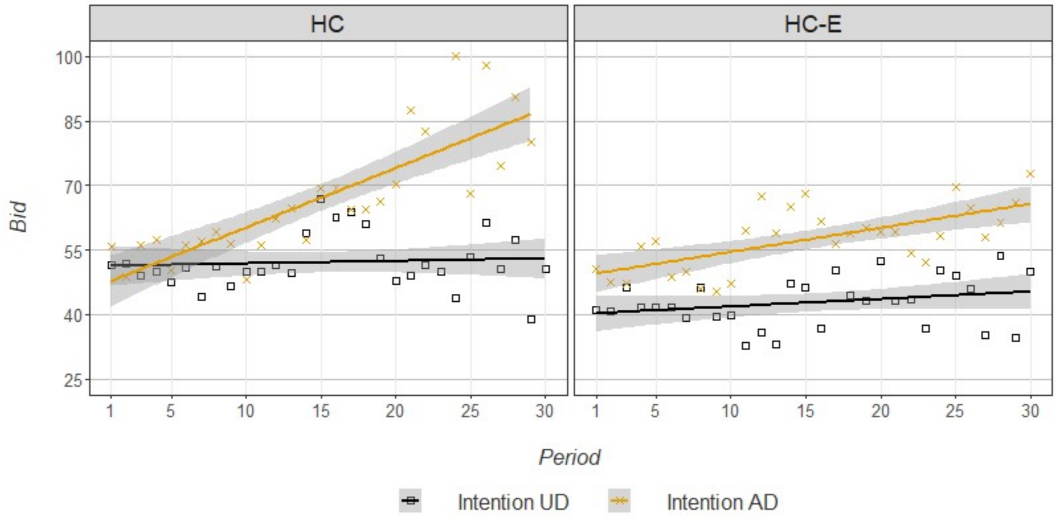
Table 4.8 – Seller entry frequencies in the HC-E treatment: binomial and runs tests.

Cohort Seller	*1*	*2*	*3*	*4*
1	1^{**++}	0.55**	1^{**++}	0.44^{++}
2	1^{**++}	0.84^{++}	0.61	0.85**
3	1^{**++}	0.65	0.46	0.78**
4	0.63	0.83	0.73**	1^{**++}
5	0.92	1^{**++}	1^{**++}	
6	0.26^{**++}	0.58	0.96^{**++}	

Notes: Displayed are the relative entry frequencies for each seller in each cohort, conditional on the modal signal received; for cohorts 1 and 2, the modal signal is *Upon Delivery*; for cohorts 3 and 4, the modal signal is *After Delivery*. ** $p < 0.01$, * $p < 0.05$, based on two-sided binomial test versus expected frequency (= 0.75 for cohorts 1 and 2, = 0.5 for cohorts 3 and 4). $^{++}$ $p < 0.01$, $^{+}$ $p < 0.05$, based on two-sided runs test.

We also observe lower-than-expected transaction prices in *HC-E*, which together with over-entry suggests more intense bidding competition when entry is endogenous. In cohorts 1 and 2, average prices are 34.53 and 38.19, below the equilibrium prediction of 40; in cohorts 3 and 4, prices are 42.1 and 56.88, below the off-equilibrium prediction of 66.67. They are also significantly below those observed in *HC* (p-value < 0.01). In Figure 4.5, we plot the time-series of bids in the *HC* and *HC-E* treatments, conditioned on the signal. Bids are markedly lower with endogenous entry and there is a significant positive time trend in both treatments off-the-path, which suggests a certain degree of learning at this information set. As a consequence, buyers earn significantly more and sellers significantly less than their counterparts in *HC* (both p-values < 0.01). Sellers earn a gross profit below the entry cost of 25 in all four cohorts, suggesting that they would have been better off staying out of the market.[87]

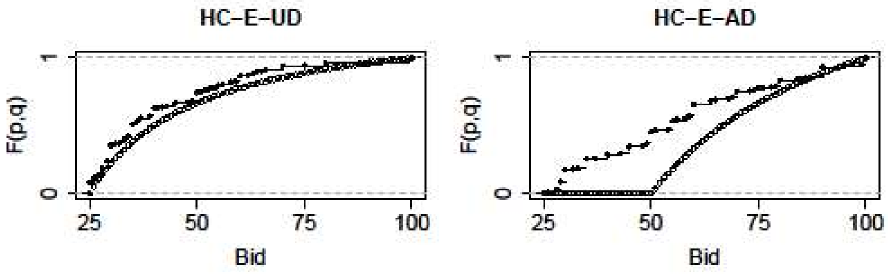
Figure 4.5. Time-series of bids in the HC and HC-E treatments (including trend line and 95% confidence intervals, UD = Upon Delivery, AD = After Delivery).



The associated bid distributions observed in *HC-E* are presented in Figure 4.6. The top (bottom) panel displays the bid c.d.f. conditional on an intention to pay upon (after) delivery. Both distributions lie to the left of the distribution associated with the mixed strategy equilibrium. On-the-path, bids track the theory and we fail to reject the null of equality of distributions (p-value = 0.104, two-sided Kolmogorov-Smirnov test); the main difference with *HC* is the larger fraction of bids below 50. Off-the-path, we reject equality with the theoretical c.d.f. (p-value < 0.01, two-sided Kolmogorov-Smirnov test).

[87] Risk aversion would require a premium and net seller profit in equilibrium. We observe no significant relationship between risk preferences and seller profits; neither is there a significant gender effect.

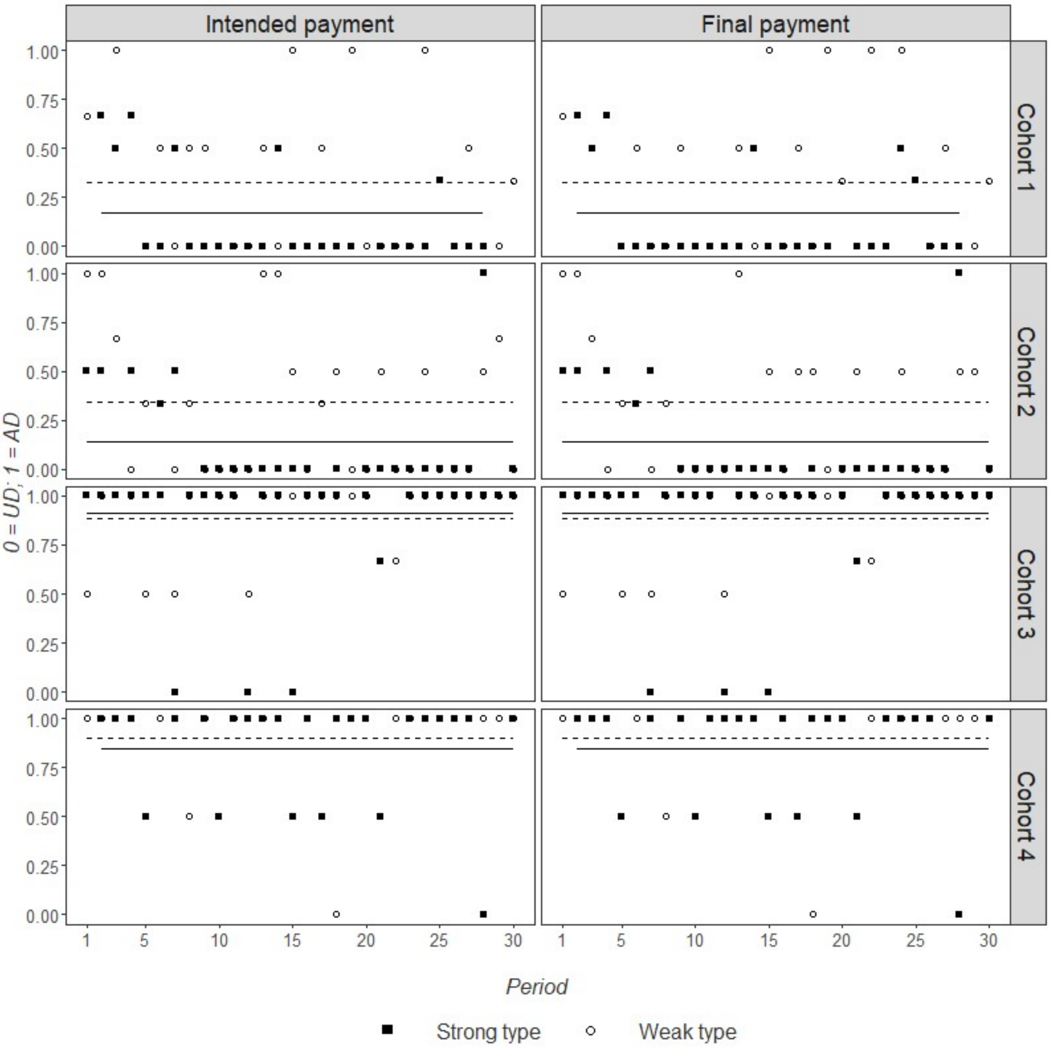
Figure 4.6. Bid c.d.f.s observed (•) versus theory (○) in the HC-E treatment conditioned on the signal (UD = Upon Delivery, AD = After Delivery).



Finally, we have seen that buyers in *HC-E* cohorts 1 and 2 followed largely different payment strategies in the Pre-Auction stage relative to buyers in cohorts 3 and 4.[88] To see this more clearly, in Figure 4.7 we present the intended and final payment decisions for each cohort and buyer type over the course of the experiment. Both types in cohorts 1 and 2 converge on the equilibrium payment strategy (UD, UD), whereas convergence in cohorts 3 and 4 is on the alternate strategy (AD, AD). These patterns emerge early on and there is no significant time trend in buyer behaviour (see formal regression analysis in the appendix). We conjecture that the differential payment paths observed are a direct result of over-entry. Unlike in *HC* and in cohorts 1 and 2 of *HC-E*, sellers in cohorts 3 and 4 did not punish buyers for offering extended payment terms: there is little difference in entry rates and bids in these cohorts conditional on the signal. Moreover, entry was significantly higher and bids lower than the theory would predict. In turn, buyers were able to create a "micro culture" in which late payments were expected and realized as a best-response to bidding behaviour.

[88] Similar heterogeneity was not observed in *HC*: the majority signal in every *HC* cohort was *UD*.

Figure 4.7. Intended and final payment decisions for the HC-E cohorts by type.



Notes: *UD = Upon Delivery, AD = After Delivery*. The solid (dashed) lines are the relative frequencies of intended and final payments after delivery in the FR treatment among Strong (Weak) buyer types.

4.6. Discussion and Conclusion

In this chapter, we harness the methodology of experimental economics to provide causal evidence that late payments feed into higher market prices and lower expected consumer surplus. Late payments, and their strategic counterpart extended payment terms, are a real liquidity risk in trading relationships and characterise a wide range of industries. Yet identifying the impact of such payment practices in the field confronts an endogeneity problem embodied in payment data. For example, in the wake of the Covid-19 pandemic, firms faced unprecedented revenue shocks (Joseph et al. 2020). In the UK, more than half of sellers reported having unpaid invoices due as a result of the crisis (Office for National Statistics 2020). The extent to which outstanding payments contribute to sellers increasing their prices or leaving the market is difficult to disentangle from the wider financing constraints and demand shortfalls of the time.

Existing models in the industrial organisation literature do not consider the horizontal effects of late payment risk on suppliers' bidding strategies. We show theoretically that the expectation of a late payment induces sellers to enter into competition less frequently and to inflate their bids conditional on entry. As a consequence, buyers end up worse off relative to a scenario in which they offer a standard payment term and make payment on-time. Our analysis suggests that introducing a regulatory penalty on those buyers who renege on a standard payment term can generate a Pareto improvement: as sellers become more confident in the payment terms received, they bid for contracts more often and compete more aggressively, leaving end consumers to benefit from lower prices.

We provide experimental evidence to back up our theoretical predictions and demonstrate that care should be taken in setting the level of any regulatory fine. If this fine is not set high enough, those firms who benefit most from paying late may free ride on the positive competition spillover effects generated by other firms who pay on-time. The experiments also inform on why certain sectors may become stuck in an inefficient equilibrium characterised by extended payment terms and prolonged unpaid invoices: if sellers are unwilling to adjust their bids in response to an offer of a standard payment term for fear of the buyer reneging (as we observed in our *LC* treatment), it is not in the buyer's interest to change their payment practices. To that end, an empirically testable prediction of the model is that the relationship between the average payment term and the level of consumer prices will be more evident in regions and/or sectors where a late payment is penalised than where it is not.

An implication of our analysis is that a regulatory penalty may be ineffective in times of economic stress and when the buyer is likely to place higher value on a late payment. An alternative would be to reward firms for prompt payment. It is not obvious, however, that this would be palatable for a regulator to implement – presumably, it would require some form of subsidy. Discounts are often offered by sellers to consumers in voluntary trade credit agreements, as a reward for early payment.[89] Late payment is, by its very nature, an *involuntary* arrangement. It is also important to recognise that implementing a penalty may not be feasible in all market settings, particularly if the cost of verifying payment performance is high. Our results suggest that buyers might be more likely to cooperate with such a regulatory scheme if they understand the horizontal price benefits from establishing a culture of prompt – rather than late – payments.

In future work, we plan to additionally consider the impact of a penalty absorbed by the creditor. For example, the EU directive 2011/7/EU on combating late payment in commercial transactions (2011) states that in the event of a late payment, statutory interest and compensation for fixed recovery costs are due to the supplier. Thus, incentives to renege on a standard payment term may also depend on the payment size. Intuitively, buyers will be less willing to incur a penalty for a small rather than a large inventory item. This could be addressed analytically by moving from a fixed-sum to proportional penalty.

A further restrictive assumption in our approach is that renegotiation to a different final payment date is always inefficient. In practice, granting an extension to the payment term may be welfare-improving if liquidity constrained buyers are reprieved from going bankrupt, and/or if flexibility in the invoice due date can mitigate a short-term

[89] See, e.g., "2/10 Net 30", in which a company extends a 30 day invoice term to customers who pay on credit and a 2% discount on the purchase price if payment is made within 10 days.

negative shock. Note that incorporating uncertainty on the buyer-side is unlikely to change the horizontal impact of a late payment on suppliers' bidding strategies (so long as sellers still discount a late payment). It may, however, alter the buyer's equilibrium incentives and so welfare comparisons. On the seller-side, it would be worthwhile to consider the implications for entry and pricing of conditional bids, i.e., offers that can be conditioned on the final payment date.

We acknowledge that the concept of welfare considered here (expected buyer surplus) is narrow. This enabled us to focus specifically on the horizontal effects of a late payment on entry and pricing, at the cost of abstracting from broader welfare implications that might arise in more complex market structures. For instance, one could examine the effects of late payment on sub-contractors in a multi-tier supply chain. Once the consequences of funding risk for downstream supplier competition are considered, the benefit of introducing a deterrent penalty for reneging on a standard payment term may be even larger. How payment terms and financial capacity interact between upstream and downstream suppliers is a possible avenue for future work. Relatedly, we limit attention to a symmetric-information (among sellers) first-price, sealed-bid auction in which preparing a bid is costly. While this is reasonable for projects in which a cost estimate can be arrived at precisely (after some effort), in other situations these assumptions are restrictive. It would be interesting to examine the influence of the auction format on buyer-side and seller-side payment strategies.

Finally, a limitation of our approach is that we consider only a one-shot interaction. In repeated interactions, reputational information may influence bidding strategies, change a buyer's incentive to pay late (e.g., due to loss of trust and with it future business), and lead to separating equilibria not captured by our model. Note that in one-shot interactions, reputational information may still be relevant if firms are required to report on prior payment practices in the public domain.[90] Note too that the consequences of repeated interactions for the propensity of late payments are far from straightforward. Small firms, for example, may be reluctant to claim against their larger

[90] For an example of payment reporting requirements in the UK, see the Department for Business, Energy and Industrial Strategy (2016).

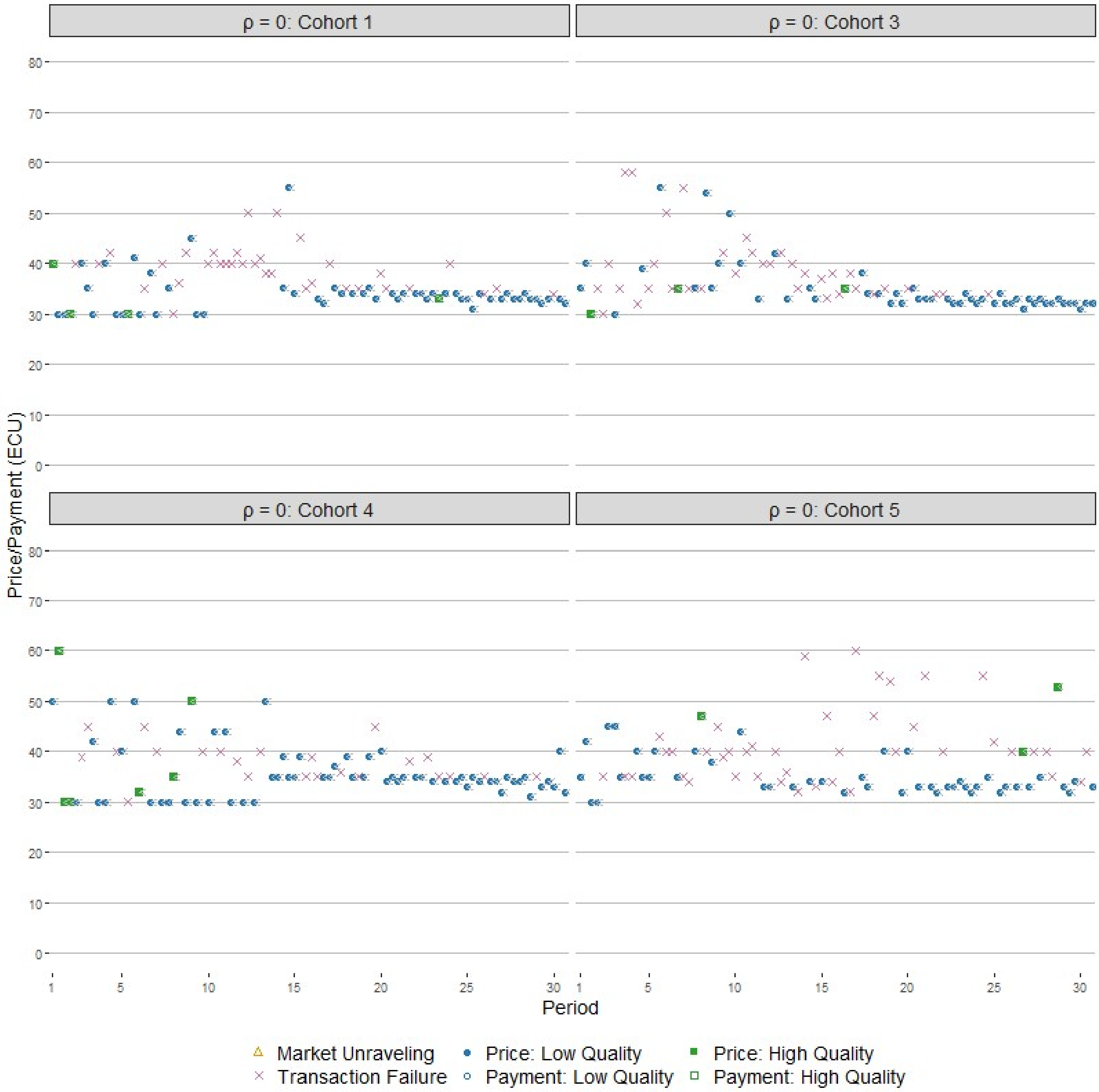
clients for fear of losing repeat business. The true welfare consequences of late payments will depend also on relative bargaining powers. Our static model predicts (intuitively) that Weak buyers benefit most from making a late payment. In a dynamic model, the buyers that benefit most are likely to be Weak buyers who possess significant bargaining power. Such buyers may be able to reap the benefits of holding onto cash without fear of repercussions from their suppliers. This was exactly the strategy of Carillion (cited in the Introduction) before it collapsed.[91] Incorporating such imperfections would be another useful extension to the model.

[91] According to Rachel Reeves MP: "It's clear that Carillion were notorious late-payers, ruthlessly exploiting their position to bully their contractors in a desperate bid to prop up their precarious business model" ("Committees publish correspondence", 2018).

A Appendix for Chapter 2

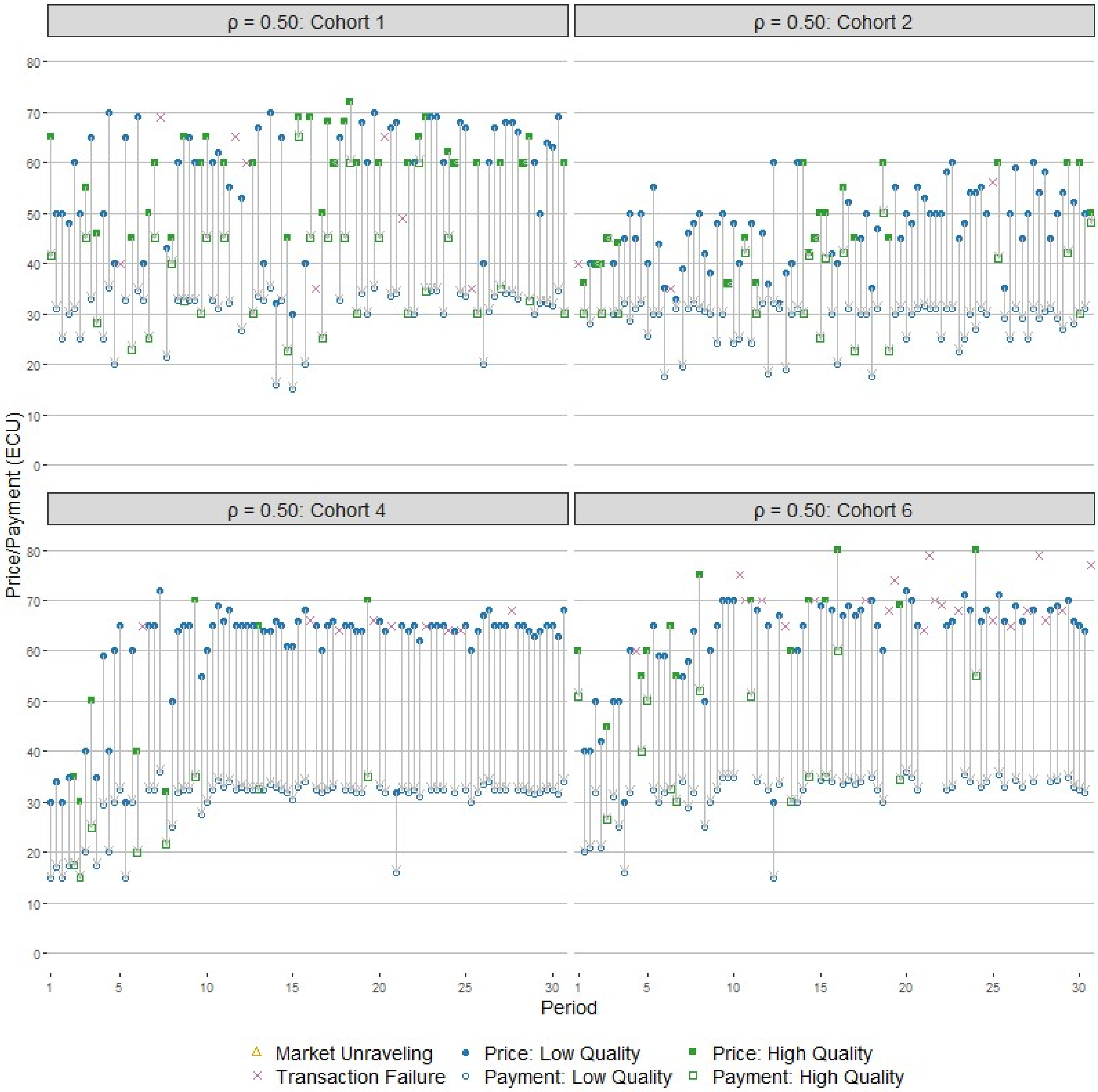
A.1. Supplementary analyses of the experimental data

Figure A. 1. Remaining cohort outcomes over time ($\rho = 0$).



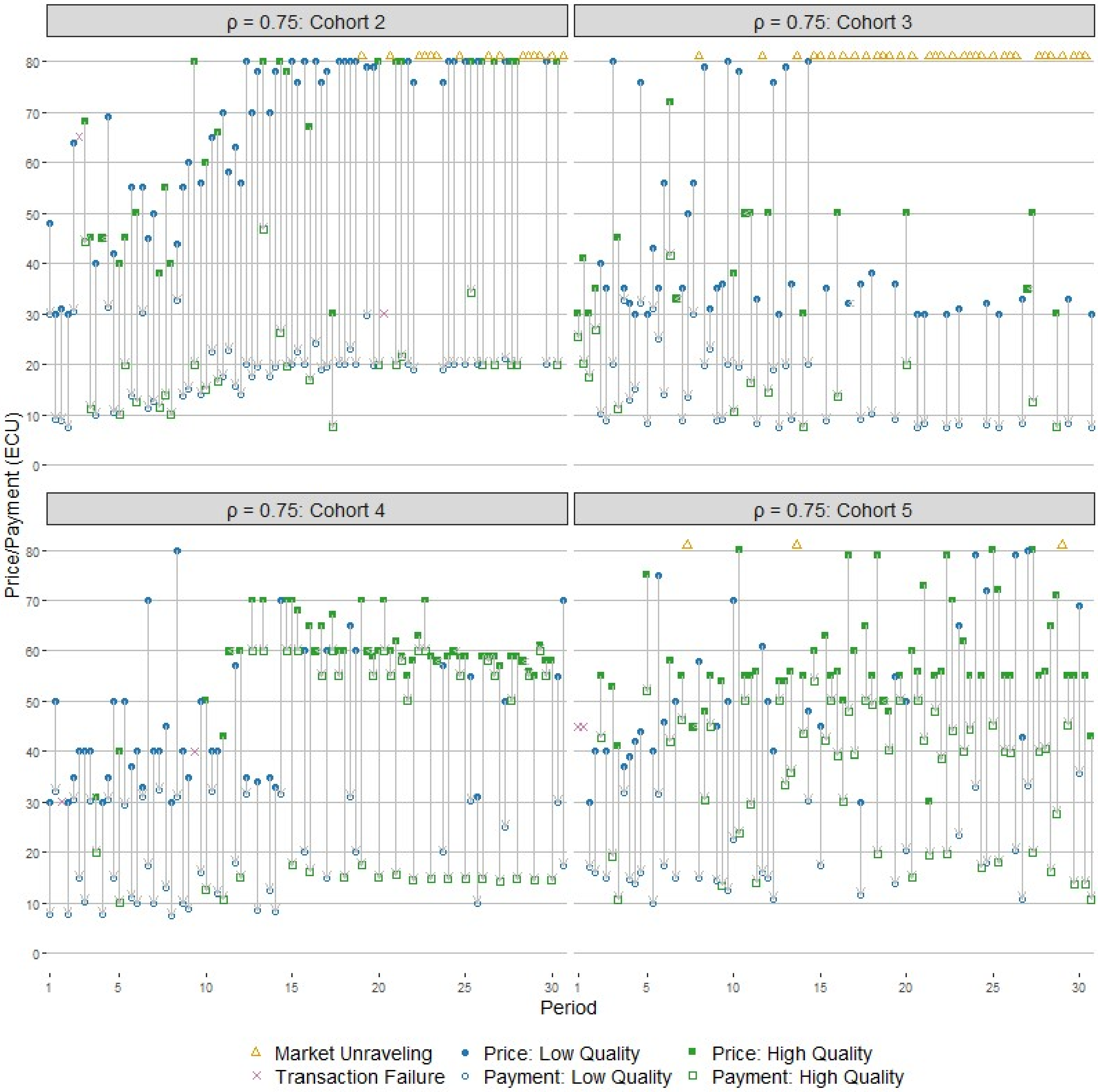
Notes: Panels display procurement outcomes over time in the specified *treatment – cohort*. Each observation corresponds to the outcome of a randomly matched interaction group in the period, with 1 buyer and 2 sellers. An open triangle is an instance of market unraveling, in which neither seller submitted a bid. A cross is a transaction failure, in which a buyer rejected the winning bid. A solid circle is an accepted winning bid at which the seller delivered low quality. A solid square is an accepted winning bid at which the seller delivered high quality.

Figure A. 2. Remaining cohort outcomes over time ($\rho = 0.50$).



Notes: Panels display procurement outcomes over time in the specified *treatment – cohort*. Each observation corresponds to the outcome of a randomly matched interaction group in the period, with 1 buyer and 2 sellers. An open triangle is an instance of market unraveling, in which neither seller submitted a bid. A cross is a transaction failure, in which a buyer rejected the winning bid. A solid circle is an accepted winning bid at which the seller delivered low quality. A solid square is an accepted winning bid at which the seller delivered high quality. An open circle and an open square are the corresponding total payments in instances where these differ from the winning bid. In such instances, the vertical arrows represent the price-payment differential.

Figure A. 3. Remaining cohort outcomes over time ($\rho = 0.75$).



Notes: Panels display procurement outcomes over time in the specified *treatment – cohort*. Each observation corresponds to the outcome of a randomly matched interaction group in the period, with 1 buyer and 2 sellers. An open triangle is an instance of market unraveling, in which neither seller submitted a bid. A cross is a transaction failure, in which a buyer rejected the winning bid. A solid circle is an accepted winning bid at which the seller delivered low quality. A solid square is an accepted winning bid at which the seller delivered high quality. An open circle and an open square are the corresponding total payments in instances where these differ from the winning bid. In such instances, the vertical arrows represent the price-payment differential.

Table A. 1 – Statistical learning analysis (I).

	Periods 1-10	Periods 11-20	Periods 21-30
		$\rho = 0$	
Bid	42.49	40.69	38.24
	(4.64)	(4.69)	(4.20)
Seller participation	0.98	0.99	0.99
	(0.03)	(0.01)	(0.01)
Buyer acceptance	0.56	0.53	0.73
	(0.20)	(0.24)	(0.22)
High quality	0.15*	0.03	0.02
	(0.07)	(0.05)	(0.04)
		$\rho = 0.50$	
Bid	57.61**	63.46	64.75
	(5.66)	(6.47)	(5.72)
Seller participation	0.99	0.99	0.95
	(0.01)	(0.01)	(0.07)
Buyer acceptance	0.94	0.90	0.89
	(0.03)	(0.08)	(0.15)
High quality	0.31	0.35	0.24
	(0.08)	(0.23)	(0.33)
Retainage return (low)	0.10*	0.05	0.03
	(0.10)	(0.08)	(0.05)
Retainage return (high)	0.40	0.27	0.33
	(0.25)	(0.17)	(0.09)
		$\rho = 0.75$	
Bid	55.22*	64.47	62.35
	(4.43)	(9.51)	(16.21)
Seller participation	0.83*	0.59	0.50
	(0.16)	(0.31)	(0.37)
Buyer acceptance	0.96^{+}	0.99	1.00
	(0.03)	(0.01)	(0.00)
High quality	0.37	0.53	0.55
	(0.15)	(0.25)	(0.37)
Retainage return (low)	0.19	0.12	0.09
	(0.09)	(0.07)	(0.09)
Retainage return (high)	0.39	0.36	0.35
	(0.18)	(0.24)	(0.20)

Table A. 2 – Statistical learning analysis (II).

	Periods 1-10	Periods 11-20	Periods 21-30
		$\rho = 0$	
Buyer profit	4.16	0.97	2.93
	(4.31)	(1.99)	(1.37)
Seller profit	5.95	5.17	2.89
	(3.19)	(2.27)	(0.75)
		$\rho = 0.50$	
Buyer profit	17.99	16.77	12.06
	(2.09)	(9.58)	(13.13)
Seller profit	-2.30	0.40	1.41
	(2.14)	(2.45)	(2.20)
		$\rho = 0.75$	
Buyer profit	29.38	30.11	33.16
	(4.40)	(4.69)	(8.31)
Seller profit	-11.34	-6.72	-8.90
	(3.06)	(7.56)	(9.13)

Notes: Displayed in this table and the previous table are mean (SD) values based on cohort averages. Matched pairs t-test (two-sided) comparing averages between adjacent blocks, * indicates significance at the 0.05 level; ** indicates significance at the 0.01 level.

A.2. Least absolute deviation (LAD) estimation of the behavioural model

In this part of the appendix, we employ a tractable LAD approach to estimating the parameters of our behavioural model on the profit data from all 30 periods of our experiment. Bolton and Ockenfels (2000) use an analogous procedure to apply their α model to behaviour in gift-exchange and trust game experiments. For comparability with Bolton and Ockenfels, we assume that trustworthy buyer types reward both low and high quality according to the fairness norm γ.

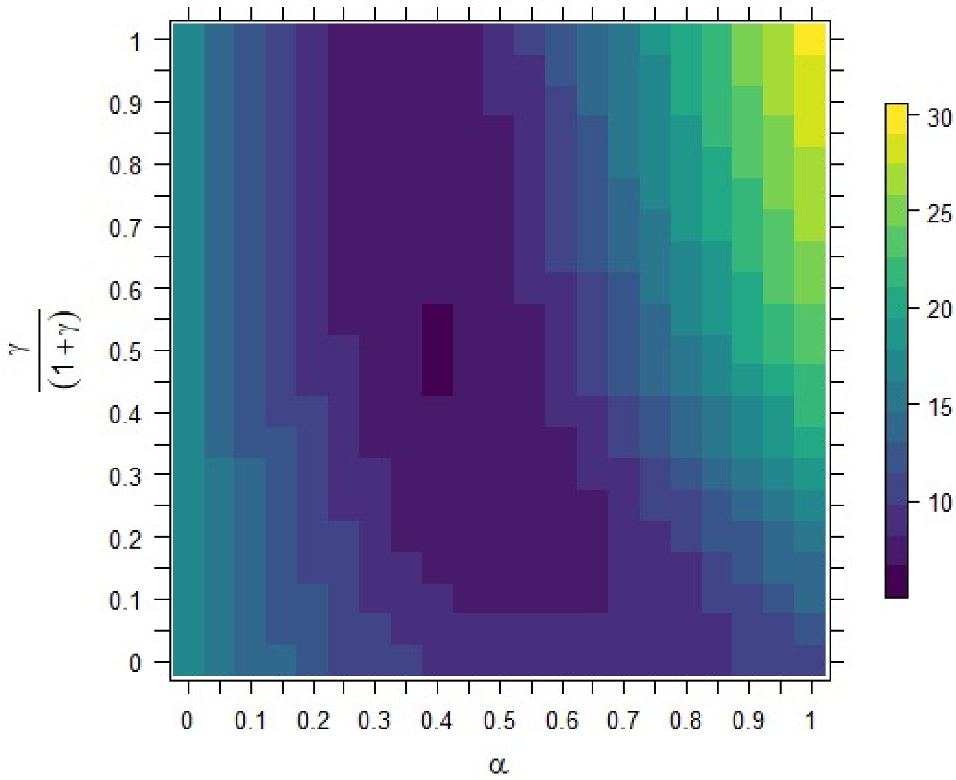
First, we define a procurement "bucket" $g_i(\rho, p, q)$ as a group of procurement interactions that share the same retainage provision and price-quality outcome. In terms of the strategic interaction, ρ determines the specification of the extensive form game, (p, q) specifies a penultimate information set and g_i identifies the subset of observations reaching this information set. Observations within this subset vary by the buyer's ultimate action. For each bucket, we compute the average actual buyer and seller profits observed in the experiment.[92] We assign each bucket a weight according to its relative empirical frequency, i.e., the number of transactions in the bucket divided by the total number of transactions across all buckets (= 907). For each bucket, we then calculate the buyer and seller profits predicted by the model of fair payment norms at all (α, γ) pairs in the unit square. This enables us to compute, for each bucket and pair, the absolute deviation between actual and predicted buyer and seller profits. Finally, we take the weighted sum of these absolute deviations across all buckets and for each (α, γ) pair. The following objective function described this calculation:

$$\sum_i \frac{\#g_i}{907} \left(|\bar{\pi}_B^{actual}(g_i) - \bar{\pi}_B^{fair}(g_i, \alpha, \gamma)| + |\bar{\pi}_S^{actual}(g_i) - \bar{\pi}_S^{fair}(g_i, \alpha, \gamma)| \right) \tag{A.1}$$

[92] Taking averages reduces noise at the cost of losing information from higher moments. The results of the LAD estimation are qualitatively unchanged if we instead conduct the exercise at the transaction level. The least deviation estimate is the same at $\alpha = 0.40$. The estimate of γ using the transaction level data is less precise, but also covers the best estimate found using the averaged data.

Results of the estimation are presented in Figure A. 4. The subset of joint parameter estimates yielding the minimum LAD are identified by the darkest shaded colour in the figure. This subset is centred around best estimates of $\alpha = 0.40$ and $\gamma = 0.50$. Our estimate of γ predicts zero retainage payment for all interim seller surplus shares greater than one half. We observed this outcome in 137 out of 139 possible cases.

Figure A. 4. LAD estimates of the behavioural model parameters.



Notes: The heatmap presents the set of feasible joint parameter estimates of α and γ, imputed from the experiment data. The colour scale indicates the least absolute deviation (LAD) between average actual profits and those predicted by the model of fair payment norms (the unit is ECU). The darkest shaded region corresponds to joint parameter estimates that yield the minimum LAD between model's predictions and realized outcomes.

A.3. Experimental instructions

This is an experiment in decision making. If you read these instructions carefully and make good decisions, you may earn a considerable amount of money. The amount of money you earn will depend on both your decisions and the decisions of other participants. The currency unit in this experiment is called Experimental Currency Unit, or ECU for short. At the end of the session, you will be asked to complete a short questionnaire. Upon completion, your total earnings from the experiment will be displayed on the screen, including your participation fee of $5, and be paid to you in private and in cash.

How you earn money

For today's session, one-third of you will be randomly assigned a Buyer role, and two-thirds of you assigned a Seller role. You will see your role at the start of the session and this role will not change for the duration of the session. This experiment will include **30 rounds**. In each round a Buyer is matched with two Sellers. The Buyer can purchase a product from one of the two Sellers. The value of the product to the Buyer, and the cost that the Seller incurs to provide the product, depends on the Seller choice to deliver a High or Low Quality product, as follows:

	Low Quality	High Quality
Buyer Value	35	80
Seller Cost	30	40

Each player will start each round with 7 ECU. Therefore, when there is a transaction between a Buyer and a Seller,

- The Buyer earns: 7 + Profit, where Profit = Buyer Value − Payment to the Seller
- The Seller earns: 7 + Profit, where Profit = Payment to the Seller − Seller Cost

How a potential Seller is determined

An auction is used to select which of the two Sellers will have the opportunity to transact with the Buyer. At the start of each round, each Seller privately submits a bid for a price at which they would be willing to deliver the product or elects not to participate in the current round's auction. The Seller submitting the lower bid wins the opportunity to potentially trade with the Buyer. In the event of a tie, the computer will choose the winner at random. Bids can be integers from 30 to 80.

Example:

Seller A bids 35 and Seller B bids 62 = A is winner.

Sellers A and B both bid 40 = A and B each have 50% probability of winning.

A Seller who does not win or elects not to participate in the auction earns 7 ECU for the round (Profit = 0). After the auction, the Buyer is presented the winning bid. The Buyer chooses whether or not to proceed with the transaction. If the Buyer does not proceed the round ends and all three parties earn 7 ECU for the round (Profit = 0).

How the payment to the Seller is determined

Participants in the $\rho = 0$ *treatment received the following instructions:*

If the Buyer chooses to proceed, the Winning Seller receives payment equal to his or her bid amount. Then the Winning Seller must decide to deliver either a Low or High Quality product.

Examples:

- Suppose the winning bid is 50 and the Seller delivers Low Quality. The round profits for the Buyer and Winning Seller are as follows:

Buyer Profit = 35 – 50 = –15*

Seller Profit = 50 – 30 = 20

- Now suppose the Seller delivers High Quality. The respective profits are:

Buyer Profit = 80 – 50 = 30

Seller Profit = 50 – 40 = 10

**Note that it is possible to lose money in a round. Make your decisions carefully.*

Participants in the $\rho = 0.50$ *treatment received the following instructions:*

If the Buyer chooses to proceed, the Winning Seller receives **50%** of his or her bid amount. This is called the **initial payment**. Then the Winning Seller must decide to deliver either a Low or High Quality product. The Buyer learns the quality level of the product and then chooses how much of the remaining 50% of the winning auction bid is paid to the Seller. This amount is called the **deferred payment**. The Buyer keeps any portion of the remaining 50% of the bid not paid to the seller.

Examples:

- Suppose the winning bid is 50, the Seller delivers Low Quality, and the Buyer sets the deferred payment at 6. The round profits for the Buyer and Winning Seller are as follows:

  Buyer Profit = 35 – 25 – 6 = 4

  Seller Profit = 25 – 30 + 6 = 1

- Now suppose the Seller delivers High Quality and the Buyer still sets the deferred payment at 6. The respective profits are:

  Buyer Profit = 80 – 25 – 6 = 49

  Seller Profit = 25 – 40 + 6 = –9*

- Now Suppose the winning bid is 70, the Seller delivers High Quality and the Buyer sets the deferred payment at 30. The respective profits are:

  Buyer Profit = 80 − 35 − 30 = 15

  Seller Profit = 35 − 40 + 30 = 25

  **Note that it is possible to lose money in a round. Make your decisions carefully.*

Participants in the $\rho = 0.75$ *treatment received the following instructions:*

If the Buyer chooses to proceed, the Winning Seller receives **25%** of his or her bid amount. This is called the **initial payment**. Then the Winning Seller must decide to deliver either a Low or High Quality product. The Buyer learns the quality level of the product and then chooses how much of the remaining 75% of the winning auction bid is paid to the Seller. This amount is called the **deferred payment**. The Buyer keeps any portion of the remaining 75% of the bid not paid to the seller.

Examples:

- Suppose the winning bid is 60, the Seller delivers Low Quality, and the Buyer sets the deferred payment at 16. The round profits for the Buyer and Winning Seller are as follows:

  Buyer Profit = 35 − 15 − 16 = 4

  Seller Profit = 15 − 30 + 16 = 1

- Now suppose the Seller delivers High Quality and the Buyer still sets the deferred payment at 16. The respective profits are:

  Buyer Profit = 80 − 15 − 16 = 49

  Seller Profit = 15 − 40 + 16 = –9*

- Now Suppose the winning bid is 72, the Seller delivers High Quality and the Buyer sets the deferred payment at 44. The respective profits are:

  Buyer Profit = 80 − 18 − 44 = 18

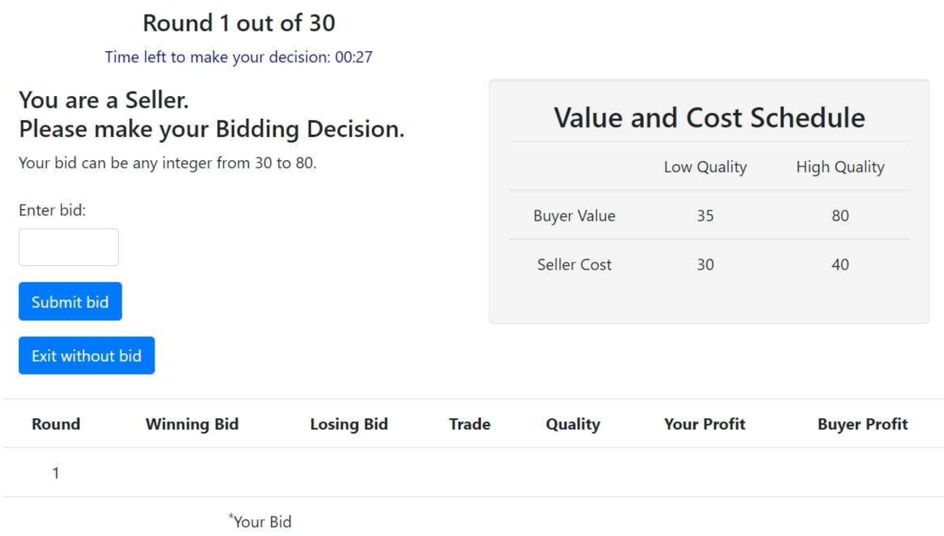
  Seller Profit = 18 − 40 + 44 = 22

**Note that it is possible to lose money in a round. Make your decisions carefully.*

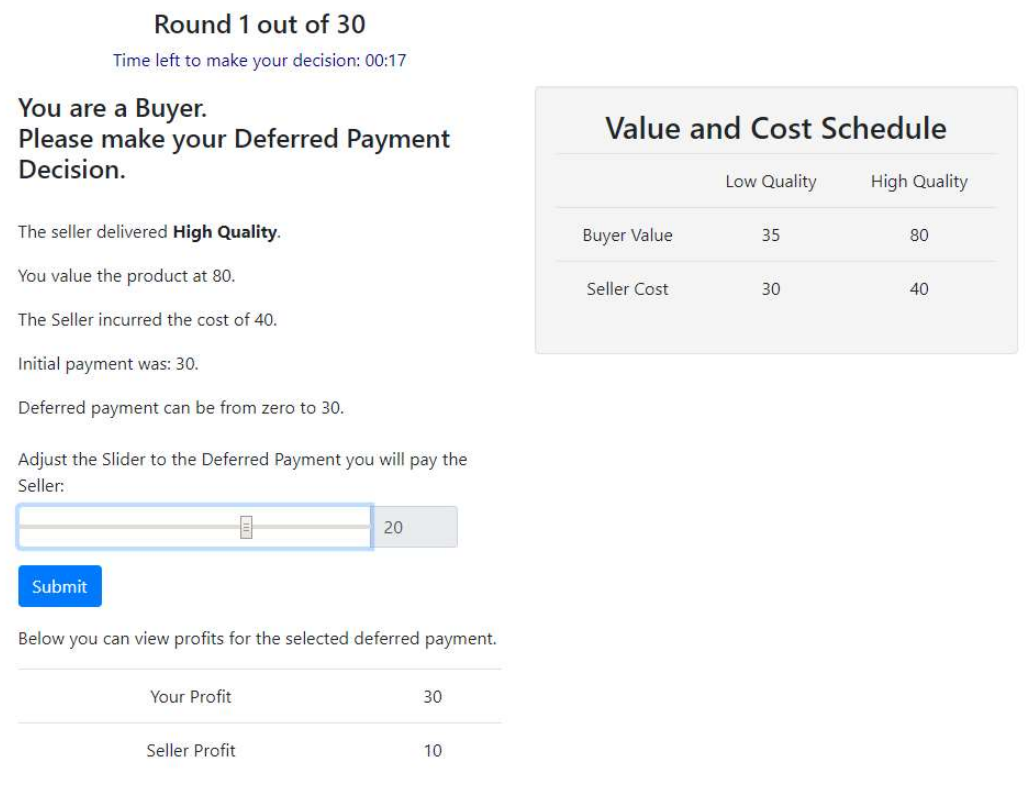
How you will be paid

At the end of the session the earnings from all rounds of the session will be converted to US dollars at the rate of 20 ECU for $1. These earnings will be added to your $5 participation fee, displayed on your screen, and paid to you at the end of the session.

Example experimental interface screenshot for sellers:



Example experimental interface screenshot for buyers:



Note: In the $\rho = 0.50$ and $\rho = 0.75$ treatments, a slider with random initial value was used to avoid anchoring bias. A calculator displayed buyer/seller profits for the different available deferred payments.

A.4. Post-experiment questionnaire and subject characteristics

Age: Interval variable.

Years.

Mean 23.89, Median 24, Standard deviation 4.74, Minimum 18, Maximum 60

Gender: Dummy variable for female.

Male 49.38%; Female 50.62%.

Field of studies: Dummy variable for economics or business major.

Arts and Education 1.24%; Economics and Finance 6.21%; Business and Management 65.22%; Law and Social Sciences 1.86%; Medicine and Health Sciences 4.35%; Engineering and Natural Sciences 20.50%; Not a Student 0.62%.

Nationality: Categorical variable:

Central and Eastern Asia 23.90%; Central and Western Africa 0%; Central, South America and the Caribbean 0.63%; Europe (excl. UK) 1.89%; Middle East and North Africa 0%; North America 11.95%; Oceania 0%; South and Eastern Africa 0.63%; South-East Asia 22.64%; Southern Asia 38.36%; UK 0%.

Income: Categorical variable.

When you were 16 years of age, what was the income of your parents in comparison to other families in your country?

Far below average 3.11%; Below average 9.32%; Average 50.93%; Above average 33.54%; Far above average 3.11%.

Risk Index: Average of six behavioural risk categories, general attitude and specific domains.

Based on the questions in Dohmen et al. (2011). Likert scale from 1 “Completely unwilling to take risks” to 10 “Completely willing to take risks”.

1) Are you generally a person who is fully willing to take risks or do you try to avoid taking risks?

Mean 6.53, Median 7, Standard deviation 2.26.

2) How would you rate your willingness to take risks while driving a car?

Mean 4.07, Median 3, Standard deviation 2.63.

3) How would you rate your willingness to take risks in financial matters?

Mean 5.55, Median 6, Standard deviation 2.26.

4) How would you rate your willingness to take risks during sports and leisure?

Mean 7.10, Median 8, Standard deviation 2.27.

5) How would you rate your willingness to take risks in job matters?

Mean 5.82, Median 6, Standard deviation 2.38.

6) How would you rate your willingness to take risks in health matters?

Mean 3.70, Median 3, Standard deviation 2.44.

Trust Index: Average of three variables.

Questions taken from the "General Social Survey" consistent with the approach of Glaeser et al. (2000).

1) Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?

"Most people can be trusted" 26.71% or "Can't be too careful" 73.29%.

2) Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?

"Would try to be fair" 26.71% or "Would take advantage of you" 73.29%.

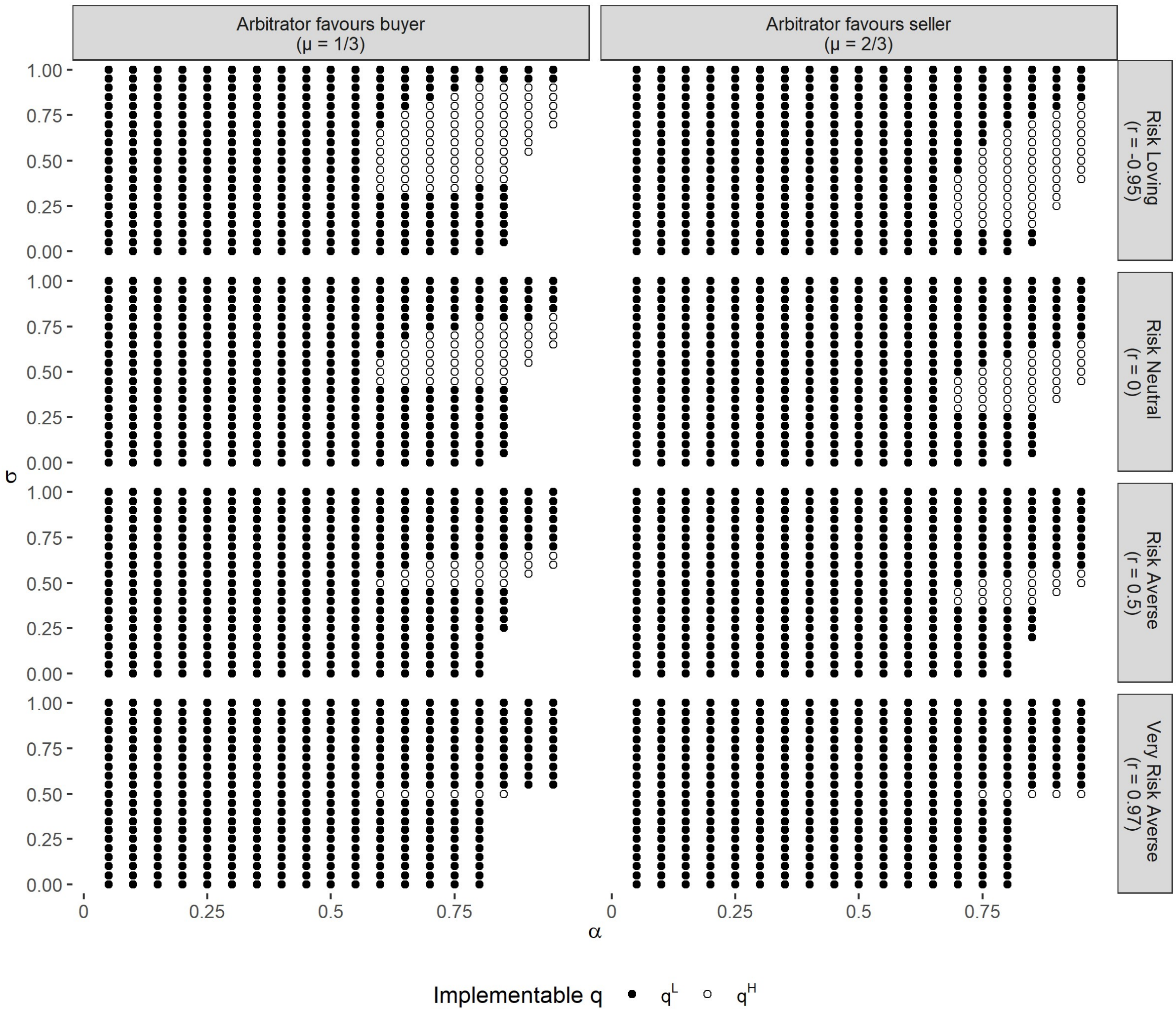
3) Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?

"Try to be helpful" 36.02% or "Just look out for themselves" 63.98%.

B Appendix for Chapter 3

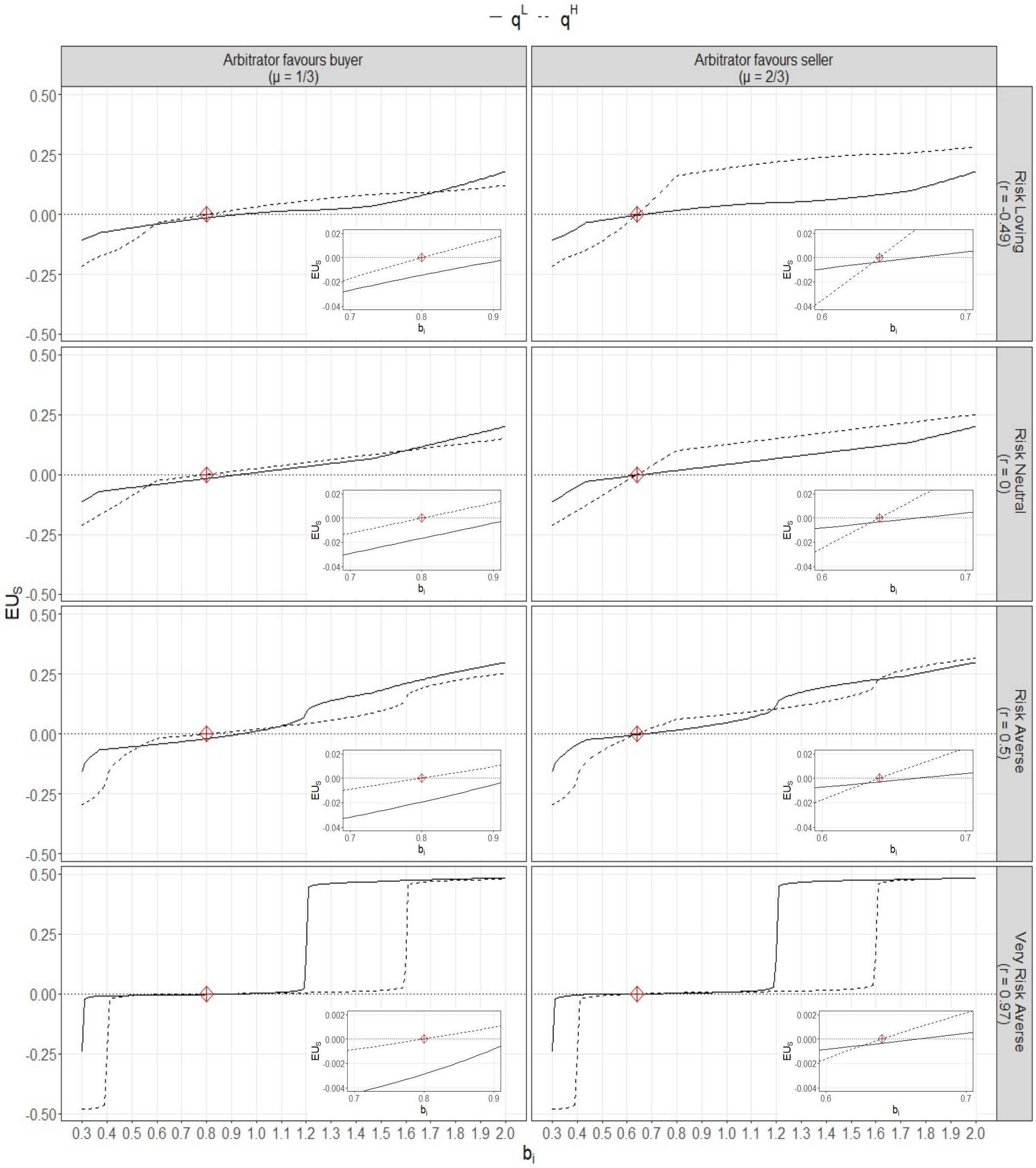
B.1. Additional figures and tables

Figure B. 1. Implementable quality level by contract for Constant Relative Risk Aversion (CRRA) utility function.



Notes: The figure assumes that potential sellers formulate their bidding strategies based on the following CRRA utility function, $u_S = {\pi_S}^{(1-r)}/(1-r)$ for $\pi_S \geq 0$ and $u_S = -(-\pi_S)^{(1-r)}/(1-r)$ for $\pi_S < 0$, where r measures the degree of relative risk aversion. A contract is defined by the probability of quality verification σ and the bid discount α. The empty regions correspond to contracts for which no profitable bidding strategy exists, given the maximum permissible bid in the experiment.

Figure B. 2. Expected seller utilities (CRRA case) conditional on submitting the winning bid in the *Arbitrator* treatment.



Notes: The figure assumes that potential sellers formulate their bidding strategies based on the following CRRA utility function, $u_S = \pi_S{}^{(1-r)}/(1-r)$ for $\pi_S \geq 0$ and $u_S = -(-\pi_S)^{(1-r)}/(1-r)$ for $\pi_S < 0$, where r measures the degree of relative risk aversion. The red diamond plus is the equilibrium bid level. Inset in each panel is the subset of the discrete price grid at the breakeven bid.

Figure B. 3. Bid c.d.f.s in the experiment (supports Results 3.1(a) and 3.1(b)).

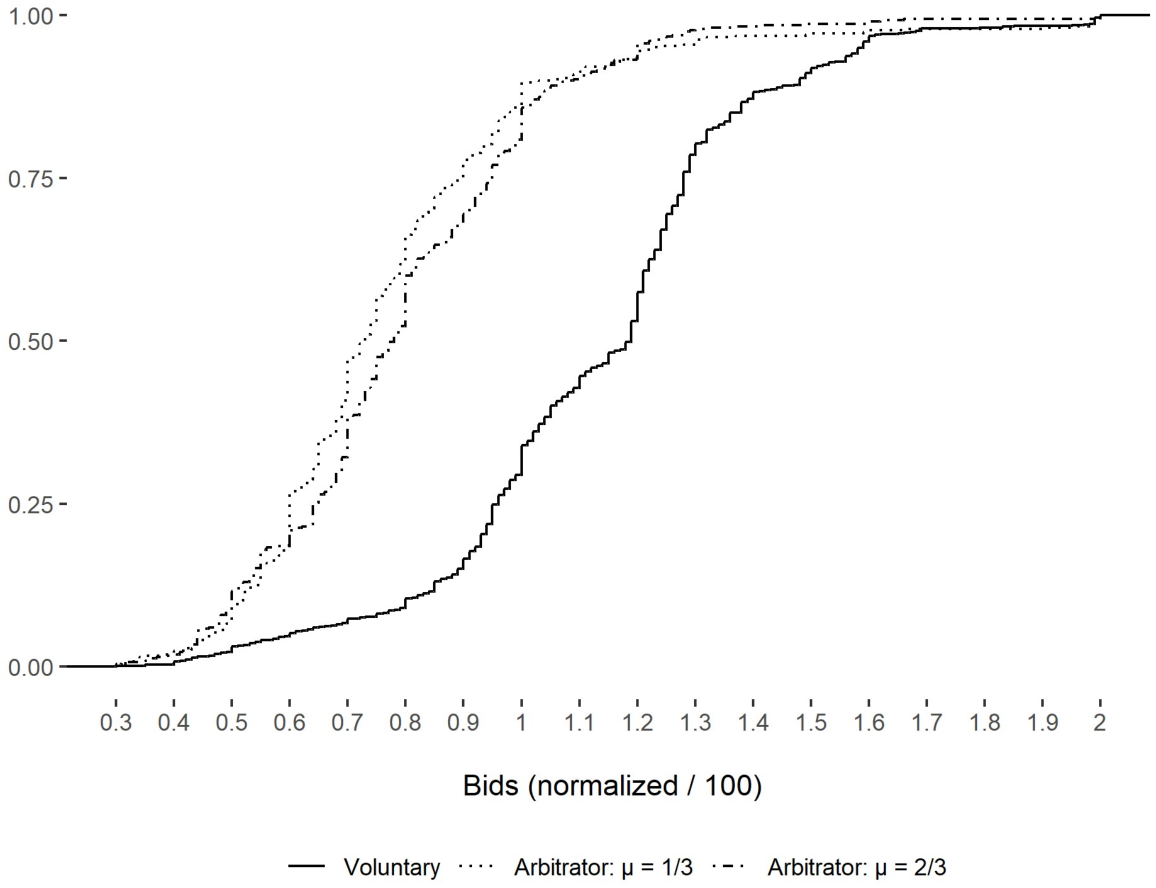
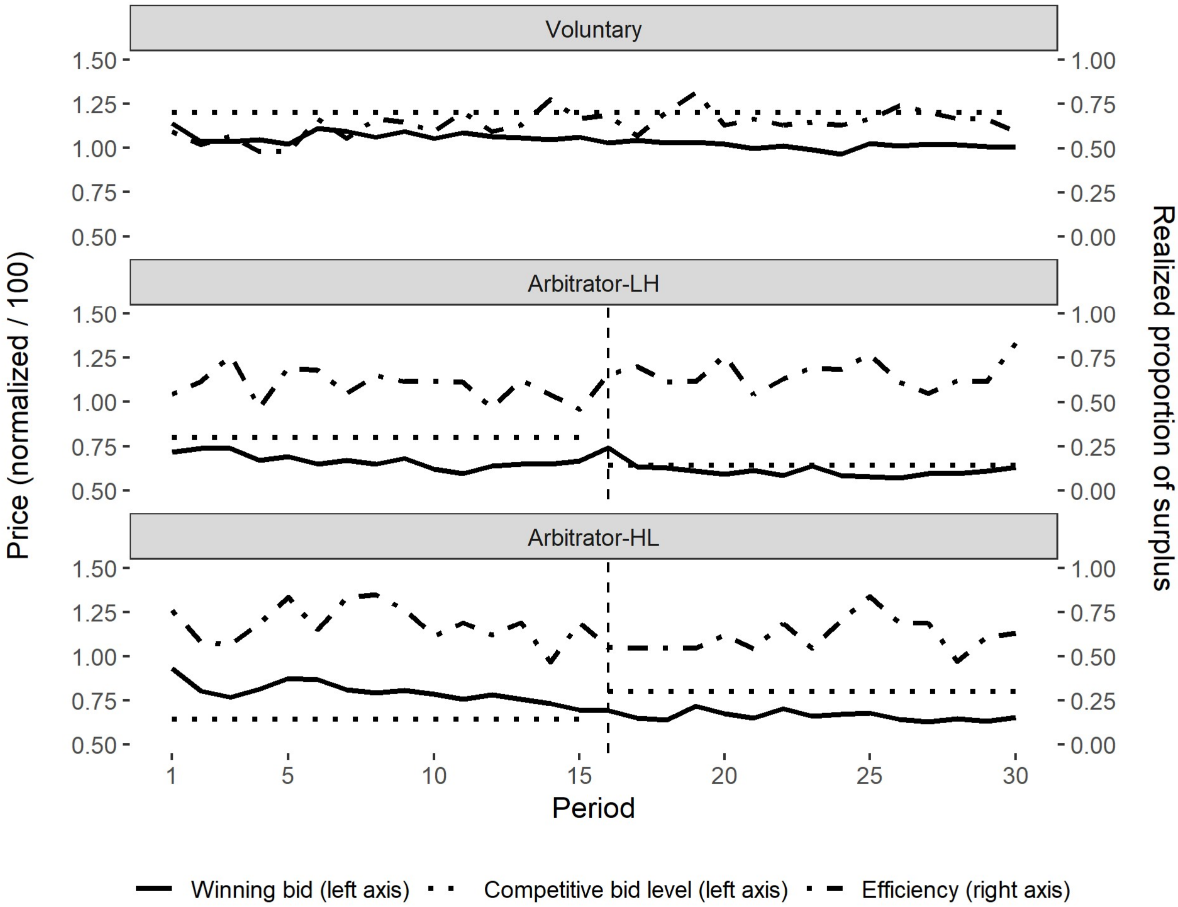
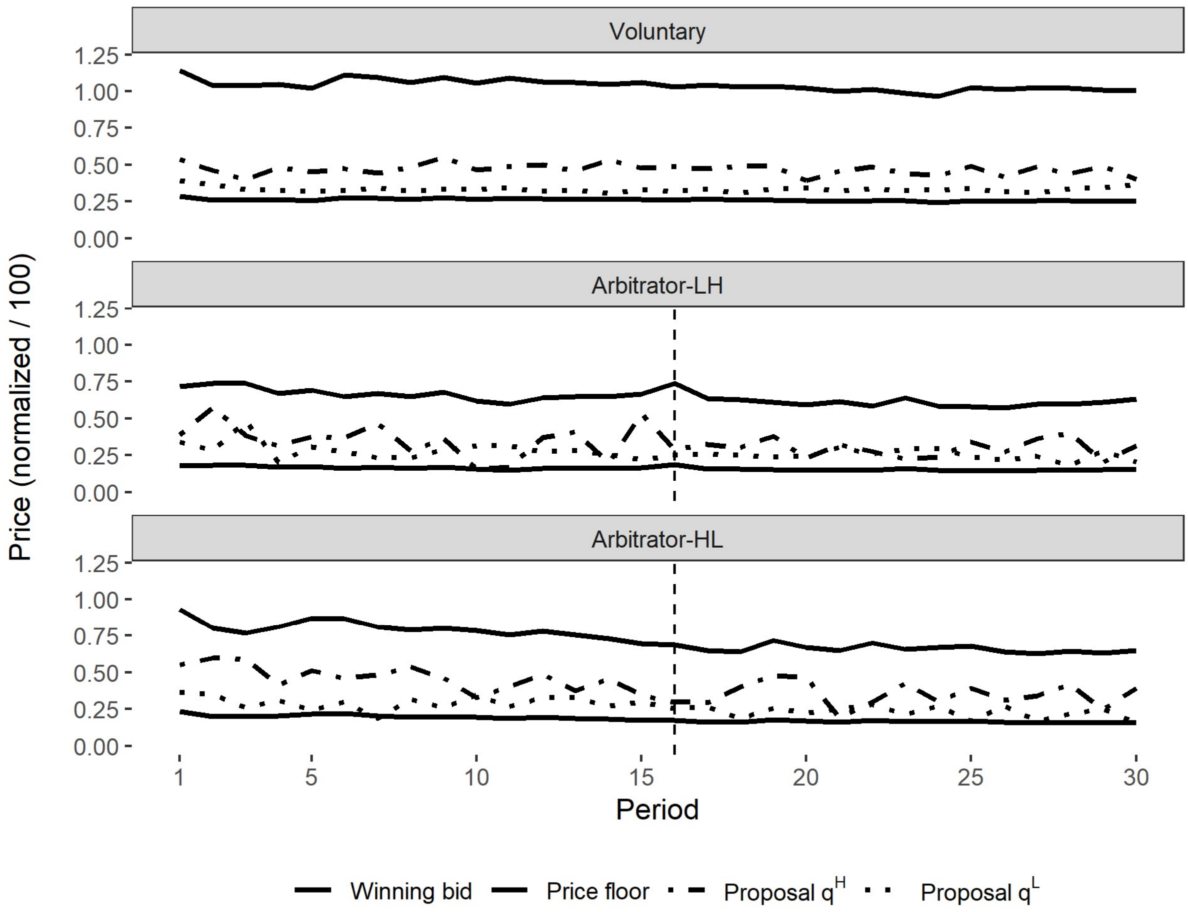


Figure B. 4. Winning bids and trade efficiency over time.



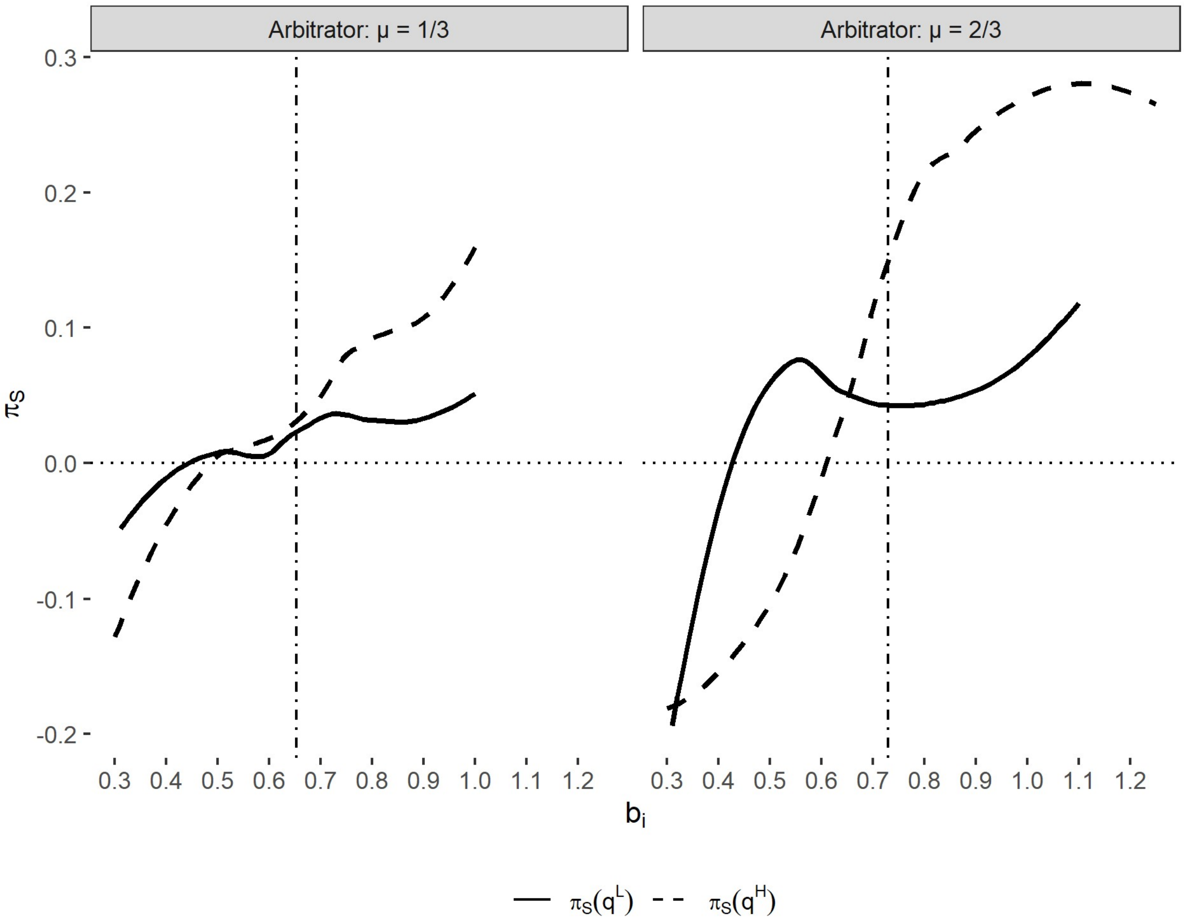
Notes: The vertical dashed line in the *Arbitrator* treatment panels represents the change in the arbitrator's preference parameter μ after period fifteen. *Arbitrator-LH* corresponds to those cohorts assigned to the sequence $\mu = \{1/3, 2/3\}$ and *Arbitrator-HL* to those cohorts assigned to the sequence $\mu = \{2/3, 1/3\}$.

Figure B. 5. Contract prices and proposals over time in the experiment.



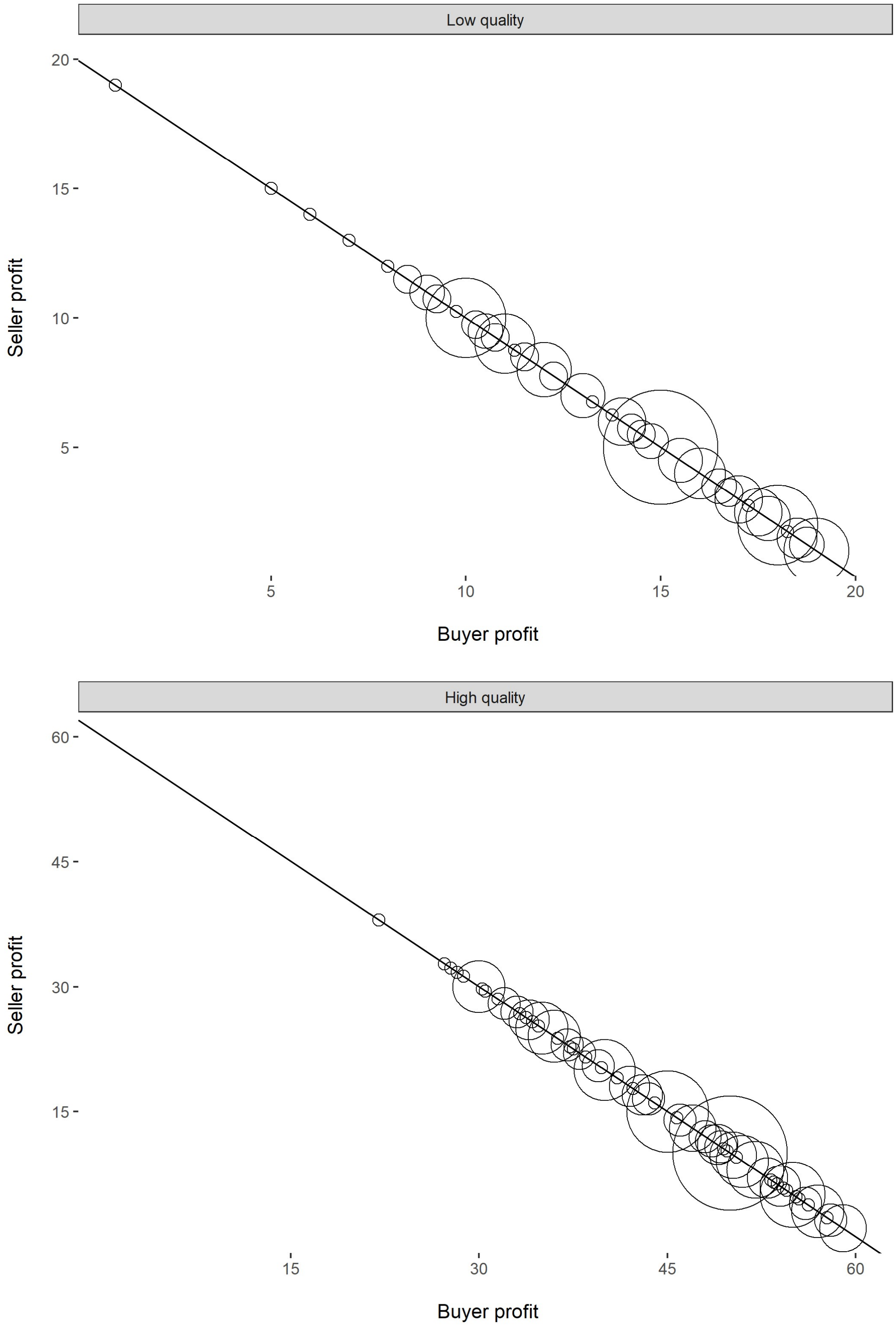
Notes: The vertical dashed line in the *Arbitrator* treatment panels represents the change in the arbitrator's preference parameter μ after period fifteen. *Arbitrator-LH* corresponds to those cohorts assigned to the sequence $\mu = \{1/3, 2/3\}$ and *Arbitrator-HL* to those cohorts assigned to the sequence $\mu = \{2/3, 1/3\}$.

Figure B. 6. Seller profits as a function of the winning bid observed in the *Arbitrator* treatment: High quality is incentive compatible on average.



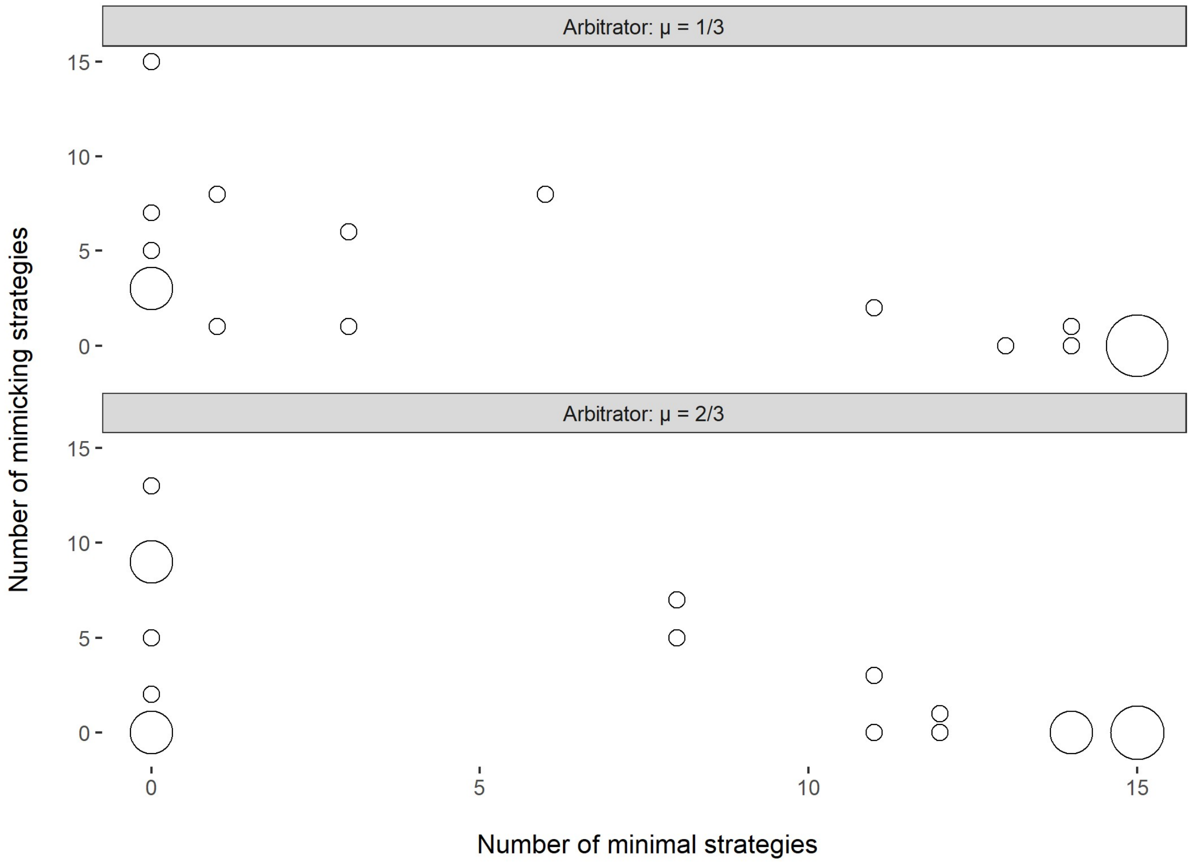
Notes: The vertical dot-dash line is the median winning bid observed in the respective experiment condition.

Figure B. 7. Buyer surplus allocations in the *Voluntary* treatment.



Notes: Based on individual buyer decisions over the course of the experiment. The sizes of the open circles are weighted by the number of surplus allocations at the coordinate.

Figure B. 8. Individual buyer strategy counts in the *Arbitrator* treatment.



Notes: Based on individual buyer decisions over the course of the experiment. The size of the open circles is weighted by the number of buyers at the coordinate. Minimal strategies are defined as buyer choices to propose the lower bound contract price. Mimicking strategies are defined as buyer choices to propose the arbitrator's preferred price +/- 2. We permit a tolerance of two points either side of this price to reflect the granularity of the buyer's price proposal slider. The slider was set at increments of 0.25 points, which did not always permit the arbitrator's preferred price to be selected exactly. The missing categories are price proposals between minimal and mimicking, and price proposals above mimicking.

Table B. 1 – Summary of experiment outcomes from periods one to fifteen only in the *Arbitrator* treatment (relates to behavioural spillover analysis in Section 3.4.1).

Measure / Treatment	*Arbitrator-LH* ($\mu = 1/3$)	*Arbitrator-HL* ($\mu = 2/3$)
Winning bid	71.16	75.00
Quality	0.42	0.56
Proposed seller share	-0.16	0.01
Low quality	-0.19	0.04
High quality	-0.12	-0.04
Buyer profit	31.80	23.19
Low quality	14.50	12.09
High quality	55.04	38.56
Seller profit	4.78	16.01
Low quality	4.86	7.24
High quality	4.12	20.35
Efficiency	60%	69%

Notes: The table summarises the main outcomes of the *Arbitrator* treatment based on data from periods one to fifteen only. *Arbitrator-LH* corresponds to those cohorts assigned to the sequence $\mu = \{1/3, 2/3\}$ and *Arbitrator-HL* to those cohorts assigned to the sequence $\mu = \{2/3, 1/3\}$. This precludes the possibility of behavioural spillover effects between the first and second half of the experiment. Values are median cohort averages, based on three independent cohorts per treatment. Variable definitions: *Winning bid* is the lowest bid submitted by potential sellers in the auction phase; *Quality* is the relative frequency that high quality is chosen for delivery by the winning seller; *Proposed seller share* is the buyer's proposal for the seller's share of the trade surplus (as inferred from the proposed price); *Buyer profit* is the trade profit per round for the buyer; *Seller profit* is the trade profit per round for the winning seller; *Efficiency* is the realised percentage of attainable trade surplus net of any verification cost incurred.

Table B. 2 – Average seller bids by quality level in the experiment.

Treatment	*Voluntary*	*Arbitrator*	
Sub-treatment		$\mu = 1/3$	$\mu = 2/3$
Low quality strategies	118.35	73.29	78.95
High quality strategies	109.71	78.66	85.38

Notes: The table presents seller bids from the strategy method data, by the quality level selected. Values are median cohort averages, based on six independent cohorts per treatment.

Table B. 3 – Frequencies of buyer choices in the *Arbitrator* treatment.

Sub-treatment	$\mu = 1/3$	$\mu = 2/3$
Minimal	0.468	0.508
Low quality	0.471	0.496
High quality	0.465	0.518
Mimicking	0.223	0.207
Low quality	0.232	0.216
High quality	0.211	0.199

Notes: The table displays relative frequencies of buyer choices to propose the lower bound contract price (minimal) and the arbitrator's preferred price +/- 2 (mimicking). We permit a tolerance of two points either side of this price to reflect the granularity of the buyer's price proposal slider. The slider was set at increments of 0.25 points, which did not always permit the arbitrator's preferred price to be selected exactly. The missing categories are price proposals between minimal and mimicking, and price proposals above mimicking.

Table B. 4 – IV first stage estimates: Determinants of quality level in seller bidding strategies (relates to Section 3.4.2).

Dependent variable	*Bid*	
Treatment	*Voluntary*	*Arbitrator*
Model	(5)	(6)
Bid $_{t-1}$	0.626***	0.476***
	(0.094)	(0.092)
Proposed bid discount (ppt) $_{t-1}$	0.286***	0.106***
	(0.073)	(0.030)
Period	-0.248***	-0.432***
	(0.071)	(0.124)
$\mu = 2/3$		-0.079
		(1.848)
OrderHL		-5.627**
		(2.73)
Constant	29.232***	47.442***
	(7.630)	(10.813)
Control variables	Yes	Yes
Cohort fixed effects	Yes	Yes
Observations	512	488
Subjects	36	36
Cohorts	6	6
F test of excluded instruments	43.50***	43.43***
Cragg-Donald Wald F statistic	238.779	181.72

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates, with robust standard errors clustered at the subject level. Cohort fixed effects are included. *Proposed bid discount (ppt)* $_{t-1}$ is one minus the ratio of the buyer's proposed price to the winning bid in the previous period, expressed in percentage points. The following control variables (not shown) are included: dummy for being female; dummy for being an economics and finance major; two Likert scales for self-reported willingness to take risks in general and in financial matters; dummy for reporting trust in strangers; and a generalised trust index. *OrderHL* is a dummy variable to indicate that the cohort followed the sequence $\mu = \{2/3, 1/3\}$ in the *Arbitrator* treatment.

Table B. 5 – Determinants of buyer price proposals.

Dependent variable	Price proposal	
Treatment	*Voluntary*	*Arbitrator*
$(1-\alpha)$ * *Winning bid*	0.493***	1.347***
	(0.077)	(0.168)
High quality	13.168***	9.389***
	(0.649)	(0.988)
Period	-0.037	-0.263***
	(0.036)	(0.062)
$\mu = 2/3$		1.428
		(0.978)
OrderHL		8.940
		(6.475)
Constant	45.201***	-62.141**
	(17.029)	(25.577)
Control variables	Yes	Yes
Random effects (subject and cohort)	Yes	Yes
Observations	535	535
Subjects	18	18
Cohorts	6	6
Wald χ^2	454.6	287.7
χ^2 *p*-value	0.000	0.000
σ_v	0.000	0.000
σ_u	6.881	9.549
σ_ε	7.118	11.09

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The models are estimated using mixed effects linear regression and include two random effects intercepts that capture intra-session correlation at the subject and cohort levels, respectively. Coefficient estimates are presented, with standard errors in parentheses. The following control variables (not shown) are included: dummy for being female; dummy for being an economics and finance major; two Likert scales for self-reported willingness to take risks in general and in financial matters; dummy for reporting trust in strangers; and a generalised trust index. *OrderHL* is a dummy variable to indicate that the cohort followed the sequence $\mu = \{2/3, 1/3\}$ in the *Arbitrator* treatment.

B.2. Robustness check of aggregate experimental findings

In this section of the appendix, we validate Results 3.1 to 3.4 of Chapter 3. using an OLS regression procedure based on disaggregated data at the matching group level. There are 540 matching groups in each treatment. Data from all periods are included and we incorporate a time trend. There are dummies for each of the *Arbitrator* sub-treatments. The *Voluntary* treatment is omitted. Robust standard errors are clustered at the cohort level to control for intra-cohort dependencies – matching groups are randomly reconstituted each round within a cohort. Since the number of clusters is small (six), standard asymptotic inference is liable to over-reject (Cameron et al. 2008). Thus, we also present p-values using wild cluster bootstrap tests of linear hypotheses (Roodman et al. 2019) and these are reported in the text below. The results are contained in Table B. 6.

Column 1 of the table supports Results 3.1(a) and 3.1(b). On average, winning auction bids are 34 to 38 points lower in the respective *Arbitrator* sub-treatments than in the *Voluntary* treatment (both p-values < 0.01). There is weak evidence of a decreasing trend in bids over time (p-value $= 0.098$). There is no significant difference in winning bids between the *Arbitrator* sub-treatments.

Result 3.2 states that there is no significant difference in trade efficiency between the contingent and the non-contingent contract. Since this insight is based on a null finding, it is important to check whether the null result still obtains in the matching group data. This is of particular interest for the comparison between *Voluntary* and *Arbitrator* when $\mu = 2/3$. A power calculation suggests that this pairwise comparison at the cohort-level has only around 13 percent power to detect an effect size of the magnitude observed in the data.[1] Column 2 of the table supports Result 3.2. There is no significant increase in trade efficiency with a contingent contract than without ($\mu = 1/3$, p-value $= 0.399$; $\mu = 2/3$, p-value $= 0.699$). This does not change significantly over time in the

[1] Mean (sd) in *Voluntary* of 0.639 (0.096) and in *Arbitrator* $\mu = 2/3$ of 0.674 (0.098). All power calculations are based on 1000 replications and assume a normal distribution.

experiment (p-value = 0.518). We fail to reject a linear hypothesis test for equality of the estimates on the *Arbitrator* sub-treatment dummies (p-value = 0.404).

Column 3 of the table supports Result 3.3. Buyer price proposals as a proportion of trade surplus are significantly lower with a contingent contract, both when $\mu = 1/3$ (p-value < 0.01) and when $\mu = 2/3$ (p-value = 0.011). This is consistent with a crowding out of buyer reciprocity. The period indicator in column 3 is also negative and significant (p-value < 0.01), which indicates decreasing levels of buyer reciprocity over time.

Table B. 6 – Regression analysis of matching group data.

Dependent variable	*Winning bid*	*Trade efficiency*	*Proposal / surplus*	*Buyer Profit*	*Seller Profit*
Validation of Result	3.1	3.2	3.3	3.4	3.4
Arbitrator ($\mu = 1/3$)	-37.63***	-0.151	-0.308***	-0.431	-1.70
	(7.56)	(0.162)	(0.039)	(2.32)	(1.05)
	[0.009]	[0.399]	[0.006]	[0.870]	[0.153]
Arbitrator ($\mu = 2/3$)	-33.52***	0.157	-0.215***	-1.84	3.96
	(8.28)	(0.335)	(0.039)	(3.33)	(2.53)
	[0.013]	[0.699]	[0.011]	[0.597]	[0.230]
Period	-0.46	0.008	-0.008***	0.373**	-0.258**
	(0.28)	(0.103)	(0.001)	(0.108)	(0.097)
	[0.081]	[0.518]	[0.005]	[0.007]	[0.077]
Constant	111.26***	0.440***	0.262***	27.48***	9.050***
	(3.51)	(0.085)	(0.050)	(1.58)	(1.89)
Observations	1,080	1,070	1,070	1,070	1,070
Wald test statistic	41.61***	26.97***	51.76***	9.40**	4.50*

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates are presented, with robust standard errors in parentheses clustered at the cohort level to correct for intra-cohort correlation. The omitted experimental condition is the *Voluntary* treatment. In square brackets are wild cluster bootstrap p-values after imposing the zero null on the bootstrap data generating process, with 999 bootstrap samples drawing from the six-point Webb distribution – which is preferred to alternative distributions when the number of clusters is less than ten – obtained using the *boottest* command in Stata (Roodman et al. 2019). The models in columns 1, 3, 4 and 5 are estimated using OLS regression. The model in column 2 was estimated using fractional logistic regression and coefficient estimates are on the logit scale. The Wald test statistic is based on a test of joint regression significance (F-statistics except for column 2, which is a chi-squared statistic). The number of observations is higher in the first column because we have auction data from all matching groups, whereas trade data from 10 matching groups was lost due to a hard time-out protocol.

Columns 4 and 5 of the table are in line with Result 3.4. Buyer profits do not differ significantly with a contingent contract than without and there is no significant difference between the *Arbitrator* sub-treatments (p-value = 0.625). Of more interest is the comparison in seller profits between *Voluntary* and *Arbitrator* when $\mu = 2/3$. There is little statistical evidence at the aggregate level to suggest that the contingent contract increases seller profits under these conditions, but this test only has around 23 percent power to detect an effect size of the magnitude observed.[2] At the matching group level, the evidence for a positive effect is stronger but still not significant at conventional thresholds (p-value = 0.230). We also fail to reject a linear hypothesis test that the coefficient estimates on the dummy variables for the *Arbitrator* sub-treatments are equal in column 5 (p-value = 0.163). The buyer-seller profit differential widens over the course of the experiment, with a significant positive time trend in the specification for buyer profits (p-value < 0.01) and a weak significant negative time trend in the specification for seller profits (p-value = 0.077).

B.3. Mixture model estimation procedure

The finite mixture model estimation procedure used here is adapted from Moffatt (2015). There are 18 buyers in the *Arbitrator* treatment, each of whom is observed over 30 trading periods. Buyer i's price proposal in period t is given by y_{it}. The variable y_{it} is constrained by the contract price range, $[\underline{p}_{it}, \overline{p}_{it}]$, which is determined in the auction phase of the period and is observed before the buyer makes her decision. The corresponding arbitrator's preferred price is:

$$p_{it}^{A} = max\{\underline{p}_{it}, min\ \{\overline{p}_{it},\ \mu_t v(q_{it}) + (1-\ \mu_t)c(q_{it})\}\}. \tag{B.1}$$

This price is determined by the seller's product quality choice q_{it} and the arbitrator's exogeneous preference parameter μ_t, which varies during the experiment.

The buyer's latent preferred price is y_{it}^{*}, which depends on her type. We assume that there are two buyer types: self-interested (*self*) and fair (*fair*). The model precludes subjects from switching between types. Self-interested buyers propose the lower bound

[2] Mean (sd) in *Voluntary* of 5.09 (3.85) and in *Arbitrator* $\mu = 2/3$ of 9.06 (9.27).

price $\underline{p}_{it}$ (i.e., the maximum bid discount). Fair buyers exhibit reference-dependent preferences. For this type, we specify the following model:[3]

$$y_{it}^{*} = \beta_0 + \beta_1 p_{it}^{A} + \varepsilon_{it,fair} \tag{B.2}$$

$$i = 1, \ldots, 18, \qquad t = 1, \ldots, 30, \qquad \varepsilon_{it,fair} \sim N(0, s^2).$$

Where the arbitrator's interim preferred price is the fair reference point in a trading period. We expect the parameter β_1 to be close to one for fair buyers. The error term is assumed to be normally distributed, with variance s^2.

A two-limit tobit model (Nelson 1976) is appropriate, where the limits correspond to the lower and upper bounds of the contract price range and are specific to each buyer and period. The relationship between the buyer's preferred price proposal and actual price proposal is determined by the following censoring rules:

For self-interested buyers:

$$y_{it} = \underline{p}_{it} \quad \forall t \tag{B.3a}$$

For fair buyers:

$$y_{it} = \begin{cases} \underline{p}_{it} & if\ y_{it}^{*} \leq \underline{p}_{it} \\ y_{it}^{*} & if\ \underline{p}_{it} < y_{it}^{*} < \underline{p}_{it} \\ \overline{p}_{it} & if\ y_{it}^{*} \geq \overline{p}_{it} \end{cases} \tag{B.3b}$$

We consider the possibility that buyers deviate on occasion from their preferred price by incorporating a tremble parameter ω. In any given period, with probability ω a buyer may lose concentration and choose a price proposal at random. The price proposal increment in the experiment is set at 0.25 points and so the cardinality of the set of possible price proposals is $|Y_{it}| = 1 + (\overline{p}_{it} - \underline{p}_{it})/0.25$. The importance of this

[3] We found no significant time trend and so a period indicator was dropped from the model.

parameter is likely decreasing during the experiment as subjects gain experience in the decision-making environment and so we employ the following specification:

$$\omega_{it} = \omega_0 \exp\left[\omega_1 (Period)\right] \tag{B.4}$$

Where ω_0 is the initial tremble probability and ω_1 is the rate of decline. The tremble parameter estimates are not statistically different from zero and so we do not report on them further here. The likelihood contributions for a single price decision are as follows, where $\Phi(.)$ and $\phi(.)$ are the standard normal cumulative distribution and probability density functions respectively:

$\boldsymbol{y = \underline{p}}$:

$$P(y_{it} = \underline{p}_{it} | i = self) = 1 - \frac{\omega_{it}}{|Y_{it}|}$$

$$P(y_{it} = \underline{p}_{it} | i = fair) = (1 - \omega_{it})\Phi\left(-\frac{\beta_0 + \beta_1 p_{it}^{A\prime}}{s}\right) + \frac{\omega_{it}}{|Y_{it}|} \tag{B.5a}$$

$\boldsymbol{\underline{p} < y < \overline{p}}$:

$$f(y_{it} | i = self) = \frac{\omega_{it}}{|Y_{it}|}$$

$$f(y_{it} | i = fair) = (1 - \omega_{it})\frac{1}{s}\phi\left(\frac{y_{it} - \beta_0 - \beta_1 p_{it}^{A\prime}}{s}\right) + \frac{\omega_{it}}{|Y_{it}|} \tag{B.5b}$$

$\boldsymbol{y = \overline{p}}$:

$$P(y_{it} = \overline{p}_{it} | i = self) = \frac{\omega_{it}}{|Y_{it}|}$$

$$P(y_{it} = \overline{p}_{it} | i = fair) = (1 - \omega_{it})\left[1 - \Phi\left(\frac{\overline{p}_{it} - \beta_0 - \beta_1 p_{it}^{A\prime}}{s}\right)\right] + \frac{\omega_{it}}{|Y_{it}|} \tag{B.5c}$$

The two buyer types are represented in the model by the mixing fractions γ_{self} and γ_{fair}. The likelihood contribution for a buyer i is:

$$L_i = \gamma_{self} \prod_{t=1}^{30} P(y_{it} = \underline{p}_{it} | self)^{I_{y_{it}=\underline{p}}} f(y_{it} | self)^{I_{\underline{p}<y_{it}<\overline{p}}} P(y_{it} = \overline{p}_{it} | self)^{I_{y_{it}=\overline{p}}}$$

$$+ \gamma_{mimic} \prod_{t=1}^{30} P(y_{it} = \underline{p}_{it} | fair)^{I_{y_{it}=\underline{p}}} f(y_{it} | fair)^{I_{\underline{p}<y_{it}<\overline{p}}} P(y_{it} = \overline{p}_{it} | fair)^{I_{y_{it}=\overline{p}}} \tag{B.6}$$

Where $I(.)$ is an indicator function for the subscripted expression and the conditional probabilities/densities are obtained from (B.5).

Finally, the sample log-likelihood is obtained as follows:

$$Log\ L = \sum_{i=1}^{18} \log\ (L_i) \tag{B.7}$$

on maximization of which we obtain maximum likelihood estimates for the five parameters $\beta_0, \beta_1, s, \omega_0,\ \omega_1$ and one of the two mixing fractions (from which the other is computed using the delta method).

Posterior probabilities for the two types are calculated as:

$$P(i = self|y_{i1,\ldots,i30}) = \frac{\gamma_{self} \prod_{t=1}^{30} P(y_{it} = \underline{p}_{it}|self)^{I_{y_{it}=\underline{p}}} f(y_{it}|self)^{I_{\underline{p}<y_{it}<\overline{p}}} P(y_{it} = \overline{p}_{it}|self)^{I_{y_{it}=\overline{p}}}}{L_i} \tag{B.8a}$$

$$P(i = fair|y_{i1,\ldots,i30}) = \frac{\gamma_{fair} \prod_{t=1}^{30} P(y_{it} = \underline{p}_{it}|fair)^{I_{y_{it}=\underline{p}}} f(y_{it}|fair)^{I_{\underline{p}<y_{it}<\overline{p}}} P(y_{it} = \overline{p}_{it}|fair)^{I_{y_{it}=\overline{p}}}}{L_i} \tag{B.8b}$$

The estimation is conducted in Stata 16, using the d0 estimator to account for the panel structure of the data. A starting value for the mixing fraction is obtained from the frequency table of buyer choices in Table B. 3. Starting values for the five parameters are (0, 1, 4, 0.11, -0.05), obtained from a process of trial and error. Bootstrapped standard errors are calculated, clustered at the cohort level based on 999 bootstrap replications.

B.4. Experimental instructions

The experiment video instructions are available online at https://github.com/mjwalker19/Trade-Contingencies-in-Procurement-Interactions. A written version is provided below. Note: Horizontal sliders were used for the suppliers' bidding decision and the shipping company's price proposal. The initial values of these sliders were set at random in each interaction to avoid anchoring bias.

You are now participating in an experiment in the economics of decision making. Based on your decisions and the decisions of other participants in the experiment, you can earn money which will be paid to you in private and in cash. Read these instructions carefully.

Please turn to silent mode your cell phone or any other electronic device that you have brought with you. These electronic devices must be stored out of sight and not used for the duration of the experiment. Please do not communicate with other participants. If at any point you have a question, raise your hand and we will answer you as soon as possible.

<u>How you earn money</u>

The experiment today consists of **30 decision-making rounds** . You will be paid for all these rounds. You earn money during the experiment by accruing points in each round. All the points that you earn during the experiment are converted to pounds sterling at the following rate:

40 points = £1

At the end of the experiment, you will be asked to complete a short questionnaire. Upon completion, you will immediately receive the monetary amount that you earned. The minimum amount that you will leave today's session with is £4.

<u>Summary of the experiment procedure</u>

For today's session, one third of you will be randomly assigned to the role of a **shipping company**, and two thirds of you assigned to the role of a **supplier**. You will be informed to which role you are assigned before the first round. You remain in the same role for the entire experiment.

In each of the 30 rounds, a shipping company is matched with two suppliers in the room. This is called a matching group. The shipping company is looking to buy an engine from **one** of the two suppliers in his/her matching group.

Note: Your matching group changes in each round. You are not matched with the same persons in consecutive rounds.

All subjects assigned to a shipping company role begin each round with an endowment of 5 points. All subjects assigned to a supplier role begin each round with an endowment of 10 points.

---- *Voluntary*----

There are two phases in each round:

1. Auction phase
2. Price-setting phase

----*Arbitrator*----

There are three phases in each round:

1. Auction phase
2. Price proposal phase
3. Arbitration phase.

In the auction phase, it is decided which supplier delivers the engine. Each supplier submits a bid. The **lower** bid wins the auction. The winning supplier at the auction delivers the engine. The losing supplier at the auction exits the round.

The engine delivered to the shipping company can be of **high or low quality**. This is decided by the winning supplier during the auction and determines the shipping company's value for the engine and the winning supplier's delivery cost (see Table 1).

Table - 1: Engine value and cost schedule

	Low quality	High quality
Shipping company's value	50	100
Winning supplier's cost	30	40

---- *Voluntary*----

In the price-setting phase, the shipping company observes the engine quality delivered. Then, the shipping company chooses a final price to pay the winning supplier. This price can be any amount from **one quarter** of the winning auction bid up to and including the winning auction bid.

At the end of the round, the points earned are calculated as follows:

Shipping company earns 5 + Profit, where Profit = value – final price Winning supplier earns 10 + Profit, where Profit = final price – cost Losing supplier earns 10 (Profit = 0)

Note: Profit can be negative. Make your decisions carefully.

We now explain the experiment procedure in detail.

Auction phase:

- The two suppliers in a matching group each submit a bid at an auction.
- This bid can be any whole number from 30 up to and including 200.
- Neither of the two suppliers can see the other's choice of bid.
- The supplier that submits the lower bid wins the auction. In the event of a tie, the computer chooses the winner at random.
- The winning bid is not necessarily the final price paid by the shipping company. Instead, the winning bid determines the price range that the winning supplier can receive:
  - ➔ the minimum price in this range is one quarter of the winning bid.
  - ➔ the maximum price in this range is the winning bid.
- At the same time as submitting a bid, each supplier also selects an engine quality (high or low) to be delivered in the event of winning the auction.
- The winning supplier at the auction proceeds to the next phase and incurs his/her engine delivery cost.
- The losing supplier at the auction exits and earns the round endowment of 10 points (a profit of zero).

Computer interface:

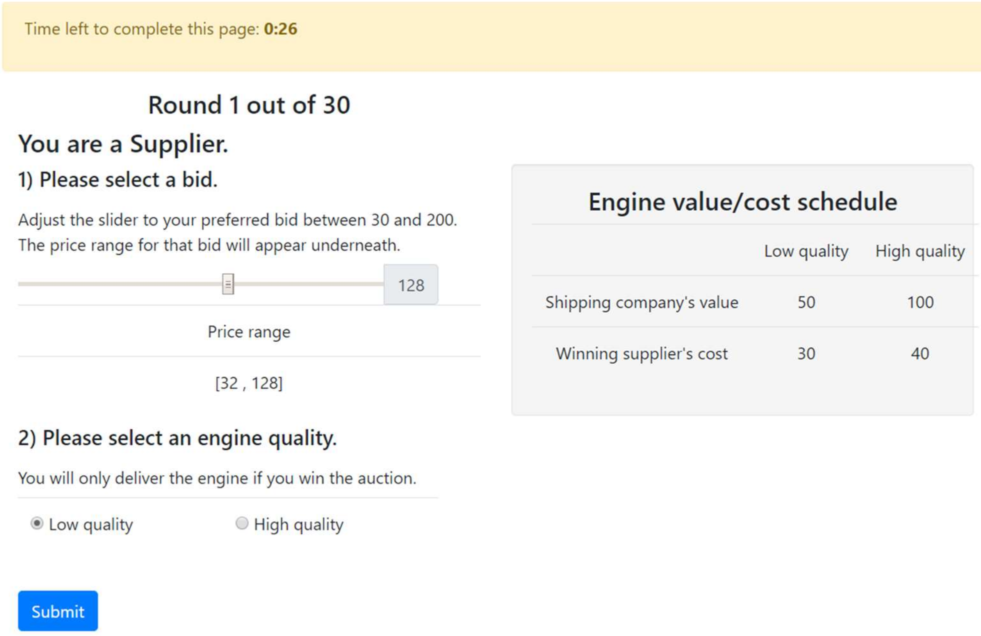
In the auction phase, each supplier submits his/her bid by adjusting a slider in the upper part of the computer screen. As you adjust the slider, underneath you see the price ranges associated with the different possible bids. The first number in the brackets is the minimum price, which is one quarter of the selected bid. The second number is the maximum price, which is the bid itself.

A supplier submits his/her quality by selecting one of two options in the lower part of the screen.

On the right-hand side of the screen appears the engine value and cost schedule.

Suppliers should attempt to make their decisions in the prescriptive time. For the first five rounds, you have 60 seconds to make your decisions. After round five, you have 30 seconds to make your decisions. If you cannot make your decisions in time, the default bid is 200 and the engine quality is selected at random.

Auction phase computer screen:



Consider the following examples for the auction phase:

1. Supplier A bids 160 and selects low quality. Supplier B bids 170 and selects high quality.

    ➔ Supplier A wins the auction.

    ➔ The winning bid is 160.

    ➔ The price range is [40, 160].

    ➔ The engine is low quality

2. Suppliers A and B both bid 128. Supplier A selects high quality and B selects low quality.

    ➔ Suppliers A and B each have 50% probability of winning the auction.

    ➔ The winning bid is 128.

    ➔ The price range is [32, 128].

    ➔ If supplier A wins, the engine is high quality

    ➔ If supplier B wins, the engine is low quality.

Price-setting phase:

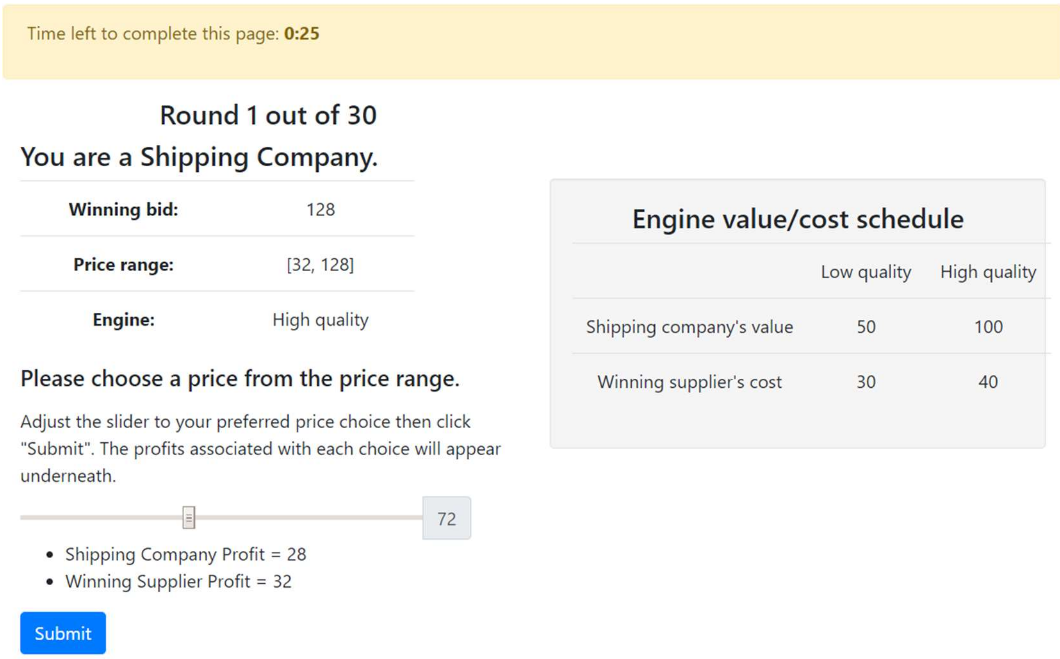
- The shipping company is informed about the price range and the engine quality selected by the winning supplier.
- The shipping company is not informed about the quality selected by the losing supplier.
- The shipping company receives his/her engine value.
- The shipping company chooses a final price from the price range to pay the winning supplier.

Computer interface:

In the price-setting phase, the shipping company's price choice is submitted by adjusting a slider. As you adjust the slider, underneath you will see the profits associated with each price.

The shipping company should attempt to make his or her decision in the prescriptive time. For the first five rounds, you have 60 seconds to make your decision. After round five, you have 30 seconds to make your decision. If you cannot make your decision in time, the default price choice is the maximum price in the price range.

Price-setting phase computer screen:



At the conclusion of the price-setting phase, the round ends and profits are realised.

Consider the following example scenarios:

1. The winning bid is 128 and so the price range is [32, 128]. The winning supplier selected a low quality engine. If the shipping company chooses a final price of 45:

   ➔ Shipping Company Profit = 50 – 45 = 5

   ➔ Winning Supplier Profit = 45 – 30 = 15

   ➔ Losing Supplier Profit = 0

2. The winning bid is 128 and so the price range is [32, 128]. The winning supplier selected a high quality engine. If the shipping company chooses a final price of 65:

   ➔ Shipping Company Profit = 100 – 65 = 35

   ➔ Winning Supplier Profit = 65 – 40 = 25

   ➔ Losing Supplier Profit = 0

Information Display

Much of the information displayed to you during the experiment is stored in a table in the lower portion of the computer screen. Here is an example of this table filled with made up numbers for demonstration purposes. The information shown is not from an actual experiment with people.

Round	Winning Bid	Losing Bid	Quality	Final Price	Profit	Balance
3	45	185*			0.00	**56.00**
2	75*	131	Low	34.00	4.00	**56.00**
1	128*	175	High	72.00	32.00	**42.00**

*Your Bid

The table provides detailed information for each round. Suppliers see the winning and losing auction bids and which of these is their own bid. The shipping company and winning supplier also see the winning supplier's engine quality delivered and the final price. All participants see their profit for the round and total accrued points balance.

Comprehension quiz

Please answer the questions below. Raise your hand if you require assistance. Once you answer all the questions correctly, you will be assigned your role and guided through two training rounds. The main part of the experiment then begins. Note: The comprehension questions and training rounds have no influence on your payment.

- Consider the following scenario.

The winning auction bid is 140 and so the price range is [35, 140]. First, suppose the winning supplier selected a low quality engine, which is valued by the shipping company at 50 and costs the winning supplier 30. The shipping company chooses a final price of 44.

1. What is the Shipping Company Profit? *6*

2. What is the Winning Supplier Profit? *14*

3. What is the Losing Supplier Profit? *0*

Now suppose the winning supplier selected a high quality engine, which is valued by the shipping company at 100 and costs the winning supplier 40. The shipping company still chooses a final price of 44.

4. What is the Shipping Company Profit? *56*

5. What is the Winning Supplier Profit? *4*

---- *Arbitrator* ----

In the price proposal phase, the shipping company observes the engine quality delivered. Then, the shipping company proposes a final price to pay the winning supplier. The shipping company's proposed price can be any amount from **one quarter** of the winning auction bid up to and including the winning auction bid.

In the arbitration phase, an appeal is triggered if the shipping company's proposed price is below the price that an arbitrator would set. The arbitrator's price depends on its favoured profit ratio, which changes during the experiment. Whenever an appeal is triggered, the arbitrator is available with **50% probability**. If the arbitrator is available, the shipping company must pay the arbitrator's price instead of the price proposal. Precise details will be given below.

At the end of the round, the points earned are calculated as follows:

Shipping company earns 5 + Profit, where Profit = value – final price
Winning supplier earns 10 + Profit, where Profit = final price – cost
Losing supplier earns 10 (Profit = 0)

Note: Profit can be negative. Make your decisions carefully.

We now explain the experiment procedure in detail.

Auction phase:

- The two suppliers in a matching group each submit a bid at an auction
- This bid can be any whole number from 30 up to and including 200.
- Neither of the two suppliers can see the other's choice of bid.
- The supplier that submits the **lower** bid wins the auction. In the event of a tie, the computer chooses the winner at random.
- The winning bid is not necessarily the final price paid by the shipping company. Instead, the winning bid determines the **price range** that the winning supplier can receive:
  - ➔ the **minimum price** in this range is **one quarter** of the winning bid.
  - ➔ the **maximum price** in this range is the **winning bid**.
- At the same time as submitting a bid, each supplier also selects an engine quality (high or low) to be delivered in the event of winning the auction.

- The winning supplier at the auction proceeds to the next phase and incurs his/her engine delivery cost.
- The losing supplier at the auction exits and earns the round endowment of 10 points (a profit of zero).

Computer interface:

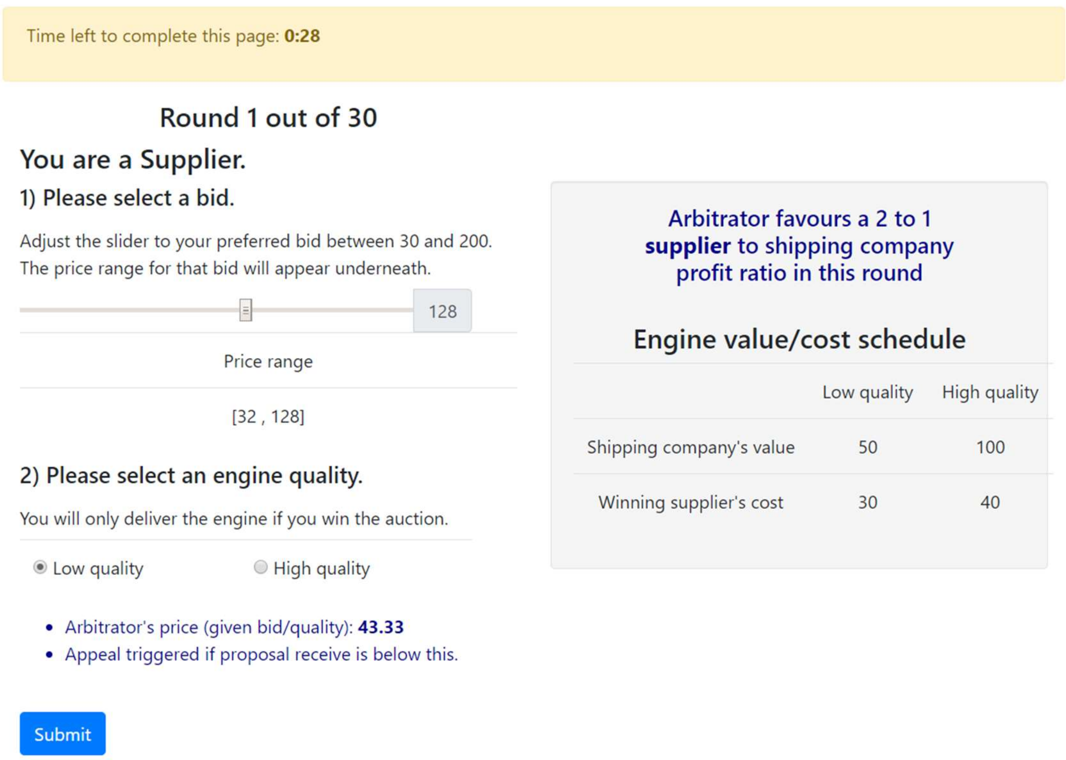
In the auction phase, each supplier submits his/her bid by adjusting a slider in the upper part of the computer screen. As you adjust the slider, underneath you see the price ranges associated with the different possible bids. The first number in the brackets is the minimum price, which is one quarter of the selected bid. The second number is the maximum price, which is the bid itself.

A supplier submits his/her quality by selecting one of two options in the lower part of the screen. As you click on each option, underneath you see the arbitrator's price associated with your combination of bid and quality level selected.

On the right-hand side of the screen appears the arbitrator's favoured profit ratio and the engine value and cost schedule.

Suppliers should attempt to make their decisions in the prescriptive time. For the first five rounds, you have 60 seconds to make your decisions. After round five, you have 30 seconds to make your decisions. If you cannot make your decisions in time, the default bid is 200 and the engine quality is selected at random.

Auction phase computer screen:



Consider the following examples for the auction phase:

1. Supplier A bids 160 and selects low quality. Supplier B bids 170 and selects high quality.

   - ➔ Supplier A wins the auction.
   - ➔ The winning bid is 160.
   - ➔ The price range is [40, 160].
   - ➔ The engine is low quality

2. Suppliers A and B both bid 128. Supplier A selects high quality and B selects low quality.

   - ➔ Suppliers A and B each have 50% probability of winning the auction.
   - ➔ The winning bid is 128.
   - ➔ The price range is [32, 128].
   - ➔ If supplier A wins, the engine is high quality
   - ➔ If supplier B wins, the engine is low quality

<u>Price proposal phase:</u>

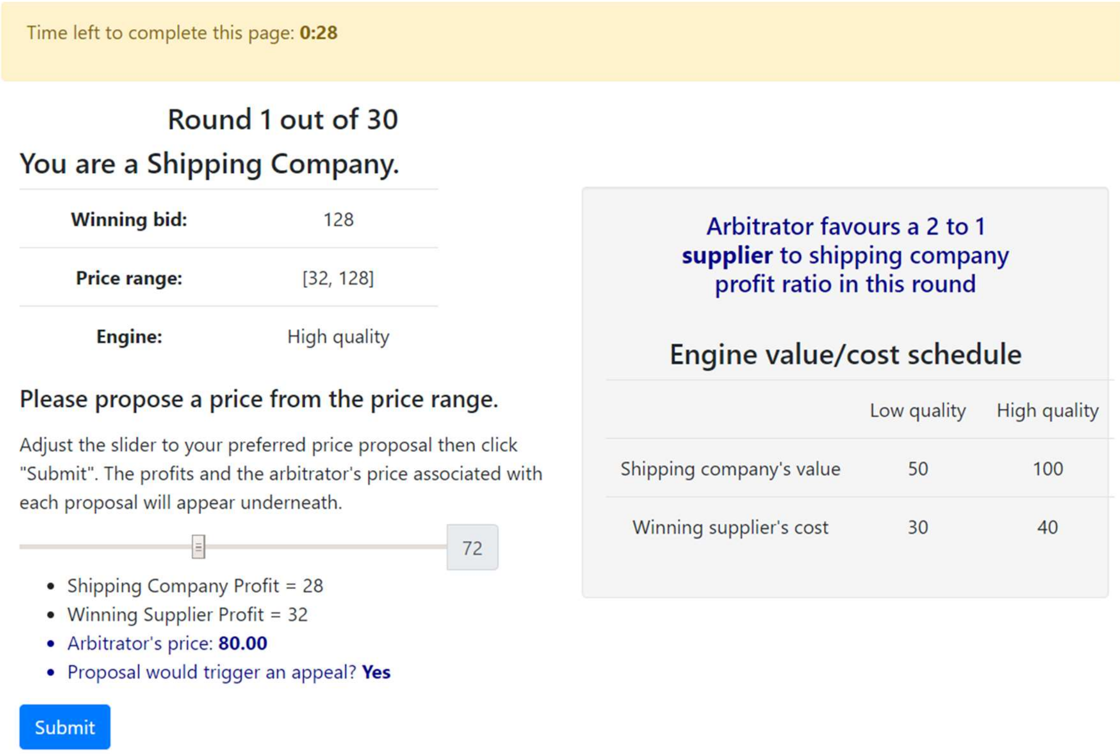
- The shipping company is informed about the price range and the engine quality selected by the winning supplier.
- The shipping company is not informed about the quality selected by the losing supplier.
- The shipping company receives his/her engine value.
- The shipping company proposes a final price from the price range to pay the winning supplier.

Computer interface:

In the price proposal phase, the shipping company's proposed price is submitted by adjusting a slider. As you adjust the slider, underneath you will see the profits associated with each price. You will also see the price that the arbitrator would set for each proposal and whether an appeal to the arbitrator would be triggered.

The shipping company should attempt to make his or her decision in the prescriptive time. For the first five rounds, you have 60 seconds to make your decision. After round five, you have 30 seconds to make your decision. If you cannot make your decision in time, the default price proposal is the maximum price in the price range.

Price proposal phase computer screen:



<u>Arbitration phase:</u>

- The outcome of the arbitration phase depends on the decisions taken in the first two phases.
- An arbitrator observes the winning bid, the shipping company's proposed price and the proposed shipping company and winning supplier profits.
- The arbitrator then sets its own price. The arbitrator cannot set any price, however. The arbitrator's price can be from between the shipping company's proposed price and the winning bid.
- The arbitrator's price is set to bring profits as close as possible to a favoured profit ratio, which varies between the first half and the second half of the experiment:
  - ➔ Either the arbitrator favours a 2 to 1 shipping company to supplier profit ratio, i.e., the arbitrator favours an outcome in which the shipping company's profit is double the winning supplier's profit.
  - ➔ Or the arbitrator favours a 2 to 1 supplier to shipping company profit ratio, i.e., the arbitrator favours an outcome in which the winning supplier's profit is double the shipping company's profit.

- You will be informed about the arbitrator's favored profit ratio immediately before the first round, and again when this ratio changes at the beginning of the sixteenth round.
- If the shipping company's proposed price is below the arbitrator's price, then an appeal to the arbitrator is automatically triggered. If no appeal is triggered, then the final price remains the shipping company's proposed price.
- Following an appeal, the computer rolls a standard six-sided die to determine whether the arbitrator is available:
  - ➔ If the die comes up 1, 2 or 3, the arbitrator is unavailable, and the final price paid by the shipping company remains the shipping company's proposed price.
  - ➔ If the die comes up 4, 5 or 6, the arbitrator is available, and the final price paid by the shipping company is the arbitrator's price. In this situation, a 2 point arbitrator fee is levied on the shipping company.

At the conclusion of the arbitration phase, the round ends and profits are realised.

Consider the following example scenarios:

1. The winning bid is 128 and so the price range is [32, 128]. The winning supplier selected a low quality engine. The shipping company proposes a price of 39.

   If the arbitrator favours a 2 to 1 shipping company to supplier profit ratio, then the arbitrator's price is also 39, because the shipping company's proposed price already gives the winning supplier more than half the shipping company's profit. No appeal is triggered.

   - ➔ Shipping Company Profit = 50 − 39 = 11
   - ➔ Winning Supplier Profit = 39 − 30 = 9
   - ➔ Losing Supplier Profit = 0

2. The winning bid is 128 and so the price range is [32, 128]. The winning supplier selected a high quality engine. The shipping company proposes a price of 65.

   If the arbitrator favours a 2 to 1 supplier to shipping company profit ratio, then the arbitrator's price is 80, at which price the winning supplier would earn double the shipping company's profit. The shipping company's proposed price is below the arbitrator's price and so an appeal is triggered.

   Suppose that the arbitrator is unavailable and so the final price remains 65:

   - ➔ Shipping Company Profit = 100 − 65 = 35
   - ➔ Winning Supplier Profit = 65 − 40 = 25

➔ Losing Supplier Profit = 0

Suppose instead that the arbitrator is available and so the final price is 80. An arbitrator fee of 2 points is subtracted from the shipping company's profit.

➔ Shipping Company Profit = 100 – 80 – 2 = 18

➔ Winning Supplier Profit = 80 – 40 = 40

➔ Losing Supplier Profit = 0

<u>Information Display</u>

Much of the information displayed to you during the experiment is stored in a table in the lower portion of the computer screen. Here is an example of this table filled with made up numbers for demonstration purposes. The information shown is not from an actual experiment with people.

Round	Bids Winning / Losing	Quality	Proposal	Arbitrator Appeal / Available	Final Price	Profit	Balance
3	45 / 131*					0.00	**75.00**
2	75* / 131	Low	45.00	No / -	45.00	15.00	**75.00**
1	128* / 175	High	72.00	Yes / Yes	80.00	40.00	**50.00**

*Your Bid

The table provides detailed information for each round. Suppliers see the winning and losing auction bids and which of these is their own bid. The shipping company and winning supplier also see the winning supplier's engine quality, the shipping company's proposed price, whether an appeal is triggered and the arbitrator's availability, and the final price. All participants see their profit for the round and total accrued points balance.

Comprehension quiz

Please answer the questions below. Raise your hand if you require assistance. Once you answer all the questions correctly, you will be assigned your role and guided through two training rounds. The main part of the experiment then begins. Note: The comprehension questions and training rounds have no influence on your payment.

- Consider the following scenario.

The winning auction bid is 140 and so the price range is [35, 140]. First suppose the winning supplier selected a low quality engine, which is valued by the shipping company at 50 and costs the winning supplier 30. The shipping company proposes a price of 44.

Let the arbitrator favour a 2 to 1 supplier to shipping company profit ratio. The arbitrator's preferred price is no higher than the shipping company's proposal (and remember it cannot be any lower), because a price of 44 already gives the winning supplier more than double the shipping company's profit. No appeal is triggered. The final price remains 44.

1. What is the Shipping Company Profit? *6*

2. What is the Winning Supplier Profit? *14*

3. What is the Losing Supplier Profit? *0*

Now suppose the winning supplier selected a high quality engine, which is valued by the shipping company at 100 and costs the winning supplier 40. The shipping company still proposes a price of 44.

Again, let the arbitrator favour a 2 to 1 supplier to shipping company profit ratio. The arbitrator's preferred price is now 80, which is higher than the shipping company's proposal and so an appeal is triggered.

The die is rolled and the arbitrator turns out to be available. The final price is now 80. An arbitrator fee of 2 points should be subtracted from the shipping company's profit.

4. What is the Shipping Company Profit? *18*

5. What is the Winning Supplier Profit? *40*

B.5. Post-experiment questionnaire and subject characteristics

Age: Interval variable.

Years.

Mean 22.71, Median 22, Standard deviation 4.03, Minimum 18, Maximum 39

Gender: Categorical variable.

Male 50.93%; Female 48.15%; Other 0.00%; Prefer not to Say 0.93%.

Field of studies: Categorical variable.

Arts and Education 16.67%; Economics and Finance 11.11%; Business and Management 13.89%; Social Sciences and Law 10.19%; Medicine and Health Sciences 14.81%; Engineering and Natural Sciences 33.33%; Not a Student 0.00%.

Nationality: Categorical variable:

Central and Eastern Asia 2.78%; Central and Western Africa 4.63%; Central, South America and the Caribbean 1.85%; Europe (excl. UK) 29.63%; Middle East and North Africa 2.78%; North America 1.85%; Oceania 0.00%; South and Eastern Africa 0.93%; South-East Asia 6.48%; Southern Asia 3.70%; UK 45.37%.

Income: Categorical variable.

When you were 16 years of age, what was the income of your parents in comparison to other families in your country?

Far below average 4.63%; Below average 15.74%; Average 37.04%; Above average 36.11%; Far above average 6.48%

Risk Indices:

Based on Dohmen et al. (2011). Likert scale from 0 “Completely unwilling to take risks” to 10 “Completely willing to take risks”.

7) Are you generally a person who is fully willing to take risks or do you try to avoid taking risks?

Mean 6.30, Median 7, Standard deviation 2.35, Minimum 0, Maximum 10

8) How would you rate your willingness to take risks in financial matters?

Mean 4.70, Median 5, Standard deviation 2.53, Minimum 0, Maximum 10

Trust Index: Average of three variables.

Questions taken from the "General Social Survey" consistent with the approach of Glaeser et al. (2000).

4) Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?

"Most people can be trusted" 29.63% or "Can't be too careful" 70.37%

5) Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?

"Would try to be fair" 31.48% or "Would take advantage of you" 68.52%

6) Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?

"Try to be helpful" 38.89% or "Just look out for themselves" 61.11%

Trust Strangers: Dummy variable.

You can't count on strangers anymore.

"More or less disagree" 38.89% or "More or less agree" 61.11%

C Appendix for Chapter 4

C.1. Supplemental theoretical results

Proof of Lemma 4.1. The proof is based on Theorems 1 and 2 in Lang and Rosenthal (1991). By assumption $\delta_S R > E$ and so, from the equilibrium expression for $\mu(q)$, the entry probability is between zero and one for all q. From the equilibrium expression for $F(p,q)$, we have $F(E/(1-q+\delta_S q), q) = 0$, $F(R,q) = 1$ and $F_p(p,q) > 0$ and so F satisfies the properties of a c.d.f.

For any potential supplier, the indifference condition for entry, required by the mixed strategy equilibrium, is

$$(1-q+\delta_S q)p[1-\mu(q)F(p,q))] = E,$$

where the left-hand-side is the sum of two terms: the expected profit when his competitor does not enter, $1-\mu(q)$; and the expected profit when his competitor does enter but submits a higher bid, $\mu(q)(1-F(p,q))$. Since R is in the equilibrium price support and $F(R,q) = 1$, we can obtain the symmetric expression for $\mu(q)$ from the indifference condition. Using $\mu(q)$, we recover $F(p,q)$ by substitution.

Next, we check the equilibrium price support. Bidding at a level above R or below a price of $E/(1-q+\delta_S q)$ would yield the potential suppliers an expected loss. Thus, the set of best responses for either supplier in the second stage must be no entry along with bids from the interval $[E/(1-q+\delta_S q), R]$.

The unconditional price c.d.f. is

$$H(p,q) = 1-[\mu(q)(1-F(p,q))]^2,$$

where in the event of no bids, R is assigned as the winning price random variable. From $\mu(q)$ and $F(p,q)$, this simplifies to $H(p,q) = 1-(E/(1-q+\delta_S q)p)^2$.[96] Conditional on at least one bid being submitted, the c.d.f. is

[96] This represents the probability that the winning bid is less than or equal to p.

$$G(p,q)=\frac{H(p,q)}{1-[1-\mu(q)]^2},$$

from which we obtain the expression in the lemma using $\mu(q)$.

By the equilibrium supplier indifference condition above, expected total welfare $W(q)$ is based on expected buyer surplus. ■

Lemma C.1. *Suppose that $\delta_1 > \delta_2$ and $\delta_2 \in (\frac{E}{R}, 1)$. There exists a unique symmetric Nash equilibrium of the entry and pricing stage characterised as follows: the potential supplier with the higher discount factor enters with probability $\mu_1 = 1$ and submits a bid according to the c.d.f. $F_1(p,q)$, the potential supplier with the lower discount factor enters with probability $\mu_2(q)$ and submits a bid according to the c.d.f. $F_2(p,q)$, where*

$$\mu_2(q) = 1 - \frac{E}{(1-q+\delta_2 q)R} \quad ,$$

$$F_1(p,q) = \begin{cases} 1 - \frac{E}{(1-q+\delta_2 q)p} & \text{for} \quad p \in [\underline{p}, R) \\ 1 & \text{for} \quad p = R \end{cases} ,$$

$$F_2(p,q) = \frac{1-\frac{E}{(1-q+\delta_2 q)p}}{1-\frac{E}{(1-q+\delta_2 q)R}} \qquad \text{for} \quad p \in [\underline{p}, R], \quad \text{and}$$

$$\underline{p} = (1-\mu_2(q))R.$$

Proof of Lemma C.1. The proof follows Marquez (1997) and Thomas (2002), who generalise the simultaneous Bertrand entry model to account for heterogeneous entry costs. Suppose that the two potential suppliers possess different discount factors and that $\delta_i > E/R$ for each supplier. Any supplier for which this is not satisfied would have no influence on the analysis.

First, order the discount factors of the two suppliers such that $\delta_1 > \delta_2$, where the supplier 1 has the (absolute) higher discount factor, and supplier 2 has the lower discount factor. Define the vector of equilibrium entry probabilities by $\boldsymbol{\mu} = \{\mu_1, \mu_2\}$. The maximum bid in the mixing distribution support of the two suppliers must cover the reserve price in equilibrium. The minimum bid in the support must win with probability one and be the same across suppliers.[97] As a result, the expected profit of any entering

[97] Technical details of why the mixing distributions of at least two firms in equilibrium must have the same support can be found in Lemma 1 of Marquez (1997).

supplier in the second stage must be the same after accounting for differences in discount factors. We denote this expected profit by $E[\pi^*(\boldsymbol{\mu})]$.

An equilibrium in which only one potential supplier enters is not stable because that supplier would submit a bid equal to R and the other supplier would have an incentive to enter with probability one and marginally undercut the price.

For supplier 2 to enter profitably in equilibrium requires $E[\pi(\boldsymbol{\mu}, \delta_2)] \geq 0$. It follows that $E[\pi(\boldsymbol{\mu}, \delta_1)] > 0$ and so supplier 1 enters with probability one. Supplier 1 could guarantee herself $(1-\mu_2)(1-q+\delta_1 q)R$ gross of the entry cost by submitting a bid equal to the reserve price and so will never submit a bid less than $(1-\mu_2)R$. The probability of winning at this price is one and so to guarantee the same expected profit for all $p \in [\underline{p}, R]$, the minimum bid cannot be above $(1-\mu_2)R$. We obtain $\underline{p} = (1-\mu_2)R$. Since in equilibrium supplier 2 is indifferent over entry, it must hold that $(1-\mu_2)(1-q+\delta_2 q)R = E$ and this can be rearranged for the expression μ_2. The following equality must hold for supplier 1,

$$(1-\mu_2)F_2(p,q)(1-q+\delta_1 q)p = (1-\mu_2)(1-q+\delta_1 q)p.$$

This solves for $F_2(p,q)$. Likewise, for supplier 2, we have

$$(1-F_1(p,q))(1-q+\delta_2 q)p = E.$$

This solves for $F_1(p,q)$.

From the resulting expressions, we can infer that Supplier 2 randomises entry and earns zero expected profit in equilibrium. By contrast, supplier 1 enters with probability one and earns an expected profit equal to:

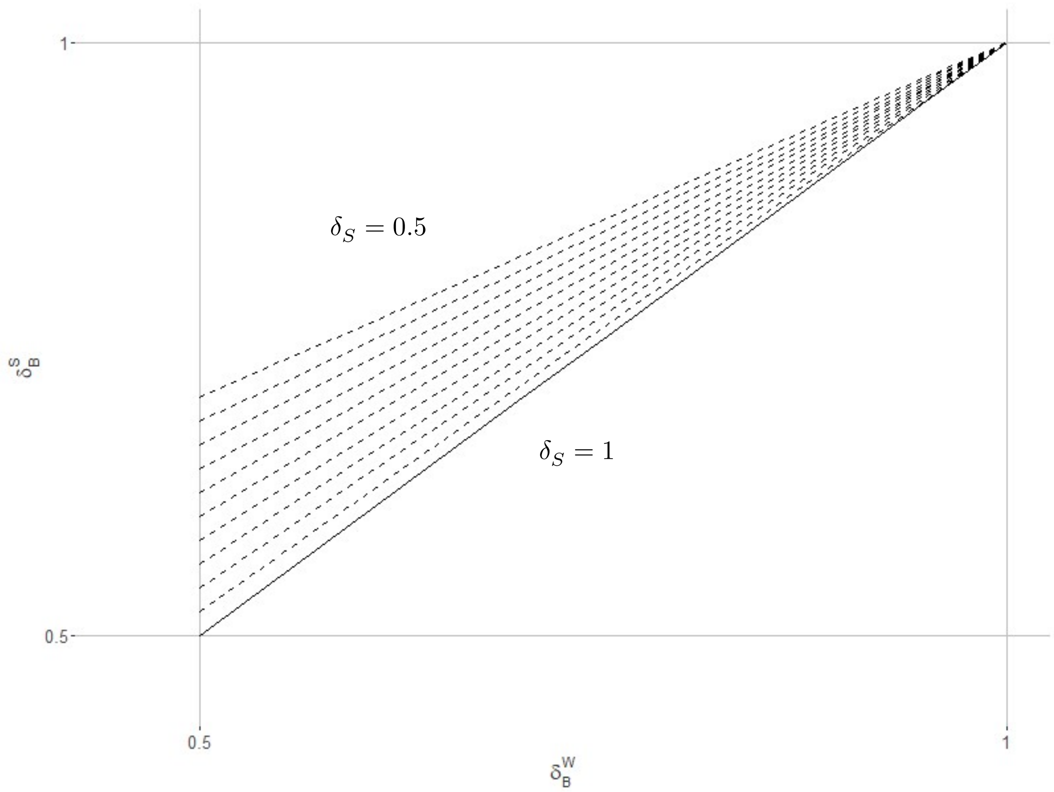
$$\left(\tfrac{(1-q+\delta_1 q)}{(1-q+\delta_2 q)} - 1\right)E. \blacksquare$$

Definition. *Sufficient Difference.* To fully characterise equilibria in the penalty space, we impose a restriction on the type distribution as follows:

$$\tfrac{(1-\delta_B^W)}{(1-\delta_B^S)} \geq \tfrac{p_1}{p_0}.$$

This restriction ensures that, in expectation, the highest threshold penalty for a Strong buyer to pay on-time is less than or equal to the lowest threshold penalty for a Weak buyer. That is, the range of the two functions $c_S(p_q)$ and $c_W(p_q)$ do not overlap. The restriction is without loss of generality. If we were to drop this restriction, then the interval $c \in [c_S(p_{1-x}), c_W(p_0)]$ may be degenerate but the remainder of the space is unaffected. To demonstrate that it is not too large to be uninteresting, consider the following parameter values: $E = 25,\ R = 100$ and $\delta_S = 0.5$. We use these parameter values in our experiment. In Figure C. 1, we graph δ_B^S against δ_B^W for discount factor values above or equal to 0.5. As the Weak buyer type's discount factor tends to one, the required Strong type's discount factor tends to the 45 degree line. For example, $\delta_B^W = 0.75$ would require a value of $\delta_B^S = 0.85$. Since, per Lemma 4.1, the difference between p_1 and p_0 (the buyer's expected benefit from making payment on-time) is also decreasing in the seller's discount factor δ_S, *Sufficient Difference* is everywhere lower in stronger seller populations.

Figure C. 1. *Sufficient Difference*.



Proof of Low Cost Proposition. In the continuation game associated with $m = 1$, both buyer types have a dominant strategy to pay at $\tau = 1$ for all $c \geq 0$. Potential suppliers anticipate this and, per Lemma 4.1, randomise at this information set according to $\mu(1)$ and $F(p, 1)$.

A. *Penalty interval*: $c < c_S(p_0)$. In this interval, both buyer types would pay late in the continuation game associated with $m = 0$, incurring the penalty c. The potential suppliers' expectation of late payment at this information set is $q = 1$. Potential suppliers' best response, independent of their posterior belief y, is to randomise according to $\mu(1)$ and $F(p, 1)$. But this is exactly the same as at $m = 1$. As a consequence, both buyer types can deviate to $m = 1$, avoid incurring the cost of reneging on a standard payment term and receive the same expected price. Thus, both types would earn a higher expected profit and the unique equilibrium involves pooling on $m = 1$ and late payment by both types. The potential suppliers' posterior belief about the buyer's type is unchanged from their prior belief, i.e., $z = x$. Beliefs are not specified off the equilibrium path.

B. *Penalty interval:* $c \in \left[c_S(p_0), c_S(p_{1-x})\right)$. Consider a separating equilibrium in which a Strong buyer sends $m = 0$ and pays on-time and a Weak buyer sends $m = 1$ and pays late. In such an equilibrium, potential suppliers would attach an expectation of late payment $q = 0$ to receipt of a message $m = 0$. Per Lemma 4.1, they would best respond by randomizing according to $\mu(0)$ and $F(p, 0)$. Given this randomization, a Strong buyer would pay on-time because in this penalty interval, $c \geq c_S(p_0)$ and so q is correct. The Weak buyer has a dominant strategy to pay late at $m = 1$. Since this is a separating equilibrium, the potential suppliers' know the buyer's type at the second stage with certainty: $y = 1$ and $z = 0$. Beliefs are not specified off the equilibrium path. A Strong buyer would prefer this separating equilibrium to a pooling equilibrium on $m = 1$ in which both types pay late if $\delta_B^S > \frac{\beta_0 p_0 - R(\beta_0 - \beta_1)}{\beta_1 p_1}$.

   Suppose that a Weak buyer were to deviate to $m = 0$ and form a pooling equilibrium. By *Sufficient Difference*, we have $c < c_W(p_0)$ and so a Weak buyer would never pay on-time in the continuation game associated with $m = 0$. If the Strong buyer still makes payment on-time, then per Lemma 4.1, potential suppliers would best respond at this information set by randomizing entry with $\mu(1-x)$ and bids on $F(p, 1-x)$. Yet given this randomization, a Strong buyer would also pay late, because $c < c_S(p_{1-x})$. Thus, potential suppliers best respond by randomizing according to $\mu(1)$ and $F(p, 1)$, independent of their posterior belief y. This is exactly the same as in the continuation game associated with $m = 1$. As a consequence, both buyer types would deviate to $m = 1$ and avoid incurring the cost of reneging on an intention to pay on-time, receive the same expected price and earn a higher expected profit. By the same logic, a separating equilibrium in which the Strong buyer sends $m = 1$ and the Weak buyer sends $m = 0$ would collapse.

Finally, note that updating of off-the-path posterior beliefs need not be consistent with Bayes' rule. Thus, in this case, a pooling equilibrium on $m = 1$ in which both types pay late may also exist when $\delta_B^S > \frac{\beta_0 p_0 - R(\beta_0 - \beta_1)}{\beta_1 p_1}$ if we specify $y \in [0, x)$. Such a specification is objectionable, however, because it implies that a Weak buyer is more likely to deviate to a standard payment term than a Strong buyer despite having less to gain. ■

Proof of Corollary 4.2. Follows from Lemma 4.1 and the Low Cost Proposition . ■

Proof of High Cost Proposition. In the continuation game associated with $m = 1$, both buyer types have a dominant strategy to pay at $\tau = 1$ for all $c \geq 0$. Potential suppliers anticipate this and, per Lemma 4.1, randomise at this information set according to $\mu(1)$ and $F(p, 1)$.

Penalty interval: $c \geq c_W(p_{1-x})$. First, consider a pooling equilibrium in which both buyer types send $m = 0$ and pay on-time. In such an equilibrium, potential suppliers would attach an expectation of late payment $q = 0$ to receipt of a message $m = 0$. Per Lemma 1, they would best respond by randomizing entry and bids according to $\mu(0)$ and $F(p, 0)$. Given this randomization, both buyer types would pay on-time because in this penalty interval, $c > c_W(p_0)$. Thus, q is correct. For neither type to have an incentive to deviate to $m = 1$, we require that the Weak buyer's discount factor is high enough, $\delta_B^W > \frac{\beta_0 p_0 - R(\beta_0 - \beta_1)}{\beta_1 p_1}$. The no-deviation condition for a Strong buyer is redundant. Since buyer types pool their messages in the first stage, the potential suppliers' posterior belief about the buyer's type on the equilibrium path is unchanged from their prior, i.e., $y = x$. The same *caveat* on multiplicity of equilibria applies as in the penalty interval B of the Low Cost Proposition. Similarly, if $\delta_B^W < \frac{\beta_0 p_0 - R(\beta_0 - \beta_1)}{\beta_1 p_1} < \delta_B^S$, then a Weak buyer would prefer to send a message $m = 1$ and pay late; the Strong buyer would prefer to send $m = 0$ and pay on-time in exchange for an identical supplier entry and bid randomization. Since this is a separating equilibrium, the potential suppliers' know the buyer's type at the second stage with certainty. Thus, $y = 1$ and $z = 0$. Finally, if $\delta_B^S < \frac{\beta_0 p_0 - R(\beta_0 - \beta_1)}{\beta_1 p_1}$ then the unique equilibrium involves pooling on $m = 1$ and the proof follows the Low Cost Proposition. ■

Proof of Corollary 4.3. Follows from Lemma 4.1 and the High Cost Proposition. ■

Proof of Free-rider Proposition. In the continuation game associated with $m = 1$, both buyer types have a dominant strategy to pay at $\tau = 1$ for all $c \geq 0$. Potential suppliers anticipate this and, per Lemma 4.1, randomise at this information set according to $\mu(1)$ and $F(p, 1)$.

Penalty interval: $c \in \left[c_S(p_{1-x}), c_W(p_{1-x})\right)$. Consider a pooling equilibrium in which both buyer types send $m = 0$, the Strong type pays on-time and the Weak type pays late. In such an equilibrium, potential suppliers would attach an expectation of late payment $q = (1 - x)$ to receipt of a message $m = 0$. Per Lemma 4.1, they would best respond by randomizing entry and bids according to $\mu(1 - x)$ and $F(p, 1 - x)$. Given this randomization, the Strong type would pay on-time and the Weak type would pay late because in this interval, $c_S(p_{1-x}) \leq c < c_W(p_{1-x})$. Thus, q is correct. For the Weak type to have no incentive to deviate to an extended payment term at any penalty in the interval, we require $\delta_B^W > \frac{\beta_{1-x} p_{1-x} - R(\beta_{1-x} - \beta_1)}{\beta_1 p_1}$. The no-deviation condition for a Strong type is redundant. Since buyer types pool their messages in the first stage, the potential suppliers' posterior belief about the buyer's type on the equilibrium path is unchanged from their prior, i.e., $y = x$. Beliefs are not specified off the equilibrium path and the same *caveat* on multiplicity of equilibria applies as in the penalty interval B of the Low Cost Proposition.

Consider the alternative pooling equilibrium in which both buyer types send $m = 0$ and pay on-time. In such an equilibrium, potential suppliers would attach an expectation of late payment $q = 0$ to receipt of a message $m = 0$. Per Lemma 1, they would best respond by randomizing entry and bids according to $\mu(0)$ and $F(p, 0)$. Given this randomization, only the Strong type would pay on-time if $c < c_W(p_0)$; thus, this equilibrium is not possible at penalties below that level. If $c > c_W(p_0)$, then both types would pay on-time and the proof follows the High Cost Proposition. Finally, note that $\beta_0(R - p_0) \geq \beta_{1-x}(R - p_{1-x})$ for all x. ■

Proof of Corollary 4.4. Follows from Lemma 4.1 and the Free-rider Proposition. ■

C.2. Supplemental empirical results

Table C. 1 – Experiment outcomes for the pooled online and lab samples.

Treatment	*LC*		*HC*		*FR*		*Test*
Panel A: Payment decisions (0 = UD; 1 = AD)							
Intended payment	0.54	(0.20)	0.36	(0.10)	0.28	(0.11)	
✓✓✓	*LC > FR: [0.022]*		*HC < LC: [0.032]*		*FR = HC: [0.229]*		***H1(i)***
Weak Buyer	0.55	(0.25)	0.44	(0.25)	0.34	(0.16)	
Strong Buyer	0.52	(0.22)	0.28	(0.12)	0.23	(0.08)	
✓✓✓	*W = S: [1.00]*		*W = S: [0.31]*		*W = S: [0.063]*		***H1(ii)***
Final payment late	0.89	(0.13)	0.35	(0.09)	0.48	(0.13)	
✓			*HC < LC: [0.001]*				***H2(i)***
Weak Buyer	0.91	(0.17)	0.47	(0.27)	0.79	(0.22)	
Strong Buyer	0.84	(0.13)	0.25	(0.13)	0.19	(0.08)	
✓✓✓	*W = S: [0.31]*		*W = S: [0.31]*		*W > S: [0.016]*		***H2(ii)***
Panel B: Bidding, prices and profits							
Bid	69.93	(9.84)	57.88	(7.61)	59.59	(6.30)	
Intention UD	69.09	(10.56)	54.94	(7.88)	56.38	(7.10)	
Intention AD	70.47	(9.73)	65.20	(10.53)	68.02	(5.67)	
✓✓✓	*UD = AD: [0.16]*		*UD < AD: [0.047]*		*UD < AD: [0.016]*		***H3(i)***
Price	65.40	(11.02)	52.27	(5.99)	55.49	(7.06)	
×✓×	*LC > FR: [0.12]*		*HC < LC: [0.032]*		*FR > HC: [0.35]*		***H3(ii)***
Buyer profit	33.75	(5.20)	44.65	(6.41)	40.35	(4.30)	
✓✓×	*LC < FR: [0.021]*		*HC > LC: [0.013]*		*FR < HC: [0.15]*		***H3(iii)***
Weak Buyer	37.85	(5.40)	46.50	(7.26)	42.17	(5.22)	
Strong Buyer	28.29	(6.60)	42.93	(5.96)	38.55	(5.86)	
Seller profit	25.68	(2.06)	26.71	(2.37)	27.18	(2.20)	
✓✓✓	*LC = FR: [0.12]*		*HC = LC: [0.12]*		*FR = HC: [0.65]*		***H3(iii)***

Notes: Displayed are mean (SD) values for the key experiment outcomes based on 6 independent cohorts per treatment, with *p*-value from specified non-parametric comparison test in square brackets. For one-sample comparisons, Wilcoxon Signed-Rank test; for two-sample comparisons, Wilcoxon-Mann-Whitney test. We employ a one-sided test when the theory predicts a direction and a two-sided test when the theory predicts equivalence. N'th ✓ or × in a row corresponds to n'th *p*-value in the same row. ✓ indicates non-parametric test supports prediction of corresponding hypothesis at 5% level; × indicates non-parametric test fails to support prediction of corresponding hypothesis at 5% level. *UD = Upon Delivery, AD = After Delivery.*

Table C. 2 – Logit regression analysis of buyer payment decisions (0 = UD; 1 = AD) with controls.

Dependent variable	*Intended payment* $\hat{\beta}$	*Final payment* $\hat{\beta}$	*Penalty = 1* $\hat{\beta}$
HC	-0.49	-3.34***	-5.49***
	(0.53)	(1.01)	(1.37)
FR	-1.68***	-3.64***	-2.82**
	(0.45)	(0.70)	(1.17)
Weak	0.72*	0.52	-0.49
	(0.38)	(0.54)	(0.63)
*HC * Weak*		0.55	2.06
		(1.09)	(1.35)
*FR * Weak*		2.10**	4.52***
		(0.90)	(1.43)
Risk general	-0.21*	-0.1	0.1
	(0.12)	(0.17)	(0.15)
Risk financial	0.18	0.03	-0.21
	(0.14)	(0.17)	(0.18)
Business & Mgmt major	-1.13***	-1.78***	0.35
	(0.41)	(0.66)	(0.58)
Female	1.42***	-0.15	-2.59***
	(0.51)	(0.42)	(0.62)
US National	1.09**	0.38	-1.24*
	(0.52)	(0.54)	(0.75)
Above average income	0.15	0.14	0.2
	(0.49)	(0.69)	(0.63)
Age	0.16**	0.25*	-0.03
	(0.08)	(0.14)	(0.10)
Period	-0.01	0.01	0.02
	(0.02)	(0.02)	(0.02)
Constant	-4.82**	-2.56	3.16
	(2.27)	(3.69)	(2.79)
Observations	850	700	700
Cohort fixed effects	Yes	Yes	Yes
Log-likelihood	-489.40	-319.31	-215.38
Wald test statistic	186.30***	309.71***	261.41***

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates on the logit scale ($\hat{\beta}$), with robust standard errors in parentheses clustered at the subject level. The omitted experimental treatment is *LC*. The models are estimated using Logistic regression. The Wald test statistic is based on a Chi-squared test of joint regression significance. *UD = Upon Delivery, AD = After Delivery.*

Table C. 3 – OLS regression analysis of bids and profits with controls.

Dependent variable	*Bid*	*Buyer profit*	*Seller profit*
HC	-16.89***	20.58***	1.54
	(3.71)	(2.80)	(1.58)
FR	-8.88***	10.52***	2.72
	(3.43)	(2.12)	(1.84)
Intention AD	(0.06)	0.4	-4.00***
	(1.85)	(1.85)	(1.29)
*HC * Intention AD*	11.64***	-13.20***	0.94
	(4.23)	(4.02)	(2.31)
*FR * Intention AD*	10.75***	-13.16***	3.06*
	(4.05)	(3.76)	(1.80)
Weak		9.13***	-0.2
		(2.57)	(0.99)
*HC * Weak*		-2.33	-3.21*
		(3.79)	(1.79)
*FR * Weak*		-1.66	-5.92***
		(3.42)	(2.08)
Risk general	-1.13	0.44	-0.47
	(0.99)	(0.55)	(0.34)
Risk financial	1.35	-0.8	0.11
	(1.05)	(0.69)	(0.34)
Business & Mgmt major	-5.14*	-2.57	-0.6
	(2.85)	(2.00)	(0.87)
Female	1.43	2.27	1.54**
	(3.20)	(2.00)	(0.76)
US National	3.42	2.11	-1.66
	(4.07)	(2.52)	(1.45)
Above average income	-0.08	-1.57	1.82**
	(2.83)	(2.72)	(0.80)
Age	1.01**	0.07	0.07
	(0.44)	(0.33)	(0.15)
Period	0.67***	-0.32***	0.09*
	(0.09)	(0.11)	(0.05)
Constant	38.33***	30.92***	27.10***
	(11.68)	(10.52)	(4.13)
Observations	1,590.00	850	1,590
Cohort fixed effects	Yes	Yes	Yes
R-squared	0.24	0.14	0.04
Wald test statistic	30.22***	7.25***	3.15***

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates are presented, with robust standard errors in parentheses clustered at the subject level. The omitted experimental treatment is *LC*. The models are estimated using OLS regression. The Wald test statistic is based on an F-test of joint regression significance. *AD = After Delivery*.

Table C. 4 – Random effects analysis of bids, transaction prices and profits.

Dependent variable	*Bid*	*Price*	*Buyer profit*	*Seller profit*
HC	-15.28***	-16.41***	20.36***	1.36
	(3.76)	(3.95)	(3.36)	(1.61)
FR	-9.11***	-9.97***	10.85***	3.14*
	(3.42)	(3.73)	(2.44)	(1.90)
Intention AD	0.79	0.76	1.29	-3.90***
	(1.34)	(2.10)	(1.76)	(1.21)
*HC * Intention AD*	9.31**	10.18***	-13.22***	0.73
	(3.87)	(3.72)	(3.99)	(2.16)
*FR * Intention AD*	9.79***	12.84***	-14.10***	3.23*
	(3.72)	(3.00)	(3.57)	(1.69)
Weak			9.43***	0.50
			(2.49)	(1.11)
*HC * Weak*			-2.62	-3.71**
			(3.83)	(1.82)
*FR * Weak*			-1.69	-6.78***
			(3.40)	(2.18)
Period	0.57***	0.48***	-0.34***	0.06
	(0.09)	(0.10)	(0.11)	(0.05)
Constant	62.44***	59.80***	30.62***	26.80***
	(2.68)	(3.14)	(2.36)	(1.38)
Observations	1,650.00	682.00	850.00	1,650.00
Cohort fixed effects	Yes	Yes	Yes	Yes
Subject random effects	Yes	Yes	Yes	Yes
R-squared	0.13	0.16	0.13	0.03
Wald test statistic	243.66***	126.86***	128.59***	47.22***

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates are presented, with robust standard errors in parentheses clustered at the subject level. The omitted experimental treatment is *LC*. The models are estimated using random effects panel regression. The Wald test statistic is based on a Chi-squared test of joint regression significance. *AD = After Delivery*.

Table C. 5 – Regression analysis of decision-making split by experimental phase.

Dependent variable	*Intended payment AD*		*Final payment AD*		*Bid*	
Periods	1-15	16-30	1-15	16-30	1-15	16-30
	(1)	(2)	(3)	(4)	(5)	(6)
HC	-0.85*	-1.07	-2.80***	-4.28***	-12.23***	-25.01***
	(0.51)	(0.66)	(0.88)	(1.15)	(3.49)	(5.02)
FR	-1.24**	-1.61**	-2.88***	-4.78***	-7.86**	-13.22***
	(0.51)	(0.67)	(0.80)	(1.22)	(3.47)	(4.37)
Weak	0.72**	0.61	0.75	-0.12		
	(0.30)	(0.48)	(0.61)	(0.85)		
*HC * Weak*			0.49	0.79		
			(1.03)	(1.26)		
*FR * Weak*			1.44*	3.08***		
			(0.86)	(1.13)		
Intention AD					0.99	-1.71
					(2.30)	(2.39)
*HC * Intention AD*					5.17	21.78***
					(3.33)	(6.66)
*FR * Intention AD*					10.40**	13.73***
					(4.94)	(4.40)
Period	-0.01	-0.01	0.01	-0.02	0.94***	0.60***
	(0.02)	(0.03)	(0.03)	(0.05)	(0.18)	(0.22)
Constant	0.2	-0.01	2.03***	3.58**	58.24***	76.93***
	(0.66)	(1.08)	(0.62)	(1.57)	(3.63)	(3.47)
Observations	510	340	419	281	990	660
Cohort fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Wald test statistic	56.59***	43.56***	144.33***	142.80***	20.34***	23.25***

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates, with robust standard errors in parentheses clustered at the subject level. The omitted experimental treatment is *LC*. Models 1 to 4 are estimated using Logistic regression. Models 5 and 6 are estimated using OLS regression. *AD = After Delivery*.

Table C. 6 – Aggregate profit statistics for the HC-E cohorts.

Cohort	*1*		*2*		*3*		*4*	
Intention UD								
Buyer profit								
Weak Buyer	65.44	(15.47)	54.06	(22.94)	60	11.73	42.5	3.54
Strong Buyer	65.54	15.36	58.27	19.91	65	4.08	44.33	12.91
Seller profit	21.39	(15.48)	23.66	(18.26)	24.44	(17.56)	32.75	(26.79)
Intention AD								
Buyer profit								
Weak Buyer	49.39	(24.32)	45.83	20.61	67.69	(14.99)	58.49	(17.89)
Strong Buyer	42.85	(27.74)	45.53	12.07	59.44	(17.91)	40.36	(18.05)
Seller profit	20.5	(13.04)	24.81	(10.5)	15.33	(12.16)	19.17	(13.76)

Notes: Displayed are mean (SD) values at the subject level in each cohort. *UD = Upon Delivery, AD = After Delivery.*

Table C. 7 – Determinants of seller entry in the HC-E treatment.

Dependent variable	*Entry*
Intention AD	-0.13*
	(0.07)
Risk general	0.01
	(0.03)
Risk financial	-0.03
	(0.03)
Business & Mgmt major	0.02
	(0.11)
Female	0.09
	(0.06)
US National	-0.39***
	(0.13)
Above average income	0.03
	(0.08)
Age	-0.01
	(0.01)
Period	-0.004
	(0.00)
Constant	1.17***
	(0.29)
Observations	630
Cohort fixed effects	Yes
R-squared	0.08
Wald test statistic	4.72***

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates are presented, with robust standard errors in parentheses clustered at the subject level. The models are estimated using OLS regression on data from the *HC-E* treatment only. The Wald test statistic is based on an F-test of joint regression significance. *AD = After Delivery*.

Table C. 8 – OLS regression analysis of bids, transaction prices and profits in the HC-E treatment.

Dependent variable	*Bid*	*Price*	*Buyer profit*	*Seller profit*
	(1)	(2)	(3)	(4)
HC-E	-10.81***	-11.17***	12.26***	-6.10***
	(3.52)	(3.56)	(4.36)	(1.94)
Intention AD	10.37***	11.33***	-11.94***	-3.05*
	(3.56)	(3.29)	(3.24)	(1.79)
*HC-E * Intention AD*	5.65	3.84	3.26	-1.57
	(5.08)	(4.66)	(5.28)	(2.42)
Weak			6.88**	-3.15**
			(2.86)	(1.40)
*HC-E * Weak*			-2.73	3.29*
			(3.87)	(1.99)
Period	0.41***	0.27**	-0.25**	0.08*
	(0.11)	(0.11)	(0.13)	(0.05)
Constant	49.32***	46.11***	49.43***	28.45***
	(2.81)	(2.88)	(2.92)	(1.41)
Observations	1,001	538	580	1,160
Cohort fixed effects	Yes	Yes	Yes	Yes
R-squared	0.16	0.2	0.14	0.05
Wald test statistic	26.69***	19.24***	10.06***	7.27***

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates are presented, with robust standard errors in parentheses clustered at the subject level. The omitted experimental treatment is *HC*. The models are estimated using OLS regression. The Wald test statistic is based on an F-test of joint regression significance. *AD = After Delivery*.

Table C. 9 – Logit regression analysis of buyer payment decisions (0 = UD; 1 = AD) in the HC-E treatment.

Dependent variable	*Intended payment*		*Final payment*	
	$\hat{\beta}$	dy/dx	$\hat{\beta}$	dy/dx
	(1)		(2)	
HC-E	0.78*	0.19*	0.90*	0.22*
	(0.45)	(0.11)	(0.48)	(0.11)
Weak	0.62	0.15	0.74	0.18
	(0.45)	(0.11)	(0.47)	(0.12)
Period	-0.02	-0.005	-0.02	-0.004
	(0.01)	(0.00)	(0.02)	(0.00)
Constant	-1.26**		-1.53**	
	(0.59)		(0.61)	
Observations	580		538	
Cohort fixed effects	Yes		Yes	
Log-likelihood	-328.46		-297.83	
Wald test statistic	144.64***		147.19***	

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficient estimates on the logit scale ($\hat{\beta}$) and average marginal effects (dy/dx), with robust standard errors in parentheses clustered at the subject level. The omitted experimental treatment is *HC*. The models are estimated using Logistic regression. The Wald test statistic is based on a Chi-squared test of joint regression significance. *UD = Upon Delivery, AD = After Delivery.*

Figure C. 2. Within-buyer relative payment frequencies (0 = UD; 1 = AD) by treatment (panel) and type (shape), weighted by the number of subjects.

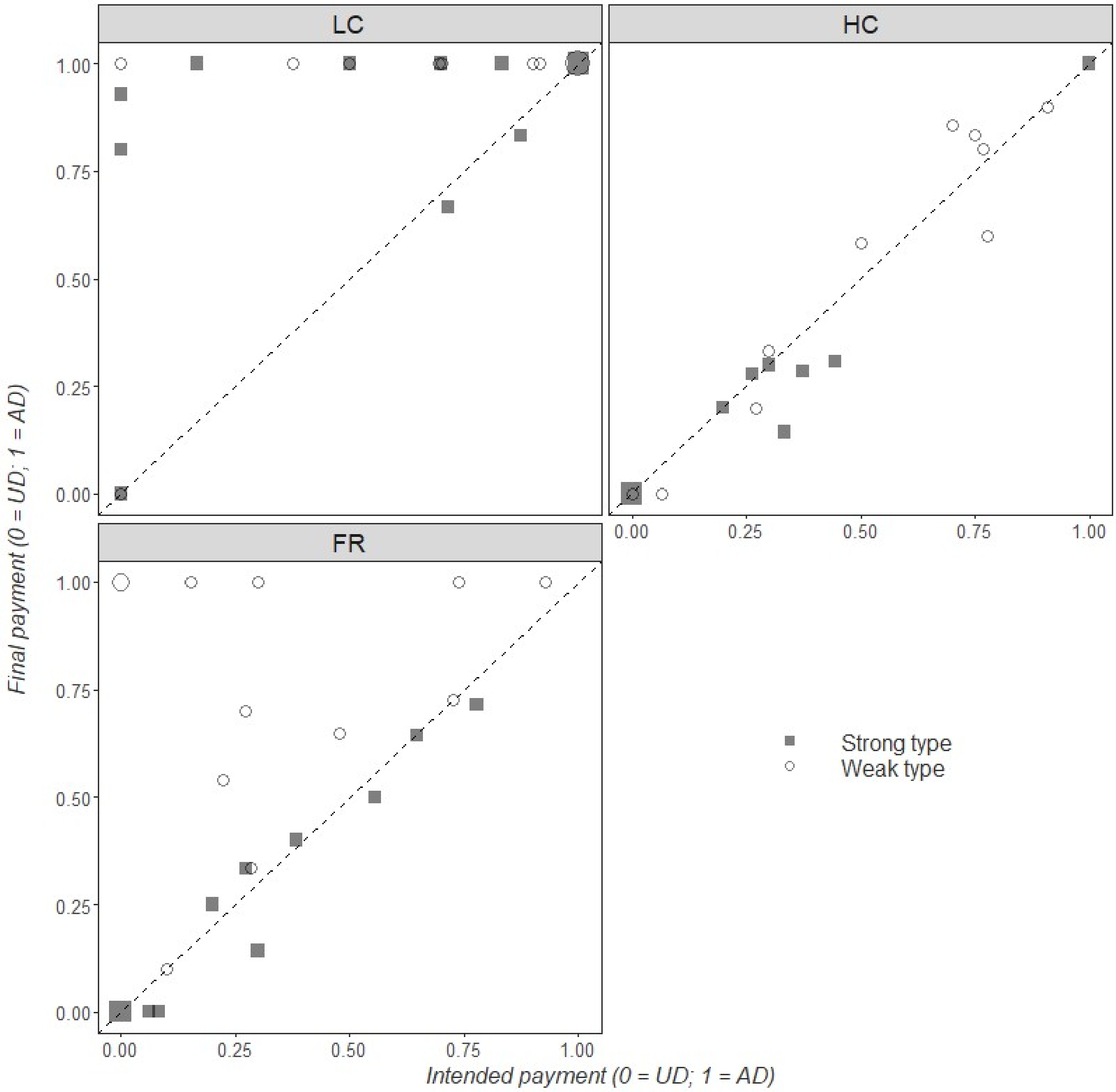
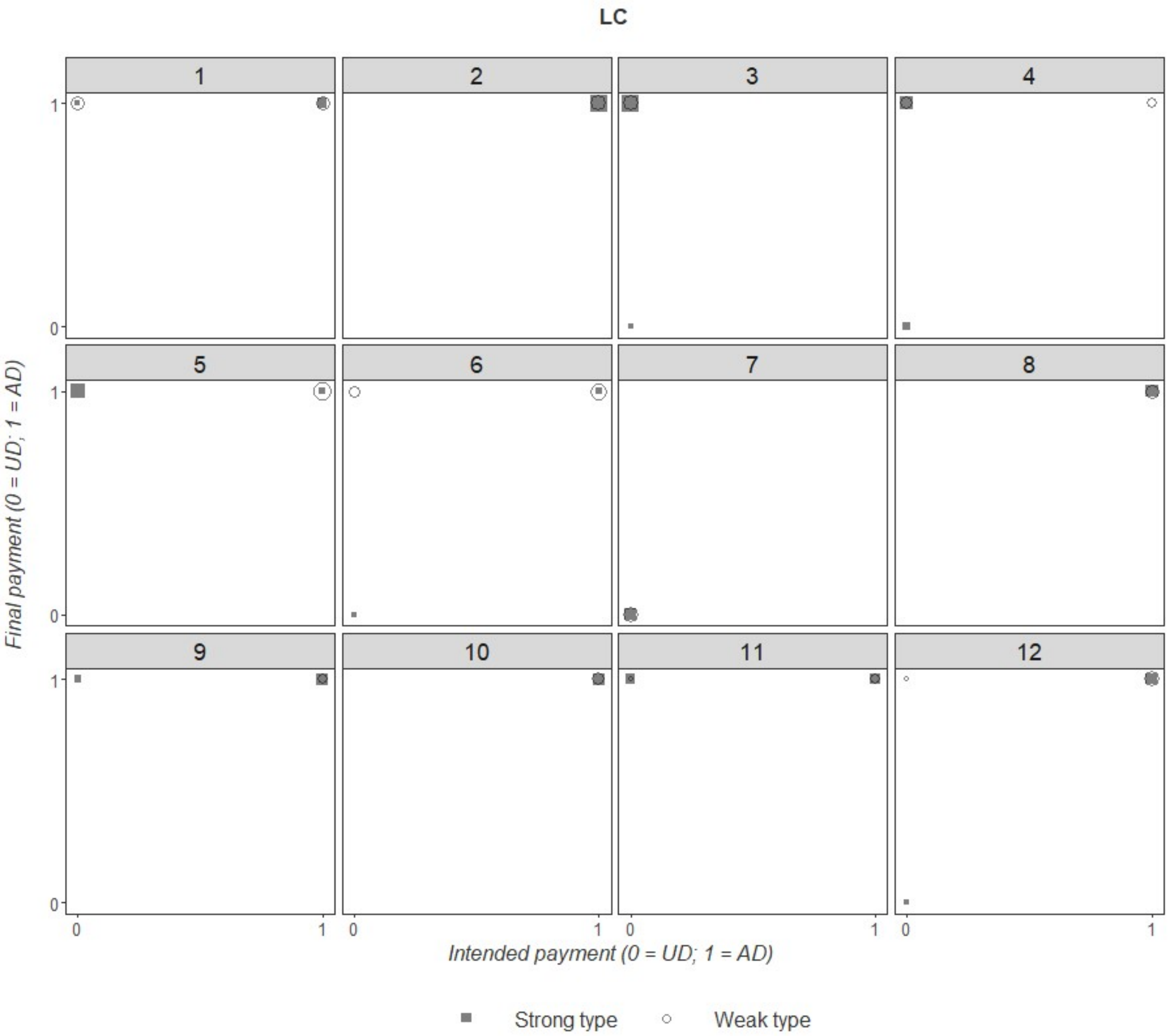
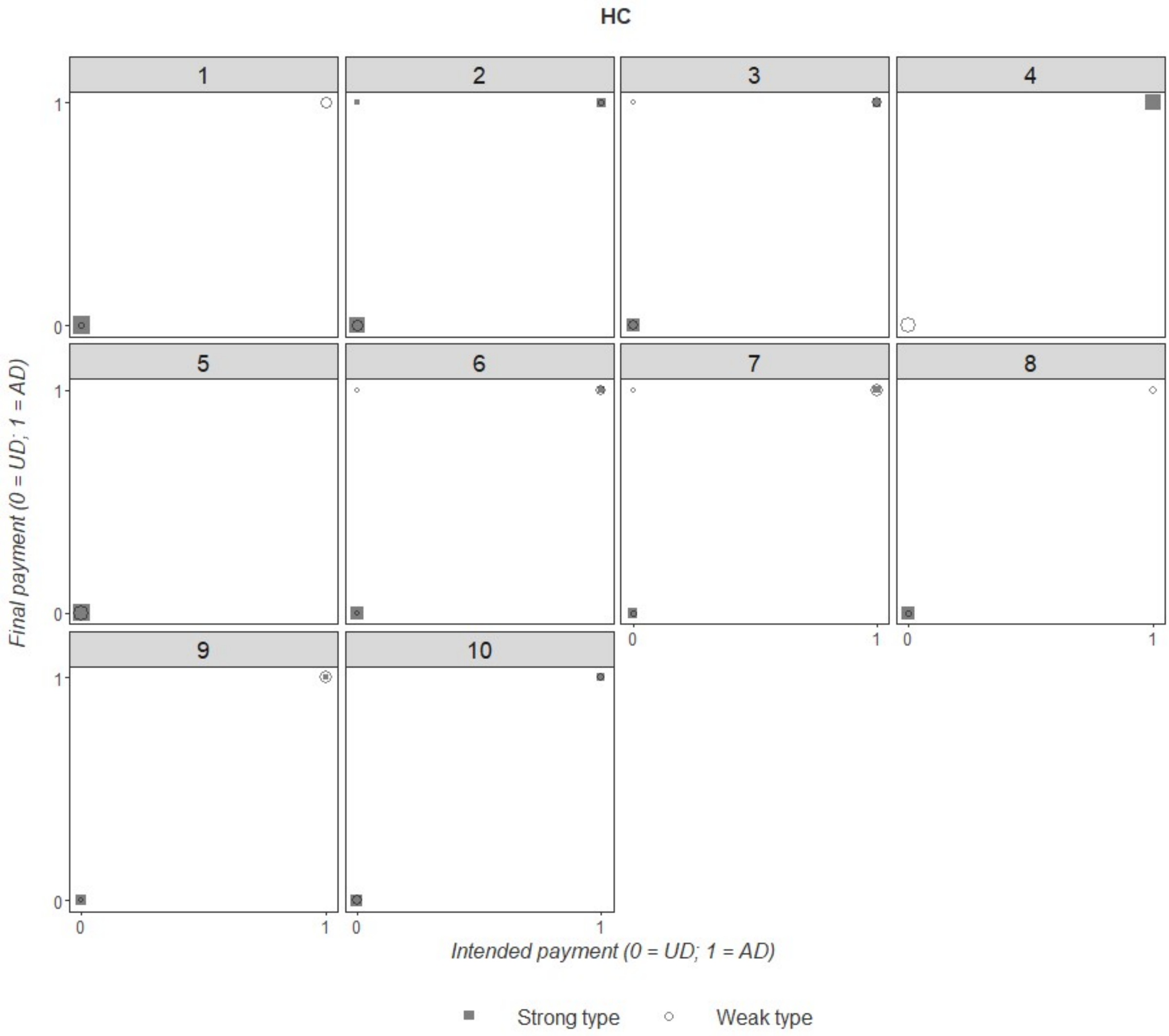
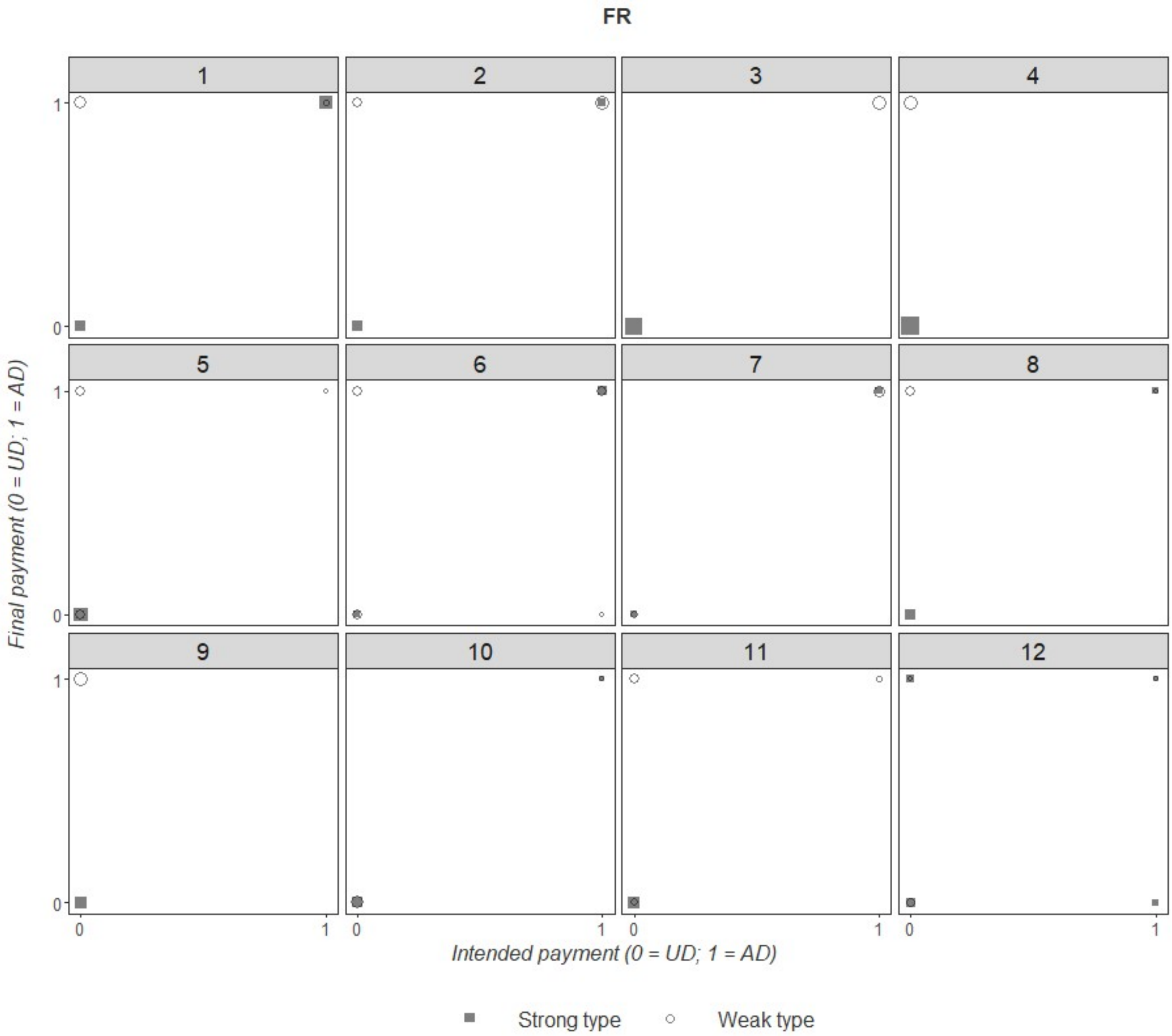


Figure C. 3. Payment strategies (0 = UD; 1 = AD) by buyer ID (panel) and type (shape) for each treatment, conditional on at least one seller being eligible, weighted by the number of market observations.




HC
1
2
3
4
5
6
7
8
9
10
Final payment (0 = UD; 1 = AD)
Intended payment (0 = UD; 1 = AD)
Strong type
Weak type


FR
1
2
3
4
5
6
7
8
9
10
11
12
Final payment (0 = UD; 1 = AD)
Intended payment (0 = UD; 1 = AD)
0
1
Strong type
Weak type

C.3. Experimental implementation and instructions

Online Implementation.

At the beginning of each experimental session, subjects were admitted into a virtual registration room one-by-one. After identity verification, their on-screen names were anonymised, videos switched off and they were moved to a second room. Once everyone had arrived in this second room, we played a pre-recorded video of the experiment instructions (see web links below), to ensure that the description of the game is common knowledge and to enhance replicability.

To preserve anonymity, while ensuring that subjects remained present and did not communicate with others, subjects were transferred to one of several breakout rooms and asked to switch on their webcam video. Each room contained an experimenter and a subset of subjects who were guaranteed never to interact with one another during the session (subjects knew this). Subjects were then randomly assigned their role and completed a short comprehension quiz, before proceeding to two practice rounds. During this time, they could ask the experimenter questions via a private chat facility. The main experiment then began.

At the end of a session, subjects' webcams were switched off and they were returned to the registration room, where we explained that their payments would be mailed out to them on pre-paid debit cards by the end of next business day. Two sellers suffered (unrelated) internet outages mid-session (one in the *LC* treatment and one in the *FR* treatment). From the point of disconnection, bids for these sellers were replaced with a random number from the available bid interval. The remaining subjects in the cohort were unaware of the outage. All bid and transaction price data involving these two sellers are excluded from the analysis.

Web links to access the experiment video instructions.

The instructions pertain to the online experimental sessions with 30 (+ 2 practice) rounds. The videos for the sessions with 20 (+ 2) rounds are identical except for the adjusted exchange rate. The instructions for the lab sessions differ only (and unavoidably) on logistical issues, such as the processing of payments. Copies of the video instructions and transcripts used in the lab sessions are available on request.

LC:

https://drive.google.com/file/d/1Wc268Aq3m5nuvtDVPH9Q0c6pVSc9HVV1/view?usp=sharing

HC:

https://drive.google.com/file/d/1cqDXuiZP5C-1lDZqYEgzsU4I5DEVlli_/view?usp=sharing

FR:

https://drive.google.com/file/d/13R0QIqoqOM-6eDwJwOTGf0V8zfMyp7RG/view?usp=sharing

HC-E:

https://drive.google.com/file/d/1NDx7EKINlumo_Jj3xo0MLIkYwkB_8fAu/view?usp=sharing

Instructions transcript with screenshots.[98]

Welcome. This is an experiment studying economic decision-making in auctions.

How you earn money

The experiment today consists of **32 decision-making rounds** . In each round of the experiment, you will have the opportunity to earn points.

The first two rounds are for practice purposes only, and you will not receive any payment for points earned in the first two rounds. You will be paid for all points earned in the remaining rounds, which are converted to US dollars at the following rate:

80 points = $1

These earnings are in addition to a show-up fee of $5. At the end of the experiment, you will complete a short questionnaire. Upon completion, you will be directed to a separate survey page on which you will be asked to provide some details to enable us to process your payment.

Preliminary information

For today's session, one third of you will be randomly assigned to the role of a buyer, and two thirds of you assigned to the role of a seller. You will be informed to which role you are assigned before the first round. You remain in the same role for the entire experiment.

In each round, a buyer is randomly grouped with two sellers. This trio participates in an auction, in which the buyer seeks to purchase a single unit of a customized good from one of the two sellers in the trio. You will not know the identity of any person in your trio. **You will be randomly re-assigned to a new trio each round.**

The buyer's valuation for a customized unit is 100 points. ----[HC-E] Each seller can choose whether or not to customize the unit. The buyer has zero value for a non-customized unit and so will only purchase from a seller who chooses to customize the unit. Any seller who chooses not to customize the unit can return it to the experimenter in exchange for 25 points. [LC, HC, FR] The sellers, however, are not always eligible to customize the unit. Specifically, sellers are only eligible to customize the unit with a certain probability (precise details to be given later). The buyer has zero value for a non-customized unit and so will only purchase from a seller who is eligible to customize the unit. Any seller who is not eligible to customize the unit can return their unit to the experimenter in exchange for 25 points. ----

[98] Where relevant, treatment-specific information is demarcated using the following format: ---- [treatment X] … information … [treatment Y] … information … ----.

If at least one seller ----[HC-E] chooses [LC, HC, FR] is eligible---- to customize the unit, a transaction takes place and the auction determines the transaction price.

The buyer can pay for the unit using either one of the following methods: **upon delivery** or **after delivery** . The seller prefers to receive payment upon delivery. The buyer prefers to make payment after delivery. How valuable payment after delivery is to the buyer depends on the buyer's type (precise details to be given later). The buyer is informed about his/her type at the beginning of each round, which can be either type 1 or type 2. The buyer's type changes between rounds and is never revealed to the sellers. The sellers only know that the buyer in their trio is either of type 1 or of type 2, with equal chance.

How does a transaction between a buyer and seller take place?

In each round, every trio of one buyer and two sellers will undergo a three stage process to determine if there is a transaction, the transaction price and the payment method. This process consists of a pre-auction stage, an auction stage and a post-auction stage.

Pre-Auction Stage: Before the auction, the buyer is randomly assigned his/her type for the round, either type 1 or type 2. The buyer then announces his/her intended payment method for the unit. Specifically, the buyer can announce an intention to pay the seller upon delivery or after delivery of the unit.

The sellers do not participate in the pre-auction stage and wait to participate in the auction stage.

----[LC] The **probability** of a seller being eligible to customize the unit in any round is **one-half (50%)** . [HC, FR] The buyer's intended payment method announced in the pre-auction stage determines the probability with which sellers are eligible to customize the unit in the round.

If the buyer **intends to pay for the unit upon delivery** , the **probability** of a seller being eligible to customize the unit is [HC] **three-quarters (75%)** [FR] **two-thirds (66.67%)** .

[HC, FR] If the buyer **intends to pay for the unit after delivery** , the **probability** of a seller being eligible to customize the unit is **one-half (50%)** . ----

Auction Stage: ----[HC-E] When the auction begins, the sellers observe the buyer's intended payment method. Each seller chooses whether or not to customize the unit and, if he/she chooses to customize it, submits a bid for which he/she would be willing to deliver it to the buyer. This bid can be any whole number from 25 to 100 (inclusive). When deciding, neither seller can see the choices of the other seller in his/her trio.

If **both sellers in the trio customize the unit** , then the buyer purchases a unit from the seller that submitted the **lower** bid. This seller acquires the right to proceed to the post-auction stage, and his/her bid becomes the transaction price. The seller that submitted the higher bid exits without selling the customized unit and earns 0 points for the round. If both sellers submit the same bid, the seller from whom the buyer purchases is chosen at random (with equal chance).

If only one seller in the trio customizes the unit , then the buyer purchases a unit from this seller. This seller acquires the right to proceed to the post-auction stage and his/her bid becomes the transaction price. The seller that chose not to customize exits the round and returns his/her non-customized unit to the experimenter for 25 points.

If neither seller in the trio customizes the unit , then the round ends without a transaction. In this case, each seller returns his/her non-customized unit to the experimenter for 25 points, and the buyer earns 0 points for the round.

[LC, HC, FR] When the auction begins, the sellers observe the buyer's intended payment method and the probability with which they are eligible. Each seller submits a bid for which they would be willing to deliver a customized unit to the buyer. This bid can be any whole number from 25 to 100 (inclusive). When submitting a bid, neither seller can see the bid of the other seller in his/her trio.

After both sellers have submitted their bids, the computer separately determines whether each seller is eligible to customize the unit, according to the probability described in the pre-auction stage. Whenever a seller is determined to be eligible, his/her unit is customized for the buyer and can no longer be returned to the experimenter.

If the computer determines that both sellers in the trio are eligible , then the buyer purchases a unit from the seller that submitted the lower bid. This seller acquires the right to proceed to the post-auction stage, and his/her bid becomes the transaction price. The seller that submitted the higher bid exits without selling the customized unit and earns 0 points for the round. If both sellers submit the same bid, the seller from whom the buyer purchases is chosen at random (with equal chance).

If the computer determines that one unique seller in the trio is eligible , then the buyer purchases a unit from the eligible seller. This seller acquires the right to proceed to the post-auction stage and his/her bid becomes the transaction price. The non-eligible seller exits the round and returns the non-customized unit to the experimenter for 25 points.

If the computer determines that neither seller in the trio is eligible , then the round ends without a transaction. In this case, each seller returns his/her non-customized unit to the experimenter for 25 points, and the buyer earns 0 points for the round. ----

The buyer does not participate in the auction stage and waits to participate in the post-auction stage.

Post-Auction Stage: This stage is only reached if at least one seller ----[HC-E] chose [LC, HC, FR] is eligible ---- to customize the unit. The buyer observes the transaction price and selects a final payment method for the customized unit. The buyer can select to make payment to the seller upon delivery or after delivery of the unit. The final payment method selected can differ from the buyer's intended payment method announced in the pre-auction stage.

The seller prefers to receive payment upon delivery. Both types of buyer prefer to make payment after delivery. The buyer types differ, however, in how valuable this is. For a given price, payment after delivery is more valuable for a type 1 buyer than for a type 2 buyer.

If the buyer selects **final payment upon delivery** , then:

- the buyer's earnings are his/her unit valuation of 100 – transaction price;
- the seller's earnings are the transaction price.

If the buyer selects **final payment after delivery** , then:

- a type 1 buyer's earnings are his/her unit valuation of 100 – (0.75 × transaction price);
- a type 2 buyer's earnings are his/her unit valuation of 100 – (0.95 × transaction price);
- the seller's earnings are 0.5 × transaction price.

A **penalty of** ----[LC] **1** [HC-E, HC] **16** [FR] **8** ---- **points is levied on any buyer who announced an intention to pay upon delivery in the pre-auction stage but selected final payment after delivery in the post-auction stage.** This penalty is subtracted from the buyer's earnings. The penalty is collected by an independent third party outside of the experiment (it is not collected by the seller).

At the conclusion of the post-auction stage, the round ends and the earnings are realized.

Summary of how earnings for the round are calculated

Seller's earnings (in points)

Customized unit & lower bid OR Only seller to customize the unit:	
If the buyer makes final payment upon delivery	transaction price
If the buyer makes final payment after delivery	0.5 × transaction price
Customized unit & higher bid:	0
Non-customized unit:	25

Buyer's earnings (in points):

At least one seller customized the unit:	*Intends to pay upon delivery*	*Intends to pay after delivery*
If the buyer makes final payment upon delivery	100 – transaction price	100 – transaction price
If a type 1 buyer makes final payment after delivery	100 – 0.75 × transaction price – [1, 16, 8]	100 – 0.75 × transaction price
If a type 2 buyer makes final payment after delivery	100 – 0.95 × transaction price – [1, 16, 8]	100 – 0.95 × transaction price
Neither seller customized the unit:	0	

Example scenarios

Consider the following baseline scenario.

Pre-Auction stage: The buyer is randomly assigned to type 1. The buyer announces an intention to make payment upon delivery. ----[LC, HC, FR] The probability of a seller being eligible to customize the unit in the round is [one-half (50%), three-quarters (75%), two-thirds (66.67%)].----

Auction stage: Seller 1 ----[HC-E] chooses to customize the unit and---- submits a bid of 48. Seller 2 ----[HC-E] chooses to customize the unit and---- submits a bid of 67. ----[LC, HC, FR] The computer determines that both sellers are eligible to customize the unit in the round.---- Seller 1 proceeds to the post-auction stage. Seller 2 exits without making a transaction. The transaction price is equal to Seller 1's bid of 48.

Post-Auction stage : The buyer makes final payment to Seller 1 upon delivery.

The earnings for the round are as follows:

- Seller 1 earns 48
- Buyer earns $100 - 48 = 52$
- Seller 2 earns 0

Alternative 1: Now suppose that in the Auction stage of the baseline scenario, ----[HC-E] Seller 2 had chosen not to customize the unit [LC, HC, FR] the computer had determined that Seller 2 is not eligible to customize the unit. ----

- Seller 1's earnings are unchanged from baseline; i.e., 48
- Buyer's earnings are unchanged from baseline; i.e., 52
- Seller 2 earns 25

Alternative 2: Now suppose that in the Auction stage of the baseline scenario, ----[HC-E] neither seller had chosen to customize the unit [LC, HC, FR] the computer had determined that neither seller is eligible to customize the unit. ----

- Seller 1 earns 25
- Buyer earns 0
- Seller 2 earns 25

Alternative 3: Now suppose that in the Post-Auction stage of the baseline scenario, the buyer had made final payment to Seller 1 after delivery.

- Seller 1 earns $0.5 \times 48 = 24$
- Buyer earns $100 - (0.75 \times 48) - [1, 16, 8] = [63, 48, 56]$
- Seller 2's earnings are unchanged from baseline; i.e., 0

Alternative 4: Now suppose that in addition to *Alternative 3* of the baseline, in the Pre-Auction stage the buyer had announced an intention to make payment after delivery.

- Seller 1's earnings are unchanged from *Alternative 3*; i.e., 24
- Buyer earns $100 - (0.75 \times 48) = 64$
- Seller 2's earnings are unchanged from *Alternative 3*; i.e., 0

Alternative 5: Now suppose that in addition to *Alternative 4* of the baseline, in the Pre-Auction stage the buyer had been assigned to type 2.

- Seller 1's earnings are unchanged from *Alternative 4*; i.e., 24
- Buyer earns $100 - (0.95 \times 48) = 54.4$
- Seller 2's earnings are unchanged from *Alternative 4*; i.e., 0

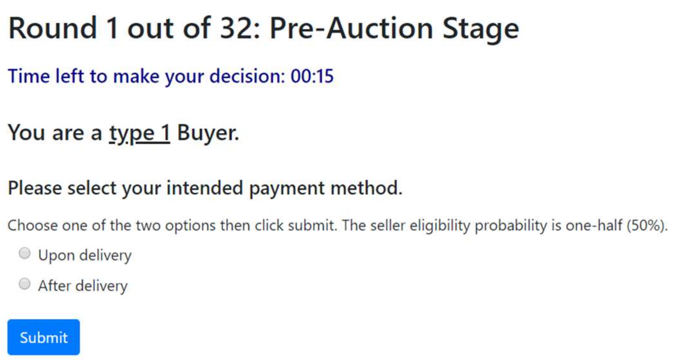
<u>How to use the computer program</u>

After all participants have successfully answered a comprehension quiz, the computerized auction rounds will begin.

Picture 1 gives an example of what your computer screen will look like in the Pre-Auction stage. Buyers observe their randomly assigned type for the round, which appears underlined ----[LC, HC, FR], and the seller eligibility probability [HC, FR] associated with each option----. Buyers must then select their intended payment method, by choosing one of the two options and clicking submit.

P. 1. Pre-Auction stage screen (for buyers only)

----[LC]----



----[HC]----

Round 1 out of 32: Pre-Auction Stage

Time left to make your decision: 00:14

You are a type 1 Buyer.

Please select your intended payment method.

Choose one of the two options then click submit. The seller eligibility probabilities associated with each are displayed underneath.

- Upon delivery
- After delivery

Intended payment	Seller eligibility probability
Upon delivery	three-quarters (75%)
After delivery	one-half (50%)

Submit

----[FR]----

Round 1 out of 32: Pre-Auction Stage

Time left to make your decision: 00:12

You are a type 1 Buyer.

Please select your intended payment method.

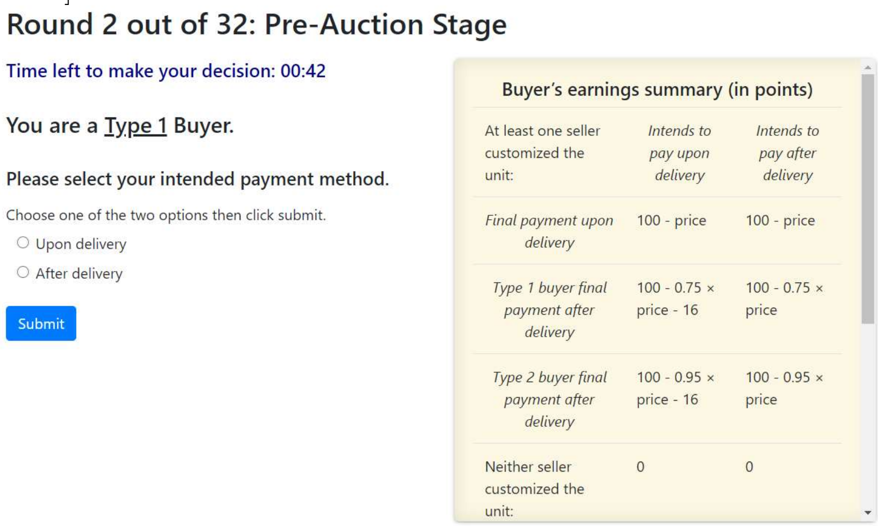
Choose one of the two options then click submit. The seller eligibility probabilities associated with each are displayed underneath.

- Upon delivery
- After delivery

Intended payment	Seller eligibility probability
Upon delivery	two-thirds (66.67%)
After delivery	one-half (50%)

Submit

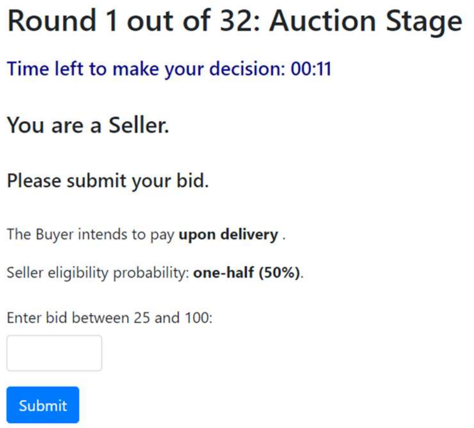
----[HC-E]----



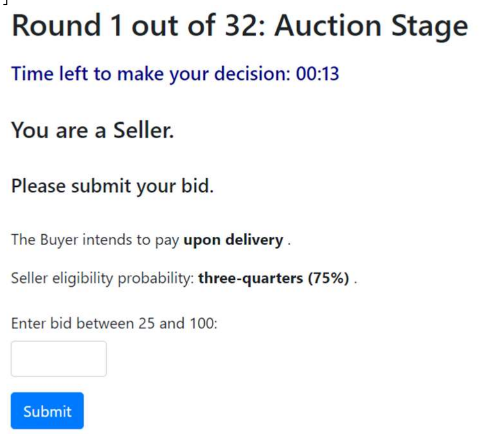
Picture 2 gives an example of what your computer screen will look like in the Auction stage. ----[HC-E] Sellers observe the buyer's intended payment method and must choose whether or not to customize the unit. If they choose to customize the unit, they must also submit a bid for the round, by entering a number between 25 and 100 in the box provided and clicking submit. [LC, HC, FR] Sellers observe the buyer's intended payment method and their eligibility probability. Sellers must then submit a bid for the round, by entering a number between 25 and 100 in the box provided and clicking submit. ----

P. 2. Auction stage screen (for sellers only)

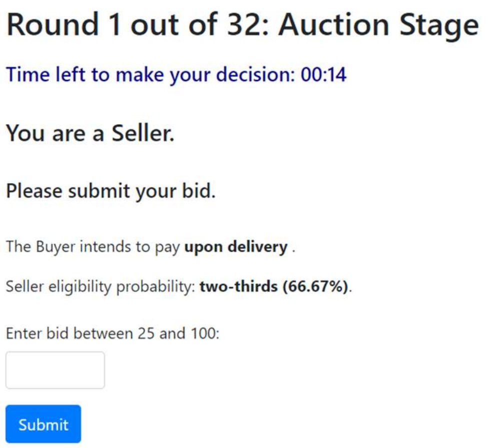
----[LC]----



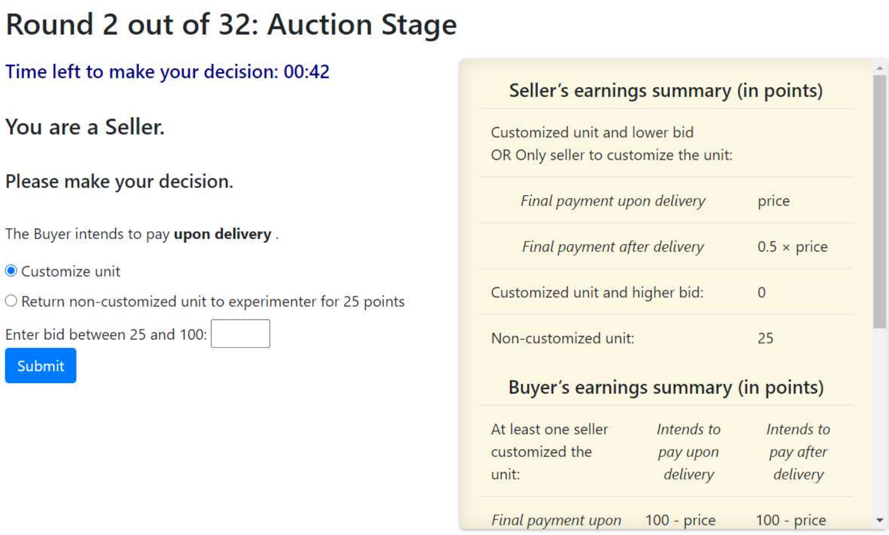
----[HC]----



----[FR]----



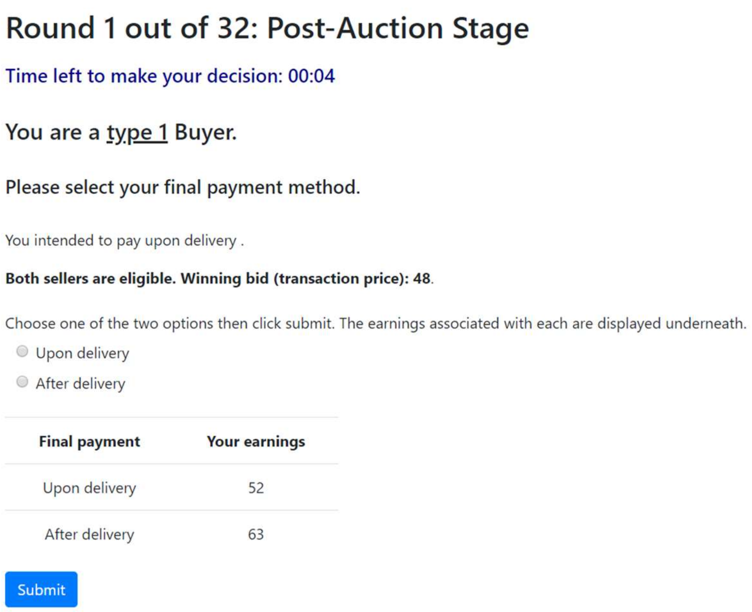
----[HC-E]----



Picture 3 gives an example of what your computer screen will look like in the Post-Auction stage. Buyers observe the outcome of the Auction stage and, if at least one seller ----[HC-E] chooses to customize the unit [LC, HC, FR] is eligible ----, must select their final payment method for the round, by choosing one of the two options and clicking submit. The buyer earnings associated with each option are displayed underneath.

P. 3. Post-Auction stage screen (for buyers only)

----[LC]----



----[HC]----

Round 1 out of 32: Post-Auction Stage

Time left to make your decision: 00:14

You are a type 1 Buyer.

Please select your final payment method.

You intended to pay upon delivery .

Both sellers are eligible. Winning bid (transaction price): 48.

Choose one of the two options then click submit. The earnings associated with each are displayed underneath.

- Upon delivery
- After delivery

Final payment	Your earnings
Upon delivery	52
After delivery	48

Submit

----[FR]----

Round 1 out of 32: Post-Auction Stage

Time left to make your decision: 00:13

You are a type 1 Buyer.

Please select your final payment method.

You intended to pay upon delivery .

Both sellers are eligible. Winning bid (transaction price): 48.

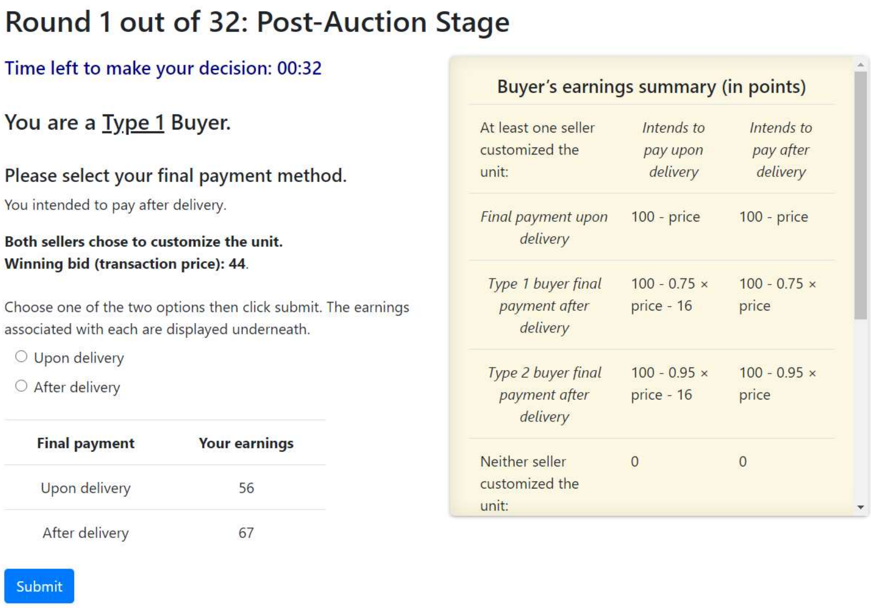
Choose one of the two options then click submit. The earnings associated with each are displayed underneath.

- Upon delivery
- After delivery

Final payment	Your earnings
Upon delivery	52
After delivery	56

Submit

----[HC-E]----



<u>Comprehension quiz</u>

Consider the following scenario.

Pre-Auction stage: The Buyer is randomly assigned to type 1. The Buyer announces an intention to make payment upon delivery.

----[LC, HC, FR]

1. What is the probability of a seller being eligible to customize the unit? Select one of the following options.
   - one-half (50%) *[LC]*
   - three-quarters (75%) *[HC]*
   - two-thirds (66.67%) *[FR]*

Auction stage: Seller 1 ----[HC-E] chooses to customize the unit and---- submits a bid of 60. Seller 2 ----[HC-E] chooses to customize the unit and---- submits a bid of 77. ----[LC, HC, FR] The computer determines that both sellers are eligible to customize the unit in the round. ---- Seller 1 proceeds to the post-auction stage. Seller 2 exits without making a transaction. The transaction price is equal to Seller 1's bid of 60.

----[LC, HC, FR]

2. What does Seller 2 earn? *[0]*

----[HC-E]

1. What does Seller 2 earn? *[0]*

2. What would Seller 2 have earned if he/she had chosen not to customize the unit? *[25]*

----[LC, HC, FR, HC-E]

Post-Auction stage : The Buyer makes final payment to Seller 1 upon delivery.

3. What does Seller 1 earn? *[60]*

4. What does the Buyer earn? *[40]*

If, in the Post-Auction stage , the buyer had instead made final payment to Seller 1 after delivery.

5. What would Seller 1 earn? *[30]*

6. Would a penalty be levied on the Buyer? *[Yes]*

C.4. Post-experiment questionnaire and subject characteristics

Age: Interval variable.

Years.

Online sample: Mean 24.31, Standard deviation 3.47, Minimum 19, Maximum 35.

Lab sample: Mean 22.72, Standard deviation 3.95, Minimum 19, Maximum 37.

Gender: Categorical variable.

Male (1); Female (2); Other (3); Prefer not to Say (4).

Online sample: 57.14%; 42.11%; 0.75%; 0%.

Lab sample: 55.77%; 42.31%; 1.92%; 0%.

Field of studies: Categorical variable.

Arts and Education (1); Economics and Finance (2); Business and Management (3); Social Sciences and Law (4); Medicine and Health Sciences (5); Engineering and Natural Sciences (6); Not a Student (7).

Online sample: 2.24%; 3.73%; 59.70%; 23.88%; 0%; 6.72%; 3.73%.

Lab sample: 11.54%; 19.23%; 13.46%; 30.77%; 11.54%; 11.54%; 1.92%.

Nationality: Categorical variable:

Central and Eastern Asia (1); Central and Western Africa (2); Central, South America and the Caribbean (3); Europe (excl. UK) (4); Middle East and North Africa (5); North America (6); Oceania (7); South and Eastern Africa (8); South-East Asia (9); Southern Asia (10); UK (11)

Online sample: 7.46%; 1.49%; 0%; 0%; 1.49%; 16.42%; 0%; 0%; 22.39%; 50%; 0.75%.

Lab sample: 0%; 3.85%; 1.92%; 23.08%; 1.92%; 0%; 0%; 1.92%; 1.92%; 7.69%; 57.69%.

Income: Categorical variable.

When you were 16 years of age, what was the income of your parents in comparison to other families in your country?

Far below average (1); Below average (2); Average (3); Above average (4); Far above average (5)

Online sample: 1.49%; 7.46%; 47.01%; 39.55%; 4.48%.

Lab sample: 9.62%; 9.62%; 38.46%; 36.54%; 5.77 %.

Risk Indices:

Based on Dohmen et al. (2011). Likert scale from 0 “Completely unwilling to take risks” to 10 “Completely willing to take risks”.

9) Are you generally a person who is fully willing to take risks or do you try to avoid taking risks?

Online sample: Mean 6.62, Standard deviation 1.91, Minimum 2, Maximum 10.

Lab sample: Mean 6.43, Standard deviation 2.14, Minimum 0, Maximum 10.

10) How would you rate your willingness to take risks in financial matters?

Online sample: Mean 5.57, Standard deviation 2.05, Minimum 0, Maximum 10.

Lab sample: Mean 5.50, Standard deviation 2.68, Minimum 0, Maximum 10.

Bibliography

Abbink, K., & Brandts, J. (2008). Pricing in Bertrand Competition with Increasing Marginal Costs. *Games and Economic Behavior, 63*(1), 1-31.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *Quarterly Journal of Economics, 97*(4), 543-569.

Akerlof, G. A., & Yellen, J. L. (1990). The Fair Wage-Effort Hypothesis and Unemployment. *Quarterly Journal of Economics, 105*(2), 255-283.

Alekseev, A., Charness, G., & Gneezy, U. (2017). Experimental methods: When and why contextual instructions are important. *Journal of Economic Behavior & Organization, 134*, 48-59.

Anderhub, V., Engelmann, D., & Güth, W. (2002). An Experimental Study of the Repeated Trust Game with Incomplete Information. *Journal of Economic Behavior and Organization, 48*(2), 197-216.

Andreoni, J. (2018). Satisfaction Guaranteed: When Moral Hazard Meets Moral Preferences. *American Economic Journal: Microeconomics, 10*(4), 159-189.

Andreoni, J., & Bernheim, B. D. (2009). Social Image and the Fifty-Fifty Norm: A Theoretical and Experimental Analysis of Audience Effects. *Econometrica, 77*(5), 1607-1636.

Arcadis. (2020). Global Construction Disputes Report. Retrieved from https://www.arcadis.com/en/united-states/our-perspectives/global-construction-disputes-report-2020/.

Argenton, C., & Müller, W. (2012). Collusion in experimental Bertrand duopolies with convex costs: The role of cost asymmetry. *International Journal of Industrial Organization, 30*(6), 508-517.

Arrow, K. J. (1972). Gifts and exchanges. *Philosophy & Public Affairs, 1*(4), 343-362.

ASA. (2018). Retainage Laws in the 50 States 2018. Alexandria, VA: Foundation of the American Subcontractors Association, Inc.

Ashenfelter, O. C., & Bloom, D. E. (1984). Models of Arbitrator Behavior: Theory and Evidence. *American Economic Review, 74*(1), 111-124.

Aycinena, D., Bejarano, H., & Rentschler, L. (2018). Informed Entry in Auctions.

International Journal of Game Theory, 47(1), 175-205.

Babich, V., & Tang, C. S. (2012). Managing opportunistic supplier product adulteration: Deferred payments, inspection, and combined mechanisms. *Manufacturing & Service Operations Management, 14*(2), 301-314.

Bajari, P., Houghton, S., & Tadelis, S. (2014). Bidding for incomplete contracts: An empirical analysis of adaptation costs. *American Economic Review, 104*(4), 1288-1319.

Bajari, P., McMillan, R., & Tadelis, S. (2009). Auctions versus negotiations in procurement: an empirical analysis. *Journal of Law, Economics, & Organization, 25*(2), 372-399.

Bajari, P., & Tadelis, S. (2001). Incentives versus Transaction Costs: A Theory of Procurement Contracts. *RAND Journal of Economics, 32*(3), 387-407.

Bartling, B., & Schmidt, K. M. (2015). Reference points, social norms, and fairness in contract renegotiations. *Journal of the European Economic Association, 13*(1), 98-129.

Bausman, D. C. (2004). *Retainage practice in the construction industry*. Alexandria, VA: Foundation of the American Subcontractors Association, Inc.

Baye, M. R., & Morgan, J. (1999). A Folk Theorem for One-Shot Bertrand Games. *Economics Letters, 65*(1), 59-65.

Baye, M. R., & Morgan, J. (2004). Price Dispersion in the Lab and on the Internet: Theory and Evidence. *RAND Journal of Economics, 35*(3), 449-466.

Bazerman, M. H., & Gillespie, J. J. (1999). Betting on the future: The virtues of contingent contracts. *Harvard Business Review, 77*, 155-162.

Bednar, J., Chen, Y., Liu, T. X., & Page, S. (2012). Behavioral spillovers and cognitive load in multiple games: An experimental study. *Games and Economic Behavior, 74*(1), 12-31.

Beer, R., Ahn, H.-S., & Leider, S. (2018). Can trustworthiness in a supply chain be signaled? *Management Science, 64*(9), 3974-3994.

Ben-Shahar, O. (2004). Agreeing to Disagree: Filling Gaps in Deliberately Incomplete Contracts Freedom from Contract Symposium. *Wisconsin Law Review, 2004*(2), 389-428.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122-142.

Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review, 71*, 117-120.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review, 90*(1), 166-193.

Bolton, G. E., Ockenfels, A., & Thonemann, U. W. (2012). Managers and Students as Newsvendors. *Management Science, 58*(12), 2225-2233.

Boone, J., Larrain Aylwin, M. J., Mueller, W., & Ray Chaudhuri, A. (2012). Bertrand Competition with Asymmetric Costs: Experimental Evidence. *Economics Letters, 117*(1), 134-137.

Bottomline. (2020). Business Payments Barometer. Retrieved from https://www.bottomline.com/uk/2020-business-payments-barometer.

Brandts, J., & Charness, G. (2004). Do Labour Market Conditions Affect Gift Exchange? Some Experimental Evidence. *Economic Journal, 114*(497), 684-708.

Brandts, J., Ellman, M., & Charness, G. (2015). Let's talk: How communication affects contract design. *Journal of the European Economic Association, 14*(4), 943-974.

Brosig-Koch, J., & Heinrich, T. (2014). Reputation and mechanism choice in procurement auctions: An experiment. *Production and Operations Management, 23*(2), 210-220.

Brown, M., Falk, A., & Fehr, E. (2004). Relational contracts and the nature of market interactions. *Econometrica, 72*(3), 747-780.

Bruttel, L. V. (2009). Group Dynamics in Experimental Studies - The Bertrand Paradox Revisited. *Journal of Economic Behavior and Organization, 69*(1), 51-63.

Cabinet Office. (2020). Guidance on responsible contractual behaviour in the performance and enforcement of contracts impacted by the Covid-19 emergency. *Government of the United Kingdom* (7 May).

Camerer, C., & Lovallo, D. (1999). Overconfidence and Excess Entry: An Experimental Approach. *American Economic Review, 89*(1), 306-318.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics, 90*(3), 414-427.

Casari, M., & Cason, T. N. (2016). Common Value Auctions with Voluntary and Qualified Entry. *Revue Economique, 67*(5), 953-976.

Cerić, A. (2016). *Trust in construction projects*. Oxon: Routledge.

Chakraborty, I., Khalil, F., & Lawarree, J. (2021). Competitive Procurement with Ex Post Moral Hazard. *CESifo Working Paper No. 8863, Available at SSRN: https://ssrn.com/abstract=3774670*.

Chakravarty, S., & MacLeod, W. B. (2009). Contracting in the shadow of the law. *RAND Journal of Economics, 40*(3), 533-557.

Chang, W.-S. (2019). The adverse effects of renegotiation in procurement auctions with endogenous liability. *Journal of Behavioral and Experimental Economics, 81*, 91-101.

Chang, W.-S., Salmon, T. C., & Saral, K. J. (2016). Procurement auctions with renegotiation and wealth constraints. *Economic Inquiry, 54*(3), 1684-1704.

Charness, G., & Dufwenberg, M. (2006). Promises and partnership. *Econometrica, 74*(6), 1579-1601.

Charness, G., & Rabin, M. (2002). Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics, 117*(3), 817-869.

Chaturvedi, A. (2021). Excessive Competition and Supplier Non-Performance Risk: Trade-offs in Reverse Auctions. *Production and Operations Management, Forthcoming*.

Che, Y.-K. (1993). Design competition through multidimensional auctions. *RAND Journal of Economics, 24*(4), 668-680.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance, 9*, 88-97.

Chen, J., Xu, L., & Whinston, A. (2009). Managing Project Failure Risk Through Contingent Contracts in Procurement Auctions. *Decision Analysis, 7*(1), 23-39.

Chen, L., & Lee, H. L. (2016). Sourcing Under Supplier Responsibility Risk: The Effects of Certification, Audit, and Contingency Payment. *Management Science, 63*(9), 2795-2812.

Choi, E. W., Özer, Ö., & Zheng, Y. (2020). Network Trust and Trust Behaviors Among Executives in Supply Chain Interactions. *Management Science, 66*(12), 5823-5849.

Committees publish correspondence on 'notorious late payers' Carillion. (2018, February 2). Retrieved from https://committees.parliament.uk/committee/164/work-and-pensions-committee/news/97648/committees-publish-correspondence-on-notorious-late-payers-carillion/.

Connell, W. (2014). *The Economic Impact of Late Payments*. European Commission Economic Papers (531).

Construction Contracts Amendment Act 2015. Retrieved from: https://www.legislation.govt.nz/act/public/2015/0092/latest/DLM4998107.html.

Conti, M., Elia, L., Ferrara, A. R., & Ferraresi, M. (2020). *Government Late Payments and Firms Survival. Evidence from the EU*. Società italiana di economia pubblica. Working Paper No. 753

Cooper, D. J., & Kagel, J. H. (2003). The impact of meaningful context on strategic play in signaling games. *Journal of Economic Behavior & Organization, 50*(3), 311-337.

Corbin, A. L. (1919). Conditions in the Law of Contract. *Yale Law Journal, 28*(8), 739-768.

Cox, J. C., Dinkin, S., & Swarthout, J. T. (2001). Endogenous Entry and Exit in Common Value Auctions. *Experimental Economics, 4*(2), 163-181.

Cox, J. C., Isacc, R. M., Cech, P., & Conn, D. (1996). Moral Hazard and Adverse Selection in Procurement Contracting. *Games and Economic Behavior, 17*(2), 147-176.

Cox, S. C., H.; Udom, K.; Lupton, S. (2011). *Which contract? Choosing the appropriate building contract* (5th ed.). London: RIBA Publishing.

Crawford, V. P. (1981). Arbitration and Conflict Resolution in Labor-Management Bargaining. *American Economic Review, 71*(2), 205-210.

Crawford, V. P. (1985). The role of arbitration and the theory of incentives. In A. E. Roth (Ed.), *Game-Theoretic Models of Bargaining*. Cambridge, MA: Cambridge Univ. Press.

Cui, T. H., Raju, J. S., & Zhang, Z. J. (2007). Fairness and channel coordination. *Management Science, 53*(8), 1303-1314.

Dastidar, K. G. (1995). On the Existence of Pure Strategy Bertrand Equilibrium. *Economic Theory, 5*(1), 19-32.

Davis, A. M. (2015). An Experimental Investigation of Pull Contracts in Supply Chains. *Production and Operations Management, 24*(2), 325-340.

Davis, A. M., & Hyndman, K. (2018). An Experimental Investigation of Managing Quality Through Monetary and Relational Incentives. *Management Science, 64*(5), 2345-2365.

Davis, A. M., & Leider, S. (2018). Contracts and Capacity Investment in Supply Chains. *Manufacturing & Service Operations Management, 20*(3), 403-421.

Department for Business, Energy, & Industrial Strategy. (2016). Duty to report on payment practices and performance. UK Department for Business, Energy & Industrial Strategy (BEIS). Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/574312/duty-to-report-on-payment-practices-and-performance-government-response.pdf.

Djankov, S., Ghossein, T., Islam, A. M., & Saliola, F. (2017). Public procurement regulation and road quality. *The World Bank*.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association, 9*(3), 522-550.

Dufwenberg, M., & Gneezy, U. (2000). Price Competition and Market Concentration: An Experimental Study. *International Journal of Industrial Organization, 18*(1), 7-22.

Dufwenberg, M., & Gneezy, U. (2002). Information Disclosure in Auctions: An Experiment. *Journal of Economic Behavior and Organization, 48*(4), 431-444.

Dufwenberg, M., Gneezy, U., Goeree, J. K., & Nagel, R. (2007). Price Floors and Competition. *Economic Theory, 33*(1), 211-224.

Dugar, S., & Mitra, A. (2016). Bertrand Competition with Asymmetric Marginal Costs. *Economic Inquiry, 54*(3), 1631-1647.

Dyer, D., Kagel, J. H., & Levin, D. (1989). A Comparison of Naive and Experienced Bidders in Common Value Offer Auctions: A Laboratory Analysis. *Economic Journal, 99*(394), 108-115.

Edgeworth, F. Y. (1925). *Papers Relating to Political Economy*. London: Macmillan and Co.

Elberfeld, W., & Wolfstetter, E. (1999). A Dynamic Model of Bertrand Competition with Entry. *International Journal of Industrial Organization, 17*(4), 513-525.

Ellingsen, T., & Johannesson, M. (2004). Promises, threats and fairness. *Economic Journal, 114*(495), 397-420.

Elmaghraby, W., & Katok, E. (2018). Behavioral research in competitive bidding and auction design. In K. Donohue, E. Katok, & S. Leider (Eds.), *The Handbook of Behavioral Operations*. Hoboken, NJ: Wiley.

Emiliani, M., & Stec, D. (2005). Commentary on "Reverse auctions for relationship marketers" by Daly and Nath. *Industrial Marketing Management, 34*(2), 167-171.

Engelbrecht-Wiggans, R. (1987). On optimal reservation prices in auctions. *Management Science, 33*(6), 763-770.

Engelbrecht-Wiggans, R., Haruvy, E., & Katok, E. (2007). A comparison of buyer-determined and price-based multiattribute mechanisms. *Marketing Science, 26*(5), 629-641.

Engelbrecht-Wiggans, R., & Katok, E. (2005). Experiments on Auction Valuation and Endogenous Entry. In J. Morgan (Ed.), *Experimental and Behavorial Economics* (Vol. 13, pp. 169-193): Emerald Group Publishing Limited.

Engelmann, D., Frank, J., Koch, A. K., & Valente, M. (2020). Second-Chance Offers and Buyer Reputation: Theory and Evidence on Auctions with Default. *CRC 190 Rationality and Competition Discussion Paper No. 237.*

Epps, T. W., & Singleton, K. J. (1986). An omnibus test for the two-sample problem using the empirical characteristic function. *Journal of Statistical Computation and Simulation, 26*(3-4), 177-203.

Ertaç, S., Hortaçsu, A., & Roberts, J. W. (2011). Entry into Auctions: An Experimental Analysis. *International Journal of Industrial Organization, 29*(2), 168-178.

European Commission. (2009). International Accounting Standard 11: Construction Contracts. Retrieved from https://www.ucetni-portal.cz/stahnout/ias-11-en_842.pdf.

Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *American Economic Review, 96*(5), 1611-1630.

Farber, H. S. (1980). An Analysis of Final-Offer Arbitration. *Journal of conflict resolution, 24*(4), 683-705.

Farber, H. S., & Katz, H. C. (1979). Interest Arbitration, Outcomes, and the Incentive to Bargain. *Industrial and Labor Relations Review, 33*(1), 55-63.

Fatas, E., Haruvy, E., & Morales, A. J. (2014). A Psychological Reexamination of the Bertrand Paradox. *Southern Economic Journal, 80*(4), 948-967.

Fehr, E., & Falk, A. (1999). Wage rigidity in a competitive incomplete contract market. *Journal of political Economy, 107*(1), 106-134.

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior, 25*(2), 63-87.

Fehr, E., & Gächter, S. (2000). Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives, 14*(3), 159-181.

Fehr, E., Gächter, S., & Kirchsteiger, G. (1997). Reciprocity as a Contract Enforcement Device: Experimental Evidence. *Econometrica, 65*(4), 833-860.

Fehr, E., Hart, O., & Zehnder, C. (2011). Contracts as reference points—experimental evidence. *American Economic Review, 101*(2), 493-525.

Fehr, E., Hart, O., & Zehnder, C. (2015). How do informal agreements and revision shape contractual reference points? *Journal of the European Economic*

Association, 13(1), 1-28.

Fehr, E., Kirchler, E., Weichbold, A., & Gächter, S. (1998). When social norms overpower competition: Gift exchange in experimental labor markets. *Journal of Labor Economics, 16*(2), 324-351.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics, 108*(2), 437-459.

Fehr, E., Klein, A., & Schmidt, K. M. (2007). Fairness and contract design. *Econometrica, 75*(1), 121-154.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics, 114*(3), 817-868.

Fischbacher, U., & Thöni, C. (2008). Excess entry in an experimental winner-take-all market. *Journal of Economic Behavior & Organization, 67*(1), 150-163.

Fonseca, M. A., & Normann, H.-T. (2008). Mergers, asymmetries and collusion: Experimental evidence. *Economic Journal, 118*(527), 387-400.

Fonseca, M. A., & Normann, H.-T. (2013). Excess Capacity and Pricing in Bertrand-Edgeworth Markets: Experimental Evidence. *Journal of Institutional and Theoretical Economics, 169*(2), 199-228.

Fouraker, L., & Siegel, S. (1963). *Bargaining Behavior*. New York: McGraw-Hill.

Fréchette, G. R. (2012). Session-effects in the laboratory. *Experimental Economics, 15*(3), 485-498.

Fréchette, G. R. (2015). Laboratory Experiments: Professionals versus Students. In G. R. Fréchette & A. Schotter (Eds.), *Handbook of Experimental Economic Methodology* (pp. 360-390). Oxford: Oxford University Press.

FSB. (2016). Time to Act. The Economic Impact of Poor Payment Practice. The (UK) Federation of Small Businesses. Report, November 2016. Retrieved from https://www.fsb.org.uk/static/517120db-2555-473f-a6ceb5c661d569fb/Time-to-Act.pdf.

FSB. (2018). Chain Reaction. Improving the Supply Chain Experience for Smaller Firms. The (UK) Federation of Small Businesses. Report, June 2018. Retrieved from https://www.fsb.org.uk/static/b51e56d3-79d9-4611-

9fb591a99ac893d8/Chain-reaction-report.pdf.

Fugger, N., Gretschko, V., & Pollrich, M. (2019). Sequential procurement with limited commitment. *ZEW - Centre for European Economic Research Discussion Paper No. 19-030*.

Fugger, N., Katok, E., & Wambach, A. (2016). Collusion in dynamic buyer-determined reverse auctions. *Management Science, 62*(2), 518-533.

Fugger, N., Katok, E., & Wambach, A. (2019). Trust in procurement interactions. *Management Science, 65*(11), 5110-5127.

Gabuthy, Y., & Muthoo, A. (2019). Bargaining and Hold-Up: The Role of Arbitration. *Oxford Economic Papers, 71*(1), 292-308.

Genus, A. (1997). Unstructuring incompetence: problems of contracting, trust and the development of the channel tunnel. *Technology Analysis & Strategic Management, 9*(4), 419-436.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring Trust. *Quarterly Journal of Economics, 115*(3), 811-846.

Gneezy, U. (2005). Step-level reasoning and bidding in auctions. *Management Science, 51*(11), 1633-1642.

Goeree, J. K., & Holt, C. A. (2005). An explanation of anomalous behavior in models of political participation. *American political science Review, 99*(2), 201-213.

Grether, D. M., Schwartz, A., & Wilde, L. L. (1988). Uncertainty and shopping behaviour: an experimental analysis. *Review of Economic Studies, 55*(2), 323-342.

Gretschko, V., & Pollrich, M. (2021). Incomplete contracts in dynamic procurement. *Management Science, forthcoming*.

Guo, R., Lee, H. L., & Swinney, R. (2015). Responsible Sourcing in Supply Chains. *Management Science, 62*(9), 2722-2744.

Hart, O. (2017). Incomplete Contracts and Control. *American Economic Review, 107*(7), 1731-1752.

Hart, O., & Moore, J. (1990). Property Rights and the Nature of the Firm. *Journal of political Economy, 98*(6), 1119-1158.

Hart, O., & Moore, J. (1999). Foundations of Incomplete Contracts. *Review of Economic Studies, 66*(1), 115-138.

Hart, O., & Moore, J. (2008). Contracts as reference points. *Quarterly Journal of Economics, 123*(1), 1-48.

Haruvy, E., & Katok, E. (2013). Increasing Revenue by Decreasing Information in Procurement Auctions. *Production and Operations Management, 22*(1), 19-35.

Herold, F. (2010). Contractual incompleteness as a signal of trust. *Games and Economic Behavior, 68*(1), 180-191.

Herweg, F., & Schmidt, K. M. (2017). Auctions versus Negotiations: The Effects of Inefficient Renegotiation. *RAND Journal of Economics, 48*(3), 647-672.

Herweg, F., & Schmidt, K. M. (2020). Procurement with unforeseen contingencies. *Management Science, 66*(5), 2194-2212.

Herweg, F., & Schwarz, M. A. (2018). Optimal Cost Overruns: Procurement Auctions with Renegotiation. *International Economic Review, 59*(4), 1995-2021.

Heymann, D., Kawamura, E., Perazzo, R., & Zimmermann, M. (2014). Behavioral heuristics and market patterns in a Bertrand–Edgeworth game. *Journal of Economic Behavior & Organization, 105*, 124-139.

Ho, T.-H., & Zhang, J. (2008). Designing Pricing Contracts for Boundedly Rational Customers: Does the Framing of the Fixed Fee Matter? *Management Science, 54*(4), 686-700.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics, 6*(2), 65-70.

Hoppe, E., & Schmitz, P. (2011). Can contracts solve the hold-up problem? Experimental evidence. *Games and Economic Behavior, 73*(1), 186-199.

Howorth, C., & Reber, B. (2003). Habitual Late Payment of Trade Credit: An Empirical Examination of UK Small Firms. *Managerial and Decision Economics, 24*(6-7), 471-482.

Hu, S., Wan, Z., Ye, Q., & Chi, W. (2017). Supplier behavior in capacity investment competition: An experimental study. *Production and Operations Management, 26*(2), 273-291.

IACCM. (2015). Payment Terms: Do Large Companies Abuse Their Power?. Report, International Association for Contract & Commercial Management. Retrieved from https://blog.iaccm.com/free-resources/payment-terms-do-large-companies-abuse-their-power.

Ivanova-Stenzel, R., & Salmon, T. C. (2008). Revenue equivalence revisited. *Games and Economic Behavior, 64*(1), 171-192.

Ivanova-Stenzel, R., & Salmon, T. C. (2011). The High/Low Divide: Self-Selection by Values in Auction Choice. *Games and Economic Behavior, 73*(1), 200-214.

Janssen, M., & Rasmusen, E. (2002). Bertrand Competition under Uncertainty. *Journal of Industrial Economics, 50*(1), 11-21.

Jap, S. D. (2007). The impact of online reverse auction design on buyer–supplier relationships. *Journal of Marketing, 71*(1), 146-159.

Joseph, A., Kneer, C., Horen, N. v., & Saleheen, J. (2020). Cash in the time of corona. *VoxEU & CEPR* (26 April), https://voxeu.org/article/cash-time-corona.

Karakostas, A., Sonntag, A., & Zizzo, D. J. (2017). Contract Choice: Efficiency and Fairness in Revenue-Sharing Contracts. *Scandinavian Journal of Economics, 119*(4), 962-986.

Katok, E. (2011). Using laboratory experiments to build better operations management models. *Foundations and Trends in Technology, Information and Operations Management, 5*(1), 1-86.

Katok, E. (2017). Designing and Conducting Laboratory Experiments. In K. Donohue, E. Katok, & S. Leider (Eds.), *The Handbook of Behavioral Operations*. Hoboken, NJ: Wiley.

Katok, E., & Pavlov, V. (2013). Fairness in supply chain contracts: A laboratory study. *Journal of Operations Management, 31*(3), 129-137.

Keser, C., & Willinger, M. (2000). Principals' principles when agents' actions are hidden. *International Journal of Industrial Organization, 18*(1), 163-185.

Konow, J. (2003). Which Is the Fairest One of All? A Positive Analysis of Justice Theories. *Journal of Economic Literature, 41*(4), 1188-1239.

Kőszegi, B., & Rabin, M. (2006). A Model of Reference-Dependent Preferences.

Quarterly Journal of Economics, 121(4), 1133-1165.

Kouvelis, P., & Zhao, W. (2012). Financing the newsvendor: supplier vs. bank, and the structure of optimal trade credit contracts. *Operations Research, 60*(3), 566-580.

Kreps, D. M. (1990). Corporate culture and economic theory. In J. E. Alt & K. A. Shepsle (Eds.), *Perspectives on Positive Political Economy* (pp. 90-143). Cambridge: Cambridge University Press.

Kreps, D. M., & Scheinkman, J. A. (1983). Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes. *Bell Journal of Economics, 14*(2), 326-337.

Kruse, J. B., Rassenti, S., Reynolds, S. S., & Smith, V. L. (1994). Bertrand-Edgeworth competition in experimental markets. *Econometrica, 62*(2), 343-371.

Kuhn, M. A. (2009). *To settle or not to settle: a review of the literature on arbitration in the laboratory*. Unpublished Manuscript, University of California, San Diego.

Lang, K., & Rosenthal, R. W. (1991). The Contractors' Game. *RAND Journal of Economics, 22*(3), 329-338.

Lee, Y. S., Ribbink, D., & Eckerd, S. (2018). Effectiveness of bonus and penalty incentive contracts in supply chain exchanges: Does national culture matter? *Journal of Operations Management, 62*, 59-74.

Lei, V., Noussair, C. N., & Plott, C. R. (2001). Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. actual irrationality. *Econometrica, 69*(4), 831-859.

Levin, D., & Smith, J. L. (1994). Equilibrium in Auctions with Entry. *American Economic Review, 84*(3), 585-599.

Li, C., Turmunkh, U., & Wakker, P. P. (2019). Trust as a decision under ambiguity. *Experimental Economics, 22*(1), 51-75.

Lo, W., Lin, C. L., & Yan, M. R. (2007). Contractor's Opportunistic Bidding Behavior and Equilibrium Price Level in the Construction Market. *Journal of Construction Engineering and Management, 133*(6), 409-416.

Loch, C. H., & Wu, Y. (2008). Social preferences and supply chain performance: An experimental study. *Management Science, 54*(11), 1835-1849.

Long, M. S., Malitz, I. B., & Ravid, S. A. (1993). Trade Credit, Quality Guarantees, and Product Marketability. *Financial management, 22*(4), 117-127.

Lu, J. (2010). Entry Coordination and Auction Design with Private Costs of Information Acquisition. *Economic Inquiry, 48*(2), 274-289.

MacLeod, W. B., & Malcomson, J. M. (1998). Motivation and Markets. *American Economic Review, 88*(3), 388-411.

Marquez, R. (1997). A Note on Bertrand Competition with Asymmetric Fixed Costs. *Economics Letters, 57*(1), 87-96.

McAfee, R. P., & McMillan, J. (1986). Bidding for Contracts: A Principal-Agent Analysis. *RAND Journal of Economics, 17*(3), 326-338.

McAfee, R. P., & McMillan, J. (1987a). Auctions with entry. *Economics Letters, 23*(4), 343-347.

McAfee, R. P., & McMillan, J. (1987b). Competition for Agency Contracts. *RAND Journal of Economics, 18*(2), 296-307.

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior, 10*(1), 6-38.

Midler, P. (2007). 'Quality Fade': China's Great Business Challenge. *Knowledge@Wharton*, Retrieved from https://knowledge.wharton.upenn.edu/article/quality-fade-chinas-great-business-challenge/.

Miller, T., & Wongsaroj, S. (2017). *The Domino Effect: The Impact of Late Payments. Prepared by Plum Consulting for Sage.* Retrieved from https://www.sage.com/en-gb/blog/wp-content/uploads/sites/10/2017/12/Domino-Effect-Late-Payments-Research-Sage.pdf.

Moffatt, P. G. (2015). *Experimetrics: Econometrics for Experimental Economics*: Palgrave Macmillan.

Moreno, D., & Wooders, J. (2011). Auctions with Heterogeneous Entry Costs. *RAND Journal of Economics, 42*(2), 313-336.

Nabi Mohamad, A., Ali, G., El-adaway Islam, H., Garza Louis, A., Tichy, S., & Girse,

J. (2021). Contractual Guidelines for Substantial Completion under National Design-Build Standard Forms of Contract. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 13*(1), 04520043.

Nelson, F. D. (1976). On a General Computer Algorithm for the Analysis of Models with Limited Dependent Variables. *Annals of Economic and Social Measurement, 5*(4), 493-509.

Office for National-Statistics. (2020). Coronavirus and the economic impacts on the UK: 10 September 2020. Retrieved from https://www.ons.gov.uk/businessindustryandtrade/business/businessservices/bulletins/coronavirusandtheeconomicimpactsontheuk/10september2020.

Onderstal, S., & Yang, Y. (2020). Cheap-talk Communication in Procurement Auctions: Theory and Experiment. *Tinbergen Institute Discussion Paper 2020-013/VII.*

Özer, Ö., & Zheng, Y. (2019). Trust and Trustworthiness. In K. Donohue, E. Katok, & S. Leider (Eds.), *The Handbook of Behavioral Operations* (pp. 489–523). Hoboken, NJ: John Wiley & Sons Inc.

Özer, Ö., Zheng, Y., & Chen, K.-Y. (2011). Trust in forecast information sharing. *Management Science, 57*(6), 1111-1137.

Özer, Ö., Zheng, Y., & Ren, Y. (2014). Trust, Trustworthiness, and Information Sharing in Supply Chains Bridging China and the United States. *Management Science, 60*(10), 2435-2460.

Palfrey, T. R., & Pevnitskaya, S. (2008). Endogenous Entry and Self-Selection in Private Value Auctions: An Experimental Study. *Journal of Economic Behavior and Organization, 66*(3-4), 731-747.

Peura, H., Yang, S. A., & Lai, G. (2017). Trade credit in competition: a horizontal benefit. *Manufacturing & Service Operations Management, 19*(2), 263-289.

Pevnitskaya, S. (2004). *Endogenous entry in first-price private value auctions: the self-selection effect* Working Paper. The Ohio State University

Pinsent Masons. (2019). International Arbitration Survey – Driving Efficiency in International Construction Disputes. Retrieved from: https://www.pinsentmasons.com/thinking/special-reports/international-arbitration-survey.

Pye Tait Consulting. (2017). Retentions in the construction industry. BEIS Research Paper 17, Department for Business, Energy and Industrial Strategy. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/654399/Retention_Payments_Pye_Tait_report.pdf.

Raina, P., & Tookey, J. (2013). *The perceptions of retentions as held by clients, contractors and subcontractors*. Mimeo.

Rauwald, C. (2016). VW Said to Pay Compensation to Supplier to End Standoff. *Bloomberg* (August 22), https://www.bloomberg.com/news/articles/2016-2008-2023/vw-said-to-pay-compensation-to-supplier-to-end-six-day-standoff.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal, 19*(1), 4-60.

Rubin, J., & Sheremeta, R. (2016). Principal–agent settings with random shocks. *Management Science, 62*(4), 985-999.

Samuelson, W. F. (1985). Competitive bidding with entry costs. *Economics Letters, 17*(1-2), 53-57.

Santamaría, N. (2015). An Analysis of Scoring and Buyer-Determined Procurement Auctions. *Production and Operations Management, 24*(1), 147-158.

Seifert, D., Seifert, R. W., & Protopappa-Sieke, M. (2013). A review of trade credit literature: Opportunities for research in operations. *European Journal of Operational Research, 231*(2), 245-256.

Seligson, H. (2011). Learn How to Collect From Slow Payers. *The New York Times* (April 6), https://www.nytimes.com/2011/2004/2007/business/smallbusiness/2007sbiz.html.

Shachat, J., & Swarthout, J. T. (2010). Procurement auctions for differentiated goods. *Decision Analysis, 7*(1), 6-22.

Shachat, J., & Tan, L. (2015). An experimental investigation of auctions and bargaining in procurement. *Management Science, 61*(5), 1036-1051.

Sharkey, W. W., & Sibley, D. S. (1993). Optimal Non-linear Pricing with Regulatory Preference over Customer Type. *Journal of Public Economics, 50*(2), 197-229.

Simon, H. (1947). *Administrative behavior*. New York: Free Press.

Singh, N., & Vives, X. (1984). Price and Quantity Competition in a Differentiated Duopoly. *RAND Journal of Economics, 15*(4), 546-554.

Skrzypacz, A. (2013). Auctions with contingent payments — An overview. *International Journal of Industrial Organization, 31*(5), 666-675.

Small Business Commissioner. (2021). Government tackles late payments to small businesses to protect jobs (19 January). Retrieved from https://www.smallbusinesscommissioner.gov.uk/news/government-tackles-late-payments-to-small-businesses-to-protect-jobs/.

Smith, G. (2016). VW's battle with contractors gets unusually messy. *Fortune* (August 22), https://fortune.com/2016/2008/2022/vw-supplier-dispute-production/.

Smith, J. L., & Levin, D. (2001). Entry Coordination in Auctions and Social Welfare: An Experimental Investigation. *International Journal of Game Theory, 30*(3), 321-350.

Smith, V. L. (1982). Microeconomic Systems as an Experimental Science. *American Economic Review, 72*(5), 923-955.

Spiliotopoulou, E., Donohue, K., & Gürbüz, M. Ç. (2016). Information Reliability in Supply Chains: The Case of Multiple Retailers. *Production and Operations Management, 25*(3), 548-567.

Stinchcombe, A. L., & Heimer, C. A. (1985). *Organisation Theory and Project Management*. Oslo: Norwegian University Press.

Strom, S. (2015). Big Companies Pay Later, Squeezing Their Suppliers. *The New York Times* (April 6), https://www.nytimes.com/2015/2004/2007/business/big-companies-pay-later-squeezing-their-suppliers.html.

Thomas, C. J. (2002). The Effect of Asymmetric Entry Costs on Bertrand Competition. *International Journal of Industrial Organization, 20*(5), 589-609.

Thomas, H. R., Smith Gary, R., & Cummings Daniel, J. (1995). Have I Reached Substantial Completion? *Journal of Construction Engineering and Management,* *121*(1), 121-129.

Tirole, J. (1986). Procurement and Renegotiation. *Journal of political Economy, 94*(2), 235-259.

Verdeaux, J.-J. (2003). Public procurement in the European Union and in the United States: a comparative study. *Public Contract Law Journal, 32*(4), 713-738.

Waehrer, K. (1995). A Model of Auction Contracts with Liquidated Damages. *Journal of economic theory, 67*(2), 531-555.

Wambach, A. (1999). Bertrand Competition under Cost Uncertainty. *International Journal of Industrial Organization, 17*(7), 941-951.

Wang, R. (2000). Bidding and Renegotiation in Procurement Auctions. *European Economic Review, 44*(8), 1577-1597.

Weitzman, M. L. (1980). Efficient Incentive Contracts. *Quarterly Journal of Economics, 94*(4), 719-730.

Wevill, J. (2015). *Law in Practice: The RIBA Legal Handbook* (2nd ed.). London: RIBA.

Williamson, O. E. (1973). Markets and Hierarchies: Some Elementary Considerations. *American Economic Review, 63*(2), 316-325.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism*: Simon and Schuster.

Williamson, O. E. (1988). The Logic of Economic Organization Conference Papers to Celebrate the Fiftieth Anniversary of the Nature Of The Firm. *Journal of Law, Economics & Organization, 4*(1), 65-94.

Zhao, S., López Vargas, K., Friedman, D., & Gutierrez, M. (2020). UCSC LEEPS Lab Protocol for Online Economics Experiments. *Available at SSRN: https://ssrn.com/abstract=3594027*.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics, 13*(1), 75-98.

Zizzo, D. J. (2013). Claims and confounds in economic experiments. *Journal of Economic Behavior & Organization, 93*, 186-195.