Bargaining for Power and Information:

An Experimental and Theoretical Analysis of the Interaction between Institutions, Externalities and Adverse Selection

Inaugural dissertation submitted by Simon Siegenthaler in fulfillment of the requirements for the degree of Doctor rerum oeconomicarum at the Faculty of Business, Economics and Social Sciences of the University of Bern

Submitted by
Simon, Siegenthaler
from Trubschachen, Bern

2014

The faculty accepted this work as dissertation on December 18th 2014 at the request of Prof. Dr. Olivier Bochet, Prof. Dr. Pedro dal Bó and Prof. Dr. Rajiv Vohra, without wishing to take position on the view presented therein.

Bargaining for Power and Information

An Experimental and Theoretical Analysis of the Interaction between Institutions, Externalities and Adverse Selection

by

Simon Siegenthaler

from Trubschachen, Bern

Submitted to the Department of Economics and
the Faculty of Business, Economics and Social Sciences
on December 1, 2014, in partial fulfillment of the
requirements for the degree of
Doctor rerum oeconomicarum

Abstract

Bargaining is the building block of many economic interactions, ranging from bilateral to multilateral encounters and from situations in which the actors are individuals to negotiations between firms or countries. In all these settings, economists have been intrigued for a long time by the fact that some projects, trades or agreements are not realized even though they are mutually beneficial. On the one hand, this has been explained by incomplete information. A firm may not be willing to offer a wage that is acceptable to a qualified worker, because it knows that there are also unqualified workers and cannot distinguish between the two types. This phenomenon is known as adverse selection. On the other hand, it has been argued that even with complete information, the presence of externalities may impede efficient outcomes. To see this, consider the example of climate change. If a subset of countries agrees to curb emissions, non-participant regions benefit from the signatories' efforts without incurring costs. These free riding opportunities give rise to incentives to strategically improve ones bargaining power that work against the formation of a global agreement.

This thesis is concerned with extending our understanding of both factors, adverse selection and externalities. The findings are based on empirical evidence from original laboratory experiments as well as game theoretic modeling. On a very general note, it is demonstrated that the institutions through which agents interact matter to a large extent. Insights are provided about which institutions we should expect to perform better than others, at least in terms of aggregate welfare.

Chapters 1 and 2 focus on the problem of adverse selection. Effective operation of markets and other institutions often depends on good information transmission properties.

In terms of the example introduced above, a firm is only willing to offer high wages if it receives enough positive signals about the worker's quality during the application and wage bargaining process. In Chapter 1, it will be shown that repeated interaction coupled with time costs facilitates information transmission. By making the wage bargaining process costly for the worker, the firm is able to obtain more accurate information about the worker's type. The cost could be pure time cost from delaying agreement or cost of effort arising from a multi-step interviewing process. In Chapter 2, I abstract from time cost and show that communication can play a similar role. The simple fact that a worker states to be of high quality may be informative.

In Chapter 3, the focus is on a different source of inefficiency. Agents strive for bargaining power and thus may be motivated by incentives that are at odds with the socially efficient outcome. I have already mentioned the example of climate change. Other examples are coalitions within committees that are formed to secure voting power to block outcomes or groups that commit to different technological standards although a single standard would be optimal (e.g. the format war between HD and BlueRay). It will be shown that such inefficiencies are directly linked to the presence of externalities and a certain degree of irreversibility in actions. I now discuss the three articles in more detail.

In Chapter 1, Olivier Bochet and I study a simple bilateral bargaining institution that eliminates trade failures arising from incomplete information. In this setting, a buyer makes offers to a seller in order to acquire a good. Whenever an offer is rejected by the seller, the buyer may submit a further offer. Bargaining is costly, because both parties suffer a (small) time cost after any rejection. The difficulties arise, because the good can be of low or high quality and the quality of the good is only known to the seller. Indeed, without the possibility to make repeated offers, it is too risky for the buyer to offer prices that allow for trade of high quality goods. When allowing for repeated offers, however, at equilibrium both types of goods trade with probability one. We provide an experimental test of these predictions. Buyers gather information about sellers using specific price offers and rates of trade are high, much as the model's qualitative predictions. We also observe a persistent over-delay before trade occurs, and this mitigates efficiency substantially. Possible channels for over-delay are identified in the form of two behavioral assumptions missing from the standard model, loss aversion (buyers) and haggling (sellers), which reconcile the data with the theoretical predictions.

Chapter 2 also studies adverse selection, but interaction between buyers and sellers now takes place within a market rather than isolated pairs. Remarkably, in a market it suffices to let agents communicate in a very simple manner to mitigate trade failures. The key insight is that better informed agents (sellers) are willing to truthfully reveal their private information, because by doing so they are able to reduce search frictions and attract more buyers. Behavior observed in the experimental sessions closely follows the theoretical predictions. As a consequence, costless and non-binding communication (cheap talk) significantly raises rates of trade and welfare. Previous experiments have documented that cheap talk alleviates inefficiencies due to asymmetric information. These findings are explained by pro-social preferences and lie aversion. I use appropriate control treatments to show that such consideration play only a minor role in our market. Instead, the experiment highlights the ability to organize markets as a new channel through which communication can facilitate trade in the presence of private information.

In Chapter 3, I theoretically explore coalition formation via multilateral bargaining under complete information. The environment studied is extremely rich in the sense that the model allows for all kinds of externalities. This is achieved by using so-called partition functions, which pin down a coalitional worth for each possible coalition in each possible coalition structure. It is found that although binding agreements can be written, efficiency is not guaranteed, because the negotiation process is inherently non-cooperative. The prospects of cooperation are shown to crucially depend on i) the degree to which players can renegotiate and gradually build up agreements and ii) the absence of a certain type of externalities that can loosely be described as incentives to free ride. Moreover, the willingness to concede bargaining power is identified as a novel reason for gradualism. Another key contribution of the study is that it identifies a strong connection between the Core, one of the most important concepts in cooperative game theory, and the set of environments for which efficiency is attained even without renegotiation.

Acknowledgments

My gratitude goes to everyone who helped me gain this invaluable experience. First of all, I would like to thank Olivier Bochet. Olivier's lectures on economic design and experimental economics sparked my interest in pursuing a doctoral degree in economics. It was his encouragement, curiosity, patience, and bright academic mind that helped me develop my thesis. I am grateful to Olivier for the many opportunities that opened up due to his efforts and involvement. It has been a privilege to be under his supervision.

The University of Bern offered me an intellectually stimulating environment. I am especially grateful to Gunter Stephan who always made sure that I would be able to overcome all the hurdles along my dissertation path. His guidance, support and encouragement was essential for the completion of my thesis. I also benefitted from Jérémy Laurent-Lucchetti, Ralph Winkler and many other faculty members of the economics department. I would like to thank Rajiv Vohra with whom I had several inspiring discussions during my visit at Brown University in spring 2014. Rajiv's impressive ability to identify relevant questions formed the basis of the last chapter of this thesis. I thank my fellow PhD students and friends for sharing thoughts and laughter: Raphael Bucher, Seraina Buob, Christin Erb, Boris Kaiser, John McNeill, Patrick Röösli, Oliver Schenker, Christian Schmid, Michael Siegenthaler, Angela Steffen and Sarina Steinmann.

Thanks to my friends and family who have enriched my life and have never been other than supportive. Thanks to my parents for their unconditional support. Last but not least, I would like to thank Nadine Josephine Saner for her love and for believing in my abilities. You mean the world to me!

Contents

List of Figures

List of Tables

Chapter 1

Better Later than Never? An Experiment on Bargaining under Adverse Selection*

1.1 Introduction

An important issue in economics is why mutually beneficial agreements are often hard to reach. While there are many possible impediments to reaching efficient agreements, an obvious obstacle is the asymmetry of information that may prevail among parties. For instance, when adverse selection is *severe*, the price mechanism fails to allocate goods efficiently and the market for high quality goods breaks down, Akerlof (1970).[1] While first-best efficiency is usually out-of-reach, institutions that differ from Walrasian markets may help alleviating the adverse selection effect. In real-life situations, where asymmetry of information is often prevalent, it is common that buyers and sellers *bargain* for some time over prices before an agreement is reached. It is also common that a buyer and a seller enter in an *exclusive* bargaining relationship in which both understand that they will talk to one another for a fixed period of time. For instance, in the housing market, a

*This chapter is joint work with Olivier Bochet.

[1] Adverse selection is severe if the buyers' expected valuation for the good falls short of the high cost of production.

potential buyer is often locked-in for several days after making an offer for a house. While he is allowed to make several successive offers for the same house during this time window, he is constrained by law not to make offers on another house. Other important examples where bargaining between uninformed buyers and informed sellers is witnessed is for hiring decisions (the worker may have superior knowledge about his level of productivity), the sale of an oil tract (the buyer may possess information about the richness of the deposit that is relevant to the owner's willingness to sell) or bargaining over the price of a software product (the buyer's knowledge about the expenses needed for the development of a new software may be limited).

Our Contribution: This paper is concerned with (i) the experimental test of a bargaining institution and its effect on trade and efficiency, and (ii) its comparison with a benchmark case in which the buyer is forced to make a single offer. Our choice of institutions is rooted in the theoretical literature. Consider first the benchmark case where the buyer commits to make a unique offer and walks away in the absence of a deal. Samuelson (1984) shows that a take-it-or-leave-it-offer is optimal from the buyer's point of view. Hence, any other case where the buyer talks more than once to the seller is detrimental to the buyer's welfare. A downside of the buyer's full commitment is the status-quo on trade failures and market breakdown. At the other end of the spectrum consider the case of a possibly infinite number of interactions between a buyer and a seller, in which the buyer makes an offer and the seller accepts or rejects. In a series of papers Vincent (1989), Evans (1989), Deneckere and Liang (2006) (henceforth DL) show the striking effect of the lack of commitment of the buyer coupled with frictions (discounting). When adverse selection is severe, trade occurs with probability one with any type of seller, and at different prices which signal qualities. Frictions drive screening and the buyer uses a monotonic price sequence to screen out low and high type sellers, while updating his belief towards the high type following each rejection along the sequence. Frictions are also a source of efficiency loss because of the delay before reaching an agreement.

We extend DL's model to the case where the number of offers is finite and provide an experimental test of this extension. We show that if the number of periods is big enough, there exists a unique sequential equilibrium in a fashion similar to the one obtained when the game is infinite.[2] When the number of periods is too low, the unique sequential equilibrium is to offer a price equal to the low quality seller's cost, having low types randomizing between acceptance and rejection until the last period of the game. We are interested in the former case where the number of periods is big enough.

In our experiment, sellers each can produce a good at different cost, high (*high quality good*) or low (*low quality good*), and this is private information to each seller. Buyers only know the probability distribution over sellers' types. Our experimental design compares two different institutions. In one set of treatments, the buyer makes repeated offers (R80 and R40). In a second set of treatments the buyer's optimal mechanism is implemented: the buyer makes a take-it-or-leave-it offer (S80 and S40). We vary the probability that the seller has a high cost of production within each set of treatments (0.4 and 0.8 respectively). The low probability is a case of market breakdown, while the high probability is a case where adverse selection does not preclude first-best efficiency.

We find that the bargaining situation (R-treatments) leads to screening of low and high type sellers, much like the qualitative predictions of the model. Rates of trade with both types of sellers are significantly boosted upwards, in particular trade failures that are common in the take-it-or-leave-it offer situations are almost eliminated with bargaining. However, buyers attempt to screen even when it would be optimal not to do so. When the production cost is high with probability 0.8, the equilibrium is to offer a single price equal to the high cost of production, and for any seller to accept this offer right away. Most importantly, we observe a significant over-delay compared to the theoretical predictions, i.e. trading pairs need longer than predicted to reach an agreement if the seller owns a high quality good. Over-delay is persistent with experience: we observe no learning effect. Delay mitigates efficiency substantially. While welfare is overall lower than predicted in both set of treatments, we find that bargaining leads to significantly lower welfare levels than in the benchmark single-offer treatments.

[2]An alternative would be to follow the recent literature on experimental repeated games (see Dal Bó (2005) and Dal Bó and Fréchette (2011)) and use a random continuation rule. We feel that using random termination rules may not be appropriate when the game involves beliefs updating at each period. Moreover, this would prevent us from observing a sufficient number of complete price sequences.

Is it better to trade later than never? If allocative efficiency is an important criterion (e.g. keeping a market "liquid" such as the housing market), then the bargaining treatments are successful in alleviating the adverse selection effect and facilitating trade. However, if total welfare is the main criterion for evaluating an institution's performance, then the observed persistent over-delay offsets the positive effects just mentioned.

What are the roots for the over-delay and its persistence in the data? First, buyers tend to start low in their price offer sequences and follow flatter price sequences than predicted. It takes them more time to reach an agreement. We show that this can be explained by loss aversion. In conjunction with loss aversion, there is an extra-delay imposed by high type sellers. At a sequential equilibrium, the buyer rips all the gains from trade with the high type seller. In practice, sellers reject offers and haggle over acceptable prices, and this even when discounting has already diluted the gains from trade. These two behavioral assumptions missing from the standard model help to reconcile data and sequential equilibrium predictions.

Related Literature: The experimental studies closest to ours are Rapoport et al. (1995) and Reynolds (2000). Both studies report on a bargaining game with the uninformed party being the proposer. Both papers analyze the case of independent valuations and discuss the Coase Conjecture, i.e., whether a declining price sequence can be observed.[3] While it is natural for us to also look at price sequences, our focus is different. We analyze a setting in which adverse selection prevails, i.e. valuations are interdependent. With interdependent values trade with high quality sellers implies an adjustment of the uninformed agent's belief. The fact that we find evidence for screening and belief updating is rather surprising in the light of a literature that states that subjects can have difficulties in inferring new information from others' actions. Eyster and Rabin (2005) refer to this slow down in information revelation as cursed equilibrium. Moreover, with independent values the uninformed party never runs the risk to make losses and thus trade failures are not a concern. In particular, trade with high quality sellers is profitable even if the buyer does not update his beliefs in the course of bargaining. In our setting, rates of trade are

[3]In their setting, the seller is the uninformed party.

an interesting object to look at. Finally, our design allows to compare the repeated offers bargaining institution to the benchmark case of a take-it-or-leave-it offer.

There is also a less recent related literature that tests the predictions of bargaining institutions[4] or of sequential equilibrium, something which our experiment also does. Roth and Malouf (1979) show that with complete information, bargaining tends to lead to equal splits of the gains from trade. By now it is also well established that bargaining power due to the bargaining protocol, as for instance in the ultimatum game, may have little impact on outcomes under complete information (see Güth and Tietz (1990) for the ultimatum game and Ochs and Roth (1989) for sequential bargaining). Roth and Murnighan (1982) and Roth and Schoumaker (1983) show that bargaining outcomes are driven away from equal division if either there is asymmetric information about valuations or bargainers have formed specific expectations about bargaining outcomes, for instance through a process of reputation building (see Embrey et al. (forthcoming)). We indeed find that subjects use their information strategically and the bargaining power of buyers is often undermined by the asymmetry of information. We also find that sequential equilibrium predicts behavior qualitatively well, and this already in the first periods.[5] In this respect, our results are in line with Embrey et al. (forthcoming) who look at reputation building in bargaining and find that subjects are strategic in the way predicted by sequential equilibrium.

The next section describes the model, provides a recap on standard adverse selection results and defines the finite game version of the bargaining model. It also characterizes the unique sequential equilibrium. In Section 3 the experimental design and the example used in the experiment and the corresponding theoretical predictions are presented. Section 4 discusses the results. Finally, Section 5 concludes.

[4]See Roth (1995) for a review of this literature.

[5]Camerer and Weigelt (1988) provide an early test of sequential equilibrium in the context of the trust game. Behavior corresponds roughly to sequential equilibrium, but only after subjects have played many repetitions of the game.

1.2 Preliminaries

1.2.1 The Model

In describing the model, we closely follow the notation used in DL. A buyer and a seller bargain over the price at which a single, indivisible good is sold. The seller's type (which determines the quality of the good) is determined by the random variable q, where q is distributed uniformly on $[0,1]$. The functions $v(q)$ and $c(q)$ represent the valuation for the object of the buyer and the cost of the seller to provide the good, respectively. It is required that $v(q) > c(q)$ for all q. Hence, it is common knowledge that there are gains from trade. The buyer's valuation and the seller's cost depend on q as follows.

$$v(q) = \begin{cases} \underline{v} \text{ if } q \in [0, \hat{q}] \\ \overline{v} \text{ if } q \in (\hat{q}, 1] \end{cases} \qquad c(q) = \begin{cases} 0 \text{ if } q \in [0, \hat{q}] \\ \overline{c} \text{ if } q \in (\hat{q}, 1] \end{cases}$$

Thus, there is a population of sellers distributed uniformly on $[0,1]$. All seller types $q > \hat{q}$ are high quality sellers (in the following we will refer to high quality sellers as H-types). All seller types $q \leq \hat{q}$ are low quality (L-type) sellers. The seller's type is private information to the seller. The buyer only knows that the seller he faces is drawn randomly from the distribution of q and is therefore uncertain about both his own valuation and the seller's cost of providing the good. Without loss of generality, $\overline{v} \geq \underline{v}$ and $\hat{q} \in (0,1)$. The assumption that there are gains from trade for all types further implies $\overline{v} > \overline{c}$ and $\underline{v} > 0$.

Despite the known gains from trade with both (payoff) types of sellers, the cutoff $\hat{q}$ drives the incentive constraints. Indeed, these may or may not preclude first-best efficient trade, as seen in the following two examples where we emphasize the equilibrium prediction of a single-price offer made by the buyer.

Example 1: *Only lemons!*
Let $c(q) = 0$ *and* $v(q) = 1750$ *if* $q \in [0, 0.6]$, *while* $c(q) = 2500$ *and* $v(q) = 3500$ *if* $q \in (0.6, 1]$. *The buyer ex-ante average valuation of* $(0.6 * 1750) + (0.4 * 3500) = 2450$ *falls short of the high cost. This precludes first-best efficient trade. If the buyer makes a*

take-it-or-leave-it offer then, at equilibrium, he offers $p = 0$, and this is accepted only by an L-quality seller.

Example 2: *Goods change hands*
Let $c(q) = 0$ and $v(q) = 1750$ if $q \in [0, 0.2]$, while $c(q) = 2500$ and $v(q) = 3500$ if $q \in (0.2, 1]$. The buyer ex-ante average valuation is $(0.2 * 1750) + (0.8 * 3500) = 3150$. This exceeds the high cost, a necessary condition for goods to change hands. Also, the buyer expected payoff from offering 2500 exceeds the one from offering 0. If the buyer makes a take-it-or-leave-it offer then, at equilibrium, trade occurs with probability one for both type of sellers at $p = 2500$. The outcome is first-best efficient.

Given the above parameters constellation, high quality goods change hands only if $\hat{q} \leq \frac{2}{7}$. Otherwise, the buyer single-price offer mechanism has a unique equilibrium in which $p = 0$.

1.2.2 Repeated Offers Game and Equilibrium Predictions

In contrast to Vincent (1989), Evans (1989) and DL, we allow the maximal number of offers to be a finite number N.[6] Let $n = N-1, N-2, \ldots, 0$ be the number of stages left before the final equilibrium stage is reached. The state variable q_n denotes the buyer's cutoff level of the seller population n stages before the final equilibrium stage. That is, n stages before the final equilibrium offer, the buyer believes that the seller's type is uniformly distributed on $[q_n, 1]$.[7] Note that $q_{N-1} = 0$. It follows that the mass of the H-quality sellers is $1 - \hat{q}$. The mass of the L-quality sellers is $\hat{q} - q_n$ when n stages are left before the final stage.

The buyer's offer is denoted by $p(q) = p_n$ for $q \in (q_n, q_{n-1}]$. The game ends if the seller accepts an offer or rejects all offers including the one in stage N. After a rejection in any other stage, the next stage is entered. The buyer updates his belief and makes a new offer.

[6] Evans (1989) also analyzes the 1 and 2 stage case, but does not provide a solution for the general finite horizon case. Also, his model differs from ours in that both trading parties have a valuation of zero for the L-quality good.

[7] The buyer's belief will always be a left truncation of the prior, i.e., q_n is non-increasing in n (see DL).

Payoffs are discounted after each stage. Let $\delta \in [0,1)$ denote the discount rate. If trade takes place n stages before the last stage N, the payoffs are $B_n(q) = \delta^{N-1-n}(v(q) - p(q))$ for the buyer and $S_n(q) = \delta^{N-1-n}(p(q) - c(q))$ for the seller. If no agreement takes place, both parties earn a payoff of 0.

We now come to the equilibrium predictions of the repeated offers game. All proofs of the results mentioned here are relegated to Appendix A. In general, the buyer has two options. The first option is to successively increase his offers to screen out the L-quality sellers. In this case, he faces the trade-off between screening less finely and delaying agreement. The second possibility is that the buyer offers 0 in all stages, focusing on the gains from trade with an L-quality seller.[8] To distinguish between these two patterns, variables belonging to the screening or the zero offer sequence are superscripted by s and z, respectively. The following lemma states that the equilibrium offers have to follow one of these two patterns. Let $k^*(q_n)$ denote the optimal number of screening stages given belief q_n.[9]

Lemma 1. *In the bargaining game with $N > 0$ stages, the final equilibrium offer is either $p_0^s = \overline{c}$ or $p_0^z = 0$.*

i) If $p_0^s = \overline{c}$, the sequence of equilibrium offers is given by $p_k^s = \delta^k \overline{c}$ for $k = k^*(0), k^*(0) - 1, \ldots, 0$.

ii) If $p_0^z = 0$, the sequence of equilibrium offers is given by $p_n^z = 0$ for $n = N-1, N-2, \ldots, 0$.

The intuition behind Lemma 1 is that L-quality sellers must be kept indifferent at equilibrium between accepting the current offer in stage, and waiting for a future offer. If the L-quality sellers rejected for sure, the buyer would delay the agreement without gaining additional information. On the other hand, certain acceptance by L-quality sellers means that rejection reveals the seller to be an H-type, implying an offer of $\overline{c}$ in the next stage.

[8]Note that if the potential number of offers is infinite, this cannot be an equilibrium pattern. Since the buyer's belief increases with each rejection, he is eventually willing to trade with the H-quality seller.

[9]If screening occurs in equilibrium, it will always start in stage 1 when $q \in [0,1]$. The reason we introduce this notation nonetheless becomes clear when discussing the zero offer sequence equilibrium.

In this case, the L-quality seller has an incentive to mimic the H-quality seller unless the offer in the current stage is $\delta\overline{c}$.

If case i) in Lemma 1 prevails, we refer to Appendix A for a derivation of the L-type seller's equilibrium behavior. Intuitively, sequential rationality requires the buyer's offers to make the L-quality sellers indifferent between a non-equilibrium offer and its subsequent offer. Further, note that a buyer's optimal offer sequence is described by Lemma 1 also after an non-equilibrium offer. These two requirements can be fulfilled jointly only if the buyer is indifferent between two different prices that belong to a sequence as described in Lemma 1: he can then mix between the two prices such that the L-type seller's expected profit in the next stage corresponds to the one he would obtain from accepting the off-equilibrium offer. This uniquely pins down the sellers' acceptance decisions.

In contrast to the infinite horizon case, the constant price sequence $(0, ..., 0)$ is a possible equilibrium if the time span given for screening is too short. However, when N is large enough, the buyer's expected profit from the zero offer sequence approaches 0 or a condition is violated such that the equilibrium is then given by the screening equilibrium.[10]

We next provide the intuition for this result. Knowing the price sequence and the acceptance decisions of the screening equilibrium allows to derive the acceptance probabilities for the zero offer sequence. The important idea here is that the sellers' acceptance decisions must render the buyer indifferent between offering the optimal screening price (given the current belief) and offering 0. Obviously, if offering a price that belongs to the screening sequence leads to a higher expected profit, the buyer would switch to the optimal screening strategy. On the other hand, if the zero price offer is the unique best offer then sequential rationality off the equilibrium path is violated. To see this, suppose that the unique best offer is zero and consider a non-equilibrium offer just slightly above 0. The L-type seller has to be indifferent between accepting and rejecting, and hence the buyer has to randomize between an offer of 0 and the optimal screening offer in the next stage. This implies a strictly larger probability of acceptance for the offer slightly above

[10]It is noteworthy that there can also be screening equilibria that are not identical to the one found in the literature for the infinite horizon game. This instance occurs if N restricts the optimal number of screening stages, but the buyer still prefers to screen rather than to follow the zero offer sequence.

0 than for an offer of 0.[11] But then this is a profitable deviation for the buyer.

The zero offer equilibrium requires that in each stage a positive fraction of the L-quality sellers accepts an offer of 0. Hence, if N is large, the delay associated with the zero offers sequence is then too large to render the buyer indifferent to the optimal screening sequence. In the appendix, we prove the following proposition. We also provide formulas to derive $\bar{N}$ as well as the equilibrium behavior of both parties.

Proposition 1. *There exists a finite $\overline{N}$ such that the unique equilibrium is the screening equilibrium for all $N \geq \overline{N}$ and the zero offer equilibrium otherwise.*

In light of the sequential equilibrium predictions, we now revisit the two examples introduced in the previous section.

Example 1 revisited: *Only lemons?*
Consider the parameters constellation of Example 1. Let the discount rate be $\delta = 0.8$ and the number of possible price offers be $N = 50$. Then the unique sequential equilibrium is the screening equilibrium with associated price sequence $p^s = (1280, 1600, 2000, 2500)$. An L-quality seller randomizes over acceptance and rejection up to $p = 2000$. At such a price, a rejection is interpreted as the seller being an H-type. Hence, at $p = 2000$, the L-type accepts for sure. The L-type's acceptance probabilities support the price sequence on the equilibrium path. Notice that $\delta^3 * 2500 = \delta^2 * 2000 = \delta * 1600 = 1280$.

Example 2 revisited: *Goods change hands*
Consider the parameters constellation of Example 2. Let the discount rate be $\delta = 0.8$ and the number of possible price offers be $N = 50$. Interestingly, the prediction coincide with the take-it-or-leave-it offer. The prior to be with an H-type seller is too high (0.8) so that the incentive to screen is too small. Indeed, the buyer trades off the cushioning of losses obtained on low types with the delay before an agreement is reached. Consider the candidate screening price sequence $p^s = (2000, 2500)$. If this is an equilibrium price sequence, a low-type seller accepts the first offer of 2000 with probability one, so that a rejection signals that the seller is a high quality one. However, cushioning the loss on

[11] A weakly smaller probability of acceptance would lead to a unique best offer of 0.

Table 1.1: Experimental Design

Treatment	Sessions	Subjects	No. Offers	Probability H-type
R80	6	70	50	0.8
R40	6	70	50	0.8
S80	4	48	1	0.4
S40	4	48	1	0.4

a low quality seller makes the buyer to trade with the high quality seller with one period delay. This is dominated by the equilibrium price $p = 2500$ *in which trade occurs right away with both type of sellers. There is thus no delay before an agreement is reached.*

If an uninformed buyer can make repeated offers, he may extract information about the quality of a good by following a specific price sequence. Through this mechanism the buyer is able to reach an agreement with an H-type seller whereas this is not be possible in a single offer setting whenever adverse selection is severe, like in Example 1. The obvious downside of making repeated offers is that delay is costly. This tradeoff determines the equilibrium number of screening stages. The lesson from Example 2 is that the *cushioning of losses* obtained from low type sellers through screening may not be optimal when the probability of an H-type is high enough. In contrast, we will show in the experimental part that the cushioning of losses is an important driver of the buyers' behavior.

1.3 Experimental Design

We now describe our experimental design. The experiment took place in the fall of 2012, and spring of 2013 at the experimental laboratory of the University of Bern. 236 students (both undergraduate and master's) from business and economics took part in the experiment. A session is in general composed of 12 participants, exception made of two session that had 10 participants. 20 sessions were run.[12] A session last approximately 70 minutes and average earnings were 32 CHF (conversion rate 0.004, including a show-up fee of 10 CHF). We run four different treatments.

[12]Sessions were run using the z-Tree software developed by Fischbacher (2007).

We used as fixed set of parameters the ones introduced in the previous examples. Namely, the buyer's valuation is given by $\overline{v} = 3500$ and $\underline{v} = 1750$. The seller's cost is $\overline{c} = 2500$ and $\underline{c} = 0$. The discount rate is given by $\delta = 0.8$.[13] Our design varies two parameters: the length of the bargaining game and the probability that the seller is an H-type seller (or, respectively $\hat{q}$). The treatments are summarized in Table 1.1. Treatments R80 and R40 allow for a maximum of 50 stages with prior probability that the seller produces an H-quality good to be 0.8 and 0.4, respectively. In the benchmark cases S80 and S40, buyers make a take-it-or-leave-it offer. The two treatments differ in probabilities in the same way as R80 and R40.

The instructions for treatment R80 are provided in Appendix B. After reading the instructions every subject had to fill out a set of control questions. Subjects were then randomly assigned to be one of the 6 buyers or one of the 6 sellers. Roles are fixed throughout the experiment. In each session, there is exclusive bargaining between a buyer and a seller. Each pair (composed of a buyer and a seller) plays a bargaining game whose rules depend on the treatment –either a repeated offers or a single-price offer game. Subjects play ten bargaining games in total. There is random re-matching after each bargaining game. Hence reputation plays little to no role due to the mitigation effect of the random matching procedure. A seller can be either an L or H-type. Sellers' types can change from one bargaining game to the next. Before each bargaining game, sellers' types are randomly determined according to the fixed probability $\hat{q}$. Each seller is informed of his own type. Buyers are not.

We give in Table 2 a summary of the predictions of the model as well as the (ex-ante) welfare level generated by each such prediction.[14] Notice that in the second row of the table, the acceptance probabilities of the L-type seller should be understood as the ex-

[13] The buyer is required to make price offers in increments of 0.1. Restricting the set of possible price offers to specific increments does not change the equilibrium as long as all price offers that are used by the buyer in equilibrium are still available.

[14] Notice that neither the single-price offer nor the bargaining institution is the mechanism which maximizes total welfare. Indeed consider a case in which the buyer offers $(\theta(p_L) = 1, p_L = 1750)$ and $(\theta(p_H) = 0.7, p_H = 2500)$, i.e. the seller chooses between transferring the good for sure and receiving 1000, or transferring the good with probability 0.7 and receiving 2500 in case the good is transferred. It can be checked that total welfare generated by this mechanism is 1330, as opposed to 1050 for S40 and 1105 for R40.

Table 1.2: Theoretical Predictions

	Price Offers	Acceptance Probabilities L	Acceptance Probabilities H	Ex-Ante Welfare
R80	2500	1	1	1150
R40	$(1280, 1600, 2000, 2500)$	$(0.5, 0.23, 0.27, 0)$	$(0, 0, 0, 1)$	1105
S80	2500	1	1	1150
S40	0	1	0	1050

ante randomization over accepting or delaying acceptance. The corresponding acceptance probabilities in each stage are then $(0.5, 0.46, 1, 0)$ –an L-type seller that is still around in stage 3 accepts $p_1 = 2000$ with probability 1. An H-type seller rejects all prices but the last one. In the single offer treatments S80 and S40, the buyer offers $p_0 = \overline{c}$ if $(1-\hat{q})\overline{v} \geq \overline{c}$ and $p_0 = 0$ otherwise. It follows that $p_0 = \overline{c}$ and both seller types accept the offer in S80. For S40 it holds that $p_0 = 0$, which is accepted only by the L-type seller. Hence, while theory predicts no trade failures in S80, S40 is an example of a situation where asymmetric information leads to unrealized gains from trade between the buyer and the H-type seller. Its counterpart repeated offers treatment R40 allows for trade with both types of sellers and yields a higher ex-ante welfare level.

The benchmark case given by S80 and S40 is important to understand the performance and limitations of the repeated offers bargaining protocol. A comparison of the outcomes between R40 and S40 allows to test whether repeated offers in conjunction with discounting indeed increases the probability to reach an agreement. Comparing R80 to S80 provides evidence on how repeated offers change behavior if adverse selection is no issue, i.e. if trade failures should be absent even in the single offer setting. The predictions on how a change in $\hat{q}$ affects behavior can be tested by comparing R80 and R40 for the repeated offer setting; S80 and S40 for the single offer setting.

In a subset of the R-treatments, subjects were presented a lottery task that allowed us to measure loss aversion. The lottery task is the same as in Fehr et al. (2013). At the beginning of the experiment, subjects were informed that the experiment would be composed of two parts but did not know what the second part would be during the first part. Subjects played first the bargaining games (Part 1), and then the lottery task (Part 2). We are interested in loss aversion, because buyers who offer a price acceptable to H-type sellers run the risk of making a loss in case they happened to be matched with an L-type seller. The lottery task will be described in Section 4.3.

1.4 Results

1.4.1 Statistics on Prices

Table 1.3: Opening Prices

		Mean	Median	SE	Min	Max
R80	H	1357	1000	809	0	3000
	L	1271	1000	833	0	1800
R40	H	818	875	303	0	1500
	L	749	850	339	0	2400
S80	H	2409	2600	578	0	3500
	L	2333	2550	653	1000	3000
S40	H	723	500	799	1	3000
	L	757	500	789	1	3300

Table 1.4: Trading Prices

		Mean	Median	SE	Min	Max
R80	H	2799	2900	390	500	3500
	L	2248	2600	745	600	3050
R40	H	2656	2750	529	650	4000
	L	1197	1000	655	100	3200
S80	H	2659	2600	205	2505	3500
	L	2385	2550	559	500	3000
S40	H	2672	2600	182	2500	3000
	L	885	750	812	1	3300

Some First Impressions: Looking at Tables 1.3 and 1.4 side-by-side is instructive as one can make inference on several possible scenarios. First notice that in R80 and R40, there is a difference between opening and accepted prices –accepted prices are between 1.6 and more than 3 times bigger than opening prices. These differences are significant for both seller types –Wilcoxon signed-rank test, all p-values between 0.03 and 0.046.[15] What information do these differences convey? In R80, recall that bargaining should stop in stage 1. Differences between median first offers (resp. mean) and median prices (resp. mean) signals delay before agreements were reached. We also see that there were most probably attempts at screening in R80, since median (and mean) prices accepted

[15]All non-parametric tests reported in this paper use session averages as the unit of observation.

differ between L and H-types –Wilcoxon signed-rank test, p-values both at 0.046. In R40, average and median price sequence start low compared to the theoretical prediction (875 and 850 vs 1280); a first indication that there could be extra-delay compared to the theoretically predicted sequence. There is an obvious attempt at screening given the sharp difference in median accepted prices (resp. mean) between both types of sellers –Wilcoxon signed-rank test, p-values both at 0.03. Standard errors on opening prices are significantly higher in R80 than in R40. In R40, buyers tend to start with an offer that splits equally the gains from trade that would be obtained with an L-type seller. In contrast in R80, there is less consensus on what the "right" first offer is. Like in R40, many subjects first offer around an equal split of the gains from trade (57% between 800 and 1000), while 18% of subjects right away announce offers acceptable by H-type sellers.

For the S-treatments, in S80 median offered and accepted prices are the same. This hints at possibly high rates of trades. Also notice that in S80, the difference between median accepted prices (resp. mean) between L-type and H-type sellers is not significant, in line with the theoretical predictions. On the other hand, in S40, the difference between median offered and accepted prices (resp. mean) for H-type sellers is large. This indicates trade failures. Surprisingly, there is also a difference for the L-type seller, from an opening median price of 500 to a median accepted price of 750. Hence, there are trade failures also with L-type sellers.

We can also make some first comparisons between the repeated and the single price offer treatments. First, in R80 the median accepted price (resp. mean) for H-type sellers is at 2900 (resp. 2799) while in S80 it is 2750 (resp. 2656). These differences are significant at the 1% level (according to a Mann-Whitney U test (MW)), indicating that H-type sellers probably use the possibility offered by R80 to delay agreement in order to trade at a higher price. There are no such differences for L-type sellers. Indeed, L-type sellers get a high informational rent in both S80 and R80. On the other hand, both types of sellers accept different prices in R40 and S40 (p-values all between 0.01 and 0.02 for both medians and means, for both types of sellers). Because buyers attempt at screening in R40, L-type sellers get a higher rent than in S40 where the equilibrium price should be 0.

Likewise, because of the possibility to delay agreement, H-type sellers are likely to reject acceptable prices in the hope to get a better subsequent offer.

Result 1. *Price Wedges*

(i) The possibility of repeated offers draws a price wedge between opening and accepted prices.

(ii) Trade with H-type sellers occurs at higher prices than trade with L-type sellers.

(iii) Accepted prices are higher in the R-treatments as compared to the S-treatments.

Conformity to Theory: In S80, accepted prices seem to be in line with the theoretical predictions (2600 with H-type sellers, and 2550 with L-types). The same is not true for accepted prices in S40 which are much higher than predicted (median accepted price of 750 with L-type sellers), and even some H-type sellers traded, contrary to the market failure prediction. For R80 and R40, these statistics are not sufficient to fully evaluate departures from the theoretical predictions. We look now at the price sequences for both treatments, restricting our attention to trades with H-type sellers. Our discussion will be in support of Result 2.

Result 2. *Conformity and Deviations*

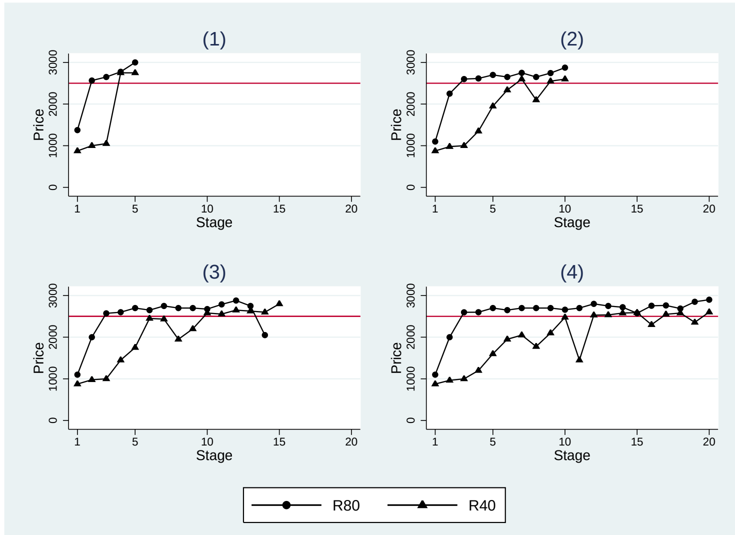
Buyers follow increasing price sequences in R80 and R40. This is in accordance with the theoretical prediction in R40. In R80, the inability of the buyer to commit not to make repeated offers drives observed behavior away from predictions.

Figure 1-1 displays four graphs of observed price sequences. Quadrants show price sequences for pairs that traded within five, ten, fifteen and twenty stages, respectively. Since we are interested in complete price sequences, only observations with an H-quality seller are used.[16]

First, we notice a strong price increase for R80. Starting from median offers clustered between 1100 and 1375, buyers who traded with an H-type seller double their offers in

[16] Higher prices are more likely to be accepted earlier by L-quality sellers. By including observations with an L-quality seller, cases involving an increasing price sequence would be underrepresented in later stages. A Kolmogorov-Smirnov test (p-value 0.845) does not reject equality of distributions of first stage offers between H and L-quality sellers, but generally rejects equality of offer distributions for stages later than stage 7. See Appendix E for a graphical representation.

Figure 1-1: Price Sequences

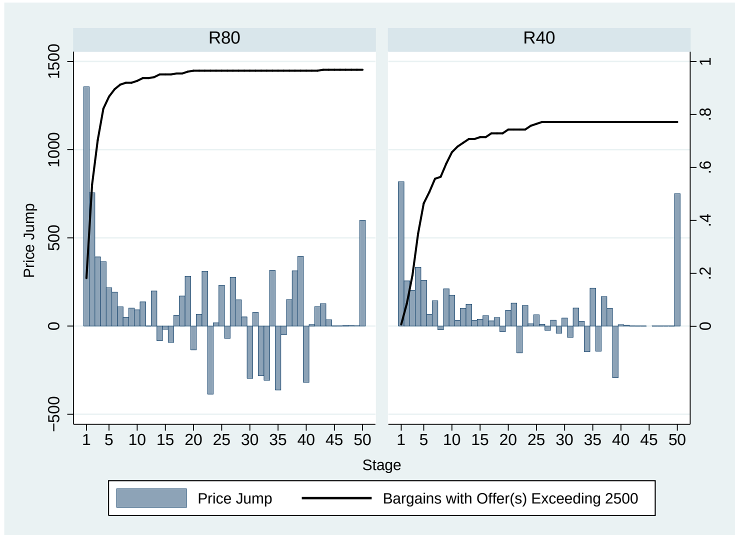


Notes: (1) Subfigures 1-4 show the median price sequence for R80 and R40 for H cases when trade was achieved within 5 stages (1), 10 stages (2), 15 stages (3), or 20 stages (4). (2) The price sequences are calculated by first taking for each buyer individually the median and then the median over all buyers. (3) The solid horizontal line corresponds to cost $\bar{c} = 2500$ of H-types.

stage 2. For pairs that traded within five stages, a median price offer higher than $\bar{c}$ is made in stage 2, while for the three other quadrants the cost of the H-type seller is always covered in stage 3, at the median offer. Except for the first quadrant, price offers stabilize around 2600. The mean price jumps are generally positive and decreasing until stage 20, as shown in Figure 1-2. Beyond stage 20, mean price jumps oscillate between positive and negative jumps –possibly because payoffs are then close to 0 due to discounting. The above observations are in stark contrast with the prediction that trade should occur immediately at a price of $\bar{c}$: the inability of buyers to commit not to make repeated offers drive observed behavior towards increasing price sequences. Figure 1-2 also shows the fraction of cases for which the buyer has offered at least one price equal to or above 2500. While only 18% of all cases start with an H-acceptable offer, this fraction increases to 85% by stage 4, and it is above 95% by stage 14.

In comparison, behavior in R40 is sluggish. Figure 1-1 shows that the median price offers are at 875 in stage 1. For pairs that traded within five stages, the first three offers are between 875 and 1050 with a sudden jump to an offer exceeding $\bar{c}$. For the three other

Figure 1-2: Price Jumps



Notes: (1) The bars indicate the mean observed price jumps. The first bar represents the mean offer in stage 1. Bars 2 to 50 represent the mean price jumps from one stage to the next. (2) The solid line shows for R80 and R40 the fraction of H cases that involve price sequences with at least one offer above 2500 in a given stage.

quadrants (10,15 and 20 stages), an offer acceptable by the H-type seller is not reached before stage 7, 10 and 12 respectively. In the same fashion, Figure 1-2 shows that in R40 the fraction of offers acceptable by H-type sellers increases but it takes ten stages to reach 60% of cases with an offer above 2500. In 23% of all cases, buyers never offer a price equal to or above 2500 and hence these cases must involve trade failure.

1.4.2 Bargaining, Adverse Selection, Trade and Efficiency

We discuss in this part our central findings regarding the performance of both bargaining protocols used in our experiment.

Trade Dominance of the Repeated Offer Treatments: Treatments S80 and S40 exhibit two different conclusions in the presence of adverse selection. In the former, the probability that the seller is an H-type is so high that the buyer is willing to take a risk and offer $\bar{c}$. In such a case, adverse selection does not cause trade failures and the outcome is predicted to be first-best efficient– goods change hands so that the full gains from trade

Table 1.5: Trade Frequency

Treatment	Type	Cases	Trade Frequency						
			≤ 50	≤ 20	≤ 10	≤ 5	≤ 3	≤ 2	≤ 1
R80	H	286	0.98	0.94	0.86	0.67	0.46	0.31	0.12
R80	L	64	0.98	0.98	0.95	0.89	0.70	0.58	0.30
R40	H	140	0.81	0.73	0.55	0.31	0.11	0.06	0.00
R40	L	210	0.99	0.93	0.84	0.75	0.66	0.53	0.35
S80	H	202							0.63
S80	L	38							0.92
S40	H	114							0.11
S40	L	116							0.82

Notes: Trades frequencies for R80 and R40 are provided separately for trades occurring within the first 50 stages (≤ 50), 20 stages (≤ 20), 10 stages (≤ 10), 5 stages (≤ 5), 3 stages (≤ 3), 2 stages (≤ 2), and 1 stage (≤ 1).

are realized. On the other hand, S40 shows a case of market failure. Buyers should always offer a price of 0 and H-type sellers never trade. In contrast, R80 and R40 predict trade with probability 1 for both type of sellers. We look at the rates of trade in light of these predictions.

Result 3. *Rates of Trade*

The possibility to make repeated offers has a strong impact. Rates of trades are boosted upward in R80-R40 compared to their respective single-price treatments S80-S40. In addition, the more likely the seller is an H-type, the higher the rates of trades.

Table 1.5 lists the observed trade frequencies for all treatments separated by seller quality, and distinguishing between different timelines over which trade occurred. For instance, the first column ≤ 50 shows rates of trade treating all trades as successful; while the second column ≤ 20 counts trades as successful only if they occur within 20 stages etc.[17]

In S80, trade should occur with probability 1 for both type of sellers. Trade with H-type sellers occur in 63% of cases and is statistically different from a rate of trade of 1 (One-Sample median test, p-value= 0.07), while trade with L-types is 92% and fits

[17] Obviously, Table 1.5 shows only one column for S80 and S40.

with the theoretical prediction. In contrast, in S40, rates of trade differ from predictions (p-values= 0.07 for both types). Some buyers make acceptable offers to both types and 11% of H-type sellers trade (13 out of 114 observation with H-type sellers). Rate of trade with L-types is then 82%. Trade failures in S40 are thus quite high. In the counterpart R-treatments, looking at column ≤ 50 shows that rates of trades are as predicted except in R40 with H-types (p-value= 0.03). Recall however our earlier comments when discussing Figure 1-2: beyond stage 20, mean price jumps become volatile and our intuition for why this is happening is that gains from trade beyond that stage fall to almost 0. If we restrict to pairs trading within 20 stages, then rates of trade with H-type sellers in both R80 and R40 are different from 1 (p-value=0.05 for R80 and p-value=0.03 for R40).

Coming to the treatment comparisons, a quick look at Table 1.5 shows differences when going from the S to the R-treatments. While there are many cases of trade failures with H-types in S80, this issue is mostly avoided in R80 where rates are 98% and 94% for the H and L-type sellers when trade occurs within 20 stages. As expected, there is little statistical significance in the change of rates with L-types between R80 and S80 (Fisher exact test, p-value=0.14). However, the gap between R80 and S80 is large for the H-types (p-value$<$ 0.01). The comparison between R40 and S40 is even more conclusive as both trades with H and L-types differ (both p-values$<$ 0.01). A large fraction of the trade failures predicted in S40 are thus avoided with R40 –even though in R40 there is still a significant rate of trade failures.

The conclusions on rates of trade are also confirmed in Table 1.6. The first column shows a linear regression on whether trade occurred.[18] Compared to S40, trade is easier in all other treatments (p-values $<$ 0.01) –respectively easier in R80 than in R40, and easier in R40 than in S80. Therefore repeated offers greatly facilitate trade. Moreover, the probability of trading with an H-type seller is significantly lower than with an L-type seller.[19]

[18] Estimating a pooled probit model with standard errors clustered on the individual level yields similar results.

[19] Combining the coefficient of "H" with each interaction term, we can reject the null hypothesis that there is no significant difference between rates of trade with H and L-quality sellers at the 5% significance level for R80 and at the 1% level for the other treatments.

Table 1.6: Random Effects Regressions

Dep. Var.:	Trade (1)	Actp. Stage (2)	Welfare (3)
R80	0.18*** (0.04)		-165* (100)
R40	0.13*** (0.04)	1.08* (0.60)	-286*** (89)
S80	0.15*** (0.06)		218.4** (99)
H	-0.70***(0.05)	2.19*** (0.47)	-1324*** (66)
R80*H	0.66*** (0.06)		576*** (98)
R40*H	0.47*** (0.08)	2.28*** (0.84)	395*** (86)
S80*H	0.37*** (0.09)		296*** (112)
Constant	0.83*** (0.04)	3.01*** (0.72)	1521*** (66)
R^2 (overall)	0.36	0.14	0.49
Observations	1170	631	1170
Individuals	117	70	117
Reference Group	S40 / L	R80 / L	S40 / L

Notes: (1) Standard errors are clustered on individuals (in parentheses). (2) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (3) All estimations include period dummies. The estimated coefficients are insignificant for all models and all periods.

We conclude on Result 3 with the impact of an increase in the probability of an H-type seller within each set of treatments. For the S-treatments, such an increase has a positive impact for trade with both types (MW p-values respectively less than 0.01 and 0.05). For the R-treatments rates of trade with L-type sellers are already not different from 1 in both R80 and R40. Not surprisingly there is no statistical difference there (p-value=0.55). On the other hand, the gap in rates of trade with H-type sellers between R80 and R40 is important and this is statistically confirmed (MW, p-value$<$ 0.01).

We close with a quick glance at the remaining columns of Table 1.5. An interesting observation there is the time it takes for trade with H-type sellers to pick up in R40 compared to R80. Trade with H-type sellers in the former reach a mere 31% within five stages whereas it is at 67% in the latter; likewise for trade within ten stages (respectively 55% and 86%). Overall, in R40, 50% of trades with H-type sellers occur where potential welfare has already fallen to a third of the gains from trade; 30% in R80. Table 1.5 gives a first snapshot of the over-delay present in both R-treatments: rates of trade are high, but agreements seem difficult to reach.

Later or Never? A Late Blooming of Trade: As Table 1.5 already hinted, trade

Table 1.7: Trading Stage

	H					L				
	Median	Mean	SE	Min	Max	Median	Mean	SE	Min	Max
R80	4	5.7	6.6	1	47	2	3	2.6	1	15
R40	7	9.7	7.6	2	36	2	3	8.2	1	50

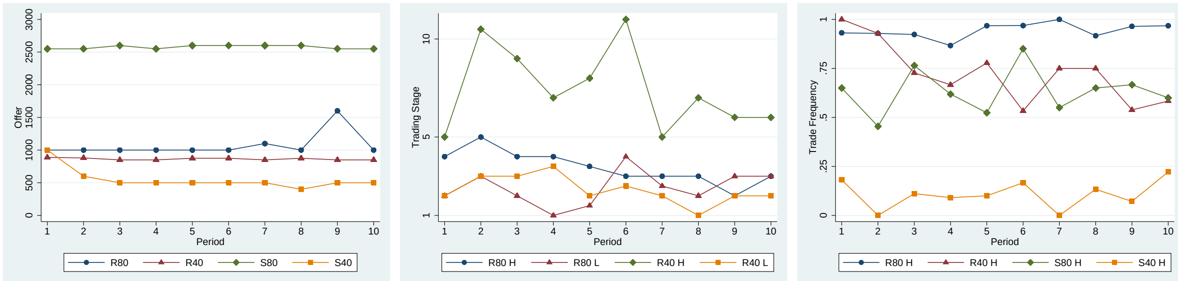
in the R-treatments takes time. Column 2 of Table 1.6 shows the acceptance stage as function of the treatments and interaction variables. The coefficient of R40 shows that trade with L-types occurs later in R40, but this is significant only at the 10% level. Striking a deal with an H-type seller takes significantly longer than with an L-type in both treatments. Coefficient R40*H confirms, however, that agreeing with an H-type seller takes longer in R40 than in R80. Finally combining estimates shows that, in R40, H-type sellers accept an offer on average 4.5 stages later than L-type sellers. We now go to the data more precisely and summarize our findings below.

Result 4. *A Long Delay*

Trade occurs significantly late in the R-treatments, in particular for H-type sellers. L-type sellers accept earlier than H-types. Also, the more likely the seller is an H-type, the earlier trade occurs. Most importantly, delay is a persistent phenomenon across bargaining games.

From Table 1.7, we get that the median acceptance stage of an H-type seller is 4 in R80 while it is 7 in R40 (MW, p-value<0.01). It takes more time to reach an agreement with an H-type when the probability that the seller is an H-type is rather low. Importantly, there is a significant over-delay in both treatments for trade with both seller types (Median test, p-values<0.01). Obviously, two sources can be at play to explain over-delay. On the one hand, buyers cause delay: they open price sequences with lower prices than expected, and they increase prices slower. On the other hand, an important channel explaining over-delay is that H-type sellers reject acceptable offers. If H-type sellers never rejected acceptable offers, the median acceptance stages are reduced in both treatments –from 4 to 2 in R80, and from 7 to 6 in R40.

Figure 1-3: Period Effects

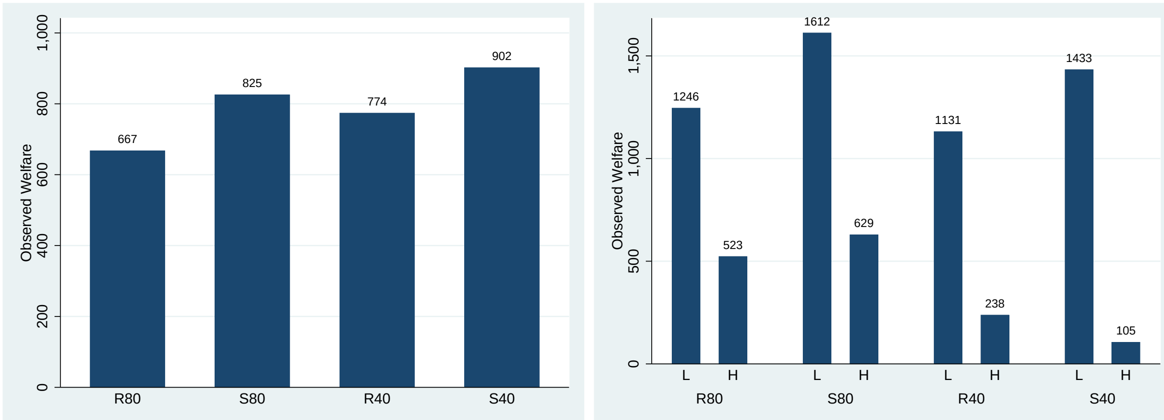


(a) Opening Offers (b) Trading Stage (c) Trade Frequency

Note that the value of the discount rate should push trade to occur sooner rather than later –recall that at a discount rate of $\delta = 0.8$, already more than half of the gains from trade with an H-type seller are gone after stage 4. In contrast, we saw that bargaining pairs seem to experience difficulties in reaching agreements in a small number of periods in both treatments. We are therefore confident that over-delay is a robust phenomenon. This is confirmed in Figure 1-3. There we see that opening offers are completely stable throughout the experiment. Trading stage is always in excess of the predictions, and hence over-delay is persistent. In addition, it is clear that delay does not decrease with experience. We see, however, that R40-H is not as quiet as the others in terms of trading stage fluctuations. Figure 1-3 (c), although not directly related to delay, highlights the relative stability of behaviors across bargaining game. In Section 4.3, we will show that over-delay can be explained by buyers' loss aversion and sellers' haggling.

Efficiency, Payoffs and Dominance of Single-Price Offers: So far we have documented the amplifying effect on rates of trade of the R-treatments over the S-treatments. However, Result 4 on delay already points at failures of the R-treatments. Indeed, possible difficulties linked to (i) the inability of buyers to commit, (ii) fear of making losses with L-type sellers, and (iii) the combination of rejection of acceptable offers by both L and H-type sellers push delay way beyond the sequential equilibrium predictions. But observing a systematic over-delay begs the question contained in our title: is it better to trade later than never? In terms of economic performance, is it better to have a significant increase in rates of trade, at the price of delay?

Figure 1-4: Observed Welfare



(a) Across types (b) By types

Result 5. *Efficiency Failures*

Efficiency falls short of the theoretical predictions for all treatments. Moreover, the R-treatments perform worse than their single-price offer counterparts.

The left panel of Figure 1-4 shows the observed average welfare levels over all bargaining pairs.[20] A first look at the figure reveals that average welfare in all treatments are lower than their respective predicted levels. Median tests confirm that the R-treatments perform worse than predicted (p-values< 0.05) although the significance is only at the 10% level for the S-treatments. Recall that, theoretically, we expect little to no difference in ex-ante welfare levels across treatments –see Table 2. Indeed, there is not enough differences across the R-treatments (MW, p-value=0.109), even though the welfare level is higher in R40 than in R80. However, the difference between S80 and S40 is confirmed at the 5% level, with higher gains from trade exploited in S40 (p-value$= 0.021$)

The main message coming out of the left panel of Figure 1-4 are the differences across set of treatments. The average welfare level trade is higher in the S-treatments than in their respective R-treatments, and this irrespective of the probability of occurrence of H-type sellers. S-treatments seem to perform better than R-treatments –for R80-S80,

[20]Note that the figure on the left panel is computed by using the observed welfare by types of the right panel, and weighted by the theoretical probabilities on L and H-types. For instance the bar for R80 gives a level of welfare of 667. This is obtained as $(0.2 * 1246) + (0.8 * 523)$, as shown on the right panel, R80-L and R80-H.

p-value= 0.055, for R40-S40, p-value= 0.011 (MW tests). Importantly, this shows that, given our parameters constellation, R-treatments fail to be second-best efficient in practice: they deliver on average lower ex-ante efficiency levels than their counterpart single-price treatments.[21] With respect to first-best efficiency, notice that the R-treatments account for only 58% in R80 and 54% in R40 of the first-best efficiency level.

Given that observed welfare falls short of the theoretical prediction, one could expect that subjects learn to adjust prices in order to reach more efficient outcomes. This is true in particular for the R-treatments where there is too much delay. However, recall that Figure 1-3 shows that opening as well as trading stages are rather stable across the 10 repetitions of the bargaining game.

Figure 1-4 also breaks down gains from trade between bargaining pairs with an H and with an L-type seller. This allows us a first attempt at disentangling where the failure of the R with respect to the S-treatments may come from. We notice first that welfare in pairs with an H-type seller does not significantly differ between S40 (105) and R40 (238)–MW, p-value=0.136. Since the rate of trade with an H-type in R40 is high (81%), it is clear that delay must account for a significant part of the loss compared to its theoretical prediction.[22] In contrast, welfare in pairs with an L-type is significantly lower in R40 than in S40 (p-value$<$ 0.05), and this is expected.[23]

Regarding R80 vs S80, realized welfare in pairs with an H-type seller differs at the 10% level (523 and 629, respectively). Here, the change in rate of trades when going to R80 is large, from 63% to 98%, yet the welfare is higher in S80. The difference between the two treatments is thus a pure consequence of over-delay, and welfare levels in pairs with an H-type are roughly at half of their theoretical prediction. Likewise, welfare levels with L-type sellers are significantly lower in R80 than in S80 (MW, p-value$<$ 0.05). Notice that

[21] Our findings here are also confirmed in the third column of Table 1.6.

[22] The unexpected rate of trade of 11% in S40 is not sufficient to explain the small difference in welfare generated in pairs with H-types in both treatments –in particular given that the theoretical gains in R40 are $1000 * 0.8^3 = 512$.

[23] On top of that, the increase in rate of trade with an L-type from 82% in S40 to 99% in R40 does not compensate for the loss due to delay. Notice that realized welfare with an L-type in S40 accounts for 82% of the expected welfare, and for 75% in R40 (and only 46% of the expected welfare when trading with an H-type).

Table 1.8: Payoffs

		R80	R40	S80	S40
Buyer	H	376 (1000)	166 (512)	528 (1000)	87 (0)
	L	-328 (-750)	435 (220)	-585 (-750)	708 (1750)
Seller	H	147 (0)	72 (0)	100 (0)	18 (0)
	L	1574 (2500)	696 (1280)	2197 (2500)	725 (0)

Notes: Average payoffs for buyers and sellers separated by sellers' types. Theoretical predictions are given in brackets.

realized welfare with an L-type in S80 accounts for 92% of the expected welfare, while only for 71% in R80. Overall, when flexibility in making offers is introduced, buyers are not able to commit to avoid this option.

Result 5 already indicates the failure of the R-treatments. In its own right, it gives a first piece of information regarding differences in welfare generated between R and S-treatments as well as between H-type and L-type bargaining pairs. What is still missing is an additional step of disaggregation of payoffs between buyers and sellers. It is important to know how gains from trade are shared, in particular compared to the theoretical predictions.

Result 6. *Cushioning of Losses and Ineffective Commitment*
In the R-treatments, buyers cushion potential losses with L-type sellers by delaying high offers. This implies that buyers bear most of the welfare losses in R80 (relative to the theoretically expected welfare), while sellers do in R40. In the S-treatments, buyers bear all the welfare losses, because commitment power cannot be used effectively.

Table 1.8 displays the buyers' and sellers' average payoffs separated by H-type and L-type cases (theoretical predictions are given in brackets). By weighting the observed payoffs according to the probability of occurrence of H-type and L-type sellers, one can derive the average payoffs over both types. For instance, in R80, the buyers' average payoff is $0.8 * 376 + 0.2 * (-328) = 235.2$. The buyers' theoretically expected payoff is $0.8 * 1000 + 0.2 * (-750) = 650$. In the same way, the sellers' average payoff in R80 is 432.4, while it should be 500 at the SE. Hence, the buyers bear 86% of the welfare loss in R80. Following the same procedure for the other treatments yields that sellers bear 97% of the welfare losses in R40 and that buyers bear 100% of the welfare losses in the S-treatments (sellers earn more than the theoretical prediction, yet welfare falls short of the theoretical prediction).[24]

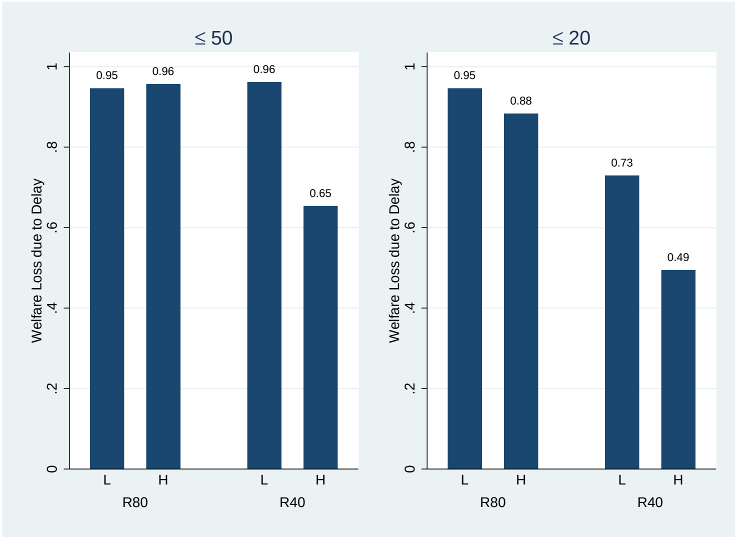
It is instructive to look at the payoffs separated by H-type and L-type cases in more detail. A common feature in R80 and R40 is that buyers trade off gains obtained with H-type sellers to get higher gains from L-type sellers, beyond the theoretical predictions (Median test p-values are all < 0.03). By shifting to different price sequences in R40 than the predicted one, buyers are able to reduce the losses made with some of the L-type sellers at the SE (those occurring with trades at 2000). Indeed, buyers are able to cushion losses obtained with L-type sellers in R80 from -750 to -328 and increase their payoffs in R40 L-type cases from 220 to 435. Both comes at the expense of lower generated gains in bargaining pairs with an H-type seller. Overall, buyers seem to be hurt by their inability to commit in the 80s-treatments (average payoffs are significantly lower in R80 than in S80) but not in the 40s-treatments.

In S80, the informational advantage of the L-type sellers is at its maximum because parameters do not prevent first-best efficiency. An L-type seller is paid the cost of an H-type seller. Because buyers attempt at screening in R80 (and cushion losses), it is not surprising to witness a significant reduction in the informational rent from an average of 2197 to 1574 (MW, p-value < 0.05). On the other hand, H-type sellers are able to extract some rent so that overall, buyers bear all the welfare losses in S80: even though most buyers make offers exceeding 2500, H-type sellers reject such offers in 27% of the cases. This reduces buyers' payoffs below the theoretical prediction.

A remarkable observation is the absence of significant difference between sellers' payoffs in R40 and S40 (MW, p-value$= 0.831$ for L-type cases). Hence, under severe adverse selection L-type sellers are equally well off if the buyer has full commitment power (single-price offer) or uses repeated offers. On the one hand, despite the commitment power of

[24]Median tests confirm that buyers earn less than theoretically predicted in all treatments (p-values < 0.03 for the R-treatments and < 0.07 for the S-treatments). In R80 and S80 sellers' payoffs are not significantly different (p-values > 0.3) from predictions and in S40 sellers earn more than theoretically expected. On the other hand, sellers' payoffs are reduced in R40 (p-value < 0.03).

Figure 1-5: Sources of Inefficiency



Notes: (1) The bars show the percentage of welfare loss due to delay relative to the ex-ante welfare predicted by sequential equilibrium. (2) In the second figure, we do the same computation but restricting to trades occurring within 20 stages, (≤ 20).

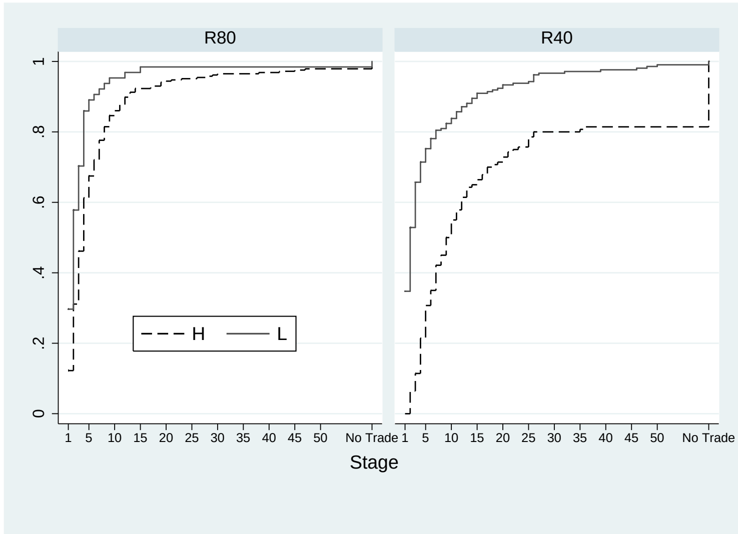
buyers in S40, buyers are unable to fully capitalize on the advantage of a single-price offer. In particular, gains from trade in S40 are roughly split in half with L-type sellers. This is reminiscent of findings from ultimatum games with complete information.[25] On the other hand, in R40, L-type sellers are only at 54% of the predicted payoff level, because buyer increases prices relatively slowly.

Two Channels of Efficiency Loss: We now close this section with a quantification of the two channels generating welfare losses. The two sources of efficiency loss are trade failures and delay. By design, any deviation from the theoretical prediction observed in the S-treatments come from trade failures. In contrast, in the R-treatments both sources can be at play. In Figure 1-5, we explicitly show the percentage of inefficiency due to delay for both types of sellers in the R-treatments. The figure shows efficiency loss by types over all 50 stages, but also for trades occurring only over the first 20 stages –thereby counting trades beyond stage 20 as unsuccessful.[26]

[25] See for instance the survey by Güth and Tietz (1990).

[26] Both figures are computed as follows. Consider R80 and the bar associated with the H-type. We first consider (i) all cases with an H*1000 (whether trade occurred or not), this gives us the potential welfare. We then compute the total efficiency loss as the potential welfare minus the achieved welfare in all these

Figure 1-6: Empirical Cumulative Distribution of Trading Stage



Not surprisingly, in R80 delay almost fully explains welfare losses. This indicates also that trades occur mostly within the first 20 stages. In contrast, both channels of efficiency loss are at play in R40. For instance, when considering trades within the first 20 stages delay and trade failures are each responsible for roughly half of the welfare loss in pairs with an H-type seller and delay is responsible for 73% of the welfare loss in pairs with an L-type seller. Contrary to R80, there is a difference between the ≤ 50 and the ≤ 20 cases, and this even with L-type sellers. This is reminiscent of Result 4 showing that trade is sluggish in R40.

1.4.3 Roots of Over-Delay

In this section we further explore the reasons for over-delay. A glance at Figure 1-6 provides a good overview. The figure shows the fraction of bargaining pairs that have traded at or before a particular stage. It is apparent that in both treatments there is more delay with H-type sellers than with L-type sellers and that delay with H-types is relatively large, in particular in R40. The separation of H and L-types is much stronger

cases. (ii) Next we do the same operation but only for cases in which trade was achieved. This gives the efficiency loss that is due to delay (no trade failure here). Dividing (ii) by (i) gives the efficiency loss that is due only to delay.

in R40 than in R80. In R40 buyers are reluctant to increase their prices even though L-types are screened out relatively fast. By stage 3 already 66 percent of the L-types have accepted. Yet, at this point only 11 percent of the H-types have traded. We show in the following that this can be explained by loss aversion.

A second source of over-delay is that sellers haggle, in particular H-types frequently reject offers above 2500. Such behavior is most common in R80. An important factor here is the expectations that are formed by sellers in response to observed price sequences. In particular, high offers trigger high expectations about future offers. Our analysis will be in support of the following result.

Result 7. *Loss Aversion and Haggling*
Buyers and sellers both contribute to over-delay. (i) Buyers exhibit loss aversion and prefer to delay potential losses by following a flatter price sequence than predicted. (ii) H-type sellers tend to delay agreement. (iii) Quick screening is complicated by the fact that both L and H-type sellers' expectation about future offers are strongly increasing in the level of the past offers.

Buyers Delay Agreement: Equilibrium price sequences in the R-treatments involve the risk of making losses. We have seen that in both treatments buyers trade off lower gains from H-types against higher gains (or reduced losses in R80) from L-types. Loss aversion seems to be a promising candidate to explain the buyers' deviation from the SE prediction. In Appendix C we show that the screening equilibrium of the bargaining model indeed implies more delay if we account for loss aversion.[27] Intuitively, starting with a high offer in R80 and screening out L-type sellers in R40 requires offers that potentially lead to a loss. A loss averse buyer may prefer to delay these losses. In particular, for a reasonable amount of loss aversion the equilibrium offer sequence becomes $(1024, 1280, 1600, 2000, 2500)$ in R40 and $(2000, 2500)$ in R80.[28]

[27] In the questionnaire conducted at the end of the experiment, some participants mentioned that they tried to avoid making losses. In Rapoport et al. (1995) the uninformed party never runs the risk to make losses. The fact that we find over-delay and they do not is therefore consistent with loss aversion.

[28] See Appendix C for a derivation of these price sequences. We use a piece-wise linear payoff function with a kink at 0, putting more weight to losses than gains. The reported sequences use a loss aversion parameter of 2, i.e., losses receive twice the weight of gains.

Table 1.9: Loss Aversion

	Opening Offer	Bargaining Length	Trade (Linear)	Trade (Probit)	Buyer Payoff
Loss Averse	-594.9**	3.6***	-0.06	-0.14	211.3
(LA)	(266.20)	(1.07)	(0.06)	(0.12)	(205.30)
LA x H		-4.15***	-0.21**	-0.27*	-295.10
		(1.18)	(0.11)	(0.15)	(286.60)
R40	-1063.50***	3.82***			1028.20***
	(229.20)	(1.00)			(174.20)
R40 x LA	563.40**	-3.21**			-154.00
	(278.80)	(1.41)			(219.30)
R40 x H		-1.25			-1224.40***
		(1.52)			(249.30)
R40 x LA x H		3.49*			239.40
		(2.04)			(297.30)
H		5.06***	-0.05	-0.09	1080.70***
		(0.96)	(0.05)	(0.13)	(237.70)
Constant	1751.10***	-0.03	0.84***		-758.30***
	(249.50)	(0.92)	(0.08)		(170.40)
R^2 (overall)	0.25	0.14	0.19		0.33
Log-Likelihood				-44.66	
Observations	340	307	170	119	340
Individuals	34	34	17	17	34
Reference Group	R80 / LA=0	R80 / LA=0 / L	R40 / LA=0	R40 / LA=0	R80 L / LA=0

Notes: (1) Columns 1-3 and 5 are random effects regressions and column 4 is a pooled probit with the average marginal effects reported. Standard errors clustered on individuals in parentheses. Significance levels * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (2) The dummy "Loss Averse" is constructed such that it is equal to 1 if a subject only accepted lotteries with a 50% probability of losing 4 CHF or less and 0 otherwise. (3) The variable "Trade" is equal to 1 if trade occurred within the first 20 stages and 0 otherwise. Columns 3 and 4 use only data from R40, since in R80 trade occurred almost always.

We gathered information on loss aversion in 6 of the 12 sessions for the R-treatments. We used the same lottery task as Fehr et al. (2013). After subjects are told their earnings from the bargaining experiment, they were presented six lotteries which they could either accept or decline. Each lottery gives a 50-50 chance between winning an additional 6 CHF or losing an amount that differs between lotteries. The amount that could be lost was 2,3,4,5,6,7 for the six lotteries. One of the six lotteries was then randomly selected and paid. In case the selected lottery was declined, no additional earnings or losses were realized.[29] [30]

[29] 34 out of 36 buyers have a unique switching point in their lottery decisions. The mean switching point is 2.6, i.e., it is between the lotteries with 50% probability of losing 3 and 4. We only use buyers with a unique switching point.

[30] In principle, the lottery task may also measure a subject's risk aversion. However, Rabin's (2000)

Table 1.9 summarizes the impact of loss aversion on opening offers, the length of the bargaining process, trade success and buyers' payoffs. The dummy "Loss Averse" is constructed such that it is equal to 1 if a subject only accepted lotteries with a 50% probability of losing 4 CHF or less. Subjects who accepted 50-50 lotteries between winning 6 CHF and losing 5 CHF or more are considered to be not loss averse.[31]

Column 1 shows that in R80 loss averse buyers on average start with an offer that is around 600 points lower than offers coming from less loss averse buyers. This effect is not present in R40 where the opening offer is generally much lower.[32] Recall that in R80 opening offers exhibit large standard errors. This high variability in R80 seems to be captured by loss aversion: while loss averse buyers are reluctant to make high offers, less loss averse buyers are willing to start with higher offers trying to increase their payoffs with H-type sellers.

The results on bargaining length show that by starting with a higher offer, buyers were indeed able to speed up the bargaining process in R80. Unfortunately, this is only true for trade with L-type sellers. Delay with H-types is not significantly lower for less loss averse buyers, possibly because H-type sellers' haggling is the main factor of delay in R80. In R40, loss aversion has no impact on bargaining length.[33] This suggest that loss aversion seems to be unimportant in R40. However, columns 3 and 4 show that loss averse buyers are responsible for a large part of the trade failures in R40.[34] If the seller is an H-type and the buyer belongs to the group with a larger loss aversion, trade rates are reduced substantially.[35] This indicates that loss averse buyers were not willing to offer high prices even if discounting has erased most gains from trade.

Finally, the last column in Table 1.9 shows that in R80 loss averse buyers incur smaller losses with L-types but also realize smaller gains with H-type sellers. These differences are,

calibration theorem shows that the rejection of lotteries for losses smaller than 6 would imply unreasonable levels of risk aversion when stakes are higher.

[31]Using this procedure 24 buyers are classified as loss averse and 10 as not loss averse. Changing the switching point does not affect results qualitatively, but the differences may become less significant.

[32]This can be seen by combining the coefficients for LA and R40*LA.

[33]This can be seen by combining the coefficients for LA and R40*LA for L-type sellers and LA, LA*H, R40*LA, R40*LA*H for H-type sellers.

[34]Only data from R40 is used, since trade failures are negligible in R80.

[35]Note, however, that the significance is only at the 10 percent level for the probit estimates.

however, not significant. Similarly, in R40 there are no significant differences in payoffs due to loss aversion. While loss aversion seems to be an important driver of buyers' behavior, less loss averse buyers could not realize higher profits. An important factor in explaining this is that H-type sellers often rejected offers above 2500, which leads to costly delay. We will refer to rejections of acceptable offers by H-types as *haggling*. Let us now turn towards sellers' behavior in more depth.

Sellers Delay Agreement: Sellers' acceptance decisions are summarized in Table 1.10. Only 22% of the offers between 2500 and 3000 were accepted by H-type sellers in R80. Similarly, in R40 this number corresponds to 27%. For the S-treatments, acceptance rates of acceptable offers for H-types are higher than in the R-treatments. This is a direct implication of the buyer's commitment power, which leaves the seller with no opportunity to haggle.[36] Acceptance rates of L-type sellers for offers between 500 and 2500 are non-negligible in R80 and R40. Hence, in both treatments it was worthwhile for buyers to start with relatively low offers to screen out L-types.

Why do H-type sellers haggle? A simple check of whether seller strive for higher profits or are motivated by other considerations is to see how often sellers accept the best possible offer. The best offer is the highest *discounted* offer in an offer sequence of a particular bargain.[37] However, we also need to take into account that sellers could accept offers too early. Therefore, we estimated price sequences and used these estimates to predict what price offers would have been made if the seller had not accepted. This allows to construct complete price sequences (for a detailed description see Appendix D Table 1.C.1). Using the predicted price sequences, the percentages of best offers accepted for H-type sellers are 38% for R80 and 27% for R40. L-types accepted the best offer in 62% of the cases in R80 and in 48% of the cases in R40. Thus, L-types accepted the best offer more often than H-types in both treatments. In contrast to H-type sellers, L-type sellers potentially

[36] Note that the acceptance rate in S80 for H-types is 73 percent for offers between 2500 and 3000. Similarly, acceptance rates in S40 for L-types is 61 percent for offers between 0 and 500. The fact that there is still a considerable fraction of rejections is in line with the literature on fairness considerations in ultimatum games.

[37] Only bargains that concluded in trade are considered.

Table 1.10: What Prices Do Sellers Accept?

Treatment	Type	Price Range 0-500	500-1000	1000-1500	1500-2000	2000-2500	2500-3000	3000-4000
R80	H	0.00	0.03	0.02	0.01	0.00	0.22	0.72
		(120)	(183)	(174)	(156)	(82)	(732)	(138)
R80	L	0.00	0.14	0.12	0.28	0.57	0.78	0.25
		(13)	(49)	(39)	(25)	(7)	(41)	(32)
R40	H	0.00	0.00	0.00	0.02	0.00	0.27	0.65
		(245)	(554)	(279)	(131)	(45)	(249)	(43)
R40	L	0.02	0.18	0.34	0.28	0.33	0.57	0.80
		(252)	(494)	(185)	(39)	(9)	(37)	(5)
S80	H	0.00	0.00	0.00	0.00	0.00	0.73	1.00
		(2)	(5)	(10)	(12)	(4)	(158)	(11)
S80	L	0.00	1.00	1.00	1.00	1.00	0.93	1.00
		(1)	(1)	(2)	(2)	(1)	(29)	(2)
S40	H	0.00	0.00	0.00	0.00	0.00	1.00	1.00
		(53)	(37)	(7)	(4)	(1)	(10)	(2)
S40	L	0.61	0.93	1.00	1.00	1.00	1.00	1.00
		(46)	(41)	(12)	(5)	(2)	(7)	(3)

Notes: (1) The acceptance rates are computed as the fraction of accepted price offers among all offers made within the corresponding price range (as given in brackets) and within the first 20 stages.

realize high profits. They thus seem to be more eager to accept the best offer.

Next, we estimate a discrete choice model of the sellers' acceptance decisions.[38] We try to distinguish between the following considerations that sellers potentially take into account when deciding to accept or reject a specific offer. First, there may have been sellers who followed simple rules of thumb that are directly linked to the current offer they face. For instance, an L-type seller decision rule may be that she never accepts less than 1000 in R40. Such rules of thumb could also be related to discounted offers. Discounted offers also cover stage effect, e.g., the same offer that is accepted in stage 2 could be rejected in later stages. Second, previously observed offers may be important, since they shape expectation about future offers. From the estimation of the price sequences, we know that there is a strong positive correlation between current and past offers. Finally, haggling is captured by the variable "Difference to Best Offer". This variable gives the difference

[38]Notice that here we use the observed offers and not the constructed price sequences.

Table 1.11: Probit Estimates of Sellers' Acceptance Decisions

	R80		R40	
	H	L	H	L
Offer_t	0.10677***	0.02113***	0.12618***	0.01954***
	(0.01334)	(0.00668)	(0.03697)	(0.00431)
Offer_{t-1}	-0.06109***	-0.00381	-0.1198***	-0.01111**
	(0.01387)	(0.00783)	(0.02752)	(0.00545)
Discounted Offer_t	0.01138***	0.00002	0.00909	0.01393**
	(0.00257)	(0.00652)	(0.00704)	(0.00584)
Difference to	-0.0013	-0.02797***	-0.01008	-0.01011**
Best Offer	(0.00283)	(0.00983)	(0.00845)	(0.00408)
Pseudo Log-Likelihood	-283.87	-57.12	-110.35	-304.76
Observations	751	148	231	978
Individuals	34	27	26	34

Notes: (1) Standard errors are clustered on individuals (in parentheses). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (2) Observations with offers below 2500 for H-type sellers are excluded. (3) Coefficients are multiplied by factor 100, i.e., they give the effect of a change in the explanatory variable of 100 points.

between the best possible offer that is observed in the respective sequence and the current (discounted) offer. If sellers try to avoid unnecessary haggling, we would expect a negative coefficient for "Difference to Best Offer", indicating that sellers try to accept an offer that is as close as possible to the best offer.

Table 1.11 presents the results of the probit estimation. The dependent variable is the binary variable "accept" which is equal to 1 if the offer was accepted and 0 otherwise. As expected the coefficient for the current offer is positive and significant for both treatments and seller types. Note that the coefficients have been multiplied by 100 for convenience. Hence, L-type sellers are 2 percent more likely to accept if the offer is increased by 100 points.[39] For H-type sellers only observations with offers above 2500 are included.[40] Therefore, a price increase of 100 points has a much stronger effect, namely 11 percent in R80 and 13 percent in R40.

The negative coefficient for "Offer_{t-1}" points towards an important difficulty that buy-

[39] At first sight this effect seems to be small. Recall however that offers often increase from around 800 to 2500 and more within a few stages.

[40] Offers below 2500 are usually not accepted.

ers had to overcome when trying to screen L-type sellers in R40. A high offer in the current stage implies that the seller is then less likely to accept a given offer in the next stage. Observing high offers, sellers seem to expect even higher offers in the future. This complicates screening. This effect is present for H-type sellers as well. Thus, a buyer that raises offers above 2500 relatively fast will face a more demanding H-type seller. The fact that high offers induce expectations for higher offers in the future may explain why less loss averse buyers are not able to realize higher profits.

Finally, the coefficient for "Difference to Best Offer" is significant only for L-type sellers. We interpret this as evidence that L-type sellers' haggling is limited. On the other hand, H-type sellers are not more likely to accept an offer that is closer to the best offer in the sequence (after controlling for the level of the offer). This is in line with the observation that L-type seller accept the best offer much more often than H-type sellers.

Summing up, our discussion draws the following picture about sellers' motivation to accept or reject an offer. H-type sellers haggle even if this implies lower payoffs. Our intuition for this is that profits for H-types are generally low, which means that other considerations dominate, such as following simple rules of thumb. L-type sellers on the other hand can generate high profits. Accordingly, they do not engage in costly haggling as much as H-type sellers and often accept the best possible offer. However, in particular in R80, rules of thumb seem to be important as well. Perhaps most importantly, behavior of sellers seems to be driven by their expectations about future offers, which directly depend on past offers.

1.5 Conclusion

Better Later than Never? A welfare-based evaluation of our experimental bargaining protocols yields that the single-price offer fares better than the repeated offer protocol, in contrast with the theoretical predictions. However, when our main concern is whether goods are traded or not, then repeated offers perform well: trade rates are boosted upwards when buyers are allowed to make a sequence of offers.

Importantly, both bargaining protocols under-performed relative to the theoretical prediction. In the S-treatments sellers rejected many offers they should have accepted if their decisions were based exclusively on monetary payoff. In the R-treatments the buyers' lack of commitment power leads to long delay before trades were reached. Both phenomena drive down efficiency substantially. We identify loss aversion as a behavioral explanation for the long delay. Another important factor is the expectations steep price sequences induce: buyers who raise prices fast are expected to raise prices even more even if the reservation price of H-type sellers is already covered. This makes it less beneficial to use steep price sequences and delays trades.

Overall, our assessment of the predictions made by sequential equilibrium is rather positive. Naturally, we observe many deviations from these predictions, most notably a substantial over-delay. This is true in particular if parameters are such that there should be no delay even if repeated offers are possible. Despite the systematic deviations from equilibrium predictions and the departure from the loss-neutrality assumption made in the standard model, the main message carries over to the experimental results: buyers use the possibility of repeated price offers to screen out L-quality sellers. This leads to trade with H-quality sellers, even though incentive constraints preclude this in the single-price offer bargaining protocol.

Extensions: Several immediate changes in our set of parameters come to mind. First, varying the discount rate seems important to evaluate the salience of the over-delay observed in our experiment. Next, the R-treatments allowed for a lengthy bargaining, possibly going to stages where payoffs become very low. It seems important to evaluate whether, under identical equilibrium predictions, shorter bargaining spans would push trade to occur faster and at the same rates.

The observed payoffs distribution indicates that in the context of exclusive bargaining, private information may be more valuable than advantages due to the specifics of the bargaining institution. It seems important to shed light on the possible differences with a set-up in which bargaining occurs in markets and the exclusivity between a buyer and a seller is only temporary. We leave these questions open for future research.

Appendix

1.A Instructions for Treatment R80

Welcome to this economic experiment. From now on you are not allowed to communicate in any other way than specified in the instructions. Please obey to this rule because otherwise we have to exclude you from the experiment and all earnings you have made will be lost. Please also do not ask questions aloud. If you have a question, raise your hand. A member of the experimenter team will come to you and answer your question in private.

In this experiment you can earn money with the decisions you make. How much you earn depends on your own decisions, the decisions of other participants as well as random events. We will not speak of CHF during the experiment, but rather of experimental points. All your earnings will first be calculated in points. At the end of the experiment the total amount of points you earned will be converted to CHF at the following rate:

$$100 \text{ points} = 0.4 \text{ CHF}$$

In addition, you will receive a show up fee of 10 CHF.

The experiment consists of two parts that are independent of one another. For each part you will receive specific instructions. These instructions will explain how you make decisions and how your decisions and the decisions of other participants influence your earnings. Therefore, it is important that you read the instructions carefully.

In case you should make losses, the show up fee of 10 CHF is used to cover for these losses. If you make losses exceeding 10 CHF, you will have the option to leave immediately and earn 0 CHF.

The General Setting

We will now describe the general setting you will face during the experiment. At the beginning of the experiment the participants will be divided into buyers and sellers. Half of the participants will be buyers and the other half will be sellers. When you are a buyer (respectively, a seller) you will stay a buyer (respectively, a seller) throughout the experiment. A decision situation (explained below) will be repeated for 10 periods. In each period a buyer and a seller are randomly matched. In other words, the participants are divided into pairs and each pair consists of one buyer and one seller. You will not get to know the identity of the buyer or seller you are paired with, neither during nor after the experiment. The participant who is paired with you will also not get to know your identity. In each period new pairs will be formed randomly.

The Decision Situation

The decision situation will be the same for all ten periods. We will now describe one such period. After the buyer and the seller have been matched, they face the following situation. The seller can be of two different types: type A or type B. A seller of type A can only produce a high quality good at cost 2500. A seller of type B can only produce a low quality good at cost 0. The buyer's valuation for the high quality good is 3500. The buyer's valuation for the low quality good is 1750.

The seller knows whether she is of type A or type B and therefore also knows how much the good is worth to the buyer. However, the buyer does not know the seller's type and hence, the buyer does neither know whether his valuation for the good is 3500 or 1750 nor whether the cost of the seller to produce the good is 2500 or 0. The type of the seller will be determined randomly according to the following probabilities at the beginning of each period: the probability that the seller is of type A (high cost / high quality good) is 0.8(80%) and the probability that the seller is of type B (low cost / low quality good) is 0.2(20%).

To acquire the good, the buyer makes offers to the seller. The offers must be between 0 and 4000 and can be as exact as to the first decimal place. If you enter an offer that is not allowed, the computer will tell you and you will have to change your offer. Upon seeing the buyer's offer, the seller can accept or reject the offer. If the seller accepts the offer, she produces the good and sells it to the buyer at the agreed price. The buyer does not make further offers and the trading pair has to wait until all other pairs have finished their trading process and buyers and sellers are rematched to form new pairs in the next period.

If the seller rejects the offer, the buyer can make a new offer to the seller which can again be accepted or rejected. There can be at most 50 stages, i.e. a buyer can make at most 50 offers to a seller. Likewise, a seller can reject up to 50 offers. If all 50 offers are rejected, the good is not produced (and not traded) and both parties earn 0.

In which stage trade takes place does matter. The buyer and the seller both discount the future at the discount rate $d = 0.8$. This means that a profit (or loss) realized in stage n is discounted according to the given discount rate. For instance, if the buyer makes a profit of x experimental points in stage 1, he earns x experimental points since there is no discounting. If the buyer makes a profit of x experimental points in stage 3, he earns $x * 0.8 * 0.8 = x * 0.8^2$ experimental points. Generally, if an offer is accepted in stage n, the payoffs are determined as follows.

$$\begin{aligned}\text{The buyer's payoff} &= (\text{Valuation of the Good} - \text{Accepted Offer}) * d^{n-1} \\ \text{The seller's payoff} &= (\text{Accepted Offer} - \text{Production Cost}) * d^{n-1}\end{aligned}$$

For convenience the valuations and costs are summarized below:

- Buyer's valuation for the high quality good = 3500
- Buyer's valuation for the low quality good = 1750
- Seller's cost of producing the high quality good = 2500
- Seller's cost of producing the low quality good = 0

Once all pairs have traded the good at some price or all offers have been rejected, the computer randomly matches buyers and sellers anew and the next period starts. The experiment ends after period 10.

1.B Characterization of the Sequential Equilibrium

Lemma 1. *In the bargaining game with $N > 0$ stages, the final equilibrium offer is either $p_0^s = \overline{c}$ or $p_0^z = 0$.*

i) If $p_0^s = \overline{c}$, the sequence of equilibrium offers is given by $p_k^s = \delta^k \overline{c}$ for $k = k^*(0), k^*(0) - 1, \ldots, 0$.

ii) If $p_0^z = 0$, the sequence of equilibrium offers is given by $p_n^z = 0$ for $n = N-1, N-2, \ldots, 0$.

Proof. The only price offered and accepted with certainty before stage N is $\overline{c}$. Offers below $\overline{c}$ are rejected with a positive probability and offers exceeding $\overline{c}$ are dominated by $\overline{c}$. The offer in the last stage is $p_0 = \overline{c}$ if $(1-\hat{q})\overline{v} \geq (1-q_0)\overline{c}$ and 0 otherwise.

To prove i), suppose by contradiction that $p_k^s \neq \delta p_{k-1}^s$ for at least one $k = k^*(0), k^*(0) - 1, \ldots, 1$. This implies that either p_k^s or p_{k-1}^s is accepted or at least one of these offers is rejected for sure by the L-quality sellers. Sure acceptance requires the buyer to offer $\overline{c}$ after a rejection, but then the L-quality sellers would not have accepted any offer below $\delta\overline{c}$. For the second to last stage, p_1^s such that $\delta\overline{c} < p_1^s < \overline{c}$ is dominated by $\delta\overline{c}$. Sure rejection of p_k^s by L-quality sellers implies $q_k^s = q_{k-1}^s$, contradicting sequential rationality, given that sellers follow a stationary strategy (see DL for a proof that the equilibrium must be stationary).

To prove ii), note that an offer of 0 cannot conclude the game for sure, unless made in the last stage. Hence, there are N equilibrium stages. Suppose $p_n^z > 0$ for at least one $n = N-1, N-2, \ldots, 1$. Then p_n^z is accepted by L-quality sellers. Observing a rejection of p_n^z implies $p_{n-1}^z = \overline{c}$. But this either contradicts the fact that there are N stages or that $p_0^z = 0$. □

Denote by $R^s(q_n)$ for $n = N-1, N-2, \ldots, 0$ the buyer's maximized ex-ante expected payoff from trading with the sellers in $[q_n, 1]$. Similarly, $R^z(q_n)$ denotes the ex-ante expected payoff of the zero offer sequence. Let $a_k^s(q_n) = q_{k-1}^s - q_k^s$ for $k = k^*(q_n)-1, k^*(q_n)-2, \ldots, 0$ denote the ex-ante probability of agreement k stages before the final equilibrium stage if the buyer follows the screening offer sequence and the buyer believes that only seller types $q \in (q_n, 1]$ are left. Finally, $a_n^z = q_{n-1}^z - q_n^z$ denotes the ex-ante probability of agreement n stages before the final equilibrium stage for the zero offers equilibrium. The ex-ante payoffs can be separated into gains in the current stage and discounted future gains, i.e.,

$$R^s(q_n) = (\underline{v} - p_{k^*(q_n)}^s) a_{k^*(q_n)}^s(q_n) + \delta R^s(q_{k^*(q_n)-1}^s) \tag{1.1}$$

$$R^z(q_n) = a_n^z \underline{v} + \delta R^z(q_{n-1}^z) \tag{1.2}$$

If the equilibrium involves screening, the buyer must be indifferent between offering p_k^s and p_{k-1}^s for $k = k^*(q_n), k^*(q_n)-1, \ldots, 1$. The intuition for this result is given in the main text. For a proof we refer to DL. Note that indifference between p_k and p_{k-2} is not possible, because then the implied cutoff level is such that the offer p_{k-1} is the preferred offer. The advantage of offering p_{k-1}^s rather than p_k^s is that the continuation surplus $R^s(q_{k-1}^s)$ is obtained one stage earlier. On the other hand, by offering the higher price the buyer loses $(p_{k-1}^s - p_k^s)$ on the seller types in $(q_k, q_{k-1}]$ that would have accepted the lower price. The gains from accelerating trade must balance out the losses, i.e., $(1-\delta)R^s(q_{k-1}^s) = (p_{k-1}^s - p_k^s)a_k^s(q_k^s)$. Using this insight, one can show that the ex-ante acceptance probabilities are given recursively by (1.3).

$$a_k^s(q_n) = \begin{cases} 1 - \hat{q} & \text{if } k = 0 \\ \frac{\overline{v} - \overline{c}}{\overline{c}} a_0^s & \text{if } k = 1 \\ \frac{\underline{v}}{\delta^{k-1}\overline{c}} a_{k-1}^s(q_n) & \text{if } 2 \le k \le k^*(q_n) - 1 \\ 1 - q_n - \sum_{i=0}^{k^*(q_n)-1} a_i^s(q_n) & \text{if } k = k^*(q_n) \end{cases} \tag{1.3}$$

The equilibrium number of screening stages is restricted by either the mass of L-quality sellers or the maximal number of stages N. More specifically, the number of screening stages is given by

$$k^*(q_n) = \min\left\{\max\left(k : \sum_{i=0}^{k} a_i^s(q_n) < 1\right), N-1\right\} \tag{1.4}$$

Knowing the price sequence and the acceptance decisions of the screening equilibrium allows to derive the acceptance probabilities for the zero offer sequence. As explained in the main text, the sellers' acceptance decisions must render the buyer indifferent between offering the optimal screening price (given the current belief) and offering 0. We use backward induction. In the last stage, the buyer must be indifferent between offering $p_0^s = \overline{c}$ and $p_0^z = 0$. This is the case when $q_0^z = 1 - \frac{\overline{v}(1-\hat{q})}{\overline{c}}$ which implies $a_0^z = \hat{q} - q_0^z = \frac{(1-\hat{q})(\overline{v}-\overline{c})}{\overline{c}}$.

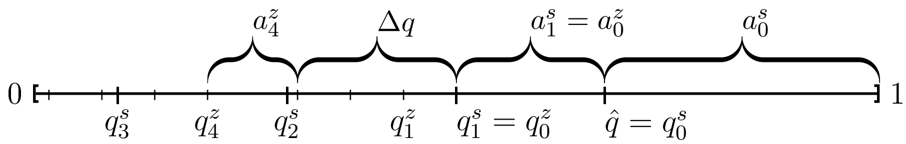
In general, two subsequent stages can either imply a belief that leads to the same number of screening stages, $k^*(q_n^z) = k^*(q_{n-1}^z)$, or the earlier stage implies one more screening stage, i.e. $k^*(q_n^z) = k^*(q_{n-1}^z) + 1$. It is easy to see that $k^*(q_n^z) < k^*(q_{n-1}^z)$ is not possible, since the buyer's belief to bargain with an H-quality seller cannot decrease over the course of the game and a higher belief implies less screening. More surprisingly, $k^*(q_n^z) = k^*(q_{n-1}^z) + 2$ can be excluded as well. Intuitively, if a change in the belief from q_{n-1}^z to q_n^z entails an increase in the optimal number of screening stages of 2 (or more), then the cost from delaying trade is greater for the screening than the zero offer sequence. But since the zero offer sequence yields a greater profit also for the current period, this is not possible.[41]

If $k^*(q_n^z) = k^*(q_{n-1}^z)$ then $a^s_{k^*(q_n^z)}(q_n^z) = a_n^z + a^s_{k^*(q_n^z)}(q_{n-1}^z)$ and thus (1.1) becomes

$$R^s(q_n^z) = (\underline{v} - p^s_{k^*(q_n^z)})a_n^z + R^s(q_{n-1}^z) \tag{1.5}$$

[41] Formally, $k^*(q_n^z) = k^*(q_{n-1}^z) + 2$ requires $a^s_{k^*(q_n^z)}(q_n^z) \leq a_n^z$ and therefore also $q^s_{k^*(q_n^z)-1} \leq q_{n-1}^z$. Writing (1.1) as $R^s(q_n^z) = (\underline{v} - p^s_{k^*(q_n^z)})a^s_{k^*(q_n^z)}(q_n^z) + (\underline{v} - p^s_{k^*(q_n^z)-1})(a_n^z - a^s_{k^*(q_n^z)}(q_n^z)) + \delta^2 R^z(q_{n-1}^z)$ and comparing it to (1.2) implies that $R^s(q_n^z) < R^z(q_n^z)$, contradicting the fact that the buyer is indifferent between the zero offer and the optimal screening sequences.

Figure 1.B.1: Cutoff Level



Since $R^z(q_n^z) = R^s(q_n^z)$ for $n = N-2, N-3, \ldots, 0$. It follows from (1.2) and (1.5) that

$$a_n^z \underline{v} - (1-\delta)R^z(q_{n-1}^z) = (\underline{v} - p_{k^*(q_n^z)}^s)a_n^z.$$

The left-hand side in the above relation is the change in expected payoff of the zero offers sequence when the belief changes from q_{n-1}^z to q_n^z. This has to be equal to the change in the expected payoff of the screening sequence given by the right-hand side of the equation. Writing $(1-\delta)R^z(q_{n-1}^z)$ as $R^z(q_{n-1}^z) - \delta\left(R^z(q_{n-2}^z) + (\underline{v} - p_{k^*(q_{n-1}^z)}^s)a_{n-1}^z\right)$ and using (1.2) to replace $R^z(q_{n-1}^z) - \delta R^z(q_{n-2}^z)$, one obtains a_n^z in terms of a_{n-1}^z. The result is shown in (1.6).

If $k^*(q_n^z) = k^*(q_{n-1}^z) + 1$ then $a_{k^*(q_n^z)}^s(q_n^z) = a_n^z + \Delta q$, where $\Delta q = q_{k^*(q_n^z)-1}^s - q_{n-1}^z = \sum_{i=0}^{n-1} a_i^z - \sum_{i=1}^{k^*(q_n^z)-1} a_i^s(q_n^z)$. Hence, (1.1) becomes

$$R^s(q_n^z) = (\underline{v} - p_{k^*(q_n^z)}^s)(a_n^z + \Delta q) + \delta R^s(q_{k^*(q_n^z)-1}^s)$$

Since $R^z(q_{n-1}^z) = R^s(q_{n-1}^z) = (\underline{v} - p_{k^*(q_n^z)-1}^s)\Delta q + R^s(q_{k^*(q_n^z)-1}^s)$, (1.2) can be rewritten as

$$R^z(q_n^z) = a_n^z \underline{v} + \delta(\underline{v} - p_{k^*(q_n^z)-1}^s)\Delta q + \delta R^s(q_{k^*(q_n^z)-1}^s)$$

Equating $R^s(q_n^z)$ and $R^z(q_n^z)$, it can be solved for a_n^z in terms of Δq.[42] The ex-ante

[42] Note that Δq is known, since $a_i^s(q_n^z)$ for $i = 1, \ldots, k^*(q_n^z) - 1$ are given by (1.3) and a_i^z for $i = 0, 1, \ldots, n-1$ are derived recursively.

acceptance probabilities for the zero offer sequence are given recursively by:

$$a_n^z = \begin{cases} \frac{(1-\hat{q})(\overline{v}-\overline{c})}{\overline{c}} & \text{if } n = 0 \\ \left(\delta + \frac{(1-\delta)\underline{v}}{\delta^{k^*(q_n^z)}\overline{c}}\right) a_{n-1}^z & \text{if } 1 \leq n \leq N-2 \text{ and } k^*(q_n^z) = k^*(q_{n-1}^z) \\ \frac{(1-\delta)\underline{v}}{(\delta^{k^*(q_n^z)}\overline{c})}\Delta q & \text{if } 1 \leq n \leq N-2 \text{ and } k^*(q_n^z) = k^*(q_{n-1}^z) + 1 \\ \hat{q} - \sum_{i=0}^{N-2} a_i^z & \text{if } n = N-1 \end{cases} \tag{1.6}$$

The zero offer equilibrium requires that in each stage a positive fraction of the L-quality sellers accepts an offer of 0. Hence, if N is large condition (1.7) fails to hold. The delay associated with the zero offers sequence is then too large to render the buyer indifferent to the optimal screening sequence.

$$\sum_{i=0}^{N-2} a_i^z \leq \hat{q} \tag{1.7}$$

If condition (1.7) holds then the buyer compares the expected profits of the screening and the zero offers strategy. The strategy implying the higher expected profit is the unique sequential equilibrium strategy.

Proposition 2. *Let p_k^s and p_n^z be defined as in Lemma 1. Let a_k^s and a_n^z be defined recursively by* (1.3) *and* (1.6), *respectively. Let $k^*(q_n)$ be defined by* (1.4). *Set $R^z(0) = 0$ if condition* (1.7) *fails to hold. Then the unique sequential equilibrium outcome if $R^s(0) \geq R^z(0)$ is*

$$p(q) \;=\; p_k^s, \quad q \in \begin{cases} [0, q_{k^*(0)-1}^s] & \textit{if } k = k^*(0) \\ (q_k^s, q_{k-1}^s] & \textit{if } k = k^*(0)-1, k^*(0)-2\ldots, 1 \\ (q_0^s, 1] & \textit{if } k = 0 \end{cases}$$

$$a(q) \;=\; a_k^s(0), \quad q \in \begin{cases} [0, q_{k^*(0)-1}^s] & \textit{if } k = k^*(0) \\ (q_k^s, q_{k-1}^s] & \textit{if } k = k^*(0)-1, k^*(0)-2, \ldots, 1 \\ (q_0^s, 1] & \textit{if } k = 0 \end{cases}$$

The unique sequential equilibrium outcome if $R^s(0) < R^z(0)$ is

$$p(q) = p_n^z, \quad q \in \begin{cases} [0, q_{N-2}^z] & \text{if } n = N-1 \\ (q_n^z, q_{n-1}^z] & \text{if } n = N-2, N-3, \ldots, 1 \\ (q_0^z, \hat{q}] & \text{if } n = 0 \end{cases}$$

$$a(q) = a_n^z, \quad q \in \begin{cases} [0, q_{N-2}^z] & \text{if } n = N-1 \\ (q_n^z, q_{n-1}^z] & \text{if } n = N-2, N-3, \ldots, 1 \\ (q_0^z, \hat{q}] & \text{if } n = 0 \end{cases}$$

For a proof of uniqueness in case of screening behavior, we refer to DL. It is then easy to see that if the zero offers sequence is an equilibrium, there can be no screening equilibrium anymore. The buyer would deviate in the first stage to offer 0.

The difference between the finite and the infinite horizon settings is that in the finite horizon case a price sequence consisting of zero offers is a possible equilibrium. However, when N is large enough, the buyer's expected profit from the zero offer sequence approaches 0 or (1.7) is violated. The equilibrium is then given by the screening equilibrium. It is noteworthy that there can also be screening equilibria that are not identical to the one found in the literature for the infinite horizon game. This instance occurs if N restricts the optimal number of screening stages through (1.4), but the buyer still prefers to screen rather than to follow the zero offer sequence.

Our Proposition 1 now follows as a corollary of Proposition 2 above.

Proposition 1. *There exists a finite $\overline{N}$ such that the unique equilibrium is the screening equilibrium for all $N \geq \overline{N}$ and the zero offer equilibrium otherwise.*

Proof. For large N the number of screening stages remains constant in N, i.e., N is irrelevant in (1.4) and $k^*(q_{N-1}^z) = k^*(q_{N-2}^z)$. Moreover, by construction the acceptance decisions in the zero offer equilibrium are such that the buyer's expected payoff is the same as the one he would obtain from optimal screening for any stage except the first one, i.e. $R^z(q_n^z) = R^s(q_n^z)$ for $n = N-2, N-3, \ldots, 0$. Hence, $R^s(0) - R^z(0) = a_{N-1}^z(\underline{v} -$

$p^s_{k^*(0)}) + R^s(q^z_{N-2}) - (a^z_{N-1}\underline{v} + \delta R^z(q^z_{N-2})) = (1-\delta)R^s(q^z_{N-2}) - a^z_{N-1}p^s_{k^*(0)}$. Note that $(1-\delta)R^s(q^z_{N-2})$ remains constant in N once the first remark in this proof holds. The same applies to $p^s_{k^*(0)}$. If $\left(\delta + \frac{(1-\delta)\underline{v}}{\delta^{k^*(q^z_n)}\overline{c}}\right) \geq 1$ then a^z_{N-1} is higher (or remains constant) the higher N becomes. In this case (1.7) is violated for a finite N. If $\left(\delta + \frac{(1-\delta)\underline{v}}{\delta^{k^*(q^z_n)}\overline{c}}\right) < 1$ then a^z_{N-1} approaches 0 as N becomes large. This implies that $R^s(0) - R^z(0) > 0$ for N large enough, since $R^s(q)$ is bounded away from zero for any q. □

1.C Additional Material

1.C.1 Screening Equilibrium under Loss Aversion

In this appendix, we present the theoretical prediction for the bargaining model when subjects' preferences exhibit loss aversion. In particular, the buyer's utility obtained from trade n stages before the final stage is now given by

$$B_n(q) = \begin{cases} \delta^{N-1-n}(v(q) - p(q)) & \text{if } v(q) \geq p(q) \\ \lambda\delta^{N-1-n}(v(q) - p(q)) & \text{otherwise} \end{cases}$$

where $\lambda \geq 1$ is the loss aversion parameter. If $\lambda = 1$ then the utility function reduces to the one used throughout the paper.

Note that the seller's utility is unaffected by loss aversion, because the seller is informed and never runs the risk of a loss. It follows from Lemma 1 that the possible equilibrium price sequences are also not changed. However, the acceptance decisions of the L-quality sellers in the screening equilibrium change. These acceptance probabilities still have to render the buyer indifferent between the current and the next price offer. Since gains from trade with the H-quality seller are always positive $a^s_0(0) = 1 - \hat{q}$ still holds. By backward induction the acceptance probability in the second to last stage solves

$$\lambda(\underline{v} - \delta\overline{c})a^s_1 + \delta(\overline{v} - \overline{c})a^s_0 = \lambda(\underline{v} - \overline{c})a^s_1 + (\overline{v} - \overline{c})a^s_0 \qquad \text{if } \underline{v} < \delta\overline{c}.$$

or

$$(\underline{v} - \delta\overline{c})a_1^s + \delta(\overline{v} - \overline{c})a_0^s = \lambda(\underline{v} - \overline{c})a_1^s + (\overline{v} - \overline{c})a_0^s \qquad \text{if } \underline{v} \geq \delta\overline{c}$$

Hence,

$$a_1^s(0) = \begin{cases} \frac{(1-\delta)(\overline{v}-\overline{c})}{\lambda(\overline{c}-\underline{v})+\underline{v}-\delta\overline{c}} a_0^s(0) & \text{if } \underline{v} \geq \delta\overline{c} \\ \frac{\overline{v}-\overline{c}}{\lambda\overline{c}} a_0^s(0) & \text{otherwise} \end{cases}$$

Similarly, in any earlier stage it holds that either

$$\lambda(\underline{v} - p_k^s)a_k^s + \delta R(q_{k-1}^s) = \lambda(\underline{v} - p_{k-1}^s)a_k^s + R(q_{k-1}^s) \qquad \text{if } \underline{v} < p_k^s$$

or

$$(\underline{v} - p_k^s)a_k^s + \delta R(q_{k-1}^s) = \lambda(\underline{v} - p_{k-1}^s)a_k^s + R(q_{k-1}^s) \qquad \text{if } \underline{v} \geq p_k^s \text{ and } \underline{v} < p_{k-1}^s$$

or

$$(\underline{v} - p_k^s)a_k^s + \delta R(q_{k-1}^s) = (\underline{v} - p_{k-1}^s)a_k^s + R(q_{k-1}^s) \qquad \text{if } \underline{v} \geq p_{k-1}^s.$$

Solving these equations yields for $k = k^*(0) - 1, k^*(0) - 2, \ldots, 2$

$$a_k^s(0) = \begin{cases} \frac{\underline{v}}{\delta^{k-1}\overline{c}} a_{k-1}^s(0) & \text{if } \underline{v} < p_k^s \text{ or } \underline{v} \geq p_{k-1}^s \\ \frac{\overline{c}\delta^k(\lambda-1)+\delta\underline{v}(1-\delta\lambda)}{\overline{c}\delta^k(\lambda-\delta)+\delta\underline{v}(1-\lambda)} a_{k-1}^s(0) & \text{if } \underline{v} \geq p_k^s \text{ and } \underline{v} < p_{k-1}^s \end{cases}$$

Hence, the calculation of the acceptance probabilities remains identical to the case without loss aversion if either an acceptance by an L-quality seller does not involve losses in two consecutive stages or it does lead to a loss in both stages. However, if the price change between two stages is such that acceptance by L-quality sellers leads to a loss in one stage and to a gain in the other stage, then the calculation of the acceptance probability differs.

For the parameters used in the experiment, it holds that $p_1^s = 2000 > 1750 = \underline{v}$ and $p_2^s = 1600 < 1750 = \underline{v}$. The acceptance probabilities if $\hat{q} = 0.2$ are therefore given by $a_0^s = 0.8$, $a_1^s = \frac{160}{750\lambda-250}$, $a_2^s = \frac{128-307.2\lambda}{120-160\lambda-600\lambda^2}$, $a_3^s = \frac{140-336\lambda}{120-160\lambda-600\lambda^2}$ and so on.

Setting $\lambda = 2$, the optimal number of screening stages is 2. If $\hat{q} = 0.6$, $\lambda = 2$ implies 5 equilibrium stages with offers $(1024, 1280, 1600, 2000, 2500)$ and L-type seller's ex-ante acceptance probabilities of $(0.46, 0.21, 0.20, 0.13)$. In general, it holds that the higher λ the more screening should be observed.

1.C.2 Constructed Price Sequence

In Table 1.C.1 we estimate price sequences. In columns 1 and 3 we use only offers that were made in stage 2. The offer for stage 2 is estimated separately since only one lagged offer can be used there. That is, for instance, the value of 0.788 in column 1 means that in R80 if the first offer was 100 points higher, the second offer increases by approximately 80. We use these coefficients to predict offers in stage 2 for price sequences that involve immediate trade in stage 1. Columns 2 and 4, on the other hand, give the predictions for all other stages in dependence of the previous two offers. It is apparent that an offer depends strongly on previous offers. Including more lags does not change results and higher lags are insignificant. Using these estimates we construct price sequences by predicting the offers that would have been made had the seller not accepted an offer. Figure 1.C.1 presents the median price sequence when using the predicted price sequences.

Table 1.C.1: Regression: Price Sequence

	R80	R80	R40	R40
Offer_{t-1}	0.788***	0.647***	1.054***	0.532***
	(0.0713)	(0.0527)	(0.0524)	(0.0675)
Offer_{t-2}		0.0883**		0.236***
		(0.0384)		(0.0393)
Constant	987.4***	722.0***	187.3***	494.0***
	(114.1)	(103.1)	(30.45)	(67.71)
R^2	0.306		0.317	
R^2 (overall)		0.549		0.658

Notes: (1) The dependent variable is Offer_t. (2) Standard errors are clustered on individuals (in parentheses). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (3) Columns 1 and 3 are OLS regressions, Columns 2 and 4 are random effects panel regressions.

Figure 1.C.1: Constructed Price Sequence

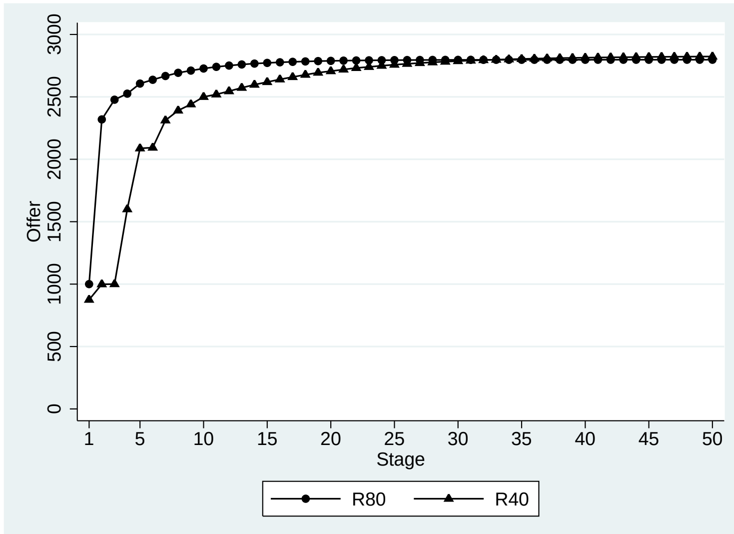
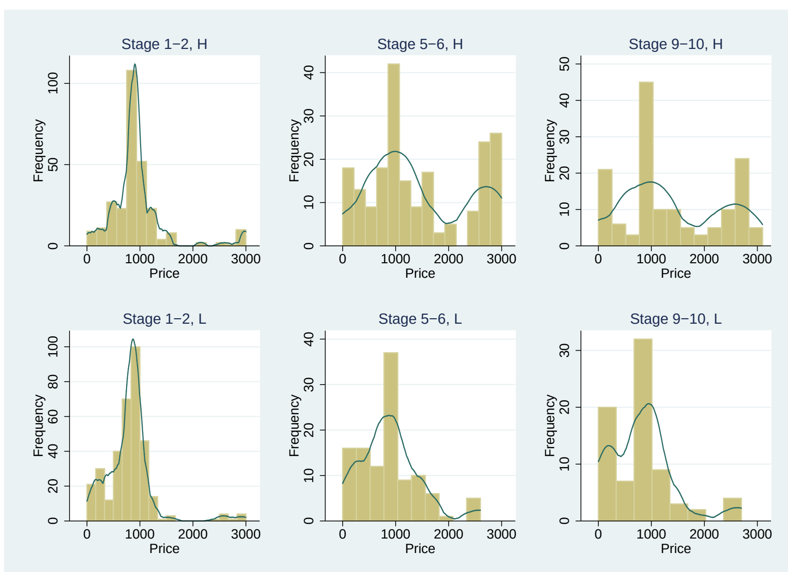


Figure 1.C.2: Price Distributions



Notes: (1) Histograms of price offers by quality and stage along with kernel density estimates. (2) Kolmogorov-Smirnov tests reject equality of offer distribution between qualities for stages 5-6 and 9-10 but not for stages 1-2.

Chapter 2

Meet the Lemons: How Cheap Talk Overcomes Adverse Selection in Decentralized Markets

2.1 Introduction

It is well-known that in the presence of incomplete information the price mechanism may fail to allocate goods optimally and markets may be inefficient due to adverse selection (Akerlof, 1970). However, when there are unrealized gains from trade, buyers and sellers have an incentive to find ways to capture this surplus. Indeed, the literature has been successful in identifying a wide range of institutional settings that alleviate the adverse consequences of information asymmetries. Examples include signalling devices such as warranties, ecolabels and building a brand name, and screening devices such as deductibles, aptitude tests and jobs with probationary periods. While these institutions successfully restore the functioning of markets, they also require agents to engage in *socially costly* activities.[1]

This article is concerned with an experimental test of a mechanism introduced in Kim

[1]For instance, there are significant costs associated with running assessment centers, including labor, physical space, and people's time. Similarly, labels have no economic value besides functioning as a signal to consumers.

(2012), which does not necessitate signalling or screening cost. Allowing for *free and non-binding communication* (cheap talk) suffices to substantially mitigate adverse selection. Communication is effective, because information is transmitted from the informed to the uninformed agents and this despite the strong incentives to misrepresent information that are usually associated with cheap talk. The fact that communication is costless and non-binding marks a stark difference to the other mentioned institutions. There is neither money-burning (e.g. aptitude tests) nor commitment (e.g. warranties).

It should be emphasized that communication is effective in a wide range of market settings. In fact, all that is required is that markets are decentralized to at least some extent in the sense that agents have some power in selecting potential trading partners. An implication of this will be that there are matching (or search) frictions: the possibility to trade is dependent on some agent of the other market side choosing you to be the receiver of the price offer.

To fix ideas, consider the following market in which an arbitrary number of buyers and sellers interact to exchange goods. Goods can be of two qualities, high or low. Each seller owns one unit of the good and is informed about its quality. Buyers are uninformed. Each buyer chooses a single seller to whom he makes a price offer to buy the good. It is possible that several buyers select the same seller and that some sellers do not receive any offer (matching friction). Finally, sellers accept at most one of their received offers. This matching technology has been employed in other contexts before (e.g. Satterthwaite and Shneyerov, 2007) and represents a decentralized version of Akerlof's original model.

Suppose we augment the market with an initial stage in which each seller announces a quality l (low quality) or h (high quality). Announcements are cheap talk, as sellers are free to send both messages at no cost. Buyers observe all messages before choosing a seller. Assume that if messages are uninformative or in absence of communication, high quality goods do not sell due to the information asymmetries. Interestingly, there is an equilibrium in which messages *do* transmit information. This equilibrium is characterized by endogenous market segmentation: a market in which only lemons sell (submarket l)

coexists alongside a market in which high quality goods are sold with positive probability (submarket h). Market segmentation is based on the observation that low quality sellers have an incentive to reveal their quality. Where does this incentive come from? For reasons familiar in the literature, we would expect low quality sellers to mimic high quality sellers whenever high prices are offered in submarket h. However, in our market buyers choose submarkets and in fact, they frequent the lemons submarket more often than submarket h. Low quality sellers thus trade off the opportunity to potentially extract high prices in submarket h against joining the lemons submarket where they tend to attract more buyers.

The reason buyers visit submarket l relatively more often than submarket h is the quality uncertainty in the latter. In submarket h, buyers either have to take the risk of making a high price offer to a low quality seller or, if low prices are offered, there is the possibility to be matched with a high quality seller who rejects the offer. In equilibrium, buyers are indifferent between the two submarkets and thus quality uncertainty is compensated for by less competition between buyers in submarket h. Of course, attractiveness of submarkets also depends on the potential gains from trade with low and high quality sellers.

Armstrong (2006) and Rochet and Tirole (2003) provide examples of two-sided markets where one group's benefit from joining a platform (or submarket) depends on the size of the other group that joins the same platform: for instance, if consumers are more likely to visit a mall where prices are generally lower, a retailer may be willing to locate in this mall even if doing so sends a negative signal about the quality of its products. Further examples of real-world institutions that seem to fit with the story of endogenous market segmentation are costless advertisement and markets where sellers post non-binding list prices such as used cars, housing and online posting sites. Naturally, different models are also in line with such institutions, for instance, Chen and Rosenthal (1996) interpret non-binding list prices as ceiling prices the seller commits to accept rather than cheap talk.

We report results from an experiment with a series of decentralized markets that puts endogenous market segmentation to a direct test and disentangles it from other potential explanations. To isolate the effect of market segmentation, we vary the availability and timing of messages. In the main treatment, messages come first and the described partially separating equilibrium exists. A priori it is, however, difficult to assess whether subjects will behave in the predicted way, as the partially separating equilibrium is quite demanding: a low quality seller is only willing to reveal her true quality if she expects that the low quality submarket is indeed heavily frequented by buyers and that high quality sellers will be truthful as well. This is further complicated by the fact that there are always pooling equilibria in which messages are uninformative. On the other hand, the market segmentation equilibrium is selected by a criterion called no incentive to separate (NITS) suggested in Chen et al. (2008). The results reported in this article will provide evidence in support of NITS.[2]

A rich experimental literature has established that private information is often communicated truthfully despite monetary incentives to lie. In these experiments, cheap talk is effective due to pro-social preferences, lie aversion or guilt. Important contributions include Gneezy (2005), Charness and Dufwenberg (2006), Vanberg (2008), Sutter (2009), and Charness and Dufwenberg (2011).[3] The approach taken in this article complements this literature by testing a mechanism in which communication alleviates adverse selection due to equilibrium incentives of pecuniary payoff maximization.[4] The challenge is to separate market segmentation from truth-telling due to non-standard preferences. To account for this, we conduct a control treatment in which the timing of messages and matching is reversed: buyers are matched to sellers first, and only then sellers send messages. Theoretically, market segmentation breaks down due to this change, because sellers cannot attract more buyers by revealing their quality. On the other hand, if the findings

[2]In the present setting, an equilibrium satisfies NITS if low quality sellers prefer the equilibrium outcome to credibly revealing their type, if they somehow could. Dickhaut et al. (1995), Blume et al. (2001) and De Groot Ruiz et al. (forthcoming) test different cheap talk equilibrium selection criteria.

[3]See also Valley et al. (1998), Valley et al. (2002), Croson et al. (2003), Lundquist et al. (2009), Charness and Dufwenberg (2010) and Erat and Gneezy (2012). Cai and Wang (2006) focus on bounded rationality as an explanation for “overcommunication”.

[4]Another important difference is that the present article explores markets, whereas the mentioned studies employ bilateral settings. Goeree and Zhang (2014) introduce competition to the model of Charness and Dufwenberg (2011). See also Cadsby et al. (1990) and Holt (1995).

were based on lie aversion and pro-social preferences, they should persist in the control treatment.

Strikingly, the experimental results closely follow the theoretical predictions of the separating equilibrium. In the main treatment, messages are informative, market segmentation can be observed frequently and rates of trade and welfare are high. Welfare is low in the control treatment mentioned above. In fact, average efficiency is not different from a treatment in which subjects do not have the possibility to communicate at all. This demonstrates that pro-social preferences and lie aversion cannot explain the success of communication in the main treatment. We also elicit a considerable degree of risk and loss aversion, but find that this does not undermine market segmentation (quite to the contrary!).

Finally, notice that market segmentation is not a coordination device in the sense that it improves the efficiency of the matching technology. In fact, the probability of high quality sellers to meet a buyer is lower in the main treatment than in the control treatments and the probability of low quality sellers to meet a buyer is identical across treatments. Hence, in the partially separating equilibrium there are *fewer* meetings between buyers and sellers in theory and this is fully reflected in the experimental data. Market segmentation works through reducing information asymmetries, not through more efficient matching.[5]

The remainder of the article is organized as follows. The next section introduces the model and characterizes equilibrium. Section 3 presents the example used in the experiment. Section 4 presents the experimental design. The experimental results are reported in Section 5, including a discussion of the model in the context of cost of lying and risk / loss aversion. Section 6 concludes.

[5]Crawford (1998) reviews a small body of experiments in which cheap talk reduces information asymmetry in bilateral settings. These models assume that agents' preferences overlap to some extent. In the present model, cheap talk only becomes effective in markets, i.e. if there is more than one seller and one buyer.

2.2 Preliminaries

Model. The model presented in the following is based on Kim (2012).[6] There are n_B buyers and n_S sellers interacting in a market for an indivisible good. Each seller can sell at most one unit and each buyer wants to buy at most one unit of the good. Goods are available in two qualities. There are n_H sellers that can sell a high (H) quality good and n_L sellers that can sell a low (L) quality good. Note that $n_H + n_L = n_S$. A seller of type $\theta = \{L, H\}$ has cost c_θ to produce a good of quality θ. A good of quality θ yields a value of v_θ to the buyer. There are gains from trade for both qualities, i.e., $v_\theta > c_\theta$ for $\theta = \{L, H\}$.

Denote the fraction of low quality sellers by $\hat{q} = \frac{n_L}{n_S}$. The focus is on markets in which adverse selection is severe: high quality goods do not trade in a pooling equilibrium. This is ensured by the assumption that the buyers' expected value for the good falls short of the high quality sellers' cost.[7]

$$\hat{q}\, v_L + (1 - \hat{q})\, v_H < c_H \tag{2.1}$$

The trading process is as follows. First, sellers simultaneously send messages $m \in \{l, h\}$.[8] Messages are cheap talk as they are sent without any direct costs. We will say that sellers who sent message l are in submarket l and sellers who sent message h are in submarket h. Second, each buyer observes the two submarkets, i.e., he learns how many sellers sent message l and h. Each buyer then chooses a seller to whom he makes a take-it-or-leave-it offer. Several buyers may select the same seller. This also implies that some sellers may not be selected by any buyer. Offers are made simultaneously and thus buyers do not observe how many competitors are making an offer to the same seller.

[6] There are several differences to Kim (2012). In order to implement the model in the laboratory, we cannot rely on a continuum of buyers and sellers. Another difference is that in our case the number of buyers in the market is fixed and buyers have no entry cost.

[7] Inequality (2.1) is sufficient but not necessary to prevent trade with high quality goods in the pooled market. As will be shown presently, the trading process implies only imperfect competition and thus, buyers may prefer to offer low prices even if their expected profit from offering high is positive.

[8] Richer message spaces are conceivable, for instance announcing non-binding selling prices. Binary messages are without loss of generality if there are only two qualities.

Third, each seller who receives at least one offer decides whether to accept or reject the offer(s). At most one offer can be accepted. A buyer whose offer p is accepted earns $v_\theta - p$ if the quality of the good is θ. A seller of type θ who accepts a price p earns $p - c_\theta$. Buyers and sellers who do not trade earn 0. All of the above is common knowledge.

Buyers can distinguish sellers only on the basis of messages. Thus, each buyer effectively chooses a submarket l or h. Let us describe a submarket by $S^m_{i,j}$, where $m = \{l, h\}$, i is the number of low and j the number of high quality sellers in the submarket. The fraction of low quality sellers in $S^m_{i,j}$ is denoted by $q(S^m_{i,j})$. Let $\beta(S^m_{i,j})$ be the probability that a buyer joins submarket $S^m_{i,j}$. Let S be the set of possible submarkets. Buyers' bidding strategies are described by a cumulative distribution function $F : \Re_+ \times S \to [0, 1]$ where $F(p, S^m_{i,j})$ is the probability that a buyer offers a price not larger than p to a seller in submarket $S^m_{i,j}$.

Equilibrium Characterization. Henceforth, a market equilibrium refers to the standard notion of sequential equilibrium of the model introduced above. A market equilibrium is thus characterized by a situation in which sellers send messages that maximize their expected payoffs and accept the highest price offer that exceeds their reservation cost. Buyers' choice of submarkets and price offers is optimal given their beliefs about the fraction of low and high quality sellers in both submarkets.

The focus is on a symmetric partially separating equilibrium. In this market equilibrium submarket l consists only of low quality sellers and submarket h contains all high quality sellers and possibly some low quality sellers. Sellers' behavior is thus fully described by the number of low quality sellers who send message l and we can refer to submarkets as S^m_i. Let α denote the probability that a low quality seller reveals his quality.[9] Under a mild condition that requires a minimal degree of competition, low quality sellers have an incentive to reveal their quality with positive probability.[10]

Proposition 3. *There exists a (partially) separating market equilibrium with $\alpha > 0$.*

[9]There may be multiple partially separating equilibria. However, in all of them there is a lemons submarket consisting only of low quality sellers. We refer to Kim (2012) for a discussion.

[10]The condition is $q(S^h_{n_L}) - q(S^h_{n_L-1}) \leq \frac{v_L - c_L}{v_H - v_L}$, see Appendix B equation (2.4). Note that with a continuum of agents, this condition always holds; the left-hand side reduces to 0. Hence, the condition requires the market to be sufficiently thick. The condition is only required if $v_L - c_L < v_H - c_H$.

The proof is relegated to Appendix B. In the introduction, we have already discussed the intuition for the result. The main point is that low quality sellers can attract more buyers in submarket l, which compensates for the forgone opportunity to extract high prices in submarket h. The example presented in the next section will provide a comprehensive picture of the incentives at work.

Let the game described above be denoted by $\Gamma^{\text{C-Sep}}$, where C-Sep stands for communication-separating. The following observation will turn out to be important for the experimental predictions. Consider a variant of $\Gamma^{\text{C-Sep}}$ where each buyer *first* chooses the seller he wants to make an offer to and only then sellers send messages. Buyers still observe all messages and make an offer to their seller. As before, sellers accept or reject offers in the last step. Call this game $\Gamma^{\text{C-Pool I}}$. A third variant of the game, $\Gamma^{\text{C-Pool II}}$, is identical to $\Gamma^{\text{C-Pool I}}$ except that buyers only observe the message sent by the seller they are matched with. Finally, Γ^{NC} refers to the game in which sellers cannot send messages.

Observation 1. *All equilibria in $\Gamma^{C\text{-}Pool\ I}$, $\Gamma^{C\text{-}Pool\ II}$ and Γ^{NC} are pooling, i.e. price offers are strictly below v_L and high quality sellers never trade.*

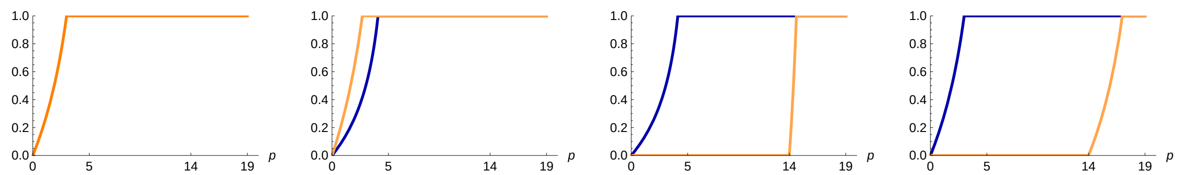
Observation 1 states that low quality sellers do not reveal their quality, if buyers cannot choose sellers conditional on observed messages. Inequality (2.1) then ensures that high quality goods are not traded. The finite number of agents again requires a mild condition that guarantees a minimal incentive for low quality sellers to misrepresent their type.[11] Notice that in $\Gamma^{\text{C-Sep}}$ there also exist "babbling" equilibria in which messages do not carry information.

2.3 A Simple Example

The following example was implemented in the experiment. Consider a market with 6 buyers and 6 sellers. There are 3 low quality sellers and 3 high quality sellers. Parameters are given by $v_H = 19$, $c_H = 14$, $v_L = 5$ and $c_L = 0$. Hence, surplus from trade is equal

[11] We need to assume that in the submarket consisting of all high quality sellers and a single low quality seller, prices that exceed c_H are offered with positive probability. See Appendix B equation (2.8).

Figure 2-1: Buyers' Bidding Strategies



(a) $\{S_0^l,\ S_3^h\}$ (b) $\{S_1^l,\ S_2^h\}$ (c) $\{S_2^l,\ S_1^h\}$ (d) $\{S_3^l,\ S_0^h\}$

The figure depicts the theoretical CDF of buyers' offers for the four market structures that are observed with positive probability in equilibrium. Submarket l is shown in blue and submarket h in orange. The corresponding probabilities of buyers to join submarket l are given by $\beta(S_0^l) = 0.00$, $\beta(S_1^l) = 0.29$, $\beta(S_2^l) = 0.59$, $\beta(S_3^l) = 0.50$.

to 5 for both qualities. Moreover, the expected value for buyers in the pooled market is 12 and falls short of the high quality sellers' cost. Without market segmentation, high quality goods do not trade.

In the partially separating equilibrium, buyers observe 4 possible pairs of submarkets: the pooled market $\{S_0^l,\ S_3^h\}$, the intermediate cases $\{S_1^l,\ S_2^h\}$ and $\{S_2^l,\ S_1^h\}$, and the completely separated market $\{S_3^l,\ S_0^h\}$. A pair of submarkets will also be referred to as market structure.

Figure 2-1 shows equilibrium bidding by means of the cumulative distribution of price offers.[12] Figure 2-1a depicts the pooled market. Here, $q(S_3^h) = \hat{q} = 1/2$ and buyers offer low prices ranging between 0 and 3. This is a situation where adverse selection leads to large inefficiencies, as high quality goods never trade. The same applies to the partially separated market structure $\{S_1^l, S_2^h\}$ shown in Figure 2-1b. We have $q(S_2^h) = 2/5$, which implies that the buyers' expected value still falls short of the high quality sellers' cost. In contrast, for the partially separated market $\{S_2^l, S_1^h\}$ (Figure 2-1c) and the completely separated market (Figure 2-1d), offers in submarket h exceed $c_H = 14$. Obviously, in all lemons submarkets buyers' price offers do not exceed v_L.

It can be shown that $\beta(S_1^l) = 0.29$ and $\beta(S_2^l) = 0.59$. Thus, in equilibrium buyers are indifferent between visiting either submarket. Moreover, the expected fraction of buyers to sellers is 1.74 vs. 0.85 in $\{S_1^l, S_2^h\}$ and 1.77 vs. 0.62 in $\{S_2^l, S_1^h\}$. The weaker competition

[12]The derivation of price offers and all other predictions follows from the proof of Proposition 3.

Table 2.1: Theoretical Predictions

		Rates of Trade		Ex Ante Efficiency			Payoffs		
	α	L	H	Total	L	H	U_B	U_L	U_H
C-Sep	0.48	0.7	0.25	14.26	10.57	3.69	1.04	2.46	0.21
C-Pool I, II / NC	0.00	0.67	0.00	9.98	9.98	0.00	1.00	1.32	0.00

between buyers in submarkets h compensates for the quality uncertainty.

Anticipating the buyers' participation and bidding decisions, let $U_L(S_i^m)$ be a low quality seller's expected payoff conditional on being in submarket S_i^m. We have $\{U_L(S_i^l)\}_{i=1}^3 = (2.75, 2.84, 1.32)$ and $\{U_L(S_i^h)\}_{i=1}^3 = (6.87, 1.03, 1.32)$. A low quality seller prefers the lemons submarket S_1^l over the pooled market S_3^h. Thus, $\alpha = 0$ is no equilibrium, because a low quality seller can unilaterally move to S_1^l. However, the market position that is by far the most attractive one is to be the only low quality seller in S_1^h. The reason is the potentially high benefit from high selling prices. Hence, $\alpha = 1$ is no equilibrium, because unilaterally moving to the high quality submarket (thereby making it a mixed quality submarket) is profitable.

What messages do sellers send? We already know that $\alpha \in (0, 1)$. Thus, α needs to be such that low quality sellers are indifferent between sending message l or h. The equilibrium is characterized by a situation in which low quality sellers' gain from their information advantage in submarket h equals the benefit from the improved competitive position in submarket l. Using equation (2.7) in Appendix B yields $\alpha = 0.48$.

Table 2.1 summarizes the theoretical predictions of the key outcome variables. C-Sep refers to the main treatment that implements $\Gamma^{\text{C-Sep}}$. The C-Pool and NC treatments represent the different control treatments corresponding to the games $\Gamma^{\text{C-Pool I}}$, $\Gamma^{\text{C-Pool II}}$ and Γ^{NC}. As implied by Observation 1, the theoretical predictions are the same for all control treatments.

Endogenous market segmentation through cheap talk (C-Sep) significantly increases rates of trade and efficiency compared to a setting without cheap talk (NC) or with cheap talk but without the possibility to choose sellers based on messages (C-Pool I, II). A remarkable finding is that cheap talk leads to trades with high quality sellers without undermining trades with low quality sellers. Table 2.1 also shows expected payoffs of buyers (U_B), low (U_L) and high (U_H) quality sellers. In C-Sep, payoffs increase for all agents relative to the pooling equilibrium, i.e., market segmentation constitutes a Pareto improvement.

If all sellers trade, a total welfare of 30 could be achieved. However, it is important to note that first-best efficiency is not the appropriate benchmark. Due to the frictions of the matching process, the first-best outcome is not attainable even with complete information. If two buyers meet the same seller, this immediately implies that another seller will not trade. In fact, we cannot go beyond expected trading rates of 67 percent for both types of sellers simultaneously. The benchmark where trade occurs in all meetings leads to an expected welfare of 19.95. Therefore, an ex ante efficiency of 14.26 constitutes a substantial improvement over the pooled market.

2.4 Experimental Design

The experiment was run in December 2013 and January 2014 at the experimental laboratory of the University of Bern. 216 students mainly from business administration and economics took part in the experiment. Each session was composed of 12 participants. 18 sessions were run, using the z-Tree software developed by Fischbacher (2007). Sessions lasted between 50 and 80 minutes and average earnings were 32 CHF including a show-up fee of 14 CHF.[13] The conversion rate was 0.6 CHF per experimental point.

We ran 4 treatments summarized in Table 2.2. The main treatment "Communication-Separating" (C-Sep) implements the example presented in the previous section for $\Gamma^{\text{C-Sep}}$. In the experiment, buyers did not choose a specific seller. Instead, buyers observed the number of l and h messages and then decided in which of the two submarkets to make their offer. The specific seller was then randomly selected by the computer and this was commonly known. Random matching within submarkets avoids potential difficulties with

[13] At the time, 1 US Dollar corresponded roughly to 0.91 CHF.

Table 2.2: Experimental Design

Treatment	Sessions	Subjects	Messages	Matching
C-Sep	6	72	Observed by all buyers	Buyers choose submarket*
C-Pool I	4	48	Observed by all buyers	Random**
C-Pool II	4	48	Observed by matched buyer	Random
NC	4	48	No messages	Random

* Buyers choose a submarket (l or h) and are randomly matched to a seller in this submarket.
** Buyers are randomly matched to one of the 6 sellers.

buyers choosing sellers based on how the choice is presented to them, e.g. the seller who is displayed on the left hand side of the screen might be selected most often.[14]

Treatment "No Communication" (NC) is implemented as a useful benchmark. In NC sellers cannot send messages to buyers. Buyers right away make offers to a randomly assigned seller in the pooled market. Theory predicts buyers to offer only prices below v_L and high quality goods never trade. The matching procedure was carefully explained to all subjects. In addition, in each period it was explicitly mentioned that everybody has now been randomly matched. This is important, since even though there is no matching decision to take, it is as important as in C-Sep for buyers to form an expectation about the number of competitors offering to the same seller.

In the light of the experimental literature on cheap talk and hidden information, differences in behavior between C-Sep and NC could also stem from subjects' preferences to tell the truth or from fairness concerns. To control for this, we implement treatments with cheap talk, but in which all equilibria are pooling. In these treatments, called "Communication-Pooling I" (C-Pool I) and "Communication-Pooling II" (C-Pool II), buyers are randomly matched to sellers *before* they send messages. The message is then either observed by all buyers (C-Pool I) or only by the buyer the seller is matched with (C-Pool II). Thus, buyers still observe messages, but they cannot choose submarkets. In absence of social preferences, messages cannot credibly transmit information in C-Pool I and C-Pool II and the theoretical predictions coincide with the ones for NC (see Observation 1). On

[14] Recall that the same seller can meet several buyers and thus the random draws of sellers are with replacement. A further advantage of random matching within submarkets is that potential considerations of a seller to reward a buyer for selecting her as the particular seller to interact with are extenuated.

the other hand, if sellers are lie averse or have pro-social preferences, messages may still be informative.

C-Pool I provides the cleanest control for C-Sep, since the only difference is the reversal in the timing of the message and matching stage. C-Pool II was introduced to give lie aversion its best shot. If buyers observe all messages, they would often see message distributions inconsistent with truth-telling (whenever there are not 3 l and 3 h messages). Buyers may then conclude not to believe the messages at all. If only one message is observed, attempts at truth-telling by some sellers cannot be frustrated as easily.

The instructions for C-Sep are provided in Appendix A. After reading the instructions every subject had to fill out a set of control questions. A brief verbal summary of the setting was given to ensure common knowledge. Subjects were then randomly assigned to be one of the 6 buyers or one of the 6 sellers. Roles were fixed throughout the experiment. Subjects played 20 periods. In each period, there were 3 H and 3 L-type sellers. Sellers' types randomly changed from one period to the next. Each seller was informed about his type at the beginning of each period. Buyers were uninformed, they only knew that there are 3 H and 3 L-type sellers. Interactions were anonymous and there were no identifiers that would allow subjects to know or guess with whom they interact in different periods.

Upon completion of the 20 periods, subjects that were assigned the role of the seller completed a short task that aims to measure lie aversion. We used a design similar to that in Gneezy (2005). Since buyers potentially suffer from large losses when offering high prices, information on subjects' risk / loss aversion was also gathered. Subjects knew that there would be two additional parts, but no details were explained to them until the previous parts had been completed. We defer a description of the lie and risk / loss aversion tasks.

2.5 Results

The discussion of the experimental results is organized around three questions. *(1) Do we observe endogenous market segmentation? (2) If market segmentation is observed, does* *it increase rates of trade and efficiency?* We would also like to understand whether the results are based on the proposed mechanism or if and to what extent truth-telling is due to other-regarding preferences and lie aversion. Thus, *(3) are the results driven by non-standard preferences?*

In the following, only data from periods 11-20 are used.[15] All non-parametric statistical tests are based on session averages as the unit of observation. Moreover, market structures have so far been denoted by $\{S_i^l, S_{n_L-i}^h\}$ where the subscripts indicate the number of low quality sellers in a submarket. In the experiment high quality sellers may sometimes send message l. The market structure is therefore denoted by, for instance, 2 l / 4 h, indicating that 2 sellers sent message l and 4 sellers sent message h. As will be shown, most high quality sellers send message h and thus 2 l / 4 h is usually equivalent to $\{S_2^l, S_1^h\}$.

2.5.1 Market Segmentation, Rates of Trade and Efficiency

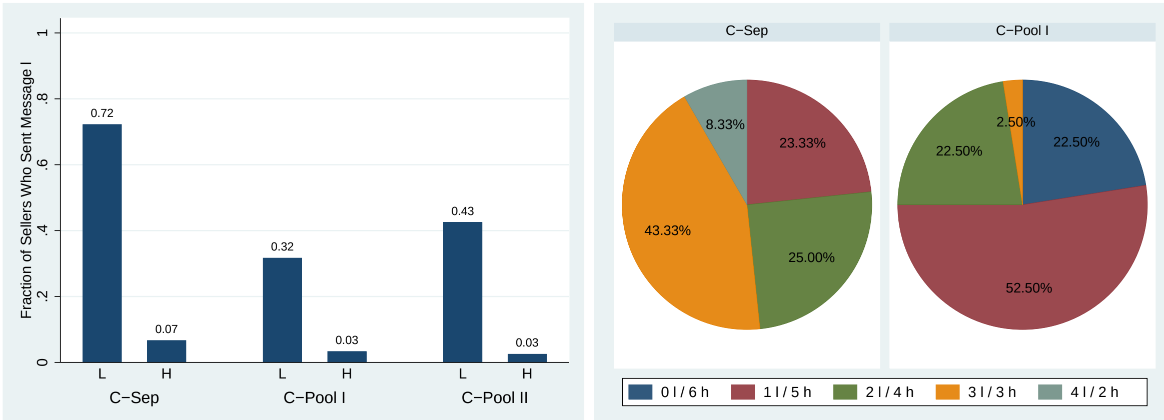
The experimental results provide clear evidence of endogenous market segmentation in C-Sep. Our discussion will be in support of the following result.

Result 8 (**Endogenous Market Segmentation**). *Behavior in C-Sep is consistent with endogenous market segmentation. Messages are informative and frequently induce market structures that permit trade with high quality sellers. Low quality sellers are willing to forgo high prices in submarket h, because by revealing their quality they on average attract twice as many offers.*

Figure 2-2a shows that messages are a good predictor of a seller's true type. A first important observation is that high quality sellers almost always send message h (in 93 percent of the cases in C-Sep). While this seems intuitive, it is also immensely important, because it allows buyers to meaningfully interpret low quality sellers' behavior. The figure further shows that low quality sellers reveal their quality in 72 percent of the cases in treatment C-Sep and in 32 and 43 percent of the cases in treatment C-Pool I and II,

[15] All qualitative results hold in an analysis that includes all periods. The discussion on rates of trade will illustrate that differences between C-Sep and the other treatments become more substantive in later periods.

Figure 2-2: Sellers' Messages and Market Segmentation



(a) Messages

(b) Market Segmentation

Figure (a) depicts the fraction of messages l among all messages sent by sellers separated by treatment and seller type. Figure (b) shows the distribution of realized market structures.

respectively.[16] Wilcoxon-Mann-Whitney (WMW) tests confirm that low quality sellers are significantly more likely to send message l in treatment C-Sep than in the C-Pool treatments (p=0.01 for both comparisons). The difference between C-Pool I and II is not significant (p=0.19). Moreover, low quality sellers' probability to reveal their type in C-Sep is significantly higher than the theoretically predicted 48 percent, according to a Wilcoxon matched-pair signed-rank (henceforth, Wilcoxon) test (p=0.03).

Buyers observe a wide range of different submarkets in C-Sep as well as C-Pool I. Figure 2-2b shows the frequency of the different market structures. The most common market structure in C-Sep is 3 l / 3 h, observed in more than 43 percent of the cases. In 85 percent of the cases this market structure corresponds to the completely separated market, i.e. all low quality sellers send message l and all high quality sellers send message h.[17] In contrast, in C-Pool I the most prominent set of messages is 1 l / 5 h (52.5 percent) and complete separation is almost never observed. Note that in 2 l / 4 h high quality goods are

[16] It is interesting to note that partial information revelation in C-Sep is not only the result of aggregating sellers. Using the 34 (out of 36) sellers who played the role of the low quality seller at least 3 times in periods 11-20, it turns out that around one third of the low quality sellers revealed their quality almost always, 44 percent revealed their quality around 70 percent of the time and the remainder sent message l in less than 50 percent of the cases.

[17] In the remaining 15 percent, submarket l contains two low quality and one high quality seller.

also expected to be traded. Overall, the observed market structure theoretically allows for trade with high quality sellers in 68 percent of the cases in C-Sep and in 25 percent of the cases in C-Pool I (assuming messages are informative). In C-Pool II the probability that a seller who sends message h is indeed of the high quality is $0.97/(0.97+0.57) = 0.63$. This translates into an expected value of 13.82, falling short of high type sellers' reservation cost.

Low quality sellers' incentive to reveal their quality in C-Sep stems from their ability to attract more buyers.[18] Simple calculations indeed reveal that in C-Sep low quality sellers receive on average 1.47 offers when sending message l and 0.74 offers when sending message h (Wilcoxon test p=0.03). More specifically, Figure 2-3a shows the buyers' decisions to enter submarket l or h for each market structure (blue) and the corresponding theoretical predictions (red). In the completely separated market structure, buyers distribute almost evenly among the two submarkets. The difference to the theoretical prediction of $\beta(S_3^l) = 0.50$ is not significant (Wilcoxon test p=0.43). This is remarkable, because buyers do not seem to fear losses in 3 l / 3 h and consider the two submarkets as equally attractive. For the other market structures, buyers are biased toward submarket l even more than theoretically expected.[19]

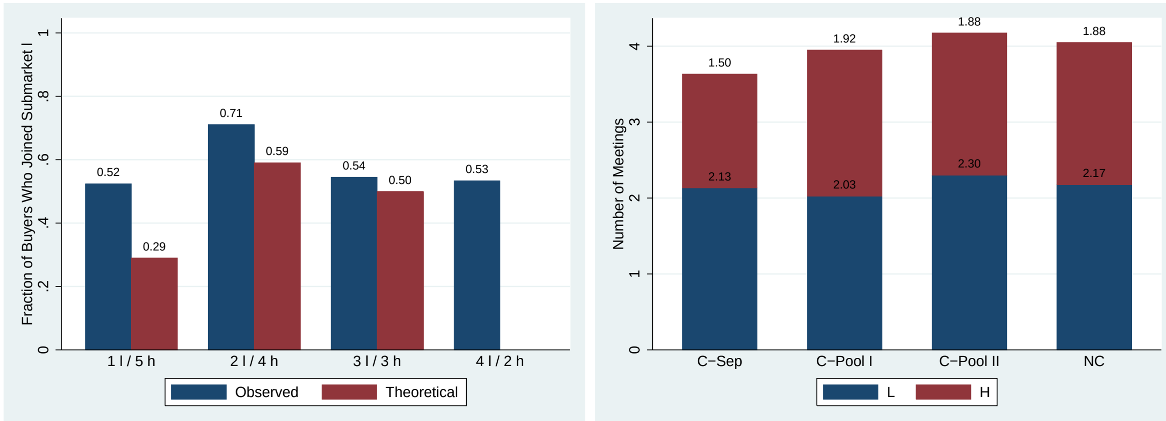
Let us sidestep a potential pitfall. It is tempting to think of the market segmentation mechanisms implemented in C-Sep as a coordination device in the sense that matching becomes more efficient. However, the opposite is true: the buyers' possibility to choose between submarkets introduces a distortion. Buyers enter the lemons submarket with a larger probability than what would be optimal in terms of matching. Figure 2-3b shows that the average number of sellers that receive at least one offer is around 4 for treatments C-Pool I, II and NC and a little lower for C-Sep. In other words, on average 2 sellers do not receive an offer. It can be seen that the number of meetings for low quality sellers is stable across treatments.[20] On the other hand, high quality sellers encounter significantly fewer

[18] Another explanation might be lie aversion. But notice that lie aversion would apply equally well to the C-Pool treatments. A discussion of lie aversion can be found in Section 5.2.

[19] We show in Section 5.2 that this can be explained by risk or loss aversion.

[20] WMW tests show that the number of meetings of low quality sellers does not differ between C-Sep and the other treatments.

Figure 2-3: Participation and Matching



(a) Buyers' Participation Decision in C-Sep

(b) Matching

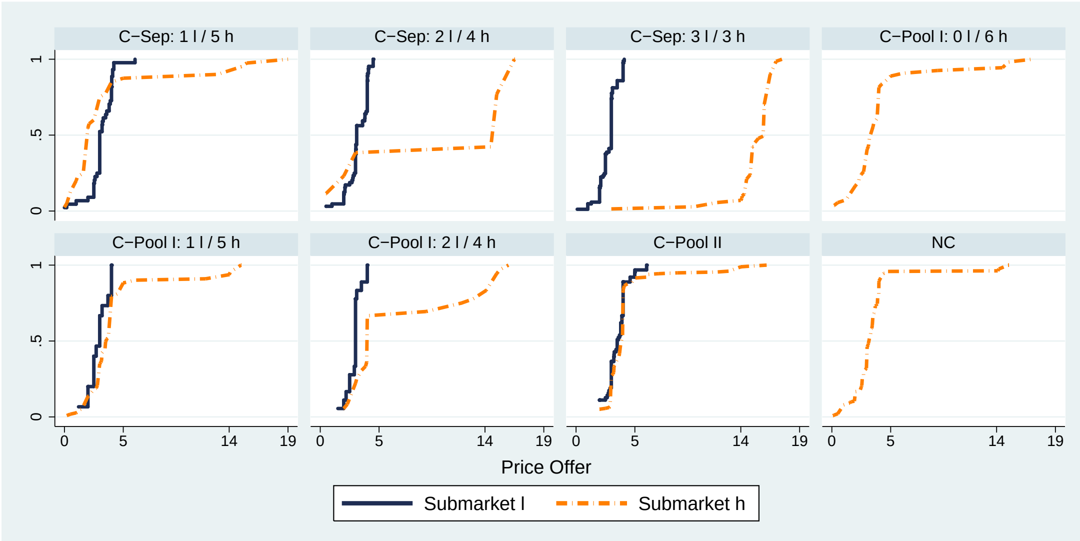
Figure (a) depicts the fraction of buyers who joined submarket l for each market structure (blue) as well as the corresponding theoretical predictions (red). Figure (b) shows the average per period number of sellers who meet at least one buyer separated by L and H-type sellers.

meetings in C-Sep than in the other treatments (WMW p $\leq$ 0.06 for all comparisons). Hence, market segmentation negatively affects the number of meetings of high quality sellers, but, as shown next, many of these meetings do not suffer from adverse selection anymore.

Figure 2-4 depicts the cumulative empirical distribution of buyers' offers for each frequently observed market structure in C-Sep and C-Pool I, for messages l and h in C-Pool II and for NC. Offers in submarket l are represented in blue (solid) and offers in submarket h in orange (dashed). In accordance with theory, in all lemons submarkets of all 4 treatments only offers are below $v_L = 5$. Moreover, price offers in C-Pool II are very similar for both messages. For treatment C-Pool I about one fourth of the prices offered in submarket h of 2 l / 4 h cover the high type sellers' production cost of 14. For other market structures in C-Pool I, price offers were low and only allow for trade with low quality sellers.

For C-Sep, theory predicts high price offers for some market structures. Indeed, in submarket h of market structure 3 l / 3 h almost all offers exceed the high type sellers' cost, and in submarket h of 2 l / 4 h 62 percent of the offers are directed at high quality

Figure 2-4: Cumulative Distribution of Buyers' Offers



The figure depicts the empirical cumulative distribution of offers for submarkets l and h by treatment and observed market structure.

sellers. In both market structures offers in submarket h are significantly larger than offers in submarket l (Wilcoxon test p=0.03 and p=0.04, respectively). Buyers correctly believe that they are likely to meet a high quality seller when joining submarket h.[21] We conclude that buyers' participation and bidding behavior reflects the informational content of the messages well.

Result 8 hints that C-Sep is successful in facilitating trade of high quality goods compared to the control treatments. The next result shows that this is indeed observed in the data.

Result 9 (**Rates of Trade and Efficiency**). *Rates of trade and efficiency in C-Sep are not significantly different from the theoretical predictions. More importantly, the rate of trade with high quality sellers is significantly larger in C-Sep than in C-Pool I, II and NC. As a result, total efficiency is by far the highest in C-Sep.*

Table 2.3 presents observed rates of trade with the theoretical predictions given in brackets. The trade frequency of high quality sellers is negligible for treatments C-Pool

[21]It is interesting to observe that, as predicted in Figure 2-1b, in 1 l / 5 h and 2 l / 4 h competition for low quality sellers is stronger in submarket l than in submarket h, as prices targeted at low quality sellers are higher in the lemons submarket.

Table 2.3: Rates of Trade and Efficiency

	Rate of Trade		Efficiency		
	L	H	L	H	Total
C-Sep	0.70 (0.70)	0.32 (0.25)	10.50 (10.57)	4.83 (3.69)	15.33 (14.26)
C-Pool I	0.67 (0.67)	0.09 (0.00)	10.00 (9.98)	1.37 (0)	11.37 (9.98)
C-Pool II	0.77 (0.67)	0.03 (0.00)	11.75 (9.98)	0.38 (0)	11.88 (9.98)
NC	0.72 (0.67)	0.04 (0.00)	10.75 (9.98)	0.62 (0)	11.37 (9.98)

Efficiency is given by the average per period surplus generated with each seller type. Total efficiency is the sum over both types. Theoretical predictions are given in parentheses.

II and NC, 9 percent for C-Pool I and 32 percent for treatment C-Sep. WMW tests confirm that the trade frequency for H-type sellers is significantly larger in C-Sep than in all other treatments (p=0.01 for all comparisons). Moreover, the trade frequency in C-Sep with high quality sellers is larger than the predicted 25 percent, but this difference is not significant (Wilcoxon test p=0.11). Trade frequencies with H-type sellers are not different between C-Pool I, C-Pool II and NC (p>0.21 for all comparisons). The trade frequency with low quality sellers is around 70 percent for all treatments and differences are insignificant except that low quality sellers trade more often in C-Pool II than in C-Sep (p=0.08) and C-Pool I (p=0.04). Recall that the matching process does not allow surpassing average rates of trade of 0.67 for low and high quality sellers simultaneously. The observed trade frequencies in C-Sep of 70 percent for low quality and 32 percent for high quality sellers should thus be considered to be relatively high.

Table 2.3 also lists generated surplus for all treatments. Total efficiency in C-Sep is significantly larger than in all other treatments (WMW p=0.01 for all comparisons). Total efficiency does not differ between treatments C-Pool I, C-Pool II and NC (p>0.37 for all comparisons). Recall that trade failures are the only source of inefficiency in our setting. Hence, the observations on rates of trade immediately imply that realized surplus with high quality sellers is significantly larger in C-Sep than all other treatments and moreover, realized surplus with low quality sellers is either not different or lower than in the control treatments. The higher total efficiency in C-Sep compared to the control treatments is thus exclusively due to higher rates of trade with high quality sellers.

Total efficiency in C-Sep is not significantly different from the theoretical prediction (Wilcoxon test p=0.17) and, not surprisingly, welfare falls short of constrained efficiency (19.95) in all treatments (Wilcoxon test p=0.03 for C-Sep). Inefficiencies due to asymmetric information are not fully eliminated.

Our third main question is whether the experimental results can be explained by non-standard preferences. If messages by themselves were sufficient to induce trade with high quality sellers, the market segmentation mechanism would be of less interest. The comparisons between C-Sep and the C-Pool treatments discussed so far provide an immediate answer.

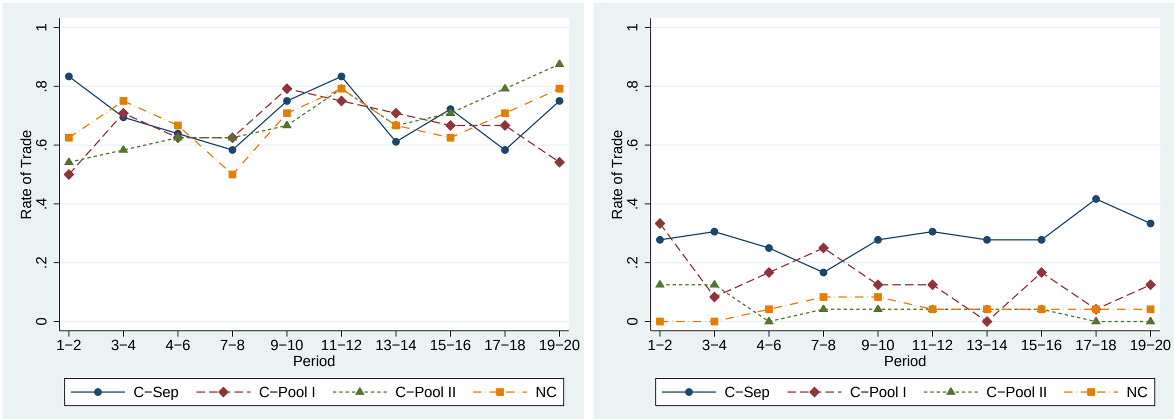
Result 10 (Non-Standard Preferences). *Non-standard preferences cannot explain the high efficiency in C-Sep.*

It has been shown that communication only makes a difference if sellers can use it to attract more buyers. If this is not the case, as in C-Pool I and II, total efficiency is not different from the setting without communication (NC) in which adverse selection is strong. This observation highlights that the timing of the message and matching stages is crucial, i.e. the buyers' possibility to choose sellers conditional on observed messages. Stated differently, comparing C-Sep and C-Pool I shows that irrespective of the type of non-standard preferences that characterize our subjects, the market for high quality goods breaks down when switching off the monetary incentives that lead to endogenous market segmentation.

C-Pool II is closer to the setting usually analyzed in the literature on cheap-talk and hidden information insofar as every buyer only observes one message. In contrast to that literature, messages do not trigger trade with high quality sellers. Subjects may still be lie averse, but the cost of lying seem to be too small to induce truth-telling. In other words, lies, if believed, are too lucrative.[22]

[22] In Appendix C it is shown that the threshold for truth-telling corresponds to a fixed cost of lying of 9.31, almost double the surplus generated by trading the good. Another explanation is that competition may lower the impact of communication and vice versa (Goeree and Zhang, 2014).

Figure 2-5: Rates of Trade



(a) Low Quality

(b) High Quality

The figure depicts the evolution of rates of trade over the 20 periods separated by low and high quality sellers. For clearer presentation, averages are taken over 2 consecutive periods.

Non-standard preferences could still act as a catalyst for market segmentation. In fact, this could explain why market segmentation seems to work better than expected. Recall that low quality sellers reveal their type more often and market structure 3 l / 3 h is more common than predicted. This is reinforced in Figure 2-5, depicting average rates of trade over the 20 periods for low and high quality sellers, respectively. The difference in rates of trade with high quality sellers between C-Sep and its control treatments becomes more pronounced in later periods.

We close this section by noting that truth-telling in C-Sep is not triggered by repeated interaction, even though the market consisted of the same 12 subjects in all periods. First, building up a personal reputation was impossible, as specific buyers and sellers could not be identified and moreover, matching was random to at least some extent. Second, if sellers' behavior had been driven by such considerations, we would expect the same to happen in C-Pool I. Finally, the absence of an end game effect in Figure 2-5 is a clear indication that truth-telling was optimal within a single period.

2.5.2 Lies, Risk and Losses

Market Segmentation under Non-Standard Preferences. In this section, we explore some implications of non-standard preferences. We focus on lie aversion, risk aversion and loss aversion. Lie aversion is an obvious candidate. Sellers may genuinely dislike lying or feel guilt when letting down buyers' expectations. In a setting of adverse selection, risk and loss aversion also seem to be of first-order importance.

For our discussion, the specifics of how to model lie, risk and loss aversion are unimportant. For concreteness, we briefly mention possible models. As in Ellingsen and Johannesson (2004), lie aversion is represented by a fixed cost subtracted from an agent's utility whenever she sends a message that does not correspond to her type. Note that lie aversion is only relevant for low quality sellers, who now earn $p - c_L - \kappa$ when sending message h, where κ is the fixed cost of lying. We use isoelastic utility with risk parameter η to model constant relative risk aversion. Finally, loss aversion captures the perception that changes in payoffs below a certain reference point have a stronger impact on utility than changes in payoff above this point. The natural reference point is the no trade outcome. Loss aversion is only relevant for buyers. We assume constant loss aversion as in Tversky and Kahneman (1991), i.e. a buyer's utility is $v_\theta - p$ if $v_\theta \geq p$ and $\mu(v_\theta - p)$ otherwise, where $\mu \geq 1$ is the loss aversion parameter and $\theta = \{L, H\}$.

Observation 2. *The probability α that a low quality seller sends message l is increasing in lie aversion (κ), risk aversion (η), and loss aversion (μ).*

We omit a formal discussion, but the intuition for the result is straightforward. Lie aversion has a direct negative effect on payoffs when misrepresenting ones type, ceteris paribus α increases in κ. For loss aversion, note that as μ increases, potential losses in submarket h receive more weight in the buyers' calculations. Loss averse buyers are therefore more likely to join submarket l. Anticipating this, submarket l becomes more attractive for sellers as well. The same argument holds for risk averse buyers, but in addition the effect is amplified by risk averse low quality sellers who value the higher probability to meet a buyer in submarket l (less risky option) relatively more than the possibility to extract high prices in submarket h (risky option).[23] Observation 2 reinforces the mechanism's relevance as a means to alleviate adverse selection. However, the effect of risk and loss aversion on efficiency is in general ambiguous. The reason is that buyers become less willing to offer high prices in the mixed quality submarket.[24]

Behavioral Measures. Following the market experiment, subjects completed a lie aversion task. The task is a variant of Gneezy (2005) and allows to categorize subjects on two dimensions, whether or not they are lie averse and whether or not they are other-regarding. Appendix C explains the task and the classification in detail. It also contains Table 2.C.1, which presents random effects regressions exploring the relation between low quality sellers' messages and being categorized as a truth-teller or liar and as other-regarding or selfish. We find no significant impact of lie aversion in C-Sep. On the other hand, other-regarding low quality sellers were more likely to reveal their quality than selfish sellers.

Upon completing the market experiment and the lie aversion task, subjects were presented 6 lotteries which they could either accept or decline. Each lottery is a 50-50 chance between winning an additional 6 CHF or losing an amount that differs between lotteries (2, 3, 4, 5, 6, 7). One of the 6 lotteries was randomly selected and paid. In case the selected lottery was declined, no additional earnings or losses were realized. We focus our discussion on treatment C-Sep. Almost all subjects (97 percent) switch at a unique point from accepting lotteries with relatively small losses to declining all lotteries that entail larger losses. Subjects are classified as loss averse if and only if they do not accept the lottery between winning 6 CHF and losing 3 CHF.[25] The lottery task may also measure a subject's risk aversion around 0. Since the theoretical predictions are qualitatively identical, the following results can be interpreted in the light of risk or loss aversion.[26]

[23]Proving these intuitions requires plugging in the new utility functions in the expressions used to derive the equilibrium in Proposition 3. Also note that high quality sellers' behavior in the separating equilibrium is unaffected by the parameters κ, η and μ.

[24]Consider submarket S_1^h for which we know that risk and loss neutral buyers offer only prices that exceed c_H. We show in Appendix C that if $\mu = 1.25$ buyers mix between low and high prices (as observed in the experiment) and with $\mu = 2$ prices never exceed v_L.

[25]24 out of the 36 buyers in C-Sep are classified as loss averse. Choosing a different threshold does not alter the qualitative results.

[26]The task does not allow to disentangle risk and loss aversion. See Fehr et al. (2013) for a thorough

Table 2.4: Loss Aversion in C-Sep

	Submarket l	Price Offers in h	Trade with H
1 l / 5 h	0.039 (0.092)	-4.563** (2.130)	-0.212 (0.234)
3 l / 3 h	-0.144* (0.078)	2.135* (1.275)	-0.110 (0.136)
4 l / 2 h	-0.183* (0.098)	2.407 (1.907)	0.129 (0.203)
Loss Averse (LA)	0.238** (0.119)	-6.087*** (2.136)	-0.444** (0.174)
1 l / 5 h x LA	-0.258** (0.106)	3.043 (2.415)	0.226 (0.244)
3 l / 3 h x LA	-0.028 (0.130)	6.832*** (2.382)	0.518*** (0.173)
4 l / 2 h x LA	-0.045 (0.165)	5.919** (2.503)	0.156 (0.224)
Constant	0.457*** (0.139)	12.890*** (1.482)	0.583*** (0.221)
R^2 (overall)	0.065	0.570	0.175
Observations (Groups)	696 (36)	280 (34)	245 (36)

Random effects regression for C-Sep using data of all periods. Standard errors in parentheses clustered on individuals. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables by column are buyers' choice of submarket (1=l, 0=h), buyers' price offers in submarket h, and trade with H-quality sellers (0=no trade, 1=trade). The baseline is LA=0, market structure 2 l / 4 h. All estimations include period dummies.

Table 2.4 displays random effects regressions on loss (or risk) aversion. Data now includes all periods to ensure a sufficient number of observations for all submarkets. The dummy *Loss Averse* is equal to 1 if the subject is classified as loss averse and 0 otherwise. The baseline are buyers who are not loss averse in market structure 2 l / 4 h. We focus on this market structure, as in theory it is the only one where loss aversion affects behavior and Figure 2-4 has shown that buyers are torn between offering low and high prices. In column 1 of Table 2.4 the dependent variable is the buyers' choice of submarkets (1=l, 0=h). In market structure 2 l / 4 h, loss averse buyers are 24 percentage points more likely to choose submarket l. Recall that low as well as high price offers were made in submarket h of 4 l / 2 h. The estimation results in column 2 suggest that most low prices were offered by loss averse buyers. As a consequence of column 1 and 2, column 3 shows that loss averse buyers are less likely to trade with a high quality seller in 2 l / 4 h. Note

discussion of the lottery task. We focus on loss aversion, since (i) sellers classified as loss (risk) averse were not more likely to send message l and (ii) subjects' comments in the questionnaire at the end of the session indicate that the fear of making losses was a first-order concern.

that this is not true for 3 l / 3 h, where most buyers are certain that submarket h consists of H-types only.[27]

Result 11 (Loss Aversion). *Sellers' anticipation of loss averse buyers has likely been conducive to endogenous market segmentation.*

Buyers' loss aversion has to be anticipated to increase low quality sellers' incentives to reveal their type. It seems plausible that over the 20 periods, sellers have learned that buyers join submarket l more often than expected and are somewhat reluctant to offer high prices in submarket h of 2 l / 4 h. This is also consistent with Figure 2-5b showing that C-Sep becomes more efficient in the course of a session. Anticipated loss aversion therefore seems to be a compelling channel that helped to establish the success of treatment C-Sep.

2.5.3 Over-Bidding and Payoffs

Comparing average observed to average predicted trade prices in Table 2.5 shows that buyers over-bid in all treatments except in submarkets h in C-Sep (Wilcoxon test $p<0.07$ for all comparisons). This is reminiscent of the experimental literature on auctions and over-bidding.[28] Potential explanations for over-bidding include risk aversion, noisy behavior, or a joy of winning (Goeree et al., 2002). Another explanation could be that buyers overestimate competition by other buyers. Because sellers reject the highest acceptable offer only in 2 percent of all cases, over-bidding is not explained by the buyers' inability to exploit the bargaining power implied by take-it-or-leave-it offers.

Figure 2-6 displays realized average payoffs of buyers and sellers as well as the theoretical predictions. Buyers' payoffs fall short of the predictions for all treatments (Wilcoxon test $p<0.07$ for all treatments). Conversely, low quality sellers earn significantly more than expected ($p<0.07$ for all treatments).

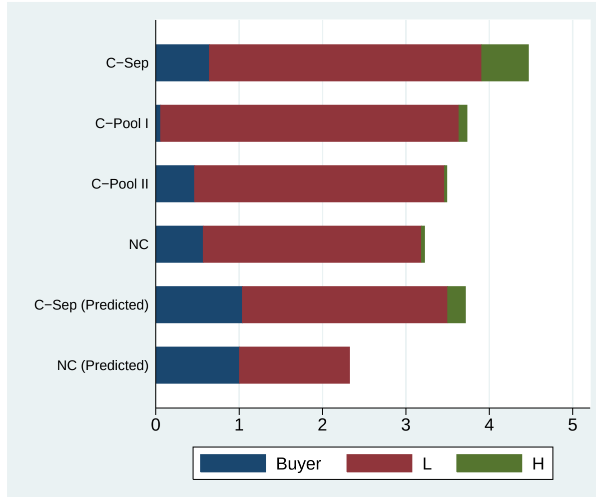
[27] Interestingly, whereas loss averse buyers are more likely to join submarket l in 3 l / 3 h, they do not offer lower prices conditional on joining submarket h.

[28] Once buyers are matched, our setting is similar to a first-price sealed-bid auction with an unknown number of competitors and a stochastic reservation value.

Table 2.5: Average Trade Prices

	Observed l	Observed h	Predicted l	Predicted h
C-Sep				
1 l / 5 h	3.17	3.80	2.69	1.55
2 l / 4 h	3.21	10.23	2.75	14.35
3 l / 3 h	2.80	15.29	1.76	15.76
C-Pool I	2.89	4.94	1.76	1.76
C-Pool II	3.51	4.21	1.76	1.76
NC		3.51		1.76

Figure 2-6: Average Payoffs



From Figure 2-6 we can also conclude that C-Sep provides a Pareto improvement over the C-Pool and NC treatments. The payoff increase is strongest for high quality sellers, who are significantly better off in C-Sep than in the other treatments (WMW tests $p<0.02$ for all comparisons). An interesting observation is that buyers in C-Pool I on average barely make positive earnings.[29] Recall that in C-Pool I there were some attempts at trading with high quality sellers: it turns out that this was a costly endeavor for buyers.

2.6 Conclusion

This article reports experimental evidence on decentralized markets with asymmetric information and matching frictions. We show that a simple form of communication – sellers can send a costless binary message – suffices to substantially alleviate adverse selection. In contrast to the existing experimental literature on cheap talk and asymmetric information, the importance of communication is not based on lie aversion or other-regarding preferences. Instead, low quality sellers have monetary incentives to reveal their type and separate themselves from high quality sellers in order to improve their competitive position by attracting more buyers.

[29]WMW tests confirm that buyers earn less in C-Pool I than in the other treatments (all $p<0.06$).

On a more general note, this article explores a setting in which inefficiencies due to one friction (incomplete information) are alleviated by exploiting the presence of additional sources of inefficiency (matching frictions). In recent years, the theoretical literature has made considerable progress in understanding what features of decentralized markets are conducive or detrimental to efficiency. Lauermann (2013) provides a general approach to such questions and emphasizes the role of competition, incomplete information and rules of bargaining. It seems worthwhile to generate more experimental insights into how different combinations of these aspects may interact and impact outcomes.

Appendix

2.A Instructions for Treatment C-Sep

Welcome to this economic experiment! In this experiment you can earn money with the decisions you make. How much you earn depends on your own decisions, the decisions of other participants as well as random events. We will not speak of Swiss Francs during the experiment, but rather of points. All your earnings will first be calculated in points. At the end of the experiment the total amount of points you earned in this part will be converted to Swiss Francs at the following rate: 1 point = 0.6 CHF. In addition, you will receive a show up fee of 14 CHF.

From now on you are not allowed to communicate in any other way than specified in the instructions. Please obey to this rule because otherwise we have to exclude you from the experiment and all earnings you have made will be lost. Please also do not ask questions aloud. If you have a question, raise your hand. A member of the experimenter team will come to you and answer your question in private.

The experiment lasts approximately 80 minutes. The experiment consists of three parts that are independent of one another. For each part you will receive specific instructions. These instructions will explain how you make decisions and how your decisions and the decisions of other participants influence your earnings. Therefore, it is important that you read the instructions carefully.

In case you should make losses, the show up fee of 14 CHF is used to cover for these losses. If you make losses exceeding 14 CHF, you will have the option to leave immediately and earn 0.

Part 1. We will now describe the general setting you will face during the experiment. At the beginning of the experiment the participants will be divided into buyers and sellers. There will be 6 buyers and 6 sellers. You will be one of these buyers or sellers. When you are a buyer (respectively, a seller) you will stay a buyer (respectively, a seller) throughout the experiment. You will not get to know the identity of the buyers or sellers you interact with, neither during nor after the experiment. Similarly, no participant will get to know your identity.

A decision situation will be repeated for 20 periods. In each period the 6 buyers and the 6 sellers can trade a good in the market. Each buyer wants to buy at most one unit and each seller can produce and sell at most one unit of this good. The seller can be of two different types: type L or type H. A seller of type L can only produce a low quality good at cost 0. The buyers' valuation for the low quality good is 5. Hence, the surplus generated from trading a low quality good is 5. A seller of type H can only produce a high quality good at cost 14. The buyers' valuation for the high quality good is 19. Hence, the surplus generated from trading a high quality good is also 5.

We will tell the seller her type (L or H) at the beginning of each period. In each period there will be 3 type L and 3 type H sellers. Which sellers are of type L or H is randomly determined. Note that a seller also knows how much her good is worth to the buyers. However, the buyers do not know the sellers' types and hence, a buyer does not know whether his valuation for the good is 5 (and the seller's cost is 0) or 19 (and the seller's cost is 14). The buyer only knows that there are 3 low quality sellers (type L) and 3 high quality sellers (type H).

Sellers and buyers interact in this market in three steps: First, sellers send messages "low" or "high" to all buyers. This generates 2 submarkets. Second, each buyer chooses a submarket "low" or "high" and makes an offer in this submarket. It is important to understand that buyers choose the submarket in which they want to make an offer and the offer they want to make. However, the computer randomly determines to which exact seller in the chosen submarket the offer goes. The implications are discussed below in detail. Third, sellers receive the offer(s) and accept at most one offer. We will now explain each step in detail.

Step 1: Sellers send a message. Before sellers and buyers potentially trade, each seller can send a message. The two possible messages are "low" and "high". The messages are sent at no costs and both types of sellers (L and H) may send both messages. That is, type L may send message "low" or "high" and likewise for type H. What happens with these messages? When buyers make their offers (see step 2 below), they are first informed about how many of the 6 sellers sent message "low" and how many sent message "high". Buyers can then choose to make an offer either to the sellers who sent "low" or to the sellers who sent "high". Therefore, the way we think about the messages is that they divide the initial market into two submarkets "low" and "high". For instance, suppose 2 sellers sent message "low" and 4 sellers message "high". Then buyers are given the choice between offering in submarket "low" with 2 sellers or submarket "high" with 4 sellers. Below you see a screen shot of the sellers' decision screen.

Step 2: Buyers choose submarkets and make offers. In this step, buyers make price offers to the sellers. Each buyer makes an offer to exactly one seller. A buyer can choose in which submarket "low" or "high" (generated by the messages in step 1) he wants to make an offer. However, to which specific seller the offer is made is randomly determined by the computer. In particular, a seller may receive an offer from several buyers or may not receive an offer at all. Let us give an example.

Suppose 2 buyers decide to make an offer in submarket "low". Also suppose that there are 2 sellers in this submarket (that is, 2 sellers sent message "low"). Thus, the 2 buyers' offers can be received only by one of the 2 sellers in the same submarket and not by a seller in submarket "high". It is randomly determined by the computer to which of the 2 sellers in submarket "low" the offer goes. In this example with 2 sellers, each buyer's offer is made to a specific seller in submarket "low" with probability 0.5 (50 percent). This means that either 1 of the sellers receives both offers or each seller receives 1 offer. More precisely, the probability that specific seller receives 2 offers is $0.5^2 = 0.25$. This corresponds to the probability that buyer 1 offers to this seller (50 percent) times the probability that buyer 2 also offers to this seller (50 percent). Of course, then the probability that a seller receives no offer is also 0.25. The probability that both sellers receive one offer is 2 * 0.5 * (1-0.5)=0.5, where the 2 occurs, because there are two ways this can happen (Buyer 1 offers to seller 1 and buyer 2 to seller 2, or buyer 1 offers to seller 2 and buyer 2 to seller 1). In summary, in a submarket with 2 sellers and 2 buyers the probability of a seller to receive no offer is 0.25, the probability of a seller to receive 1 offer is 0.5, and the probability to receive 2 offers is 0.25.

These probabilities depend of course on the number of buyers and sellers in a submarket. A submarket may contain a different number of buyers and sellers than in the above example. The idea is not that you calculate all these probabilities in detail (although you can do some calculations if you like). What is important is that given you are in a specific submarket (a group of sellers who sent the same message together with a group of buyers who chose to make an offer to these sellers), your offer as a buyer only goes to one of the sellers and each seller has the same probability to receive your offer.

The above implies in particular that if you are a buyer and there are a lot of buyers in the same submarket as you, the seller who receives your offer is likely to also receive other offers. On the other hand, if you are the only buyer in a submarket, you are certain that your offer will be the only one. Of course, you do not know how many buyers make offers in the same submarket when you make your offer.

A similar remark holds for sellers. If you are a seller, the more sellers are in the same submarket as you, the lower your probability to receive many offers and the higher your probability to receive no offer. If you are the only seller in a submarket and there is at least one buyer who makes an offer in this submarket, you are certain to receive this offer.

Let us give one more example. Suppose 1 seller sends message "high" and 5 sellers send message "low". Also suppose that, after observing the sellers' messages, 5 buyers choose to offer in submarket "high" and 1 buyer chooses to offer in submarket "low". Then the seller in submarket "high" is certain to receive 5 offers and each of the 5 buyers competes with 4 other offers. On the other hand, in submarket "low" only 1 of the 5 sellers will receive an offer from the buyer and the buyer will not compete with any other offer.

Finally, note that offers have to be between 0 and 19 and can be as exact as to the second decimal place. Hence, offers of 1, 7.9, 16.11 are possible. Offers of -3, 5.557, 19.2 are not possible. Below you are shown a screen shot of the buyers' decision screen in step 2: buyers choose a submarket and an offer.

Step 3: Sellers accept or reject offers. In this final step, sellers decide which offer (if any) to accept. If a seller does not receive an offer, she cannot trade. If a seller receives 1 or more offers (see step 2 to understand how more than one offer can be received) she can accept at most one of these. A seller can also reject all offers. See the screen shot below for an example where a seller received 2 offers. If the seller accepts an offer, she produces the good and sells it to the buyer at the agreed price. The payoffs of the seller and the buyer who has made the offer are determined as follows.

- Seller's payoff = Accepted Offer - Production Cost
- Buyer's payoff = Valuation of the Good - Accepted Offer

To calculate payoffs, recall the valuations and costs. Seller's production cost: low quality good 0, high quality good 14. Buyer's valuation: low quality good 5, high quality good 19. As an example, consider a buyer who offers a price of 6 and a seller who accepts this offer. If the seller is a type L (low quality) seller, his payoff is (Accepted Offer - Production Cost) = 6-0 = 6. The buyer's payoff is (Valuation - Accepted Offer) = 5-6 = -1. On the other hand, if the seller is a type H (high quality) seller, his payoff if he accepts the offer is (Accepted Offer - Production Cost) = 6-14 = -8. The buyer's payoff in this case is (Valuation - Accepted Offer) = 19-6 = 13.

The sellers who did not receive an offer or rejected all offers earn a payoff of 0. The buyers whose offers were rejected also earn a payoff of 0.

Once sellers have decided which offers to accept (if any) and the goods are traded, you are shown your earnings in this period. Then the next period starts (there are 20 periods). The setting is the same in all periods. As a seller you may sometimes be type L and sometimes type H.

2.B Proofs

2.B.1 Proposition 1

The sellers' acceptance decision is trivial: accept the highest offer as long as it covers the reservation cost. In the putative symmetric partially separating equilibrium, sellers' behavior is thus fully described by α and buyers can infer $q(S_i^m)$.

The probability that a buyer competes with k other buyers for the same seller when going to S_i^m is denoted by $\lambda(k, S_i^m, \beta(S_i^m))$. A buyer's expected payoff is then

$$\begin{aligned} U_B(S_i^m, \lambda(0, S_i^m, \beta(S_i^m))) &= \lambda(0, S_i^m, \beta(S_i^m)) \\ &\max\{q(S_i^m)(v_L - c_L), q(S_i^m)(v_L - c_H) + (1 - q(S_i^m))(v_H - c_H)\}. \end{aligned} \tag{2.2}$$

To understand (2.2), note that buyers must follow a mixed strategy. In fact, $F(\cdot, S_i^m)$ has no atom, because in a symmetric equilibrium deviating to a slightly higher offer would be profitable. This entails that the lowest offer over which buyers are mixing corresponds to the offer that is optimal conditional on being the only bidder ($k = 0$). Whether a monopsonist offers c_L or c_H depends on $q(S_i^m)$ as in (2.2).

Suppose there is only one buyer in the market and he faces market structure $\{S_1^l, S_{n_L-1}^h\}$. He will strictly prefer to join submarket S_1^l if and only if

$$q(S_{n_L-1}^h) > \frac{v_H - c_H - (v_L - c_L)}{v_H - v_L}. \tag{2.3}$$

This is obviously satisfied if $v_L - c_L > v_H - c_H$. Otherwise, from (2.1) we have $q(S_{n_L}^h) > \frac{v_H - c_H}{v_H - v_L}$. Hence, (2.3) holds if we assume (2.4).

$$q(S_{n_L}^h) - q(S_{n_L-1}^h) \leq \frac{v_L - c_L}{v_H - v_L} \tag{2.4}$$

Under (2.4), $\beta(S_1^l) > 0$ for any number of buyers (competition between buyers in submarket h will make it even more profitable to deviate from $\beta(S_1^l) = 0$). If $\beta(S_1^l) = 1$, $\alpha > 0$ is obvious. For $\beta(S_1^l) \in (0, 1)$, buyers are indifferent between submarkets and $\beta(S_1^l)$ is given by

$$U_B(S_i^l, \lambda(0, S_i^l, \beta(S_i^l))) = U_B(S_{n_L-i}^h, \lambda(0, S_{n_L-i}^h, \beta(S_{n_L-i}^h))) \tag{2.5}$$

for $i = 1$. Because of (2.2) and (2.3), we need $\lambda(0, S_{n_L-1}^h, \beta(S_{n_L-1}^h)) > \lambda(0, S_1^l, \beta(S_1^l))$ for (2.5) to hold.

Note that $\lambda(k, S_i^h, \beta(S_i^h)) = \sum_{b=k}^{n_B-1} \beta(S_i^h)^b (1 - \beta(S_i^h))^{n_B-1-b} \binom{n_B-1}{b} (\frac{1}{i+n_H})^k (1 - \frac{1}{i+n_H})^{b-k} \binom{b}{k}$ and $\lambda(k, S_i^l, \beta(S_i^l)) = \sum_{b=k}^{n_B-1} \beta(S_i^l)^b (1-\beta(S_i^l))^{n_B-1-b} \binom{n_B-1}{b} (\frac{1}{i})^k (1-\frac{1}{i})^{b-k} \binom{b}{k}$, where $i > 0$ for the latter and using the convention that $0^0 = 1$. It follows that $\lambda(0, S_1^l, \beta(S_1^l)) = (1 - \beta(S_1^l))^{n_B-1}$. Using the Binomial Theorem we also obtain $\lambda(0, S_{n_L-1}^h, \beta(S_{n_L-1}^h)) = (1 - \frac{1-\beta(S_1^l)}{n_S-1})^{n_B-1}$. Hence, $\lambda(0, S_{n_L-1}^h, \beta(S_{n_L-1}^h)) > \lambda(0, S_1^l, \beta(S_1^l))$ implies $\beta(S_1^l) > \frac{1}{n_S}$.

Let $U_L(S_i^m)$ be a low quality seller's expected payoff conditional on being in submarket S_i^m. If we can show that $U_L(S_{n_L}^h) < U_L(S_1^l)$, then there is an equilibrium with $\alpha > 0$. Since it is optimal for a buyer to offer c_L, a buyer's expected payoff is $U_B^{\alpha>0} \equiv (1 - \beta(S_1^l))^{n_B-1}(v_L - c_L)$ in S_1^l and $U_B^{\alpha=0} \equiv (1 - \frac{1}{n_S})^{n_B-1}(v_L - c_L)$ in $S_{n_L}^h$. The probability that a low quality seller trades is $x_L^{\alpha>0} \equiv 1 - (1 - \beta(S_1^l))^{n_B}$ in S_1^l and $x_L^{\alpha=0} \equiv 1 - (1 - \frac{1}{n_S})^{n_B}$ in $S_{n_L}^h$. Since the sum of the expected payoffs of the expected number of buyers plus the sum of the expected payoffs of the sellers has to equal the total expected gains generated in a submarket, we obtain $U_L(S_1^l) = x_L^{\alpha>0}(v_L - c_L) - \beta(S_1^l) n_B U_B^{\alpha>0}$ and $U_L(S_{n_L}^h) = x_L^{\alpha=0}(v_L - c_L) - \frac{n_B}{n_S} U_B^{\alpha=0}$. It follows that $U_L(S_1^l) > U_L(S_{n_L}^h)$ if and only if $(1 - \frac{1}{n_S})^{n_B-1}(1 + (n_B - 1)\frac{1}{n_S}) > (1 - \beta(S_1^l))^{n_B-1}(1 + (n_B - 1)\beta(S_1^l))$. The latter holds if $\beta(S_1^l) > \frac{1}{n_S}$. QED.

Equilibrium Derivation. For completeness, we provide the remaining expressions needed to calculate $F(\cdot, S_i^m)$ and α. The probability that p is a winning offer in submarket S_i^m is $\pi_{S_i^m}(p) = \sum_{k=0}^{n_B-1} \lambda(k, S_i^m, \beta(S_i^m)) F^k(p, S_i^m)$. The expected payoff of a buyer who bids p is equal to $\pi_{S_i^m}(p) q(S_i^m)(v_L - p)$ if $p < c_H$ and $\pi_{S_i^m}(p)(q(S_i^m)(v_L - p) + (1 - q(S_i^m))(v_H - p))$ if $p \geq c_H$. Buyers' bidding strategies can be derived by setting these expressions equal to (2.2). One also finds

Lemma 2. *Let* $\underline{q} = (v_H - c_H)/(v_H - c_L)$ *and* $\overline{q}(S_i^m) = (v_H - c_H)/(v_H - v_L +$

$\lambda(0, S_i^m, \beta(S_i^m))(v_L - c_L))$. *Let* $\overline{p}(S_i^m)$ *and* $\underline{p}(S_i^m)$ *be the maximum and minimum offer in the support of* $F(p, S_i^m)$.

(i) If $q(S_i^m) \geq \overline{q}(S_i^m)$ *then* $\underline{p}(S_i^m) = c_L$ *and* $\overline{p}(S_i^m) < v_L$.

(ii) If $\underline{q} < q(S_i^m) < \overline{q}(S_i^m)$ *then* $\underline{p}(S_i^m) = c_L$ *and* $\overline{p}(S_i^m) > c_H$.

(iii) If $q(S_i^m) \leq \underline{q}$ *then* $\underline{p}(S_i^m) = c_H$ *and* $\overline{p}(S_i^m) > c_H$.

Low quality sellers' expected payoff conditional on being in submarket S_i^m is

$$U_L(S_i^m) = \sum_{b=1}^{n_B} \Big[\beta(S_i^m)^b (1 - \beta(S_i^m))^{n_B - b} \binom{n_B}{b} \sum_{k=1}^{b} (\frac{1}{i + \mathbf{I_h} n_H})^k (1 - \frac{1}{i + \mathbf{I_h} n_H})^{b-k} \tbinom{b}{k} \int_{\underline{p}(S_i^m)}^{\overline{p}(S_i^m)} (p - c_L) dF^k(p, S_i^m)\Big], \quad (2.6)$$

where $m = \{l, h\}$ and $\mathbf{I_h} = 1$ if $m = h$ and 0 otherwise.

Note that $\alpha = 1$ is possible if $U_L(S_{n_L}^l) \geq U_L(S_1^h)$. Otherwise, $\alpha \in (0, 1)$ is given by setting equal the expected payoffs from sending message l (LHS) and h (RHS):

$$\sum_{i=0}^{n_L - 1} \alpha^i (1-\alpha)^{n_L - 1 - i} \binom{n_L - 1}{i} U_L(S_{i+1}^l) = \sum_{i=0}^{n_L - 1} \alpha^i (1-\alpha)^{n_L - 1 - i} \binom{n_L - 1}{i} U_L(S_{n_L - i}^h). \quad (2.7)$$

2.B.2 Observation 1

In every (partially) separating equilibrium there is a submarket that exclusively consist of low quality sellers (see Kim, 2012). Sending message l thus reveals a seller to be of the low type. Moreover, messages cannot impact buyers' matching decisions. Hence, low quality sellers are at best indifferent between l and h. If $\alpha > 0$, there is a positive probability that all other low quality sellers send message l. Assuming

$$q(S_1^h) < \overline{q}(S_1^h) \quad (2.8)$$

guarantees that in S_1^h prices above c_H are offered with positive probability (see Lemma 2). Sending message h is then a strictly profitable deviation. QED.

2.C Lie and Loss Aversion

2.C.1 Lie Aversion Task and Analysis

The lie aversion task is a variant of Gneezy (2005). A sender communicates one of two possible messages to a receiver. The message is either "Option A will earn you a higher payoff than option B" or "Option B will earn you a higher payoff than option A." The sender is informed about the payoff consequences of both options. The receiver is not informed and observes only the message. Payoffs depend exclusively on the option chosen by the receiver. The list of payoffs if option A is chosen is: (9, 11), (8, 12), (7, 13), (6, 14), (5, 15), (4, 16), (3, 17), (2, 18), where the first entry corresponds to the sender's payoff and the second entry to the receiver's payoff. Option B gives the same payoffs except that the receiver now earns the lower amounts. Thus, Message A is always the truth. One of the 8 decisions was randomly selected and paid. Receivers only observed their own payoff. Note that total surplus is always 20 and the induced inequality is always the same for option A and B. Preferences for efficiency and pure inequality aversion therefore do not affect a sender's decision. Option B is the senders preferred message if he exhibits no lie aversion and the incentives to lie increase as differences in payoffs grow.

71 percent of the senders have a unique switching point. We keep the remaining subjects in the sample and use the most unequal payoff pair for which the subject is truthful as truth-telling index. Receivers followed the senders' advice in 75 percent of the cases. Senders are also asked to state their beliefs on whether receivers will follow their advice, and are paid for a correct guess. Only 54 percent believed the receiver would follow their advice. This calls for a careful categorization of senders. Subjects who send message B for payoff distribution 7-13 (and all more unequal distributions) are classified as liars. We further divide subjects into selfish and other-regarding. Consider a liar who believes that the receiver will not follow his advice. Clearly, he must care about the gains of the other, because he expects the receiver to choose option A in response to receiving message B. In other words, he is an other-regarding liar. A liar who expects the other to follow his advice is referred to as a selfish liar. A non-liar who believes that the other

Table 2.C.1: Does Lie Aversion Explain Truth-Telling by Low Quality Sellers?

	All Treatments	C-Sep	C-Pool I	C-Pool II
Non-Liar	-0.155	-0.114	0.169	-0.158
	(0.105)	(0.133)	(0.182)	(0.189)
Other-Regarding	-0.086	0.258**	-0.083	-0.134
	(0.147)	(0.101)	(0.159)	(0.240)
Non-Liar x Other-Regarding	0.430**	-0.010	0.218	0.697**
	(0.176)	(0.159)	(0.248)	(0.297)
Constant	0.525***	0.682***	0.166	0.361***
	(0.076)	(0.0845)	(0.141)	(0.133)
R^2 (overall)	0.069	0.060	0.109	0.195
Observations (Groups)	420 (84)	180 (36)	120 (24)	120 (24)

The table presents random effects regressions for low quality sellers. The dependent variable takes value 1 if the seller sends message l and 0 if h. Standard errors in parentheses are clustered on individuals. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. To allow for a direct interpretation of the constant, no period dummies are used. Including period dummies does not affect variables other than the constant. Probit regressions yield similar results.

will follow his advice is an other-regarding non-liar. Finally, there were truth-tellers who anticipated not to be believed, i.e. in some cases saying the truth may be misguiding (see also Sutter, 2009). Thus, selfish non-liars are those who send message A but expect the receiver to choose option B in response. In total there are 22 selfish liars, 36 selfish non-liars, 13 other-regarding liars and 23 other-regarding non-liars.

Table 2.C.1 reports results of random effects regressions. The dependent variable is the low quality sellers' messages (1=l, 0=h). The dummies *Non-Liar* and *Other-Regarding* follow the classification described above. Notice that in C-Sep other-regarding (liar and non-liar) low quality sellers are more likely to reveal their quality than selfish sellers. Lie aversion, on the other hand, has no significant impact in C-Sep. Looking at the results over all treatments indicates that it was mostly other-regarding non-liars who were willing to send message l (a t-test for *Non-Liar* + *Other-Regarding* + *Non-Liar x Other-Regarding*=0 yields p=0.06).

2.C.2 Predictions with Lie and Loss Aversion

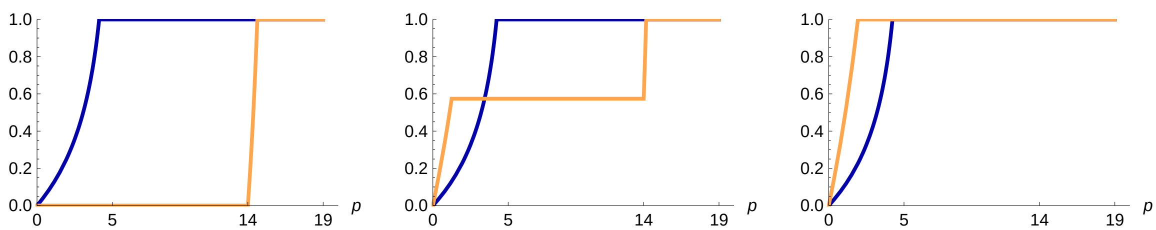
Table 2.C.2 shows predictions for different combinations of lie and loss aversion. Loss aversion leads to more market segmentation. Lie aversion leads to information disclosures for the C-Pool treatments if $\kappa \geq 9.31$. In this case full separation is obtained. Since there is either full separation or pooling, loss aversion plays no role in the C-Pool treatments.

Table 2.C.2: Theoretical Predictions with Lie and Loss Aversion

				Rates of Trade		Ex Ante Efficiency			Payoffs		
	κ	μ	α	L	H	Total	L	H	U_B	U_L	U_H
C-Sep	0	1	0.48	0.70	0.25	14.26	10.57	3.69	1.04	2.46	0.21
	0	1.25	0.71	0.70	0.32	15.34	10.57	4.77	1.24	2.15	0.48
	2	1	0.72	0.70	0.45	15.66	10.57	4.80	1.30	2.05	0.56
	2	1.25	1.00	0.67	0.67	19.95	9.98	9.98	2.01	1.32	1.32
C-Pool I, II	<9.31	$[1, \infty)$	0.00	0.67	0.00	9.98	0.00	9.98	1.00	1.32	0.00
	>9.31	$[1, \infty)$	1.00	0.67	0.67	19.95	9.98	9.98	2.01	1.32	1.32

Figure 2.C.1 depicts the bidding behavior for market structure $\{S_2^l, S_1^h\}$ for $\mu = \{1, 1.25, 2\}$. Bidding in other market structures is unaffected by loss aversion.

Figure 2.C.1: Buyers' Bidding Strategies in $\{S_2^l,\, S_1^h\}$ with Loss Aversion



(a) $\{S_2^l,\, S_1^h\}$ (b) $\{S_2^l,\, S_1^h\}$, $\mu = 1.25$ (c) $\{S_2^l,\, S_1^h\}$, $\mu = 2$

Chapter 3

Gradual Coalition Formation with Externalities

3.1 Introduction

A remarkable result in the literature on coalition formation is that despite the possibility to write binding agreements, equilibrium outcomes often fail to be efficient. At the same time, it has been shown that a larger flexibility in renegotiating agreements may restore efficiency. This latter finding is in line with the Coase Theorem, which states that if (re-)negotiation frictions are negligible, the efficient outcome should eventually be reached. Behind this assertion lies the idea that moving to a more efficient state sets free additional resources that can be used to compensate potential losers.

This article analyzes an environment in which the degree of renegotiation is endogenous. In particular, after forming a coalition players have two options: either they stay available for future renegotiation or they irrevocably leave the negotiation table. This modeling approach is suitable for situations that involve decisions that are irreversible or very costly to reverse. Examples include the declaration of a war, currency unions, the adoption of a technological standard, the decision to build environmentally friendly facilities, mergers between firms, or the position a political party takes on important issues during an election campaign. In all these situations, alliances form to steer outcomes in the direction preferred by its members. The question implicit in the Coase Theorem is then whether the incentives to form coalitions are aligned with the socially optimal outcome.

The previous literature on endogenous renegotiation has identified two main sources of inefficiency. The first one is linked to the so-called "Outside Option Principle", which refers to the result in the bargaining literature that outside options merely serve as a constraint on payoffs (Sutton, 1986). In our context, this implies that if a coalition is already in a position that guarantees a high payoff, it has little incentives to further expand cooperation, even if this is socially desirable. Intuitively, such coalitions prefer to simply "walk away", because they are unable to capture a share of the gains realized by moving to a more efficient outcome.[1] The second reason inefficiencies may occur is the presence of externalities between coalitions. However, to the best of our knowledge, it has so far remained unanswered what types of externalities prevent cooperation.

In this article, we propose a coalition formation model that eliminates inefficiencies linked to the Outside Option Principle. By focussing on externalities, we demonstrate that *successful cooperation through renegotiation may only be forestalled in environments that feature free riding opportunities.* This is an important insight, because for a broad class of games – which includes characteristic function games– efficiency is always attained through renegotiation.

Having established this result, we continue to explore free riding as an obstacle to efficiency and find that the notion of gradualism is key. Gradualism refers to coalition formation processes in which players do not immediately form the comprehensive agreement, but cooperation ensues in several steps. What are the roots of gradualism? A convincing mechanism is explored in Seidmann and Winter (1998): partial coalitions form to increase their bargaining leverage in future negotiations.[2] For instance, in 2010 and 2011 Ethiopia, Kenya, Uganda, Rwanda and Tanzania signed a Cooperative Framework Agreement to seek more water from the River Nile. This move seems to have shifted

[1] In Seidmann and Winter (1998), this is indeed the major reason for inefficiencies. They also hint at a third potential source of inefficiency, based on coordination failures within a coalition.

[2] This also hints at the fact that full cooperation does not necessarily entail a fair (and certainly not equal) division of surplus.

the relative bargaining power in water politics between the Nile riparian states in favor of the signatories, as it was strongly opposed by Egypt and Sudan. In particular, they demanded to omit the qualification "significantly" in Article 14b on water security, which requires member countries to avoid to "significantly affect the water security of any other Nile Basin State."[3]

This line of explanation is, however, inapplicable for agreements on public good provision. In this case, players within a coalition tend to internalize the positive externalities on the other members and thus, it is the outsiders who are better off, as they equally benefit from the increased provision levels. Consequently, the players who initiated cooperation will have less leverage in subsequent negotiations. Indeed, it will be shown that *in environments with free riding opportunities, gradualism can never occur in order to improve the own coalition's bargaining position.* Yet, it is these environments in which cooperation is important and gradualism is frequently observed. For instance, in the context of climate change, "The Kyoto Protocol is seen as an important first step towards a truly global emission reduction regime that will stabilize GHG emissions, and can provide the architecture for the future international agreement on climate change."[4]

An extension of the coalition formation model allows us to explain gradualism in public good settings by uncovering the other side of the coin: coalitions may form to *concede* bargaining power. Players are willing to do so in order to provide others with an incentive to enter into negotiations with them. Parties who initiate cooperation weaken their position relative to the ones who do not concede bargaining power, but the size of the cake grows such that everybody is better off. Forming a coalition can thus be interpreted as a deliberate commitment to not make use of free riding opportunities. Indeed, it seems plausible that the commitments observed in climate change negotiations were made to keep negotiations going, in particular with developing countries.[5]

[3]Our analysis further suggest that the game of water politics between Nile riparian states has a non-empty (strict) Core, as we will show that gradualism occurs if and only if the strict Core is empty.

[4]The statement is taken from the UNFCCC website.

[5]As another illustration, consider the federal elections in Germany, which are typically followed by extensive negotiations on the formation of coalitions between the winning party and parties which the winning party needs to achieve the required majority to form the government. In these negotiations, it is common that parties make public concessions early on. Concessions weaken the bargaining position

This paper is organized as follows. The next section presents the model and clarifies the connection to the existing literature. In Section 3, equilibrium is characterized. We derive our central results on gradual coalition formation, efficiency and renegotiation in Section 4. Section 5 applies the findings to the public goods case. Section 6 concludes.

3.2 Model

3.2.1 The underlying cooperative game

Let $N = \{1, 2, \ldots n\}$ be the set of players. A coalition structure π is a partition of N. The set of all coalition structures is denoted by Π. Let the restrictions to $S \subset N$ be $\pi(S)$ and $\Pi(S)$, respectively. The value of a coalition S in coalition structure π is summarized by a TU partition function $v(S, \pi)$. Let $v \equiv \{v(S, \pi)_{S \in \pi}\}_{\pi \in \Pi}$. Thus, v determines for all possible coalition structures the value of all coalitions. We normalize the minimum payoff a player can guarantee itself to be bounded away from 0, i.e. $v(\{i\}, \pi) > 0$ for all $i \in N$, $\pi \in \Pi$.

The partition function v is the primitive of our setting. However, v could in general be derived from a strategic form game (we will do so in Section 5). The interpretation is then that when a coalition leaves the formation process, it chooses its action as part of a non-cooperative game between coalitions.

3.2.2 Bargaining with irreversible actions

We model coalition formation as infinite horizon bargaining with the possibility to write binding agreements. There are two distinct phases, a *bargaining phase* and an *implementation phase*. We refer to the lapse of both phases as a *negotiation round*, or simply *round*. The game starts with the bargaining phase of the first negotiation round. In the bargaining phase, players make, accept, and reject proposals to determine which coalitions form

when it comes to agreeing on the government's position on minimum wage, tax raises, and so on. On the other hand, parties who commit not to bargain on these issues once the government has formed are more attractive to cooperate with and can thus avoid negotiation breakdowns.

and to pin down how the value of a coalition will be shared among its members. After contract(s) have been signed, the implementation phase starts. Coalitions now sequentially choose between implementing the current contract or remaining available for negotiations in future rounds. If a coalition chooses to implement, it effectively leaves the game by executing an irreversible action. When doing so, it will predict the final coalition structure and in particular how this structure depends on the fact that the coalition leaves. Once all coalitions have completed the implementation phase, the next negotiation round starts. We now turn to a formal description.

Negotiation rounds

Negotiation rounds are indexed by $\tau = 1, 2, \ldots$ At the beginning of each negotiation round there is a set $\mathcal{N} \subseteq N$ of players who *control* a coalition, the meaning of which will become clear presently. There is also a set $\mathcal{A} \subseteq \mathcal{N}$ of *active* players who have not yet implemented their contracts. Finally, as the negotiation round unfolds, there is a set $\mathcal{B} \subseteq \mathcal{A}$ of *negotiating* players who have not yet signed a contract in the current round. A state is described by $\omega = (\mathcal{N}, \mathcal{A}, \mathcal{B})$. Let Ω be the set of all possible states.

Proposals and counter-proposals in the bargaining phase

The bargaining phase begins with some player, say i, proposing a contract (S, t) to $S \subseteq \mathcal{B}$ such that $i \in S$. Thus, proposals can only be made to negotiating players. Note that a coalition can make a proposal to itself, thereby leaving the set of negotiating players without merging.

The second part of a proposal is a vector of transfers t satisfying $\sum_{j \in S} t_j = 0$. It is interpreted as the amount i offers to each j to obtain control over j's resources. When a proposal is accepted, player i becomes the controlling player of the newly formed coalition S. Players who accepted the proposal receive their transfers and will never be able to take another decision (nor will they be affected by the resulting coalition structure). We therefore use $\mathcal{N} \subseteq N$ to refer to the current set of players who control a coalition.

This interpretation of a proposal follows the one used in Bloch and Gomes (2006).[6] Its advantage is that gradual coalition formation becomes tractable: a coalition can always be identified with a single player and thus we can abstract from potential disagreements within coalitions. To be sure, who the non-controlling players in a coalition are matters, because v is defined on the initial set of players N. Also note that coalitions can never disintegrate.[7]

After proposal (S,t) by player i is made, all $j \in S_{-i}$ sequentially decide whether to accept or reject. Coalition S (with controlling player i) only forms if all $j \in S_{-i}$ accept the proposal. If a proposal is rejected, the coalition structure remains unaltered and the *rejector seizes the initiative*.[8] Notice that players can pass the initiative by making unacceptable proposals. At the start of the game and after any acceptance, the bargaining protocol ρ selects a player in $\mathcal{B}$ to make the next proposal. Let $\rho(i,\omega)$ be the probability that i is selected at ω. We assume $\rho(i,\omega) > 0$ for all $i \in \mathcal{B}$ and all $\omega \in \Omega$.[9]

Time $t = 0, 1, \ldots$ runs discretely. It is assumed that there is a geometric time cost δ (as in Rubinstein (1982)) incurred on all players *only if a rejection is followed by a counter-proposal*, where a counter-proposal is defined as follows.

Counter-Proposal. A proposal (S,t) by player i at ω is a *counter-proposal* if and only if at least one $j \in S$ has previously made a proposal at ω that was rejected by i.

By linking time costs to counter-proposals, we depart from the standard assumption that *every* rejection entails time costs. This departure is well motivated. Discounting in bargaining models functions as a technical device to i) force players to reach an agreement at some point and ii) reduce the set of equilibria by introducing a minimal degree of asymmetry between players. As will be shown, a model of costly counter-proposals is

[6] Ray and Vohra (1999) allow for more flexible sharing rules that depend on realized coalition structures. One could also let players renegotiate sharing rules. This, however, leads to coordination failures inside coalitions such as in Lemma (i) of Seidmann and Winter (1998), p. 808.

[7] See Gomes and Jehiel (2005) and Hyndman and Ray (2007) for contributions that allow for disintegration.

[8] This is in accordance with most of the coalition formation literature discussed in the next section. For an alternative approach see Okada (1996).

[9] The protocol also pins down the order in which players respond to a proposal, which turns out to be inconsequential. The assumption $\rho(i,\omega) > 0$ could be replaced by assuming that whenever a player is indifferent in the implementation stage, it chooses to remain active.

fully capable of assuming this role of discounting.[10] Time costs are also widely applied because they are intuitively convincing: it seems natural that formulating offers requires time and effort. We believe that the model of costly counter-proposals does not lose this intuitive appeal. It corresponds to the view that approaching another player per se is free of cost, but that it is haggling that makes bargaining costly. Formulating a counter-proposal takes more effort, because players know that they are in conflict about how to share the gains from cooperation. Moreover, from a psychological perspective, haggling with the same player over a long period of time seems more exhausting than initiating new potential cooperations. Finally, our model is a natural generalization to n players of the two-player bargaining model presented in Sutton (1986) and Osborne and Rubinstein (1990). In these articles, it is assumed that a rejector can consume its (exogenous) outside option before discounting sets in. In a similar vein, costly counter-proposals guarantee each player its (endogenous) outside option. This last point is crucial and will become clear when discussing Example 1.

Implementation phase: three models of renegotiation

Three different models of renegotiation will be considered. Model Γ^{NR} assumes that agreements cannot be renegotiated.

No Renegotiation. In Γ^{NR} an accepted agreement (S, t) implies that coalition S leaves the game immediately. There is thus no need for an implementation phase, as a coalition is forced to leave.

In the remaining two models, renegotiation is possible. In the implementation phase, all players in $\mathcal{A}$ are asked sequentially whether they want to implement their current contract. If player i implements in negotiation round τ, it is removed from $\mathcal{A}$ for all future negotiation rounds. If player i does not implement in round τ, i returns to the set $\mathcal{B}$ in $\tau + 1$.

[10] All studies discussed in the next section minimize the asymmetry in the bargaining protocol by looking at the outcomes when discounting frictions are negligible. In fact, our model further reduces the asymmetry in the bargaining process, because players are not forced to suffer time costs from rejecting proposals of players they have no interest in cooperating with, but who (perhaps arbitrarily) move earlier.

Renegotiation. In Γ^R an accepted agreement (S, t) immediately triggers the implementation phase. The player who controls S chooses whether to implement the contract or stay available for further negotiations. Players who control a coalition $S' \neq S$ do not have the possibility to implement in this round, but have to wait until they are the ones signing a contract in the bargaining phase of a future round.[11]

In Γ^R implementation decisions are taken before observing cooperation efforts of other coalitions. For instance, in committees or boards of directors it may be unclear what other members are going to do and the very fact that a coalition forms may crucially affect the decisions of the remaining parties. In other contexts, a more natural assumptions seems to be that when a coalition decides to leave the bargaining table, it is aware of other ongoing negotiations. For instance, in climate change negotiations countries have a good understanding of all potential partnerships. This motivates a model of renegotiation rounds. It differs from Γ^R with respect to the timing of the implementation phase.

Renegotiation Rounds. In Γ^{RR} the implementation phase is entered when there are no negotiating players left, i.e. $\mathcal{B} = \emptyset$. In other words, each active player signs one (and only one) contract in the bargaining phase of each round. In the implementation phase, the order in which players take decisions is the same as the order in which contracts were written in the bargaining phase of the same negotiation round.

The game ends if and when all coalitions have implemented their contracts. Payments are realized when the coalition formation process ends. If the coalition formation process never ends, all players are assumed to receive 0.[12]

[11] Recall that the contract could also be the singleton contract, i.e. S' does not need to grow to be implementable.

[12] Assuming that transfers are consumed immediately does not affect any of the results. That is, we could allow players who have accepted an offer to receive a positive amount even if the bargaining process is indefinite.

3.2.3 Relation to the literature

Our choice of negotiation models is rooted in the existing literature.[13] Model Γ^{NR} views all contracts as final. Important contributions that have applied this approach are Chatterjee et al. (1993), Bloch (1996), Ray and Vohra (1999), and Ray and Vohra (2001). A central conclusion in this literature is the persistence of inefficiency. The reason for such inefficiencies is that (Ray 2007, p. 85) "the very act of making a proposal opens the door to possible counteroffers" and hence, "the proposer must give away part of the social surplus when a group is formed. This drives a wedge between the proposer's incentives and the socially efficient outcome." We contribute to this literature by clarifying the connection between the grand coalition and the core (Theorem 2).

Acknowledging the incentives to collect rents at the expense of efficiency, are proposers able to do so through intermediate contracts, which are eventually renegotiated until the socially efficient outcome prevails? In order to provide an answer Perry and Reny (1994) and Seidmann and Winter (1998) introduce endogenous renegotiation, i.e. after signing contracts, coalitions can choose to continue negotiations or may credibly end negotiating. Interestingly, the latter paper shows that renegotiation can lead to gradual formation of coalitions, but even absent externalities, efficiency is not guaranteed.[14] In contrast, we find in Corollary 1 that renegotiation always leads to the efficient outcome for characteristic functions. This is a consequence of the assumption that only counter-proposals entail time costs. Corollary 1 is in accordance with Bloch and Gomes (2006), who present a model in which inefficiencies are explained exclusively by externalities. We confirm this finding, but in addition identify conditions on externalities that guarantee efficiency (Theorem 1). Moreover, Theorem 3 shows that in environments with strong free riding incentives, renegotiation is inconsequential, i.e. equilibrium outcomes in Γ^R and Γ^{NR} coincide. This finding links coalition formation with non-renegotiable contracts to the

[13] Naturally, this section cannot cover the vast literature on coalition formation. We refer to Ray (2007) and Ray and Vohra (2014) for comprehensive discussions.

[14] Model Γ^R is closely linked to Seidmann and Winter (1998)'s model. The most important difference is that Seidmann and Winter assume that after an acceptance or rejection of a proposal all players who have signed at least one contract can choose to implement. Because we allow for externalities, this would render the order in which contracts can be implemented an important object.

literature on endogenous renegotiation.

Finally, there is a literature on reversible actions with on-going negotiations, i.e. players cannot terminate the negotiation process. A remarkable result in Hyndman and Ray (2007) is that if the grand coalition is the efficient outcome, then irrespective of externalities, players will eventually end up forming the grand coalition.[15] Hence, while a lot remains to be explored in the context of reversible actions – in particular how the gains of the grand coalition will be distributed – the basic message is in accordance with the Coase Theorem.

3.3 Equilibrium Characterization

3.3.1 Equilibrium concept

We restrict attention to subgame perfect equilibria in stationary strategies. Recall that a state $\omega = (\mathcal{N}, \mathcal{A}, \mathcal{B})$ is composed of the current controlling players, the active players, and the negotiating players. In the bargaining phase, a strategy requires a player to make a proposal whenever it is asked to do so, conditioned only on the state ω. As a responder, a player's decision to accept or reject a proposal depends also on the nature of the proposal. In the implementation phase, a strategy specifies, conditional only on ω, whether to implement the current contract or to enter the next negotiation round.

Equilibrium coalition structures will be compared in terms of their efficiency properties. $\Gamma^i(v, \delta) \succsim \Gamma^j(v, \delta)$ indicates that all equilibria in $\Gamma^i(v, \delta)$ are weakly more efficient than the most efficient equilibrium in $\Gamma^j(v, \delta)$, depending on δ and partition function v. For instance, we could say that the comparison holds for all δ above a certain value $\hat{\delta}$ and for all partition functions v that are also characteristic functions. If $\Gamma^i(v, \delta) \sim \Gamma^j(v, \delta)$, then for each equilibrium coalition structure in $\Gamma^i(v, \delta)$ there is an equilibrium coalition structure in $\Gamma^j(v, \delta)$ that is equally efficient (and vice versa), given δ and v.

The following proposition guarantees existence of equilibrium in all three models. The

[15] This conclusion is true without the commonly imposed restriction to stationary strategies. Other important contributions include Seidmann and Winter (1998), Gomes and Jehiel (2005), and Gomes (2005).

proof is relegated to the Appendix.

Proposition 4. Γ^{RR}, Γ^{R} *and* Γ^{NR} *admit a stationary subgame-perfect equilibrium.*

3.3.2 Optimal proposals

Fix a state ω. Let $x_i(\omega, \mathcal{P}(R))$ be the payoff i obtains at ω given that he is the next offerer and his proposal must be to a coalition $S \in \mathcal{P}(R)$, where $\mathcal{P}(R)$ is the power set of R.[16] If $R = \mathcal{B}$ we simply write $x_i(\omega)$. We also adopt the convention $x_i(\omega, \mathcal{P}(\mathcal{B}_{-i})) = x_i(\omega)$. Importantly, $x_i(\omega)$ is interpreted as the payoff to i net of the payments he has made to the non-controlling players in his coalition. This does not affect i's behavior, since the payments are sunk costs. Let $y_i(\omega, j)$ be i's equilibrium response value to j at ω. It is the offer of j that is just accepted by i, knowing that every player acting after i accepts the proposal. From the definition of a counter-proposal, it follows that

$$y_i(\omega, j) = \begin{cases} x_i(\omega, \mathcal{P}(\mathcal{B}_{-j})) & \text{if } x_i(\omega, \mathcal{P}(\mathcal{B}_{-j})) > \delta x_i(\omega), \\ \delta x_i(\omega) & \text{otherwise.} \end{cases} \tag{3.1}$$

A proposal (S, t) is *optimal* for i if it yields a payoff of $x_i(\omega)$. Fix a player i with an optimal proposal to coalition S. We must have

$$x_i(\omega) \geq w_S(\omega) - \sum_{j \in S_{-i}} y_j(\omega, i) = w_S(\omega) - \sum_{j \in K(i,S)_{-i}} x_j(\omega, \mathcal{P}(\mathcal{B}_{-i})) - \delta \sum_{j \in \overline{K}(i,S)} x_j(\omega), \tag{3.2}$$

where $w_S(\omega)$ denotes the continuation value of coalition S. The weak inequality holds, because i can guarantee acceptance by offering $y_j(\omega, i)$ to every $j \in S_{-i}$. Expression (3.2) holds with equality if i's offer is acceptable. The set $K(i, S) \subseteq S$ consists of all $j \in S$ for which $x_j(\omega, \mathcal{P}(\mathcal{B}_{-i})) \geq \delta x_j(\omega)$. Note that $i \in K(i, S)$. According to (3.1), $y_j(\omega, i) = x_j(\omega, \mathcal{P}(\mathcal{B}_{-i}))$ for $j \in K(i, S)$ and $y_j(\omega, i) = \delta x_j(\omega)$ for $j \in \overline{K}(i, S) = S \backslash K(i, S)$. The set $K(i, S)$ is pinned down uniquely by the following condition.[17] We have $j \in K(i, S)$

[16] Note that in principle $x_i(\omega, \mathcal{P}(R))$ also depends on the set of players who have already made an offer to i at ω. However, Lemma 3 will show that we can safely ignore this.

[17] To see that the solution to (3.3) is unique, take $K(i, S)$ and $K'(i, S)$ and let $|K(i, S)| < |K'(i, S)|$.

for all $j \in S_{-i}$ if and only if

$$x_j(\omega, \mathcal{P}(\mathcal{B}_{-i})) \geq \overline{y}(i,S) = \frac{\delta \left(w_S(\omega) - \sum_{k \in K(i,S)_{-j}} x_k(\omega, \mathcal{P}(\mathcal{B}_{-i})) \right)}{1 + \delta \left(|S_{-j}| - |K(i,S)_{-j}| \right)}, \tag{3.3}$$

where $\overline{y}(i,S)$ is the equilibrium response value of $j \in \overline{K}(i,S)$ obtained by solving (3.2) for fixed outside options of $j \in K(i,S)_{-i}$. We now turn to a powerful result.

Lemma 3. *There exists $\hat{\delta} \in (0,1)$ such that for $\delta \geq \hat{\delta}$ the following holds. If i proposes to S and is accepted, then for every $j \in S$ it is also optimal to propose to S. If i strictly prefers to propose to S, proposing to S is strictly optimal for all $j \in S$.*

Proof. We start with two observations. If i strictly prefers S, then $x_i(\omega, \mathcal{P}(\mathcal{B}_{-j})) < x_i(\omega)$ and thus

$$y_i(\omega, j) = \delta x_i(\omega) \quad \forall j \in S_{-i}, \tag{3.4}$$

because there is $\delta \geq \hat{\delta}$ for which $x_i(\omega, \mathcal{P}(\mathcal{B}_{-j})) < \delta x_i(\omega)$.[18] By the same reasoning it follows that

$$y_j(\omega, i) = x_j(\omega, \mathcal{P}(\mathcal{B}_{-i})) = x_j(\omega) \Leftrightarrow j \in K(i,S). \tag{3.5}$$

We prove the second statement of the lemma. Consider $j \in K(i,S)$ and suppose j has an alternative (weakly or strictly) better than S. We have $y_j(\omega, i) = x_j(\omega)$ by (3.5) and hence $y_j(\omega, i) \geq w_S(\omega) - y_i(\omega, j) - \sum_{k \in S_{-ij}} y_k(\omega, j)$. Using (3.4), it follows that $x_i(\omega) > w_S(\omega) - y_j(\omega, i) - \sum_{k \in S_{-ij}} y_k(\omega, j)$. Combining this with (3.2) for i yields an immediate contradiction for $|S| = 2$ and otherwise, we obtain $\sum_{k \in S_{-ij}} y_k(\omega, i) < \sum_{k \in S_{-ij}} y_k(\omega, j)$.

Note that $K(i,S) \subset K'(i,S)$. Let $J = K(i,S)' \backslash K(i,S)$. For all $j \in J$, we have $j \notin K(i,S)$ and $j \in \overline{K}'(i,S)$. Using (3.3) for both cases implies $\delta(w_S(\omega) - \sum_{k \in K(i,S)_{-j}} x_k(\omega, \mathcal{P}(\mathcal{B}_{-i})))/(1 + \delta(|S| - 1 - |K(i,S)|)) < \sum_{k \in J_{-j}} x_k(\omega, \mathcal{P}(\mathcal{B}_{-i}))/|J_{-j}|$. Hence, there exists at least one $j \in J \cap K(i,S)$, a contradiction.

[18] Because gradualism does not necessarily induce discounting, a coalition may build up gradually, even if it is optimal for all $i \in S$ to form S in one step. Thus, the fact that $x_i(\omega, \mathcal{P}(\mathcal{B}_{-j})) < x_i(\omega)$ if i strictly prefers S is not obvious. We show that if at ω it is certain that S will form eventually, it is strictly optimal to form S immediately. Notice that at some state ω' coalition S will form. Because it was optimal to form S at the initial state, $x_i(\omega) = x_i(\omega')$. As a responder at ω', i obtains $\max\{x_i(\omega', \mathcal{P}(\mathcal{B}'_{-j})), \delta x_i(\omega')\} < x_i(\omega')$, because $x_i(\omega', \mathcal{P}(\mathcal{B}'_{-j})) < x_i(\omega')$ holds as S is now the only optimal proposal. But because $\rho(\omega', i) < 1$, i strictly prefers to offer S immediately.

Hence, for at least one $k \in S_{-ij}$, $y_k(\omega, i) < y_k(\omega, j)$. But k could have rejected i's offer and obtain at least $\delta x_k(\omega)$, which is either larger than or arbitrarily close (for $\delta \geq \hat{\delta}$) to $y_k(\omega, j)$. It follows that $\delta x_k(\omega) > y_k(\omega, i)$, which contradicts (3.1).

Consider now $j \in \overline{K}(i, S)$. Let S' be the proposal that j (weakly or strictly) prefers to S. We have $i \in S'$ by (3.5). Thus, $x_j(\omega) = w_{S'}(\omega) - \sum_{k \in S'_{-j}} y_k(\omega, j)$. Moreover, $x_i(\omega) > w_{S'}(\omega) - \sum_{k \in S'_{-i}} y_k(\omega, i)$ because S is strictly optimal for i. If $x_i(\omega) \leq x_j(\omega)$ then $y_i(\omega, j) \leq y_j(\omega, i)$, because of (3.4) and $y_j(\omega, i) = \delta x_j(\omega)$, where the latter is implied by $j \in \overline{K}(i, S)$. For $|S| = 2$, the contradiction is obvious. Otherwise, we need $\sum_{k \in S'_{-ij}} y_k(\omega, i) > \sum_{k \in S'_{-ij}} y_k(\omega, j)$. By the same argument as above, at least one $k \in S'$ should reject j's offer. Hence, $x_i(\omega) > x_j(\omega)$. But then (3.2) for i (holding with equality), (3.4) and $y_j(\omega, i) = \delta x_j(\omega)$ imply $x_j(\omega) < w_S(\omega) - \sum_{k \in S_{-j}} y_k(\omega, j)$, which contradicts (3.2) for j. This completes the proof of the second statement.

Assume now i's proposal to S is weakly optimal, but some $j \in S$ strictly prefers a different proposal S'. By the first part of the proof, $i \notin S'$ and hence, $x_j(\omega) = x_j(\omega, \mathcal{P}(\mathcal{B}_{-i})) = y_j(\omega, i)$. Thus, $y_j(\omega, i) > w_S(\omega) - \sum_{k \in S_{-j}} y_k(\omega, j)$. Combining the latter with (3.2), it follows that $x_i(\omega) + \sum_{k \in S_{-ij}} y_k(\omega, i) < y_i(\omega, j) + \sum_{k \in S_{-ij}} y_k(\omega, j)$. By the same reasoning as above, we obtain a contradiction for $|S| = 2$ and otherwise, at least one $k \in S$ should reject i's offer. Also note that S cannot be strictly preferred by j due to the first part of the proof. Hence, S is weakly optimal for j. $\square$

Lemma 3 and expressions (3.1) - (3.3) describe the nature of proposals that are accepted. Are proposals sometimes rejected?

Lemma 4. *There exists $\hat{\delta} \in (0, 1)$ such that for $\delta \geq \hat{\delta}$ no counter-proposals are made along the equilibrium path.*

Proof. Suppose there is a counter-proposal. By definition $\exists\ i,\ j,\ \omega$ such that i) i has an optimal proposal (S, t^S), where $j \in S$, and ii) for j it is optimal to turn down t_j^S and offer (R, t^R), where $i \in R$. Moreover, j is the player who actually rejects (S, t^S) with positive probability.

First, i) and ii) imply that i finds it (weakly) optimal to pass the initiative to j. This holds because either j rejects t_j^S for sure, or, if j is indifferent between accepting and rejecting, a slight increase in t_j^S would eliminate the risk of being rejected by j.

Second, we show that (R, t^R) is accepted with probability 1. Suppose by contradiction that there is a $k \in R$ who rejects j's proposal with probability $p \in (0, 1]$. If $k \neq i$, j finds it (weakly) optimal to pass the initiative to k (this again holds, because any mixing by k could be turned into acceptance by a slight increase in t_k^R) using a proposal that includes i. But this cannot be true, since by excluding i, j's equilibrium payoff increases by factor $1/\delta$ (assuming stationarity). If $k = i$, using the same reasoning as above, it is (weakly) optimal for i and j to indefinitely pass the initiative to each other, which contradicts the fact that equilibrium payoffs strictly exceed 0.

Combining the first and second observation, we conclude that (S, t^S) yields i an expected payoff of δt_i^R. Moreover, because (R, t^R) is accepted, it follows that $t_i^R = y_i(\omega, j)$ and by Lemma 3 that proposing to R must also be optimal for i. Hence, (R, t^R) yields $x_i(\omega) \geq t_i^R$. But $x_i(\omega) > \delta t_i^R$ means that (S, t^S) is not optimal. $\square$

To be sure, it may well be that proposals are rejected.[19] However, it follows from Lemma 4 that the full set of equilibrium outcomes can be identified by considering only acceptable proposals. To see this, suppose $\rho(\omega)$ selects i to propose and j rejects. Since there is no delay, the resulting equilibrium outcome must be identical to the one in which j was selected to be the next proposer at ω.

3.4 Gradual Coalition Formation

3.4.1 Endogenous outside options

We start with an example that illustrates how renegotiation helps to reach efficient outcomes and clarifies the role of counter-proposals. The notation $v(S_1, \ldots, S_M) = (v_1, \ldots, v_M)$ is used throughout the paper, where v_m is the worth of coalition S_m in coalition structure $\{S_1, \ldots, S_M\}$. When convenient we write ij instead of $\{i, j\}$.

[19] See Seidmann and Winter (1998)'s Example 1 for a case that involves a rejection.

Example 1. This example is due to Seidmann and Winter (1998). $N = \{1, 2, 3\}$. Let the characteristic function be given by $v(1, 2, 3) = (0, 0, 0)$, $v(ij, k) = (z, 0)$ and $v(N) = 1$, where $z > 2/3$ and $i, j, k = 1, 2, 3$.

There cannot be immediate formation of the grand coalition. To see this note that $x_i(\omega^0, k) + x_j(\omega^0, k) \geq z$ where ω^0 is the singleton structure, i.e. each two-player coalition can obtain a worth of at least z. However, an offer to the grand coalition must allocate an aggregate payoff below z to at least one pair of players. In equilibrium some player proposes a two-player coalition that is accepted, say, player 1 proposes coalition $\{12\}$. In Γ^{NR} the equilibrium coalition structure is thus $\{12, 3\}$.

In both models with renegotiation, players could enter round 2 where they face a two-player bargaining game. Consider player 1's behavior.[20] Suppose player 3 makes the first offer in round 2. If player 1 (who controls coalition $\{12\}$) enters the second round and player 3 were to offer less than z, player 1 would reject and consume his guaranteed "outside option" *without suffering any time cost.* Applying (3.3) shows that player 1 accepts exactly z. However, because $\rho(1, \omega) > 0$ for all ω, there is a positive probability that player 1 is selected to make the first proposal in round 2. In this case he secures $z + (1 - \delta)(1 - z) > z$. It is therefore strictly optimal to form the grand coalition.

In contrast, the grand coalition does not form *if we assumed that time costs are incurred after any rejection.* Optimal behavior in the first negotiation round is unaltered. If player 1 does not implement its contract in round 1, he is not guaranteed his outside options, since player 3's offer of δz must be accepted. Player 1 leaves in round 1 to obtain z (minus his payment to player 2). Interestingly, this is true even if player 1 is almost certain to offer first in round 2. Note that player 1's offer must be $\delta(1 - \delta z)$. Hence, player 1 chooses to leave in round 1 if $z > 1 - \delta(1 - \delta z)$, which holds for $\delta > (1 - z)/z < 1/2$. In Example 1 the outcome for Γ^{NR} is the same for both approaches to modeling time costs. This does not hold in general, as will be shown in Example 3.

We believe that neither the grand coalition nor the inefficient outcome should be dis-

[20] For large δ, player 3 obtains the main share of the gains realized by forming the grand coalition and will thus not implement in round 1.

missed as unrealistic in Example 1 – indeed in some cases "walking away" may be preferred to keep negotiating for negligible gains. However, the notion of counter-proposals saves us from simultaneously dealing with negotiation breakdowns attributable to coalitions trying to avoid being pushed below status quo payoffs and breakdowns due to externalities.

3.4.2 Defining free riding incentives

It will turn out that free riding incentives (or absence thereof) are central for effective renegotiation. Our goal in this section is to make precise what we mean by free riding. We start with a standard condition on v (see Yi, 1997). Under positive (negative) externalities, coalitions that are not involved in a merger are better (worse) off after the merger.

Positive Externalities. $v(S,\pi) \geq v(S,\pi')$ where $S \subset \pi, \pi'$ and $\pi \backslash \{S\}$ can be derived from $\pi' \backslash \{S\}$ by merging coalitions in $\pi' \backslash \{S\}$.

The next condition –combined with positive externalities– captures free riding: a merger increases (decreases) the worth of each coalition not involved in the merger by more (less) than the aggregate worth of the merging players.

Free Riding. Let $\{S_1, \ldots, S_M\} \subset \pi$ and $R \in \pi$, $R \neq S_m$ for all $m = 1, \ldots, M$. Define $S = \cup_{m=1}^{M} S_m$. Let $\pi' = \pi \backslash \{S_1, \ldots, S_M\} \cup \{S\}$. Then, $v(R,\pi') - v(R,\pi) \geq v(S,\pi') - \sum_{m=1}^{M} v(S_m,\pi)$.

Note that Free Riding neither implies nor is implied by Positive Externalities. In a symmetric game v depends only on the numeric coalition structure. For symmetric games, Free Riding implies that smaller coalitions enjoy higher per member payoffs than larger coalitions.[21]

Symmetric Free Riding. $v(S,\pi)/|S| \geq v(S',\pi)/|S'|$ if and only if $|S| \leq |S'|$.

Games of public good provision represent an important class of games that typically satisfy Positive Externalities and Free Riding or Symmetric Free Riding. We will verify this in Section 5. Another example is cartel formation in Cournot oligopolies.

[21] The reverse is false. Consider the example $v(i,j,klm) = (\epsilon,\epsilon,0)$, $v(ij,klm) = (1,0)$, all other partitions yield payoffs of 0 to all players, to convince yourself that Symmetric Free Riding holds but Free Riding does not.

Importantly, games for which Free Riding fails to hold may still allow for free riding opportunities. The class of games for which there are no free riding incentives for *any* merger satisfies the following condition.

No Free Riding. Let $\{S_1, \ldots, S_M\} \subset \pi$ and $R \in \pi$, $R \neq S_m$ for all $m = 1, \ldots, M$. Define $S = \cup_{m=1}^{M} S_m$. Let $\pi' = \pi \backslash \{S_1, \ldots, S_M\} \cup \{S\}$. Then, $M\left(v(R, \pi') - v(R, \pi)\right) < v(S, \pi') - \sum_{i=1}^{M} v(S_m, \pi)$.

Under No Free Riding, a merger implies a larger increase of the average payoff of the merging players than the payoff increase of each outsider. That is, to fully exclude free riding incentives, a coalition must be able to simultaneously guarantee *all* its members a larger increase in payoff (with appropriate transfers) than the outsiders obtain.

Finally, Grand Coalition Superadditivity (GCS) states that the grand coalition is strictly efficient.

Grand Coalition Superadditivity. $\sum_{S \in \pi} v(S, \pi) < v(N, \{N\})$ for all $\pi \in \Pi$.

3.4.3 Efficient negotiations

One of the central questions we attempt to answer in this paper is whether endogenous renegotiation results in an efficient outcome. If efficiency cannot be obtained, what are the reasons for this? Our first set of results links inefficiency to the presence of free riding externalities. It will also be shown that efficiency for characteristic functions games is guaranteed if either renegotiation is possible, or the strict core is non-empty.

Theorem 1. *Let v satisfy GCS, Positive Externalities and No Free Riding. There exists $\hat{\delta} \in (0, 1)$ such that for $\delta \geq \hat{\delta}$, $\Gamma^{RR}(v, \delta)$ and $\Gamma^{R}(v, \delta)$ always result in the grand coalition. Moreover, $\Gamma^{RR}(v, \delta) \sim \Gamma^{R}(v, \delta) \succsim \Gamma^{NR}(v, \delta)$.*

Intuitively, the conditions in Theorem 1 imply that coalitions draw their bargaining power from being involved in mergers which improve their position relative to outsiders. Stated differently, a coalition's bargaining power is *not* based on threats to leave the bargaining table, and thereby forcing others to cooperate. It is this absence of free riding incentives that allows for efficient renegotiation. To be sure, there can still be a gradual process, but Theorem 1 establishes that all players remain active in order to collect some of the gains that are realized by forming the grand coalition.

Proof. To see that $\Gamma^{NR}(v, \delta)$ may be inefficient, note that Example 1 satisfies Positive Externalities and No Free Riding and yet we have shown that the grand coalition does not form.

We now show that Γ^{R} and Γ^{RR} are efficient. Let $v(i, \mathcal{A})$ be the value of coalition i if the set of active players is $\mathcal{A}$ (the coalition structure of $\mathcal{N} \backslash \mathcal{A}$ is fixed). We write $v(\mathcal{A}, \mathcal{A})$ simply as $v(\mathcal{A})$. Positive Externalities and No Free Riding jointly imply weak superadditivity,

$$v(\mathcal{A}) \geq \sum_{i \subset \mathcal{A}} v(i, \mathcal{A}) \qquad \text{for any } \mathcal{A} \subseteq \mathcal{N}. \tag{3.6}$$

The inequality in (3.6) is strict whenever at least one player in $\mathcal{N} \backslash \mathcal{A}$ strictly benefits from the merger of $\mathcal{A}$ (GCS implies strict superadditivity for mergers to the grand coalition).

If $|\mathcal{A}| = 2$ each player earns at least its status quo worth, i.e. $y_j(\mathcal{A}, i) \geq v(j, \mathcal{A})$, where we abuse notation by writing $y_j(\mathcal{A}, i)$ instead of $y_j(\omega, i)$. Equation (3.6) and $\rho(\omega, j) > 0$ imply that the two-player coalition forms unless the singleton structure is also efficient (hence, it forms for sure if $n = 2$). Moreover, all players $k \in \mathcal{N} \backslash \mathcal{A}$ earn at least $v(k, \{\pi(\mathcal{N} \backslash \mathcal{A}) \cup \{ij\}\})$, $i, j \in \mathcal{A}$.

Suppose we have shown that no player leaves the negotiations before $\mathcal{A}$ has formed for $|\mathcal{A}| = r$. Showing that the same holds for $|\mathcal{A}| = r + 1$ inductively proves that the grand coalition forms when $|\mathcal{A}| = n$. Suppose by contradiction that at the implementation stage of Γ^{R} or Γ^{RR}, there is a set $J(\mathcal{A}) \neq \emptyset$, where for $j \in J(\mathcal{A})$ it is (weakly) optimal to terminate negotiations. By the previous inductive step we know that if any j leaves, $M_j = \mathcal{A} \backslash \{j\}$ forms (for $|\mathcal{A}| = 3$ the following follows from the discussion of $|\mathcal{A}| = 2$). Thus,

$$x_j(\mathcal{A}) = y(\mathcal{A}, i) = v(j, \{j, M_j\}), \qquad \forall j \in J(\mathcal{A}) \text{ and } i \in \mathcal{A}_{-j}. \tag{3.7}$$

We select a particular state, whose existence is guaranteed whenever $J(\mathcal{A}) \neq \emptyset$. Fix $j \in J(\mathcal{A})$ and consider some players who trigger a sequence of mergers which do not include j. Let $\mathcal{A}'$ be the resulting state. If $J(\mathcal{A}') \neq \emptyset$, choose $\mathcal{A}'$ to be the state under consideration. Suppose therefore $J(\mathcal{A}') = \emptyset$. We must have $x_j(\mathcal{A}) \leq x_j(\mathcal{A}')$, because the previous inductive steps imply that $v(j, \{j, M_j\})$ is still obtainable. Next, $x_j(\mathcal{A}) < x_j(\mathcal{A}')$ is only possible if there exists $\mathcal{A}''$ with $J(\mathcal{A}'') \neq \emptyset$, where $\mathcal{A}'' \neq \mathcal{A}, \mathcal{A}'$ is some state along the sequence of mergers. To see this, note that Positive Externalities imply $v(j, \{j, M_j\}) \geq v(j, \mathcal{A}')$ for all possible states $\mathcal{A}'$ and thus, if leaving is optimal at $\mathcal{A}$, the same will be true at $\mathcal{A}'$, unless there was some $i \in \mathcal{A}''$ who either left, or decided not to leave but leaving was a weakly optimal strategy. Hence, either $x_j(\mathcal{A}) = x_j(\mathcal{A}')$ or, if not, pick state $\mathcal{A}''$ to be the state under consideration and repeat the chain of arguments. Without loss of generality, choose $\mathcal{A}$ such that there is a j for which leaving is (weakly) optimal and all other agents do not want to merge.

Suppose now that $\sum_{j \in J(\mathcal{A})} v(j, \{j, M_j\}) + \sum_{i \notin J(\mathcal{A})} v(i, \mathcal{A}) < v(\mathcal{A})$. Consider $k \in J(\mathcal{A})$ proposing $(\mathcal{A}, t)$ with $t_j = x_j(\mathcal{A}) + \epsilon = v(j, \{j, M_j\}) + \epsilon$ for $j \in J(\mathcal{A})_{-k}$ and $t_i = y(\mathcal{A}, k) + \epsilon$ for $i \notin J(\mathcal{A})$, where $y(\mathcal{A}, k)$ is pinned down by (3.3). Clearly, this offer is accepted. This offer is also feasible, since we chose $\mathcal{A}$ such that all $i \notin J(\mathcal{A})$ neither leave the negotiations nor have an incentive to form other coalitions, and $t_i > v(i, \mathcal{A})$. Hence, once k is selected to be the next proposer, $(\mathcal{A}, t)$ is a profitable deviation from leaving. But then it is also not optimal to leave in the implementation phase, because $v(k, \{k, M_k\})$ is guaranteed and with a positive probability k will be the proposer. We conclude that $\sum_{j \in J(\mathcal{A})} v(j, \{j, M_j\}) + \sum_{i \notin J(\mathcal{A})} v(i, \mathcal{A}) \geq v(\mathcal{A})$. Moreover, since $v(i, \{i, M_i\}) \geq v(i, \mathcal{A})$,

$$\sum_{i \in \mathcal{A}} v(i, \{i, M_i\}) \geq v(\mathcal{A}). \tag{3.8}$$

We now use condition No Free Riding to arrive at a contradiction. Let $\bar{i}$ be the player identified with coalition M_i. By (3.6),

$$v(i, \{i, M_i\}) + v(\bar{i}, \{i, M_i\}) \leq v(\mathcal{A}), \qquad \forall i \in \mathcal{A}. \tag{3.9}$$

By No Free Riding, we have

$$|M_i|\left(v(i,\{i,M_i\}) - v(i,\mathcal{A})\right) < v(\bar{i},\{i,M_i\}) - \sum_{j\in M_i} v(j,\mathcal{A}), \quad \forall i \in \mathcal{A}. \tag{3.10}$$

Combining (3.9) and (3.10) one obtains $(|M_i|+1)v(i,\{i,M_i\}) - |M_i|v(i,\mathcal{A}) < v(\mathcal{A}) - \sum_{j\in M_i} v(j,\mathcal{A})$ for all $i \in \mathcal{A}$. Summing over all $i \in \mathcal{A}$ and noting that $|M_i| = |\mathcal{A}|-1$ for all $i \in \mathcal{A}$, it follows that $|\mathcal{A}|\sum_{i\in\mathcal{A}} v(i,\{i,M_i\}) - (|\mathcal{A}|-1)\sum_{i\in\mathcal{A}} v(i,\mathcal{A}) < |\mathcal{A}|v(\mathcal{A}) - (|\mathcal{A}|-1)\sum_{i\in\mathcal{A}} v(i,\mathcal{A})$. Hence,

$$\sum_{i\in\mathcal{A}} v(i,\{i,M_i\}) < v(\mathcal{A}). \tag{3.11}$$

Expressions (3.8) and (3.11) yield a contradiction. This completes the inductive step. □

In the next section there will be ample opportunity to explore the consequences of dropping No Free Riding. For now we stay in a world without free riding incentives but allow for negative externalities.

Example 2. $N = \{1,2,3\}$. Let the v be given by $v(1,2,3) = (z,z,0.6)$, $v(12,3) = (0,0.4)$, and $v(N) = 1$, with $z \in (0,0.2)$. Coalitional worths are 0 in all other coalition structures. Player 3 moves first.

For Γ^{NR} it can be shown that the grand coalition forms with equilibrium payoffs $y_i(\omega^0,3) = \delta 0.4/(1+\delta)$ for $i = 1,2$ and $x_3(\omega^0) = (1+\delta 0.2)/(1+\delta)$. Player 3 obtains strictly more than 0.6. On the other hand, in Γ^R players 1 and 2 can secure themselves an aggregate payoff of approximately 0.5 once they are asked to respond to player 3's offer. This is achieved by forming the two-player coalition $\{12\}$ and subsequently enter into negotiations with player 3. Anticipating this, player 3 leaves the negotiations at the start, enforcing the singleton coalition structure. Hence, Γ^R performs *worse* than Γ^{NR} in terms of efficiency.[22] Interestingly enough, Γ^{RR} predicts again the grand coalition! However, gains are distributed differently than in Γ^{NR} (player 3 earns approximately 0.5 and players 1 and 2 each approximately 0.25).

[22] To be sure, there are games with negative externalities for which renegotiation is efficiency-enhancing. Let $N = \{1,2,3\}$ and $v(i,j,k) = (0.1,0.1,0.1)$, $v(i,jk) = (0,0.7)$, $v(N) = 1$. It is easy to verify that the first proposer will propose a two-player coalition, followed by the grand coalition. Γ^{NR} is inefficient.

Example 2 is linked to the Ubiquitous Bad Partnership example of Gomes and Jehiel (2005). In their setting, coalitions can renege their contracts if all members agree to do so. Players may then have an incentive to form coalitions that are unprofitable in the short term, if such a move reduces payoffs of outside players even more. This allows to extract ransoms from outsiders who urge to move back to the more efficient state. However, in Example 2 the merging coalition loses more than the outsider whenever $z > 0.1$. It is the threat to level out bargaining power once the initiative is seized that hinders efficiency.

Since characteristic functions by definition abstract from free riding incentives, given Theorem 1 the following observation is hardly surprising.

Corollary 1. *Let v be a characteristic function that satisfies GCS. There exists $\hat{\delta} \in (0,1)$ such that for $\delta \geq \hat{\delta}$ $\Gamma^{RR}(v,\delta)$ and $\Gamma^{R}(v,\delta)$ result in the grand coalition.*

Proof. Consider a state at which $\mathcal{A} = \mathcal{N}$, i.e. all players are still active. For characteristic functions, $v(i,\mathcal{N})$ is guaranteed for all $i \in \mathcal{N}$. Assuming that $J(\mathcal{N}) \neq \emptyset$ and applying the same reasoning that lead to (3.8) in Theorem 1 implies $\sum_{i\in\mathcal{N}} v(i,\mathcal{N}) \geq v(\mathcal{N})$, a contradiction with GCS. $\square$

In absence of externalities, GCS is sufficient to obtain efficiency when renegotiation is possible. This result is intimately connected to Bloch and Gomes (2006), where efficiency is also guaranteed, if there are no externalities. We can, however, say something more about the case when renegotiation is not possible.

Core. The core $C(N,v)$ of a characteristic function v consists of all allocations z for which $\sum_{i\in N} z_i = v(N)$ and $S \subset N \Rightarrow \sum_{i\in S} z_i \geq v(S)$.

Denote the interior of the Core by $C^\circ(N,v)$. Interestingly, excluding a special case to be made precise in the following, the grand coalition forms in Γ^{NR} if and only if $C^\circ(N,v) \neq \emptyset$. A direct implication of this is that in the models with renegotiation the grand coalition forms immediately only if the interior of the Core is non-empty. Otherwise gradualism should be observed.

Theorem 2. *Let v be a characteristic function. There exists $\hat{\delta} \in (0,1)$ such that for $\delta \geq \hat{\delta}$ the following holds for $\Gamma^{NR}(v,\delta)$. If $C^\circ(N,v) \neq \emptyset$ the grand coalition is the unique* *equilibrium structure. If $C(N,v)=\emptyset$ the grand coalition is not an equilibrium structure. In the remaining case $C^{\circ}(N,v)=\emptyset$ and $C(N,v)\neq\emptyset$, the grand coalition does not form if $v(N)\neq\sum_{j\in N}x_j(\omega^0,\mathcal{P}(\mathcal{B})\backslash N)$ and is weakly optimal otherwise.*

Proof. For the following, note that $y_j(\omega,i)$ is independent $k\in S_{-ij}$ as long as $t_k^S\geq y_k(\omega,i)$ for all k.

We first show that a non-empty strict core leads to the grand coalition. Suppose the grand coalition does not form. Let (S_i,t^{S_i}) with $S_i\subset N$ be the proposal that is optimal for i as the first proposer, but is not allowed to offer to the grand coalition. Let Ψ be the set of all distinct S_i. A player i earns at most (he may earn less if he first passes the initiative to a different player) $x_i(\omega^0,\mathcal{P}(\mathcal{B})\backslash N)=v(S_i)-\sum_{j\in S_{i,-i}}y_j(\omega^0,i)$. It follows that

$$v(N)=\sum_{j\in S_i}z_j>\sum_{j\in S_i}x_j(\omega^0,\mathcal{P}(\mathcal{B})\backslash N)=\sum_{j\in S_i}x_j(\omega^0)\quad\forall S_i\in\Psi. \tag{3.12}$$

The inequality holds, because i) by (3.1) and (3.2), $\delta\sum_{j\in S_i}x_j(\omega^0,\mathcal{P}(\mathcal{B})\backslash N)\leq v(S_i)$ and ii) $C^{\circ}(N,v)\neq\emptyset$ implies that there exists a vector of payoffs z such that $\sum_{j\in S_i}z_j>v(S_i)$ for all $S_i\in\Psi$ and thus there also exists $\delta\geq\hat{\delta}$ such that $\delta\sum_{j\in S_i}z_j>v(S_i)$. The last equality holds, because the grand coalition is assumed to not be strictly optimal for any player (Lemma 3). Let i's best offer to N be (N,t^N). Hence i earns $x_i(\omega^0,N)=v(N)-\sum_{j\in N_{-i}}y_j(\omega^0,i)$. Because $x_j(\omega^0)\geq y_j(\omega^0,i)$, it follows from (3.12) that $x_i(\omega^0,N)>x_i(\omega^0,\mathcal{P}(\mathcal{B})\backslash N)$. Hence, (S_i,t^{S_i}) is not optimal.

Assume now that the core is empty and the grand coalition forms. There exists S_i such that $\sum_{j\in S_i}z_j<v(S_i)$. By Lemma 3, it is optimal to offer the grand coalition for all players and thus also for $i\in S_i$. It follows that

$$v(S_i)=x_i(\omega^0,\mathcal{P}(\mathcal{B})\backslash N)+\sum_{j\in S_{i,-i}}y_j(\omega^0,i)>\sum_{j\in S_i}x_j(\omega^0,N)=\sum_{j\in S_i}x_j(\omega^0). \tag{3.13}$$

The inequality holds, because i) by (3.1) and (3.2), $\delta\sum_{j\in S_i}x_j(\omega^0,N)\leq\sum_{j\in S_i}z_j$ and ii) there exists $\delta\geq\hat{\delta}$ such that $\sum_{j\in S_i}z_j<\delta v(S_i)$. The last equality holds, because N is optimal. Since $x_j(\omega^0)\geq y_j(\omega^0,i)$, it follows that $x_j(\omega^0,\mathcal{P}(\mathcal{B})\backslash N)>x_j(\omega^0,N)$. Proposal

(N, t^N) is not optimal.

For the remaining case, $C(N, v) \neq \emptyset$ requires $\sum_{j \in S_i} z_j \geq v(S_i)$ for all $S_i \in \Psi$ and $C^\circ(N, v) = \emptyset$ requires $\sum_{S_i \in \Psi} \sum_{j \in S_i} z_j \leq \sum_{S_i \in \Psi} v(S_i)$. Thus, the vector of payoffs z induced by (N, t^N) satisfies

$$\sum_{j \in S_i} z_j = v(S_i) \quad \forall S_i. \tag{3.14}$$

We now show that N can only form if z satisfies $z_j = x_j(\omega^0, \mathcal{P}(\mathcal{B}) \backslash N)$ for all $j \in N$. This implies that N is at best weakly optimal. Moreover, it implies that $C^\circ(N, v) = \emptyset$, N is optimal if and only if $v(N) = \sum_{j \in N} x_j(\omega^0, \mathcal{P}(\mathcal{B}) \backslash N)$. To prove this, suppose N is optimal for i and suppose there are $j, k \in S_i$ such that $x_j(\omega^0, \mathcal{P}(\mathcal{B}_{-k})) < x_j(\omega^0, \mathcal{P}(\mathcal{B}) \backslash N)$. By Lemma 3, we can pick a player not in S_i who optimally offers to N with $t_l^N \geq x_l(\omega^0, \mathcal{P}(\mathcal{B}) \backslash N)$ for all $l \in S_i$. But $\sum_{j \in S_i} x_j(\omega^0, \mathcal{P}(\mathcal{B}) \backslash N) > v(S_i) = x_k(\omega^0) + \sum_{j \in S_{i,-k}} y_k(\omega^0, \mathcal{P}(\mathcal{B}_{-k}))$, because we know that $x_j(\omega^0, \mathcal{P}(\mathcal{B}_{-k})) < x_j(\omega^0, \mathcal{P}(\mathcal{B}) \backslash N)$ for at least one $j \in S_i$ and $x_j(\omega^0, \mathcal{P}(\mathcal{B}_{-k})) \leq x_j(\omega^0, \mathcal{P}(\mathcal{B}) \backslash N)$ for all $j \in S_i$. This contradicts (3.14). Thus, if N forms, $x_j(\omega^0, \mathcal{P}(\mathcal{B}_{-k})) = x_j(\omega^0, \mathcal{P}(\mathcal{B}) \backslash N)$ for all $j, k \in S_i$ and all $S_i \in \Psi$. Thus, $t_j^N = z_j = x_j(\omega^0, \mathcal{P}(\mathcal{B}) \backslash N)$. □

This result contrasts with the previous literature.[23] The difference stems from the fact that in our model only counter-proposals are time-consuming. We revisit the Employer-Employee Game of Chatterjee et al. (1993) to highlight this point.

Example 3. $N = \{1, 2, 3\}$. Let v be given by $v(1, 2, 3) = (0, 0, 0)$, $v(12, 3) = v(13, 2) = (1, 0)$, $v(1, 23) = (0, \epsilon)$, and $v(N) = 1 + \mu$, $0 < \mu < 0.5$. This game has a non-empty strict Core.

Chatterjee et al. show that when agreements are non-renegotiable and there are time costs after every rejection, the equilibrium coalition structures for large δ are $\{12, 3\}$ or $\{13, 2\}$. On the other hand, Theorem 2 implies that the unique equilibrium for Γ^{NR} is the grand coalition. To illustrate the difference, suppose player 3 makes the first proposal. In Chatterjee et al., player 3 chooses between proposal $(\{13\}, t)$ with $t_1 = \delta/(1 + \delta)$ and

[23] Seidmann and Winter (1998) and Chatterjee et al. (1993) find that a non-empty (strict) Core does not in general imply efficiency.

proposal $(\{123\}, t')$ with $t_1' = t_2' = \delta(1+\mu)/(1+2\delta)$. For large δ, player 3 prefers the two-player coalition (yielding a payoff of t_1/δ) to the grand coalition (yielding a payoff of t_1'/δ). If time costs are only incurred for counter-proposals, *player 1 would reject player 3's offer* $\delta/(1+\delta)$ and subsequently offer to player 2. Player 3 therefore compares proposal $(\{13\}, t)$ with $t_1 = 1/(1+\delta)$ to proposal $(\{123\}, t')$ with $t_1' = 1/1+\delta$ and $t_2' = \delta(1+\mu-t_1')/(1+\delta)$. The grand coalition is strictly preferred. Note how it is impossible for player 3 to force time costs upon player 1, because player 1's option to sign an agreement with player 2 functions as endogenous outside option.

3.4.4 Conceding bargaining power

So far renegotiation has been discussed in settings that satisfy No Free Riding. Strikingly, we show next that in the presence of free riding incentives, renegotiation as in Γ^R is unable to promote cooperation.

Theorem 3. *Let v satisfy Positive Externalities and Free Riding. For symmetric games, let v satisfy Symmetric Free Riding. There exists $\hat{\delta} \in (0,1)$ such that for $\delta \geq \hat{\delta}$, $\Gamma^R(v,\delta)$ and $\Gamma^{NR}(v,\delta)$ have the same set of equilibrium coalition structures (implying $\Gamma^R(v,\delta) \sim \Gamma^{NR}(v,\delta)$) with the same distribution of payoffs.*

Proof. We start with the following Lemma.

Lemma 5. *Consider $\Gamma^R(v,\delta)$ at state ω. Suppose Positive Externalities and Free Riding holds. For symmetric games, suppose symmetric Free Riding holds. There exists $\hat{\delta} \in (0,1)$ such that for $\delta \geq \hat{\delta}$, if coalition S is part of the equilibrium coalition structure, then for each $i \in S$ it is optimal to make an acceptable proposal to S (referred to as one-step proposal) at ω.*

Assume S is part of the equilibrium coalition structure in $\Gamma^R(v,\delta)$ and the current state is ω. Lemma 5 shows that proposing S is optimal for all $i \in S$ already at ω. Hence, no subset of S can increase its payoff by forming intermediate coalitions. Moreover, the behavior of a player $i \notin S$ can only depend on whether S forms in one step or gradually if i is indifferent between some optimal proposals. Hence, all equilibrium outcomes also exist in $\Gamma^{NR}(v, \delta)$ for the different optimal behaviors of i. Given S forms, we can abstract from behavior of $i \notin S$. It follows that gradual build ups of S in $\Gamma^{R}(v, \delta)$ do not lead to equilibrium outcomes not also present in $\Gamma^{NR}(v, \delta)$. Moreover, Lemma 5 also proves the reverse. Suppose S does not form in $\Gamma^{R}(v, \delta)$ but does in $\Gamma^{NR}(v, \delta)$. By stationarity, one $i \in S$ must have a strictly better proposal to $S' \neq S$ that is accepted. Because one-step equilibrium outcomes in $\Gamma^{R}(v, \delta)$ also exist in $\Gamma^{NR}(v, \delta)$, coalition S' is part of multiple step coalition formation process. But this is excluded by Lemma 5.

We now prove Lemma 5. Let $\mathcal{S}$ be the reduction of $\mathcal{A}$ to the players eventually forming S. Let $v(i, \mathcal{S})$ be the value of coalition $i \in \mathcal{S}$ (taking the rest of the coalition structure as given). It needs to be shown that S cannot form if the one-step proposal is not optimal. The latter implies that there is a k such that

$$x_k(\mathcal{S}) > v(S) - \sum_{i \in \mathcal{S}, i \neq k} y_i(\mathcal{S}, k). \tag{3.15}$$

Thus, k proposes to a proper subset of $\mathcal{S}$ and the proposal is accepted. Collect k in set R. Let $\mathcal{S}_k$ be the resulting set of active players. Denote by $M(\mathcal{S}_k)$ the set of players who merge from $\mathcal{S}$ to $\mathcal{S}_k$. If (3.15) holds with equality at $\mathcal{S}_k$ (if not repeat the same argument until it is true), and since $x_k(\mathcal{S}_k) \geq x_k(\mathcal{S})$ for the proposer and all other $j \in M(\mathcal{S}_k)$ earn $y_j(\mathcal{S}, k)$, it must be that

$$x_i(\mathcal{S}_k) < x_i(\mathcal{S}) \quad \text{for at least one } i \notin M(\mathcal{S}_k). \tag{3.16}$$

Pick a player k' for whom (3.16) is true and consider the initial state $\mathcal{S}$. There must be a merger $M(\mathcal{S}_{k'}) \subset \mathcal{S}$, $M(\mathcal{S}_{k'}) \neq M(\mathcal{S}_k)$ which k' is able to induce with positive probability such that $x_{k'}(\mathcal{S}_{k'}) \geq x_{k'}(\mathcal{S})$. Collect k' in set R. If (3.15) holds with equality at $\mathcal{S}_{k'}$ (if not repeat the above reasoning for $\mathcal{S} = \mathcal{S}_{k'}$ until it is true), there must be a $l \neq k'$ for which (3.16) holds for $M(\mathcal{S}_{k'})$. Repeat this process until the first instance at which $l \in R$, which is guaranteed if the number of players is finite. Hence, there exists a set R such that (i) $|R| > 1$, (ii) (3.15) holds with equality at $\mathcal{S}_k$ for all $k \in R$, and (iii) (3.16) holds for all $k \in R$ for (at least) one $M(\mathcal{S}_{k'})$, $k' \in R$. This implies $x_k(\mathcal{S}) > x_k(\mathcal{S}_{k'})$

for all $k \in R$ and one $k' \in R$, where $x_k(\mathcal{S})$ and $x_k(\mathcal{S}_{k'})$ are given by

$$x_k(\mathcal{S}) = v(k, \mathcal{S}_k) - \sum_{i \in M(\mathcal{S}_k), i \neq k} y_i(\mathcal{S}, k) + \alpha_k^{\mathcal{S}_k} \left[v(S) - \sum_{i \in \mathcal{S}_k} v(i, \mathcal{S}_k) \right], \tag{3.17}$$

$$x_k(\mathcal{S}_{k'}) = v(k, \mathcal{S}_{k'}) + \alpha_k^{\mathcal{S}_{k'}} \left[v(S) - \sum_{i \in \mathcal{S}_{k'}} v(i, \mathcal{S}_{k'}) \right]. \tag{3.18}$$

Note that (3.17) reflects the fact that if k proposes to $M(\mathcal{S}_k)$ and the proposal is accepted, k controls the coalition of worth $v(k, \mathcal{S}_k)$, pays the acceptors their equilibrium response values, and obtains some share $\alpha_k^{\mathcal{S}_k}$ of the surplus from moving to coalition $\mathcal{S}$ in the next step. In (3.18), $v(k, \mathcal{S}_{k'})$ is guaranteed by (3.1) and Positive Externalities. Moreover, it can be shown that $k \in K(i, \mathcal{S}_k)$ for all $i \in \mathcal{S}_k$.[24]

Since $x_k(\mathcal{S}_k, i) = v(S) - \sum_{i \in K(k, \mathcal{S}_k)_{-k}} v(i, \mathcal{S}_k) - \sum_{i \in \overline{K}(k, \mathcal{S}_k)} \overline{y}(k, \mathcal{S}_k) \rightarrow v(\mathcal{S}_k)$ as $\delta \rightarrow 1$, we get $\alpha_k^{\mathcal{S}_k} \rightarrow 0$. Summing (3.17) and (3.18) over all $k \in R$, yields $\sum_{k \in R} \left(v(k, \mathcal{S}_k) - \sum_{i \in M(\mathcal{S}_k), i \neq k} y_i(\mathcal{S}, k) \right) > \sum_{k \in R} v(k, \mathcal{S}_{k'})$. Positive Externalities imply $y_i(\mathcal{S}, k) \geq v(i, \mathcal{S})$ for all i. Thus,

$$\sum_{k \in R} \left(v(k, \mathcal{S}_k) - \sum_{i \in M(\mathcal{S}_k), i \neq k} v(i, \mathcal{S}) \right) > \sum_{k \in R} v(k, \mathcal{S}_{k'}). \tag{3.19}$$

Next, condition Free Riding gives

$$v(k, \mathcal{S}_k) - \sum_{i \in M(\mathcal{S}_k)} v(i, \mathcal{S}) < v(k', \mathcal{S}_k) - v(k', \mathcal{S}) \quad \forall k \in R, k' \notin M(\mathcal{S}_k).$$

[24] Let $R = \{k' : x_{k'}(\mathcal{S}_k) < x_{k'}(\mathcal{S})\}$. For each $k' \in R$, $\exists\, j \in M(S_k) \cap M(S_{k'})$. Otherwise k' at S_k still has an outside option $v(k', \mathcal{S}_{k'} \backslash \{i : i \in M(\mathcal{S}_k)\} \cup M(\mathcal{S}_k)) \geq v(k', \mathcal{S}_k)$ due to Positive Externalities. This implies that the players $j \in M(S_k) \cap M(S_{k'})$ will extract all the gains $\Delta_R \equiv \sum_{k' \in R}(x_{k'}(\mathcal{S}) - x_{k'}(\mathcal{S}_k))$. Also note that $i \in \overline{K}(k, \mathcal{S}_k) \Rightarrow i \in \overline{K}(k, \mathcal{S})$. Thus for large δ, $(v(S) - \sum_{i \in K(k, \mathcal{S}_k)_{-k}} x(i, S_k) - \Delta_R)/((1 + \delta(|M(\mathcal{S}_k)| - 1))(1 + \delta(|\mathcal{S}_k| - |K(k, \mathcal{S}_k)| - 1))) < (v(S) - \sum_{i \in K(k, \mathcal{S})_{-k}} x(i, S))/(1 + \delta(|\mathcal{S}| - |K(k, \mathcal{S})| - 1))$, where we also used $|\mathcal{S}| = |\mathcal{S}_k| + |M(\mathcal{S}_k)| - 1$. The LHS is the maximum payoff obtained by k if $k \in \overline{K}(i, \mathcal{S}_k)$. The RHS is the payoff k obtains if he makes an acceptable proposal to $\mathcal{S}$ at $\mathcal{S}$. Hence, the merger $M(\mathcal{S}_k)$ would not be optimal.

Summing over all $k \in R$ and noting that $\sum_{k' \in R} v(k', \mathcal{S}) = \sum_{k \in R} v(k, \mathcal{S})$,

$$\sum_{k \in R} \left(v(k, \mathcal{S}_k) - \sum_{i \in M(\mathcal{S}_k), i \neq k} v(i, \mathcal{S}) \right) < \sum_{k \in R} v(k, \mathcal{S}_{k'}),$$

a contradiction to (3.19).

For symmetric games, Symmetric Free Riding guarantees each player a payoff of at least $\underline{x} = \delta v(S)/(1 + \delta(|S| - 1))$, which is obtained by staying a singleton (if some other players merge, the payoff of the singleton will increase). Consider player j who controls the (weakly) largest coalition S_j at $\mathcal{S}$ and suppose j proposes to form $\mathcal{S}$ (as a respondent he earns weakly less). The maximum payoff player j can obtain is $v(S)/(1+\delta(|\mathcal{S}|-1)) - (|S_j| - 1)\delta v(S)/(1 + \delta(|S| - 1))$, because by expression (3.3), j cannot extract more than $v(S)/(1+\delta(|\mathcal{S}|-1))$ at $\mathcal{S}$ and each singleton in S_j has earned at least $\underline{\mathrm{x}}$. As δ approaches 1, the latter expression only exceeds $\underline{x}$ if $|\mathcal{S}| > |S_j||\mathcal{S}|$, a contradiction. □

Notice that we do not need GCS for Theorem 3. The key observation is that players cannot extract rents from others by following a gradual formation process. Remarkably, in environments with free riding incentives abstracting from renegotiation as in Γ^R is without loss of generality.[25] This raises important questions. For instance, are the repeated international meetings and efforts to agree on joint measures against global warming in vain? In general, should we expect gradualism to play no role in games with free riding incentives?

Model Γ^{RR} is motivated by the negative result of Theorem 3. It reestablishes the importance of gradualism for games with free riding incentives.

Example 4. $N = \{1, 2, 3\}$. The partition function is defined by $v(i, j, k) = (0, 0, 0)$, $v(i, jk) = (z, \epsilon)$, and $v(N) = 1$ for $i, j, k = 1, 2, 3$, ϵ small, and $z \in (1/3, 1)$.

This example satisfies Symmetric Free Riding. Without the possibility to renegotiate, the initial proposer decides to leave immediately. Since the remainder prefers to merge,

[25] The equivalence between Γ^{NR} and Γ^R also provides a valuable short cut when searching for equilibrium outcomes.

the singleton obtains a payoff of z. Any offer to the grand coalition would be accepted only if each responder obtains no less than z, which implies that the proposer earns strictly less than z. By Theorem 3, Γ^R predicts the same outcome. What happens in Γ^{RR}? Again, the first-mover, say player 1, will sign the singleton contract in the first negotiation round. However, before player 1 gets to the implementation phase, players 2 and 3 form a coalition. This eliminates player 1's incentive to leave the negotiations, as he can capture some of the gains set free when moving to the grand coalition. Efficiency is restored. Player 1 obtains approximately $\max\{0.5, z\}$, players 2 and 3 each earn approximately $(1 - \max\{0.5, z\})/2$.

The difference between Example 1 (in which Γ^R is efficient) and Example 4 is the motivation to form the two-player coalition in the first negotiation round. In Example 1, the initial mover is part of the two-player coalition, which forms to increase its bargaining power in subsequent negotiations. In Example 4, the initial mover is not part of the two-player coalition, which forms to concede bargaining power to the initial mover. The crucial point is that before player 1 implements, players 2 and 3 can credibly commit to not make use of their free riding possibilities, and they are willing to do so because player 1 will be the one who moves first in the implementation phase.

Unfortunately, strong free riding externalities may prevent efficiency also in Γ^{RR}.

Example 5. Let $N = \{1, 2, 3, 4\}$ and the partition function be given by $v(i, j, kl) = (0.4, 0.4, 0)$, $v(i, jkl) = (0.55, 0.4 + \epsilon)$, $v(N) = 1$ for $i, j, k, l = 1, 2, 3, 4$ and ϵ small. All other coalition structures result in a payoff of 0 for everyone.

In Example 5, players benefit if others form coalitions, but only as long as they themselves remain singletons. This could represent a setting where the formation of a coalition entails high fixed costs. For similar reasons as in the previous example, in Γ^{NR} and Γ^R the equilibrium coalition structure is $\{i, j, kl\}$. Are players willing to concede bargaining power in Γ^{RR}? We describe equilibrium behavior. In round 1, the first two proposers, say players 1 and 2, sign the singleton contract, players 3 and 4 form a coalition. In the implementation phase, player 1 leaves, predicting correctly that player 2 remains active to obtain some of the gains obtained from the merger to the three-player coalition in round

2. The final coalition structure is thus $\{1, 234\}$. The possibility to concede bargaining power helps to some extent, but full efficiency is not obtained. The reason is that the second proposer prefers free riding on players 3 and 4 to inducing the grand coalition. The fact that Example 5 features four players is no coincidence.

Corollary 2. *If* $n \leq 3$, Γ^{RR} *is efficient.*

Proof. Follows by Example 4 and exhaustively discussing all cases. See Appendix. ☐

These insights raise the question whether it is possible to rank the different negotiation protocols in terms of efficiency. Example 2 has already shown that Γ^{NR} and Γ^{R} cannot be ranked in general. Perhaps surprisingly, the same conclusion applies to the comparison between Γ^{NR} and Γ^{RR}.[26] Moreover, in all examples we have discussed, $\Gamma^{RR}(v, \delta) \succsim \Gamma^{R}(v, \delta)$. We conjecture that this holds in general, but leave the question for future work.

3.5 Public Goods

This section applies our findings to a model of public good provision discussed in Ray and Vohra (2001). There are n symmetric regions negotiating over the level of pollution control z a region should undertake. Reducing emissions involves a private cost $c(z)$, taken to be increasing and strictly convex in z. Let $Z = \sum_{i=1}^{n} z_i$ be the total amount of pollution control. The payoff to a region with control level z is

$$Z - c(z) \tag{3.20}$$

Because regions are symmetric, the pollution control level will only depend on the size of coalitions. Because of the strict convexity of $c(\cdot)$, $z_i = z_s$ for all players that are part of a coalition S of size s. The payoff of S is thus $s[sz_s - c(z_s) + Z_{-S}]$, where Z_{-S} denotes the aggregate pollution control of all other coalitions. Observe that the optimal choice of z_s is independent of the behavior of other coalitions, because of the linearity of the external

[26] The counter-example involves 6 players and features both, mergers for which Free Riding holds and mergers for which No Free Riding holds. The example is available from the author upon request.

effects (of course, payoffs do depend on the actions of other regions). A coalition of size s solves

$$\max_{z_s} sz_s - c(z_s) \tag{3.21}$$

Binding agreements allow players to internalize benefits of pollution control within but not across coalitions. To what extent do players make use of this possibility? For negotiation model Γ^{NR}, Ray and Vohra show that efficiency is generally not attained.[27] Interestingly, Theorem 3 implies that the same is true for Γ^{R}.

Corollary 3. *In the public goods model introduced above, Γ^{NR} and Γ^{R} have the same (unique) equilibrium outcome.*

Proof. Note that according to (3.20), players' payoffs only differ in the cost of pollution control. Solving (3.21) shows that members of larger coalitions undertake larger efforts. It follows that smaller coalitions enjoy higher per member payoffs than larger coalitions. Symmetric Free Riding is satisfied. Theorem 3 applies. Ray and Vohra (2001) show that the outcome is unique. $\square$

In contrast, the possibility to concede bargaining power (as made possible in Γ^{RR}) affects predictions.

Example 6. Consider the public goods model with $n = 3$. Let $c(z) = z^3/3$. It follows that $z_s = \sqrt{s}$ and the aggregate payoff of a coalition S of size s is $s\,[sz_s - 1/3z_s^3 + Z_{-S}] = s\left[2/3s^{2/3} + Z_{-S}\right]$. The partition function is thus given by $v(1,2,3) = (2.\overline{6}, 2.\overline{6}, 2.\overline{6})$, $v(i,jk) = (2\sqrt{2} + 2/3, 2(1 + 2\sqrt{8}/3))$ where $i,j,k = 1,2,3$, and $v(123) = 6\sqrt{3}$.

Ray (2007) discusses this example for Γ^{NR}. We omit a detailed discussion, as in terms of incentives, Example 6 is identical to Example 4. The reader can easily convince herself that the equilibrium coalition structure in Γ^{NR} and therefore also in Γ^{R} is $\{i, jk\}$. On the other hand, Γ^{RR} leads to the grand coalition. Ray and Vohra (2001) provide bounds on the maximal amount of inefficiency observable in Γ^{NR} (as the number of players increases).

[27] Ray and Vohra (2001)'s model differs from Γ^{NR} in the way proposals are made, but their arguments directly apply to Γ^{NR}.

It would be interesting to see to what extent Γ^{RR} shifts these bounds towards the efficient outcome. We were not able to solve this question.

3.6 Conclusion

This paper studies coalition formation with endogenous renegotiation. In accordance with the Coase Theorem, we find that renegotiation leads to efficiency even in the presence of widespread externalities. Only if externalities involve free riding incentives, the fully cooperative outcome may not be reached. We also propose an extension of the coalition formation model that uncovers the incentive to concede bargaining power as a novel explanation for gradualism. On the methodological side, it is shown that a bargaining model in which time costs are only incurred for counter-proposals allows to isolate the effects of externalities on equilibrium outcomes.

We provide a set of testable predictions. Does the grand coalition form in absence of free riding incentives? Is the strict Core a good predictor of outcomes without renegotiation? In games of public good provision, does renegotiation indeed only play a limited role? There is a vast empirical literature on coalition formation in international negotiations on environmental or trade issues, but only few studies make the link to the theoretical coalition formation literature.[28] We believe that there is also a role for experiments on coalition formation. For instance, by focussing on externalities and the degree of renegotiation, do we miss some other important features of a bargaining environment?

This study has not discussed incomplete information, which should be expected to play a role in explaining gradualism. Incomplete information in multilateral bargaining is difficult to analyze, because of the multiple ways information may get revealed in the process of coalition formation. Three channels that come to mind are signalling and screening via proposals, learning by observing the evolution of cooperation, and information sharing within coalitions.

We have concluded that renegotiation and binding agreements cannot fully eliminate

[28] An exception is Esteban et al. (2012).

inefficiencies if actions are irreversible. In contrast, for reversible actions and ongoing negotiations, Hyndman and Ray (2007) show that as long as the grand coalition is efficient, it is guaranteed to form for arbitrary externalities. Ultimately, it would be insightful to have a model of costly reversible actions. By subsuming reversible and irreversible actions such a model would allow to tackle new questions. For instance, is the reluctancy of many countries to substantially curb carbon dioxide emissions part of a reversible process in which players try to extract rents, or are incentives such that renegotiation will be unable to eventually bring about cooperation?

Appendix

3.A Proofs

3.A.1 Proposition 4

The proof adapts the proofs of existence in Ray and Vohra (1999) and Bloch and Gomes (2006). For a given strategy profile, let $\phi_i^j(\omega)$ denote the continuation value of player i at the bargaining phase when the state is ω and the proposer is player j. Note that $\phi_i^i(\omega) = x_i(\omega)$. Let $\varphi_i(\omega)$ denote the continuation value of player i at the implementation phase.

Let $\sigma^2 = (\sigma_i^2)_{i \in \mathcal{A}}$ be a strategy profile at the implementation phase. Clearly, σ_i^2 is a probability distribution over $\{\text{implement}, \text{remain}\}$ for each $\omega \in \Omega$. In equilibrium $\sigma_i^2(\omega)$ maximizes the continuation value $\phi_i^j(\omega^2)$, where ω^2 is the state after the implementation phase as implied by $\sigma^2(\omega)$.

We describe the optimal behavior of proposers and respondents in the bargaining phase when the state is ω. Let $\Pi_i(\mathcal{B})$ be the set of all possible coalitions containing player i. Let Σ_i^1 be the set of probability distributions over $M_i = (\Pi_i(\mathcal{B}), (\{j\})_{j \in \mathcal{N} \setminus \{i\}})$. This means that i can either make a proposal to a set of active coalitions that includes itself or make an unacceptable proposal, say to player j. Let $\sigma_i^1(S, \omega)$ be the probability with which i makes an acceptable proposal to $S \in \Pi_i(\mathcal{B})$ at ω. Similarly, $\sigma_i^1(\{j\}, \omega)$ is the probability with which i makes an unacceptable offer.

Define $\Sigma^1 = \prod_{i \in \mathcal{B}} \Sigma_i^1$ and fix a proposer strategy profile $\sigma^1 \in \Sigma$. This profile describes for all players their proposer choices for each possible state. Let $\alpha_S(\omega)$ be the probability distribution induced by σ^1 over the set of possible states $\omega^1 \in \Omega$ at the end of the bargaining phase. Thus, $\sigma^1(\omega)$ and $\alpha^S(\omega)$ fix a vector of expected continuation values $\varphi(\omega^1)$.

According to (3.1) a respondent j's minimal acceptable offer is $y(\omega, i) \in [\delta\phi_j^j, \phi_j^j]$. A proposer i in the bargaining phase has two options. First, i can name a coalition $S \in \Pi_i(\mathcal{B})$ and make an acceptable proposal (S, t). If the proposal is accepted, it must be given by

$$S \in \arg\max_{R\in\Pi_i(\mathcal{B})} \sum_{\omega^1\in\Omega} \alpha_R(\omega^1)\varphi_i(\omega^1) - \sum_{j\in R; j\neq i} t_j \tag{3.22}$$

$$t_j = y_j(\omega, i) \qquad \text{for all } j \in S, j \neq i \tag{3.23}$$

Denote by $g(S, x, \varphi)$ the maximal payoff i can obtain by solving this problem.

Second, i can make an unacceptable proposal to j. For a fixed i, the value player i receives when player j proposes is

$$\phi_i^j(\phi, \varphi, \sigma^1) = W_i^j + \sum_{k\neq j} \sigma_j^1(\{k\})\phi_i^k(\phi, \varphi, \sigma^1)$$

for all j and k, where $W_i^i \equiv \sum_{S\in\Pi_i(\mathcal{B})} \sigma_i^1(S)g(S, \phi, \varphi)$ and for $j \neq i$,

$$W_i^j \equiv y_i(\omega, j)\left[\sum_{S\in\Pi_j(\mathcal{B}); i\in S} \sigma_j^1(S)\right] + \sum_{l\in\mathcal{B}} \rho(l, \omega) \sum_{S\in\Pi_j(\mathcal{B}); i\notin S} \sigma_j^1(S)\phi_i^l(\phi, \varphi, \sigma^1),$$

where player l is determined by ρ. Ray and Vohra (1999) show that ϕ_i^j is continuous in σ^1, ϕ and φ for all j. Now define a function on $\Phi \times \Phi \times \Sigma^1 \times \Sigma_i^1$ by

$$\phi_i(\phi, \varphi, \sigma_{-i}^1, \sigma_i^{1'}) \equiv \sum_{S\in\Pi_i(\mathcal{B})} \sigma_i^{1'}(S)g_i(S, \phi, \varphi) + \sum_{j\neq i} \sigma_i^{1'}(\{j\})\phi_j^i(\sigma^1, \phi, \varphi) \tag{3.24}$$

and maximize with respect to $\sigma_i^{1'} \in \Sigma_i$. Let the set of maximizers to this problem be $\hat{\sigma}_i^1(\phi, \varphi, \sigma^1)$. The implied payoff is denoted by $\hat{\phi}_i^1(\phi, \varphi, \sigma^1)$. Using the maximum theorem and the fact that $\phi_i(\phi, \varphi, \sigma_{-i}^1, \sigma_i^{1'})$ is continuous, one can see that $\hat{\phi}_i^1(\phi, \varphi, \sigma^1)$

is a continuous function and that $\hat{\sigma}_i^1(\phi, \varphi, \sigma^1)$ is a convex-valued, upper hemicontinuous correspondence.

Since $v(\{i\}, \pi) > 0$ for all i and $\pi \in \Pi$, for all $(\phi, \varphi, \sigma^1, \sigma^2) \in \Phi \times \Phi \times \Sigma^1 \times \Sigma^2$, $\phi_i^1(\phi, \varphi, \sigma^1) \in [0, v(N)]$ for all i. Thus $\prod \phi_i^1$ maps from $\Phi \times \Phi \times \Sigma^1 \times \Sigma^2$ into Φ.

Define now a correspondence $F : \Phi \times \Phi \times \Sigma^1 \times \Sigma^2 \rightarrow\rightarrow \Phi \times \Phi \times \Sigma^1 \times \Sigma^2$. A fixed-point of F is an equilibrium (we still need to describe the behavior of respondents). Recall that Φ is a closed, convex interval of a finite-dimensional Euclidean space. Σ^1 is the set of proposers' strategies σ^1 in the bargaining phase. We have seen that $\sigma_i^1(\cdot, \omega)$ is a probability distribution over the finite set $\{M_i\}_{i \in B}$. Σ^2 is the set of strategies σ^2 at the implementation stage and σ_i^2 is a probability distribution over a binary choice for each state. Both Σ^1 and Σ^2 are thus convex and compact subsets of a finite-dimensional Euclidean space. Thus, $Z = \Phi \times \Phi \times \Sigma^1 \times \Sigma^2$ is a compact and convex subset of a finite dimensional Euclidean space. Moreover, $F(Z) \subset Z$. $F(z)$ is a convex and non-empty set for all $z \in Z$. The graph of F is closed. Kakutani's fixed point theorem guarantees that a fixed point exists. It is now possible to construct a stationary equilibrium using the derived fix point. This is done as in Ray and Vohra (p. 311f.), except that the argument has to be repeated for each negotiation round.

3.A.2 Proof of Corollary 2

Suppose players i, j, k enter the implementation phase in $\tau = 1$ in coalition structure $\{ijk\}$. The grand coalition must be efficient, for if $v(i, \{ij, k\}) + v(k, \{ij, k\}) > v(ijk)$ we know that one player has not received its equilibrium response value. Suppose players enter the implementation phase in coalition structure $\{ij, k\}$. The two-player coalition ij forms only if $v(i, \{ij, k\}) \geq v(i, \{i, j, k\}) + v(j, \{i, j, k\})$. Moreover, if the grand coalition is efficient, no player implements because ρ select both coalitions with positive probability and outside options are safe. Suppose therefore, players enter the implementation phase as singletons. If the singleton structure is efficient, all players leave in $\tau = 1$. If structure $\{ij, k\}$ is efficient, the two-player coalition will form in one of the future rounds and it is unimportant whether k leaves before this happens. Suppose the grand coalition is efficient. If there is a two-player coalition ij that can extract some payoff from k by first inducing $\{ij, k\}$, this will be done in a future round. Moreover, the third player does not leave, because once the two-player coalition forms, ρ selects both coalitions with positive probability and outside options are safe. If i can extract payoff by leaving as a singleton, he signs the singleton contract securing at least the payoff for $\{i, jk\}$. If j, k expect i to leave if the implementation phase is entered as singletons, jk forms already in the bargaining phase of $\tau = 1$, because in expectation both obtain a positive share of the efficiency gains when merging in the next negotiation round.

Bibliography

Akerlof, George A., "The market for lemons: Quality uncertainty and the market mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Armstrong, Mark, "Competition in two-sided markets," *The RAND Journal of Economics*, 2006, *37* (3), 668–691.

Bloch, Francis, "Sequential formation of coalitions in games with externalities and fixed payoff division," *Games and Economic Behavior*, 1996, *14* (1), 90–123.

_ **and Armando Gomes**, "Contracting with externalities and outside options," *Journal of Economic Theory*, 2006, *127* (1), 172–201.

Blume, Andreas, Douglas V DeJong, Yong-Gwan Kim, and Geoffrey B. Sprinkle, "Evolution of communication with partial common interest," *Games and Economic Behavior*, 2001, *37* (1), 79–120.

Bó, Pedro Dal, "Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games," *The American Economic Review*, 2005, *95* (5), 1591–1604.

_ **and Guillaume R. Fréchette**, "The evolution of cooperation in infinitely repeated games: Experimental evidence," *The American Economic Review*, 2011, *101* (1), 411–429.

Cadsby, Charles B., Murray Frank, and Vojislav Maksimovic, "Pooling, separating, and semiseparating equilibria in financial markets: Some experimental evidence," *Review of Financial Studies*, 1990, *3* (3), 315–342.

Cai, Hongbin and Joseph Tao-Yi Wang, "Overcommunication in strategic information transmission games," *Games and Economic Behavior*, 2006, *56* (1), 7–36.

Camerer, Colin and Keith Weigelt, "Experimental tests of a sequential equilibrium reputation model," *Econometrica*, 1988, *56* (1), 1–36.

Charness, Gary and Martin Dufwenberg, "Promises and partnership," *Econometrica*, 2006, *74* (6), 1579–1601.

_ **and** _ , "Bare promises: An experiment," *Economics Letters*, 2010, *107* (2), 281–283.

_ **and** _ , "Participation," *The American Economic Review*, 2011, *101* (4), 1211–1237.

Chatterjee, Kalyan, Bhaskar Dutta, Debraj Ray, and Kunal Sengupta, "A noncooperative theory of coalitional bargaining," *The Review of Economic Studies*, 1993, *60* (2), 463–477.

Chen, Ying, Navin Kartik, and Joel Sobel, "Selecting cheap-talk equilibria," *Econometrica*, 2008, *76* (1), 117–136.

Chen, Yongmin and Robert W. Rosenthal, "Asking prices as commitment devices," *International Economic Review*, 1996, *37* (1), 129–155.

Crawford, Vincent, "A survey of experiments on communication via cheap talk," *Journal of Economic Theory*, 1998, *78* (2), 286–298.

Croson, Rachel, Terry Boles, and J. Keith Murnighan, "Cheap talk in bargaining experiments: lying and threats in ultimatum games," *Journal of Economic Behavior & Organization*, 2003, *51* (2), 143–159.

Deneckere, Raymond and Meng-Yu Liang, "Bargaining with interdependent values," *Econometrica*, 2006, *74* (5), 1309–1364.

Dickhaut, John W., Kevin A. McCabe, and Arijit Mukherji, "An experimental study of strategic information transmission," *Economic Theory*, 1995, *6* (3), 389–403.

Ellingsen, Tore and Magnus Johannesson, "Promises, threats and fairness," *Economic Journal*, 2004, *114* (495), 397–420.

Embrey, Matthew S., Guillaume R. Fréchette, and Steven F. Lehrer, "Bargaining and reputation: An experiment on bargaining in the presence of behavioral types," *The Review of Economic Studies*, forthcoming.

Erat, Sanjiv and Uri Gneezy, "White lies," *Management Science*, 2012, *58* (4), 723–733.

Esteban, Joan, Laura Mayoral, and Debraj Ray, "Ethnicity and conflict: An empirical study," *The American Economic Review*, 2012, *102* (4), 1310–1342.

Evans, Robert, "Sequential bargaining with correlated values," *The Review of Economic Studies*, 1989, *56* (4), 499–510.

Eyster, Erik and Matthew Rabin, "Cursed equilibrium," *Econometrica*, 2005, *73* (5), 1623–1672.

Fehr, Ernst, Holger Herz, and Tom Wilkening, "The lure of authority: Motivation and incentive effects of power," *The American Economic Review*, 2013, *103* (4), 1325–1359.

Fischbacher, U., "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 2007, *10* (2), 171–178.

Gneezy, Uri, "Deception: The role of consequences," *The American Economic Review*, 2005, *95* (1), 384–394.

Goeree, Jacob K. and Jingjing Zhang, "Communication & competition," *Experimental Economics*, 2014, *17* (3), 421–438.

— , **Charles A. Holt, and Thomas R. Palfrey**, "Quantal response equilibrium and overbidding in private-value auctions," *Journal of Economic Theory*, 2002, *104* (1), 247–272.

Gomes, Armando, "Multilateral contracting with externalities," *Econometrica*, 2005, *73* (4), 1329–1350.

— **and Philippe Jehiel**, "Dynamic processes of social and economic interactions: on the persistence of inefficiencies," *Journal of Political Economy*, 2005, *113* (3), 626–667.

Güth, Werner and Reinhard Tietz, "Ultimatum bargaining behavior: A survey and comparison of experimental results," *Journal of Economic Psychology*, 1990, *11* (3), 417–449.

Holt, Charles A., "Industrial organization: A survey of laboratory research," in "Handbook of Experimental Economics, ed. by J. Kagel and A. Roth" Princeton University Press 1995, pp. 349–443.

Hyndman, Kyle and Debraj Ray, "Coalition formation with binding agreements," *The Review of Economic Studies*, 2007, *74* (4), 1125–1147.

Kim, Kyungmin, "Endogenous market segmentation for lemons," *The RAND Journal of Economics*, 2012, *43* (3), 562–576.

Lauermann, Stephan, "Dynamic matching and bargaining games: A general approach," *The American Economic Review*, 2013, *103* (2), 663–689.

Lundquist, Tobias, Tore Ellingsen, Erik Gribbe, and Magnus Johannesson, "The aversion to lying," *Journal of Economic Behavior & Organization*, 2009, *70* (1), 81–92.

Ochs, Jack and Alvin E. Roth, "An experimental study of sequential bargaining," *The American Economic Review*, 1989, *79* (3), 355–384.

Okada, Akira, "A noncooperative coalitional bargaining game with random proposers," *Games and Economic Behavior*, 1996, *16* (1), 97–108.

Osborne, Martin J. and Ariel Rubinstein, *Bargaining and markets*, Vol. 34, Academic press San Diego, 1990.

Perry, Motty and Philip J. Reny, "A noncooperative view of coalition formation and the core," *Econometrica*, 1994, pp. 795–817.

Rabin, Matthew, "Risk aversion and expected-utility theory: A calibration theorem," *Econometrica*, 2000, *68* (5), 1281–1292.

Rapoport, Amnon, Ido Erev, and Rami Zwick, "An experimental study of buyer-seller negotiation with one-sided incomplete information and time discounting," *Management Science*, 1995, *41* (3), 377–394.

Ray, Debraj, *A game-theoretic perspective on coalition formation*, Oxford University Press, 2007.

_ **and Rajiv Vohra**, "A theory of endogenous coalition structures," *Games and Economic Behavior*, 1999, *26* (2), 286–336.

_ **and** _ , "Coalitional power and public goods," *Journal of Political Economy*, 2001, *109* (6), 1355–1385.

_ **and** _ , "Coalition formation," *Working Paper, Brown Univeristy*, 2014.

Reynolds, Stanley S., "Durable-goods monopoly: Laboratory market and bargaining experiments," *The RAND Journal of Economics*, 2000, *31* (2), 375–394.

Rochet, Jean-Charles and Jean Tirole, "Platform competition in two-sided markets," *Journal of the European Economic Association*, 2003, *1* (4), 990–1029.

Roth, Alvin E., "Bargaining experiments," *Handbook of experimental economics*, 1995, *195*, 243–348.

_ **and Francoise Schoumaker**, "Expectations and reputations in bargaining: An experimental study," *The American Economic Review*, 1983, *73* (3), 362–372.

_ **and J. Keith Murnighan**, "The role of information in bargaining: An experimental study," *Econometrica*, 1982, *50* (5), 1123–1142.

_ **and Michael W. Malouf**, "Game-theoretic models and the role of information in bargaining," *Psychological review*, 1979, *86* (6), 574–594.

Rubinstein, Ariel, "Perfect equilibrium in a bargaining model," *Econometrica*, 1982, *50* (1), 97–109.

Ruiz, Adrian De Groot, Theo Offerman, and Sander Onderstal, "An experimental study of credible deviations and ACDC," *Experimental Economics*, forthcoming.

Samuelson, William, "Bargaining under asymmetric information," *Econometrica*, 1984, *52* (4), 995–1005.

Satterthwaite, Mark and Artyom Shneyerov, "Dynamic matching, two-sided incomplete information, and participation costs: Existence and convergence to perfect competition," *Econometrica*, 2007, *75* (1), 155–200.

Seidmann, Daniel J. and Eyal Winter, "A theory of gradual coalition formation," *The Review of Economic Studies*, 1998, *65* (4), 793–815.

Sutter, Matthias, "Deception through telling the truth?! Experimental evidence from individuals and teams," *Economic Journal*, 2009, *119* (534), 47–60.

Sutton, John, "Non-cooperative bargaining theory: An introduction," *The Review of Economic Studies*, 1986, *53* (5), 709–724.

Tversky, Amos and Daniel Kahneman, "Loss aversion in riskless choice: A reference-dependent model," *The Quarterly Journal of Economics*, 1991, *106* (4), 1039–1061.

Valley, Kathleen, Joseph Moag, and Max Bazerman, "A matter of trust: Effects of communication on the efficiency and distribution of outcomes," *Journal of Economic Behavior & Organization*, 1998, *34* (2), 211–238.

_ **, Leigh Thompson, Robert Gibbons, and Max Bazerman**, "How communication improves efficiency in bargaining games," *Games and Economic Behavior*, 2002, *38* (1), 127–155.

Vanberg, Christoph, "Why do people keep their promises? An experimental test of two explanations," *Econometrica*, 2008, *76* (6), 1467–1480.

Vincent, Daniel R., "Bargaining with common values," *Journal of Economic Theory*, 1989, *48* (1), 47–62.

Yi, Sang-Seung, "Stable coalition structures with externalities," *Games and Economic Behavior*, 1997, *20* (2), 201–237.

Selbstständigkeitserklärung

"Ich erkläre hiermit, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetzes vom 5. September 1996 über die Universität zum Entzug des aufgrund dieser Arbeit verliehenen Titels berechtigt ist."

Simon Siegenthaler	Ort und Datum