

Schumacher, Heiner; Karle, Heiko; Volund, Rune

Conference Paper

Settlement Offers

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel -
Session: Bargaining and Participation Constraints, No. B04-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Schumacher, Heiner; Karle, Heiko; Volund, Rune (2016) : Settlement Offers, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Bargaining and Participation Constraints, No. B04-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/145772>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

SETTLEMENT OFFERS

BY HEIKO KARLE, HEINER SCHUMACHER, AND RUNE VØLUND*

We analyze how early settlement offers affect Nash bargaining outcomes when agents are asymmetrically informed about the distribution of bargaining powers and exhibit expectations-based loss-aversion preferences. Before the start of the bargaining process, the sender has private information about the receiver's bargaining power. His settlement offers may convey this private information so that they change the receiver's expectations and, by loss aversion, preferences. We show that the sender can exploit this mechanism by making non-binding settlement offers, i.e., the receiver is free to accept them at any time, even after learning her true bargaining value.

I. INTRODUCTION

Individuals are often uncertain of their true bargaining power—and therefore the final outcome—before they enter a bargaining situation. In contrast, established organizations with a historical background and years of experience are often able to come up with a precise assessment of their bargaining position. As an example, consider a single worker who needs to negotiate work conditions with a large firm. The worker does not necessarily know how many qualified applicants there are for the job or how desperate the firm is to fill the position. The firm, however, may possess knowledge of the labor pool and the internal needs for hiring.

Since individuals are generally assumed to dislike uncertainty, it could be presumed that they would go to great lengths in trying to remove the intrinsic risk associated with bargaining. Organizations, being aware of this distaste

*Contact information: **Heiko** - ETH Zurich, hkarle@ethz.ch; **Heiner** - KU Leuven, heiner.schumacher@kuleuven.be; **Rune** - Aarhus University, rune@econ.au.dk.

for risk, could therefore potentially benefit by removing the individual's uncertainty through settlement. However, if individuals are able to rationally draw inference from an observed settlement offer and organizations cannot commit to an offer-strategy, such settlement schemes become more subtle. This is the case since an organization would have no incentive to propose a lucrative settlement offer if it knew that it had a favorable bargaining position. In this way, a settlement offer works as a signal that, at least partially, reveals the relative bargaining relation, and individuals should therefore be able to update their beliefs over potential bargaining outcomes after observing such an offer.

With settlement offers working as strategic instruments that contain private information, these strategic interactions can be formulated in the more abstract setting of a *signaling game*.¹ The form of signaling then follows the same basic structure as in [Leland and Pyle, 1977], where an original owner is selling shares in her company: By selling a larger or smaller fraction of the firm's shares the owner may signal her knowledge about its intrinsic value. In other words, the signal is embodied in the contract phase where it is manifested through the proposed contract. But how a receiver responds to the information contained in a settlement offer, and how this affects the relationship between settlement and bargaining outcomes, depends crucially on the risk-concept that best describes the receiver.

Following much of recent theoretical and empirical work, the receiver is assumed to follow the expectation-based model of reference-dependent preferences formulated by [Kőszegi and Rabin, 2006]. With this risk-concept, the receiver obtains intrinsic utility from outcomes, but he also derives *gain-loss utility* from comparing outcomes with reference-points that are determined by lagged expectations. As will be thoroughly elaborated in the next section, the degree to which the sender can manipulate the receiver's expectations through the settlement offer is going to constitute the main strategic dynamics of the game.

Building on the observations expounded in the previous paragraphs, this paper develops a bilateral bargaining model where the players face uncertainty

¹ For expository purposes, the rest of the paper uses the arbitrary convention of referring to the sender as "she" and the receiver as "he."

over their bargaining relation. This model extends the asymmetric Nash bargaining solution of [Kalai, 1977], by introducing an initial stage where an informed sender proposes a settlement offer to a loss-averse receiver. Lasting for two periods, 0 and 1, the chronicle of the game can more precisely be stated as follows. Before the game begins, both players are uncertain of the receiver's bargaining power. In period 0, the sender observes this bargaining power and then uses her private information to formulate a settlement offer to the receiver. Observing this offer, the receiver updates his expectations over possible outcomes. In period 1, the receiver discovers his true bargaining power and then decides whether he wants to accept the settlement offer or bargain. What the receiver ends up choosing ultimately determines the outcome, and therefore the welfare, of both players.

As is generally the case for signaling games, this model suffers from a plethora of possible equilibria, which makes it difficult to perfectly predict the path of play. This problem stems not only from the many degrees of freedom available in choosing the receiver's beliefs off-the-equilibrium-path, but is also ingrained in the basic structure of the model. The analysis in this paper will therefore focus its attention on a particular economic issue: To what extent is it possible for the sender to obtain a payoff above what she would be able to get through fair bargaining? In other words, is it possible for the sender to manipulate the receiver's expectations in such a way that the receiver is worse off than he would be in fair bargaining? It is shown that all equilibria that answer positive to this question are in some sense economically equivalent: The sender partitions the support of the probability distribution of the bargaining power, and, in effect, introduces noise into her signal by reporting settlement offers with the same value for any bargaining power in a given interval.

The rest of the paper is organized as follows. Section II offers a rigorous specification of the model and defines equilibrium, while Section III focuses on the derivation of equilibria. Section IV illustrates the mechanisms of the former sections by providing a simple example. Section V provides a discussion of potential complications, and section VI concludes the paper.

II. THE MODEL

Consider a model with a privately informed sender and a loss-averse receiver, who need to agree on a decision $d \in [0, 1]$ through bilateral bargaining. The game lasts for two periods, 0 and 1, and before the game begins the players face uncertainty over the receiver's bargaining power. In period 0, before the actual bargaining process begins, the sender observes the receiver's bargaining power b , whose differentiable probability distribution, $F(b)$, with density $f(b)$, is supported on $[0, 1]$. The sender then proposes a settlement offer $\langle d_o, t_o \rangle \in [0, 1] \times [0, 1]$ to the receiver, consisting of a decision d_o and a transfer t_o . Observing this offer, the receiver uses any available information to update his expectations over possible outcomes. In period 1, the receiver observes his true bargaining power and then decides whether he should accept the sender's offer or bargain. All aspects of the game except b are common knowledge.²

The players have standard risk-neutral preferences. In addition, the receiver is expectation based loss-averse as formulated by [Kőszegi and Rabin, 2006]; that is, he has reference-dependent preferences. In this formulation the utility function has two components: Intrinsic value from the actual outcome and comparison of this outcome with a reference point. The intrinsic utility can simply be thought of as the classical notion of outcome based utility. But in addition, the receiver derives *gain-loss utility* from the comparison of the actual outcome with the outcome he expected.

Letting the sender and the receiver prefer respectively the smallest and largest d possible, the normalized utility functions can, for any fixed outcome pair $\langle d, t \rangle$ and reference pair $\langle d_r, t_r \rangle$, be written as

$$U_s(d, t) = 1 - d - t$$

$$U_r(d, t | d_r, t_r) = d + t + \mu(d - d_r) + \mu(t - t_r)$$

where the gain-loss function $\mu(\cdot)$ is assumed to be piecewise linear

² Of course, when the receiver discovers his true bargaining power in period 1, every aspect of the game will be common knowledge. The process in which the receiver discovers his true bargaining power can simply be thought of as a learning process in which he obtains information in the market. Some thoughts on this process can be found in section V.

$$\mu(x) = \begin{cases} \eta x & \text{if } x \geq 0 \\ \eta \lambda x & \text{if } x < 0 \end{cases}$$

with $\lambda > 1$ being the coefficient of loss aversion and $\eta > 0$ the relative weight the receiver attaches to gain-loss utility.³ By having a constant marginal utility from gains and a constant, but larger, marginal dis-utility from losses, this specification formalizes [Kahneman and Tversky, 1979]’s prospect theory.

To incorporate the full range of possible expectations the receiver can form, his utility function is extended by allowing references to be a pair of probability distributions $\langle R^d, R^t \rangle$. This formulation captures the notion that the sense of gain and loss from a given outcome derives from comparing it with all possible values in the support of the reference lottery. For example, if the receiver expects either a decision $d = 0$ or $d = 1$ under the reference lottery, the decision $d = 0.5$ would feel like a gain relative to 0 and a loss relative to 1. Henceforth, for any outcome pair $\langle d, t \rangle$ and probability distributions $\langle R^d, R^t \rangle$, total utility for the receiver is defined by

$$U_r(d, t | R^d, R^t) = d + t + \int_0^1 \mu(d - d_r) dR^d(d_r) + \int_0^1 \mu(t - t_r) dR^t(t_r)$$

An essential feature captured by this utility specification is the assumption that people assess gains and losses in separate dimensions. For example, suppose the receiver had expected to accept a settlement offer $\langle d_o, t_o \rangle$, but it turns out that the decision from bargaining, d_b , is greater than d_o . If he chooses to bargain he would then evaluate it as a gain in the decision dimension ($d_b > d_o$) and a loss in the transfer dimension ($0 < t_o$). This interpretation is not only consistent with much experimental evidence, see for example [Kahneman et al., 1991], but is also crucial for the main conclusions of this paper.

The bargaining outcome is going to be determined by the *asymmetric Nash bargaining solution*. Formulated by [Nash, 1950] and [Nash, 1953], a bargaining problem is a pair consisting of a choice set in the utility possibility

³ That the sender and the receiver prefer respectively the smallest and largest d possible is without loss of generality. Furthermore, normalizing the utility functions does not qualitatively affect any results of the paper. This is the case since the receiver’s gain-loss function is piecewise linear and therefore decision-theoretical invariant to affine transformations.

set and a threat point. By postulating a list of properties (axioms) that a reasonable bargaining solution ought to satisfy, Nash showed how to assign a unique payoff allocation to each compact and convex bargaining problem. By removing the axiom of symmetry, [Kalai, 1977] generalized the work of Nash by defining the asymmetric Nash bargaining solution as the unique maximizer of the generalized Nash product

$$\max_{U_r, U_s \geq 0} U_r^b U_s^{1-b} \quad \text{s.t.} \quad U_r, U_s \in U$$

for U being the utility possibility set and the threat point normalized to zero.⁴ This solution picks out the utility pair $\langle U_r, U_s \rangle$ that maximizes the weighted product of the players' utilities, with the domain of optimization being the utility possibility set characterized by the players' full utility functions.

Determining what constitutes the receiver's full utility function at the time of bargaining is not a trivial task. This is because the receiver is expectation based loss-averse, which means that his utility function depends on the expectations he holds at a given point in time. How these expectations behave intertemporally can affect the bargaining outcome in non-negligible ways, since the gain-loss part of the receiver's utility can make up a large proportion of the total utility. Restrictions on how the receiver's expectations behave over time are therefore necessary (for a full discussion on this matter see section V).

In this paper, I will make use of the following interpretation. If the receiver chooses the settlement offer in period 1, the outcome is given by $\langle d_o, t_o \rangle$ and the game ends. If he rejects the settlement offer and chooses to bargain instead, there will be no transfer, and the decision is determined by the asymmetric Nash bargaining solution at some point in the future. Now, when the receiver faces the choice between settlement and bargaining in period 1, he still holds the expectations he formed in period 0. But if he decides to bargain, then his expectations will be revised to match the new prospect of the future, since the bargaining outcome is not obtained immediately. The receiver's preferences

⁴ The model could easily be extended by allowing the threat point to take different values. But since such a generalization does not bring anything to the analysis, except cumbersome notation, I assume throughout the paper that the threat point is given by the zero vector.

therefore collapses to the standard risk-neutral preferences at the time of bargaining. Hence, the relevant utility possibility set can be written as

$$U = \{ \langle d, 1 - d \rangle \in [0, 1] \times [0, 1] \mid d \leq \bar{d} \text{ and } 1 - d \leq 1 - \bar{d}, \text{ for } \bar{d} \in [0, 1] \}$$

Since this is a compact and convex set, this bargaining problem satisfies the necessary and sufficient conditions for a unique solution. It follows that the asymmetric Nash bargaining solution is given by

$$b = \operatorname{argmax}_{d \in [0, 1]} d^b (1 - d)^{1-b} \quad (1)$$

for any bargaining power $b \in [0, 1]$. Since the choice between bargaining and settlement is made in period 1—where the receiver still holds his period 0 expectations—the receiver evaluates the outcome $\langle b, 0 \rangle$ from bargaining with respect to the expectations he formed in period 0. To sum up, when the receiver needs to decide whether he should bargain in period 1, he evaluates the payoff from bargaining by the intrinsic value, $\langle b, 0 \rangle$, but he also compares this outcome to the expectations he formed in period 0.

In order to close the model there need to be some restrictions on the players' behaviour. To make the receiver's behaviour endogenous it is necessary to specify how his expectations are formed. Applying the ideas from [Kőszegi and Rabin, 2006], I assume that expectations must be consistent with rationality; that is, the receiver correctly predicts the implications of his plans, and makes the best plan he knows he will carry through. Hence, a credible plan must be optimal given the expectations it generates. If the receiver is not willing to follow the plan through when period 1 comes, given he expected to do it in period 0, he should not have rationally formulated the plan in the first place. A plan that has the property that it is optimal given the expectations generated by itself (a credible plan), is from now on referred to as a *personal equilibrium* (PE). Since the receiver is constrained to choosing among the set of PE plans, he maximizes his ex-ante utility by choosing the best PE plan. Such a credible plan is referred to as a *preferred personal equilibrium* (PPE).

To facilitate the formalization of the concepts PE and PPE, I now reason as follows. Observing any settlement offer $\langle d_o, t_o \rangle$, the receiver updates his beliefs according to $F(b \mid \langle d_o, t_o \rangle)$ and then forms his self-fulfilling expectations. Now,

notice that for any beliefs held by the receiver, a credible plan must have the following cut-off structure: Bargain if $b > b^*$ and settle if $b \leq b^*$ (assuming that any potential indifference is broken in favour of settlement). This follows from the fact that, given any period 0 expectations, the utility the receiver derives from bargaining is strictly increasing in b , while the settlement offer remains unchanged. Hence, if the receiver prefers to bargain with bargaining power b , he also prefers to bargain with $b_\epsilon = b + \epsilon$, for any arbitrary $\epsilon > 0$. Any credible plan must therefore induce expectations $R_{b^*}^d[F(b|\langle d_o, t_o \rangle)]$ of decision d_o with probability $F(b^*|\langle d_o, t_o \rangle)$ or a decision $b \in (b^*, 1]$ distributed according to $F(b|\langle d_o, t_o \rangle)$; and expectations $R_{b^*}^t[F(b|\langle d_o, t_o \rangle)]$ of receiving transfer t_o with probability $F(b^*|\langle d_o, t_o \rangle)$ or transfer 0 with probability $1 - F(b^*|\langle d_o, t_o \rangle)$. A credible plan can then be defined as follows.

Definition 1. A cut-off bargaining power b^* is a personal equilibrium (PE), given the beliefs $F(b|\langle d_o, t_o \rangle)$, if for the induced expectations $R_{b^*}^d[F(b|\langle d_o, t_o \rangle)]$ and $R_{b^*}^t[F(b|\langle d_o, t_o \rangle)]$, it follows that

$$U_r(d_o, t_o | R_{b^*}^d, R_{b^*}^t) \geq U_r(b, 0 | R_{b^*}^d, R_{b^*}^t) \quad \text{for } b \leq b^*$$

and

$$U_r(d_o, t_o | R_{b^*}^d, R_{b^*}^t) < U_r(b, 0 | R_{b^*}^d, R_{b^*}^t) \quad \text{for } b > b^*.$$

Utility maximization in period 0, and the fact that the receiver is able to choose among the different PE plans, implies that the receiver acts optimally by choosing the PE plan that maximizes his expected utility.

Definition 2. A cut-off bargaining power b_p^* is a preferred personal equilibrium (PPE), given the beliefs $F(b|\langle d_o, t_o \rangle)$, if it is a PE and for any PE cut-off bargaining power b^* , it is true that

$$\mathbb{E}_{b|\langle d_o, t_o \rangle} \left[U_r(d, t | R_{b_p^*}^d, R_{b_p^*}^t) \right] \geq \mathbb{E}_{b|\langle d_o, t_o \rangle} \left[U_r(d, t | R_{b^*}^d, R_{b^*}^t) \right].$$

The two concepts just defined (PE and PPE) are restrictions on the receiver's behaviour. With strategic interactions between the sender and the receiver, it is, however, also necessary to restrict how they respond to each others' strategies. Because the receiver has imperfect information when forming his expectations, the equilibrium notion applied is that of *perfect Bayesian equilibrium* (henceforth

referred to as *equilibrium*). Because this solution concept requires sequential rationality, the players are not allowed to commit to any strategy ex-ante. The receiver is therefore forced to use all the information he can extract from the settlement offer and use this information to make rational inference over the possible bargaining powers, before he forms his expectations. Likewise, the sender is not allowed to commit herself to any offer-strategy before observing the receiver's bargaining power.

Formally, an equilibrium consists of a class of offer-strategies for the sender, denoted $q(\langle d_o, t_o \rangle | b) \subseteq \Delta(\langle d_o, t_o \rangle)$, and a pair of induced expectations for the receiver, $R_{b^*}^d[F(b|\langle d_o, t_o \rangle)]$ and $R_{b^*}^t[F(b|\langle d_o, t_o \rangle)]$, such that:

- (1) for each $b \in [0, 1]$, $\int_0^1 \int_0^1 q(\langle \varepsilon_d, \varepsilon_t \rangle | b) d\varepsilon_d d\varepsilon_t = 1$, and if $\langle d_o^*, t_o^* \rangle$ is in the support of $q(\cdot | b)$, then $\langle d_o^*, t_o^* \rangle$ solves

$$\max_{\langle d_o, t_o \rangle \in [0, 1] \times [0, 1]} U_s(d, t)$$

given the receiver's equilibrium behaviour; and

- (2) given any settlement offer $\langle d_o, t_o \rangle$, the receiver follows his preferred personal equilibrium, b_p^* , where $F(b|\langle d_o, t_o \rangle)$ is derived by applying Bayes' rule, whenever possible.

Condition (1) says that, given the receiver's equilibrium behaviour, the sender's offer-strategy yields a utility maximizing outcome for each bargaining power she observes. Condition (2) says that the receiver responds optimally to each possible settlement offer, using Bayes' rule to update his prior beliefs, $F(b)$, taking into account the offer he observes and the sender's offer-strategy. Given these posterior beliefs, $F(b|\langle d_o, t_o \rangle)$, the receiver chooses his preferred personal equilibrium.

The model basically describes a signaling game with signaling through the contract phase. When the sender proposes a settlement offer she contemporaneously reveals some information about the receiver's bargaining power; something that fosters non-trivial strategic interactions between the sender and the receiver. As was motivated in the introduction of the paper, the main question now is whether it is possible for the sender to obtain a payoff above what she could get from bargaining. The intricacy formulated by

this inquiry is what the rest of the paper will try to dissect. As will be shown in the next section, for the sender to be able to benefit from the receiver's uncertainty she will need to include some noise in her offer-strategy, but not too much. In fact, I will provide sufficient conditions for the existence of a class of semi-separating equilibria where the sender's payoff dominates the one from bargaining. Furthermore, I will also show that any equilibrium with this property can be thought of as being economically equivalent to one in this class.

III. EQUILIBRIUM

This section starts by characterizing the structure of the full class of equilibria where the sender benefits from her informational advantage. These equilibria are partition equilibria of a special kind, in which the sender introduces noise into the settlement offer by not discriminating perfectly among the bargaining powers she may observe. A particularly interesting subclass is then constructed, and it is argued that the equilibria in this subclass are sufficient to provide a complete economic picture of the basic structure in the full class. It is instructive to notice that for the sender to benefit above the bargaining payoff, the receiver has to accept a settlement offer where $d_o + t_o < b$.⁵ This observation leads to the following useful definition.

Definition 3. *Any equilibrium where the receiver always accepts a settlement offer that satisfies $d_o + t_o < b$ is said to have property Ω .*

It is straight forward to see that no equilibria with property Ω exist in the classes of pooling or separating equilibria. First, in any pooling equilibrium, the sender has to propose the same settlement offer for all $b \in [0, 1]$. But the only settlement offer that the sender would ever propose for all $b \in [0, 1]$ is $\langle 0, 0 \rangle$. Even if the receiver expected to accept this settlement offer, he would always want to deviate for $b > 0$. Second, in any separating equilibrium,

⁵ Notice that this condition can never be satisfied when $b = 0$. In other words, the sender cannot "extract" any surplus from the receiver if she already has all the surplus. But since this is a zero-probability event, I have chosen to leave this minor detail out of the general statement. Definition 3 should therefore be interpreted with this in mind.

the sender has to propose a distinct settlement offer(s) for each $b \in [0, 1]$. However, in any equilibrium where the sender follows such a strategy, the receiver would be able to map each settlement offer to his true bargaining power. For any $b \in [0, 1]$, the receiver would therefore prefer to expect bargaining with probability one. From these observations it follows that any equilibrium with property Ω is to be found in the class of semi-separating equilibria (or at least a subclass thereof).

In order to facilitate exposition, it is useful to first provide some additional structure. Therefore, define V as the set of all strictly decreasing sequences $(v^k)_{k=1}^\infty$ in $(0, 1)^\mathbb{N}$, and J as the set of all sequences $(J^k)_{k=1}^\infty$ of (possible uncountable) index sets J^k . For any $(v^k)_{k=1}^\infty \in V$ and $(J^k)_{k=1}^\infty \in J$, let $\langle d_o^{k,j}, t_o^{k,j} \rangle$ be the j 'th settlement offer that satisfies $d_o^{k,j} + t_o^{k,j} = v^k$ for $j \in J^k$, and define $\beta^{k,j} \subseteq (v^k, 1]$ as the set of bargaining powers where $q(\langle d_o^{k,j}, t_o^{k,j} \rangle | b) > 0$ for all $b \in \beta^{k,j}$. These sets are useful due to the fact that the sender is only willing to mix among settlement offers $\langle d_o^{k,j}, t_o^{k,j} \rangle$ for $j \in J^k$. In order to keep the notation consistent, define $\beta^{0,j} = \emptyset$ for all j in some J^0 , and notice that $\inf \emptyset = \infty$.

Lemma 1. Fix any $(v^k)_{k=1}^\infty \in V$ and $(J^k)_{k=1}^\infty \in J$. In any equilibrium with property Ω , it is the case that for all $k \in \mathbb{N}$, $j \in J^k$, and $j' \in J^{k-1}$,

- (A) $d_o^{k,j} + t_o^{k,j} \left(\frac{1+\eta\lambda}{1+\eta} \right) \leq \inf \beta^{k-1,j'}, \quad d_o^{k,j} + t_o^{k,j} \left(\frac{1+\eta\lambda}{1+\eta} \right) \geq \sup \beta^{k,j}$
- (B) $\bigcup_{j \in J^k} \beta^{k,j} = \beta^k$ constitutes a non-degenerate interval,
- (C) $\beta^{k+1} \cap \beta^k = \emptyset, \quad \sup \beta^{k+1} = \inf \beta^k, \quad \text{and} \quad \bigcup_{k \in \mathbb{N}} \beta^k = (0, 1].$

Proof. Fix any $(J^k)_{k=1}^\infty \in J$ and $v^k \in (0, 1)$. Suppose there exists an equilibrium with property Ω where β^k does not constitute a non-degenerate interval for some $k \in \mathbb{N}$. Then there exist $b_L, b_H \in \beta^k$ such that $q(\langle d_o^{k,j}, t_o^{k,j} \rangle | b_L) > 0$ and $q(\langle d_o^{k,j'}, t_o^{k,j'} \rangle | b_H) > 0$ for some $j, j' \in J^k$, but $q(\langle d_o^{k,j''}, t_o^{k,j''} \rangle | b_M) = 0$ for all $j'' \in J^k$ for some $b_M \in (b_L, b_H)$ (assuming $b_L < b_H$). Now, fix any $v^{k+1}, v^{k-1} \in (0, 1)$ such that $v^k \in (v^{k+1}, v^{k-1})$.

- (I) Suppose that $q(\langle d_o^{k+1,j''}, t_o^{k+1,j''} \rangle | b_M) > 0$ for some $j'' \in J^{k+1}$. In this case, if the sender observes b_L she would want to deviate to the settlement offer $\langle d_o^{k+1,j''}, t_o^{k+1,j''} \rangle$, since this offer would have been accepted on the equilibrium-path if the true bargaining power was $b_M > b_L$.

- (II) Suppose that $q(\langle d_o^{k-1,j''}, t_o^{k-1,j''} \rangle | b_M) > 0$ for some $j'' \in J^{k-1}$. In this case, if the sender observes b_M she would want to deviate to the settlement offer $\langle d_o^{k,j'}, t_o^{k,j'} \rangle$, since this offer would have been accepted on the equilibrium-path if the true bargaining power was $b_H > b_M$.

Hence, β^k needs to constitute a non-degenerate interval. It follows trivially that these intervals need to be disjoint, adjacent, and decreasing in k .

Now, any collection of intervals that supports an equilibrium with property Ω has to be countable infinite. First, suppose there exists an equilibrium with property Ω that has finite number of intervals. Then there exists a finite set $K \subset \mathbb{N}$ such that there are $|K| < \infty$ intervals β^k that satisfy

$$\bigcup_{k \in K} \beta^k = (0, 1]$$

But then, for $\bar{k} = \max K$, the settlement offer $\langle d_o^{\bar{k},j}, t_o^{\bar{k},j} \rangle = \langle 0, 0 \rangle$ would have to be proposed for all $j \in J^{\bar{k}}$. This clearly contradicts property Ω . Second, since every collection of disjoint intervals in \mathbb{R} is countable, it follows that any equilibrium with property Ω should consist of a countable infinite collection of intervals. This further implies that in any such equilibrium, it is possible to construct a bijection that maps each interval to a settlement value in the sequence $(v^k)_{k=1}^\infty$.

The last two paragraphs have established (B) and (C). Now, for any $(v^k)_{k=1}^\infty \in V$ and $(J^k)_{k=1}^\infty \in J$, fix a collection $(\beta^k)_{k=1}^\infty$ that satisfies (B) and (C). Given that the receiver expects to bargain with probability one, the following incentive compatibility constraints need to be satisfied: For all $k \in \mathbb{N}$, $j \in J^k$, and $j' \in J^{k-1}$,

$$d_o^{k,j} + t_o^{k,j} \geq \sup \beta^{k,j} + \eta(\sup \beta^{k,j} - d_o^{k,j}) - \eta \lambda t_o^{k,j}$$

and

$$d_o^{k,j} + t_o^{k,j} \leq \inf \beta^{k-1,j'} + \eta(\inf \beta^{k-1,j'} - d_o^{k,j}) - \eta \lambda t_o^{k,j}$$

If these conditions are not satisfied, then the receiver would want to deviate and bargain when he learns his true bargaining power in period 1, and the sender would want to deviate to a settlement offer with a lower value than the one specified in equilibrium. Simple algebra reveals that these two conditions are exactly (A). ■

These restrictions have the following interpretation. For a given equilibrium with property Ω to exist, (A) describes the incentive compatibility constraints that need to be satisfied, (B) shows that the sets of bargaining powers that make the sender propose settlement offers with the same value v^k constitute an interval, and (C) says that these intervals are disjoint, adjacent, decreasing in k , and together with $\{0\}$ constitute an infinite partition of the closed unit interval. In order to formalize the full class of equilibria with property Ω , some further notation is helpful.

For each $(v^k)_{k=1}^\infty \in V$ and $(J^k)_{k=1}^\infty \in J$, define β as the set of all sequences $(\beta^k)_{k=1}^\infty$ that satisfy (A)-(C) for atleast one pair $\langle \eta, \lambda \rangle \in (0, \infty) \times (1, \infty)$. Also, because the sets that contain the receiver's personal equilibria are going to differ for each $\beta^{k,j}$, I define $B^{k,j}$ to be the set of the receiver's personal equilibria after he observes settlement offer j in the k 'th interval. From these definitions the validity of theorem 1 follows trivially from the conclusions of Lemma 1, and a proof is therefore omitted all together.

Theorem 1. Fix $(v^k)_{k=1}^\infty \in V$, $(J^k)_{k=1}^\infty \in J$, and $(\beta^k)_{k=1}^\infty \in \beta$. Any equilibrium with property Ω has the following structure: For each $\eta \in (0, \infty)$ there exists a set $\Lambda(\eta) \subseteq (1, \infty)$, such that for $\lambda \in \Lambda(\eta)$, it is the case that for all $k \in \mathbb{N}$ and $j \in J^k$,

$$(A) \quad q(\langle d_o^{k,j}, t_o^{k,j} \rangle | b) > 0 \text{ for all } b \in \beta^{k,j},$$

$$(B) \quad \text{the receiver's PPE is } b_p^* = \sup \beta^{k,j}.$$

In particular, for each $\eta > 0$,

$$\Lambda(\eta) = \{ \lambda \in (1, \infty) \mid \forall k \in \mathbb{N}, d_o^{k,j} + t_o^{k,j} \geq \mathbb{E}_{b | \langle d_o^{k,j}, t_o^{k,j} \rangle} [U_r(d, t | R_{b^*}^d, R_{b^*}^t)]$$

$$\forall b^* \in B^{k,j}, \forall j \in J^k \}.$$

Remark. Note that the restrictions on the set $\Lambda(\eta)$ depend on the size of η . The eligible parameter values can therefore be seen as a two-dimensional solution set in an $\langle \eta, \lambda \rangle$ space.

This result does *not* guarantee that such an equilibrium exists, but instead shows how an equilibrium with property Ω would look like *if* it existed. Although the structure of these (potential) equilibria can appear hard to decipher, they actually have a very similar economic structure: For each bargaining

power in interval β^k , the sender proposes a settlement offer that has value v^k , and the receiver then accepts this offer. To get a more intuitive understanding of this structure—and to prove general existence—I now construct a specific subclass of theorem 1, where the infinite partitions are derived from the following recursive sequences: Fix $t_o^0 = 1$ and define, for each $t_o^1 \in [\frac{1+\eta}{1+\eta\lambda}, 1)$, the recursive sequence

$$t_o^k = t_o^1 \left(\frac{1+\eta}{1+\eta\lambda} \right)^{k-1} \quad \text{for all } k \in \mathbb{N}$$

In addition, define T as the set of all sequences $((t_o^k, t_o^{k-1}])_{k=1}^\infty$. These sequences of half-open intervals will be shown to be derived directly from the incentive compatibility constraints of Lemma 1-(A). Also, it is trivial to check that these half-open intervals actually satisfy all the conditions of Lemma 1—and as will be demonstrated in proposition 1 below—give rise to a class of equilibria that can be thought of as a neat representation of the basic structure that was found in theorem 1.

Proposition 1. *Fix any $((t_o^k, t_o^{k-1}])_{k=1}^\infty$ from T . For each $\eta > 0$ there exists $\lambda(\eta) \in (1, \infty)$, such that for all $\lambda \geq \lambda(\eta)$ there exists an equilibrium, such that for all $k \in \mathbb{N}$,*

$$(A) \quad q(\langle 0, t_o^k \rangle | b) = 1 \quad \text{for all } b \in (t_o^k, t_o^{k-1}],$$

$$(B) \quad \text{the receiver's PPE is } b_p^* = t_o^{k-1}.$$

In particular, for each $\eta > 0$,

$$\lambda(\eta) = \min\{\lambda \in (1, \infty) \mid \forall k \in \mathbb{N} \quad t_o^k \geq \mathbb{E}_{b|\langle 0, t_o^k \rangle} [U_r(d, t | R_{b^*}^d, R_{b^*}^t)] \quad \forall b^* \in B^k\}.$$

Proof. It is useful to first outline the broad steps of the proof. First, by assuming that the receiver expects to bargain with probability one, *Part 1* shows how the sequences in T are derived from two incentive compatibility constraints. Having constructed infinite partitions from the sequences in T , *Part 2* characterizes the set of personal equilibria, B^k , for each interval. Using these sets, *Part 3* establishes that expecting to settle with probability one can always be supported as the receiver's preferred personal equilibrium, for a high enough degree of loss aversion. To round off, *Part 4* comments on beliefs off-the-equilibrium-path.

Part 1. Suppose the receiver expects to accept the settlement offer with probability one. The sequences from T are then constructed from the following restrictions: First, for each $k \in \mathbb{N}$ it must be the case that for all $b \in (t_o^k, t_o^{k+1}]$ the receiver weakly prefers the settlement offer when he observes the true bargaining power in period 1. Second, the sender should have no incentive to propose the settlement offer $(0, t_o^{k+1})$ whenever $b \in (t_o^k, t_o^{k+1}]$. That is, the sender should have no incentive to pretend that the true bargaining power is in a different interval than it actually is. By letting $t_o^0 = 1$, these restrictions are formalized by proceeding from right to left in the unit interval with the following steps:

1 :

$$t_o^1 \geq t_o^0 + \eta(t_o^0 - 0) - \eta\lambda t_o^1$$

$$\Rightarrow t_o^1 \in \left[\frac{1+\eta}{1+\eta\lambda}, 1 \right)$$

2 :

$$t_o^2 \geq t_o^1 + \eta(t_o^1 - 0) - \eta\lambda t_o^2 \quad \text{and} \quad t_o^1 + \eta(t_o^1 - 0) - \eta\lambda t_o^2 \geq t_o^2$$

$$\Rightarrow t_o^2 = t_o^1 \frac{1+\eta}{1+\eta\lambda}$$

·
·
·

k :

$$t_o^k \geq t_o^{k-1} + \eta(t_o^{k-1} - 0) - \eta\lambda t_o^k \quad \text{and} \quad t_o^{k-1} + \eta(t_o^{k-1} - 0) - \eta\lambda t_o^k \geq t_o^k$$

$$\Rightarrow t_o^k = t_o^{k-1} \frac{1+\eta}{1+\eta\lambda}$$

This procedure produces a monotonically decreasing and bounded sequence $(t_o^{k-1})_{k=1}^\infty \in [0, 1]$, for any choice of $t_o^1 \in [\frac{1+\eta}{1+\eta\lambda}, 1)$. Now, for each $k \in \mathbb{N}$, construct a half-open interval with endpoints equal to the settlement value for the receiver, $(t_o^k, t_o^{k-1}]$, and take the countable union of these intervals

$$\bigcup_{k \in \mathbb{N}} (t_o^k, t_o^{k-1}] = (0, 1]$$

Since these intervals are all disjoint, it follows that $\{0\} \cup ((t_o^k, t_o^{k-1}])_{k=1}^\infty$ constitutes an infinite partition of the closed unit interval.

Part 2. I will now characterize the sets of personal equilibria. First, fix an arbitrary interval $(t_o^k, t_o^{k-1}]$ from any sequence in T . When the receiver observes the settlement offer $\langle 0, t_o^k \rangle$ on the equilibrium-path, he updates his beliefs according to

$$f(b|\langle 0, t_o^k \rangle) = \frac{f(b)q(\langle 0, t_o^k \rangle|b)}{\int_0^1 q(\langle 0, t_o^k \rangle|\varepsilon_b)f(\varepsilon_b)d\varepsilon_b} = \begin{cases} \frac{f(b)}{F(t_o^{k-1})-F(t_o^k)} & \text{if } b \in (t_o^k, t_o^{k-1}] \\ 0 & \text{if } b \notin (t_o^k, t_o^{k-1}] \end{cases}$$

from which $F(b|\langle 0, t_o^k \rangle)$ is trivially derived. Now, notice that by the differentiability of $F(b)$, it follows that $F(b|\langle 0, t_o^k \rangle)$ is differentiable and therefore continuous. With $F(b|\langle 0, t_o^k \rangle)$ being continuous, any cut-off bargaining power b^* that constitutes a personal equilibrium must per definition 1 be defined by the equality

$$\begin{aligned} b^* - F(b^*|\langle 0, t_o^k \rangle)\eta\lambda t_o^k + F(b^*|\langle 0, t_o^k \rangle)\eta b^* + \eta\lambda \int_{b^*}^{t_o^{k-1}} (b^* - b_r)dF(b_r|\langle 0, t_o^k \rangle) \\ = \\ t_o^k + \left(1 - F(b^*|\langle 0, t_o^k \rangle)\right)\eta t_o^k - \eta\lambda \int_{b^*}^{t_o^{k-1}} b_r dF(b_r|\langle 0, t_o^k \rangle) \end{aligned}$$

That is, since $F(b|\langle 0, t_o^k \rangle)$ is continuous, the utility the receiver gets from bargaining should be equal to the utility he gets from settlement when $b = b^*$. Rewriting this equality, a solution b^* can be formulated as a fixed point for the function

$$g(b) = t_o^k \left[\frac{1 + \eta + F(b|\langle 0, t_o^k \rangle)\eta(\lambda - 1)}{1 + \eta\lambda - F(b|\langle 0, t_o^k \rangle)\eta(\lambda - 1)} \right] \quad (2)$$

Since this is a continuous self-map on the closed, bounded, and convex set $[t_o^k \frac{1+\eta}{1+\eta\lambda}, t_o^k \frac{1+\eta\lambda}{1+\eta}] \subset \mathbb{R}$, there exists at least one $b^* \in [t_o^k \frac{1+\eta}{1+\eta\lambda}, t_o^k \frac{1+\eta\lambda}{1+\eta}]$ such that $g(b^*) = b^*$, by the *Brouwer fixed point theorem*. Notice that for any b such that $F(b|\langle 0, t_o^k \rangle) \leq \frac{1}{2}$, the only fixed point is $b^* = t_o^k \frac{1+\eta}{1+\eta\lambda} \notin (t_o^k, t_o^{k+1}]$, and therefore $F(b^*|\langle 0, t_o^k \rangle) = 0$; that is, the receiver expects to bargain with probability one. At the other extreme, notice that there always exists a fixed point $b^* = t_o^{k-1}$ where $F(b^*|\langle 0, t_o^k \rangle) = 1$; that is, the receiver expects to settle with probability one. Whether or not there exist one or more fixed points depends on $F(b)$.

Part 3. It was established in *Part 1* that, given the receiver expects to settle with probability one, there exists a class of infinite partitions that are incentive compatible. Further, in *Part 2* it was shown that expecting to settle with probability one is in fact a personal equilibrium; however, not the only one. It therefore needs to be shown that, for a high enough loss aversion coefficient, expecting to accept the settlement offer with probability one is the receiver's preferred personal equilibrium. In other words, there should exist, for each $\eta \in (0, \infty)$, a $\lambda(\eta) \in (1, \infty)$ such that $b^* = t_o^{k-1}$ is the receiver's preferred personal equilibrium, given the settlement offer $\langle 0, t_o^k \rangle$. The receiver's expected utility for any personal equilibrium, defined by the cut-off bargaining power b^* , can be written as

$$\begin{aligned} & F(b^*|\langle 0, t_o^k \rangle) \left[t_o^k + \left(1 - F(b^*|\langle 0, t_o^k \rangle)\right) \eta t_o^k - \eta\lambda \int_{b^*}^{t_o^{k-1}} b_r dF(b_r|\langle 0, t_o^k \rangle) \right] + \\ & \int_{b^*}^{t_o^{k-1}} \left[b - F(b^*|\langle 0, t_o^k \rangle) \eta \lambda t_o^k + F(b^*|\langle 0, t_o^k \rangle) \eta b \right] dF(b|\langle 0, t_o^k \rangle) + \\ & \int_{b^*}^{t_o^{k-1}} \left[\eta \int_{b^*}^b (b - b_r) dF(b_r|\langle 0, t_o^k \rangle) + \eta\lambda \int_b^{t_o^{k-1}} (b - b_r) dF(b_r|\langle 0, t_o^k \rangle) \right] dF(b|\langle 0, t_o^k \rangle) \end{aligned}$$

Notice that for any personal equilibrium—except $b^* = t_o^{k-1}$, that is, expect to settle with probability one—the expected utility contains gain-loss utility. With losses being scaled by $\eta\lambda$ but gains by only η , increasing λ will decrease the receiver's expected utility. However, by increasing λ the expected utility from settlement also decreases, since $t_o^k = t_o^1 \left(\frac{1+\eta}{1+\eta\lambda}\right)^{k-1}$ is decreasing in λ . But

while gain-loss utility is unbounded below, t_o^k is bounded below by 0. Further, the loss from gain-loss utility is increasing linearly in λ while $\frac{\partial^2 t_o^k}{\partial \lambda^2} < 0$. It therefore follows that, for each $\eta \in (0, \infty)$, it is always possible to find a $\lambda(\eta) \in (1, \infty)$, such that for all $\lambda \geq \lambda(\eta)$, it holds that for all $k \in \mathbb{N}$,

$$t_o^k \geq \mathbb{E}_{b|\langle 0, t_o^k \rangle} \left[U_r(d, t | R_{b^*}^d, R_{b^*}^t) \right] \quad \forall b^* \in B^k$$

That is, it is always possible to find a lower bound on loss aversion such that $b_p^* = t_o^{k-1}$, given the settlement offer $\langle 0, t_o^k \rangle$, for all $k \in \mathbb{N}$.

Part 4. Any beliefs off-the-equilibrium-path where the receiver chooses to bargain, when the sender deviates from the equilibrium-path, is seen to support equilibrium as constructed above. ■

The equilibria just constructed were seen to partition the support of the probability distribution of the bargaining power into disjoint intervals (as was expected). In some sense, these equilibria can be thought of as categorizing different bargaining relationships into different settlement schemes. In other words, the strategic dynamics foster a form of coarse reasoning that makes the uninformed part (in this case, the receiver) worse off. It is readily observed that for each equilibrium with property Ω , there exists a bijection between $(\beta^k)_{k=1}^\infty$ and some $((t_o^k, t_o^{k-1}])_{k=1}^\infty$ that can be thought of as an economic structure preserving map (in analogy to the homomorphisms of abstract algebra). Since the structure and not the exact number value of the settlement offers are the important factor, the equilibria of proposition 1 can be thought of as characterizing the economic relevant essence of the more general theorem 1.

The basic structure of these equilibria was seen to proceed from the following logic: For the receiver to be willing to accept a settlement offer where $d_o + t_o < b$, the sender has to include some noise in the settlement offer. That is, because the receiver is loss-averse, and therefore dislikes uncertainty, the sender can benefit by leaving enough uncertainty in the settlement offer to discourage the receiver from expecting to bargain. But it is not enough that the receiver expects to settle with probability one; the structure of the settlement offers has to be incentive compatible. First, because the receiver observes his true bargaining power before he has to make his choice, the utility he obtains from the offer has to be high enough to discourage him from deviating once he learns his true bargaining power. Second, it should be credible for the

receiver to deviate, if he finds out in period 1 that the sender has deviated and proposed a lower settlement offer than specified in equilibrium.

The single most important feature that made settlement offers beneficial to the sender, was seen to be the sender's ability to attach the receiver to the transfer dimension. If the sender can make the receiver expect to accept a sufficiently high transfer in period 1, then the receiver has no incentive to deviate, even if he finds out that his true bargaining power could get him a high outcome in the decision dimension. This was the case since the expectational loss from loosing the transfer is weighted by $\eta\lambda$ while the gain from a larger decision is only weighted by η . The interval length of the elements in the sequences from T could therefore be seen to be decreasing from right to left—and with an immense speed for reasonable values of η . When the receiver holds a favorable bargaining position, the sender is willing to propose a high transfer. And because the loss in the transfer dimension is scaled by $\eta\lambda$, the largest bargaining power that would make the receiver stay on the equilibrium-path is relatively large compared to a situation where the sender would only offer a low transfer.

An infinite number of different equilibria with property Ω was to be supported for at least some parameter values. Fortunately, these different equilibria were also seen to have a uniform structure, which in turn makes the complication of consistent economic analysis non-critical in itself. In continuation of the this problem is the notion of beliefs off-the-equilibrium-path. These beliefs were handled rather casual, and it therefore seems evident to consider whether complete agnosticism is the case. As is common in signaling games, the set of strategy profiles that can be supported as equilibria was seen to be very large, and an obvious disadvantage could be that some of these equilibria may predict insensible behavior from the players. It has been common in the literature to handle such complications by using refinement criteria such as the *Intuitive Criterion* from [Cho and Kreps, 1987]. It can indeed be shown that this refinement criterion does not rule out any of the equilibria, but that it does provide some extra structure on the possible beliefs off-the-equilibrium-path. These complications will, however, not be discussed any further in this paper. Instead I will move on to an example that illustrates the mechanisms of proposition 1 in a somewhat less abstract way.

IV. AN EXAMPLE

This section elucidates the structure of proposition 1 by providing a simple example where $F(b)$ is taken to be uniform on $[0, 1]$. As was shown in the last section, an infinite partition equilibrium can always be supported for a high enough degree of loss aversion. But even though proposition 1 provides a neat existence result, it does not tell much about how large this parameter, λ , actually has to be. As will be shown in the next few pages, for λ to support an infinite partition equilibrium it may need to be very large if η is small; however, not necessarily unreasonably large.⁶

Without loss of generality, consider the infinite partition equilibrium with the lowest settlement offers, that is, the equilibrium where $t_o^1 = \frac{1+\eta}{1+\eta\lambda}$. In this case, the infinite partition is derived from the recursive sequence $t_o^k = (\frac{1+\eta}{1+\eta\lambda})^k$ for all $k \in \mathbb{N}$. When the receiver observes a settlement offer $\langle 0, t_o^k \rangle$ he therefore updates his beliefs according to

$$f(b|\langle 0, t_o^k \rangle) = \begin{cases} \left(\frac{1+\eta\lambda}{1+\eta}\right)^k \left(\frac{1+\eta}{\eta(\lambda-1)}\right) & \text{if } b \in (t_o^k, t_o^{k-1}] \\ 0 & \text{if } b \notin (t_o^k, t_o^{k-1}] \end{cases}$$

Now, fix an arbitrary settlement offer, $\langle 0, t_o^k \rangle$, from the partition described above. To characterize the set of personal equilibria, B^k , note that two candidates have already been described in the proof of proposition 1. That is, expect to bargain with probability one and then bargain, and expect to accept the offer with probability one and then accept the offer. Any other personal equilibrium must be defined by a cut-off bargaining power $b^* \in (t_o^k, t_o^{k-1})$ that satisfies equation (2). Inserting the uniform distribution and rewriting equation (2), a candidate for a personal equilibrium in (t_o^k, t_o^{k-1}) is seen to satisfy

$$b^2(1+\eta) \left(\frac{1+\eta\lambda}{1+\eta}\right)^k - b(1+\eta\lambda) = 0$$

Solving this polynomial of degree two shows that no solution exists in (t_o^k, t_o^{k-1}) , and it can therefore be concluded that the set of personal equilibria can be written as $B^k = \{t_o^{k+1}, t_o^{k-1}\}$.

⁶ A discussion of the required size of λ can be found in section V.

In order to ensure that expecting to accept the settlement offer is the receiver's preferred personal equilibrium, λ needs to satisfy the condition in proposition 1. For this condition to be satisfied, the expected utility obtained when expecting to bargain with probability one should be smaller than the expected utility from expecting to settle with probability one. Thus, the expression

$$\int_{t_o^k}^{t_o^{k-1}} \left(b + \eta \int_{t_o^k}^b \left(\frac{b - b_r}{t_o^{k-1} - t_o^k} \right) db_r + \eta \lambda \int_b^{t_o^{k-1}} \left(\frac{b - b_r}{t_o^{k-1} - t_o^k} \right) db_r \right) \left(\frac{1}{t_o^{k-1} - t_o^k} \right) db$$

should be less than or equal to t_o^k for all $k \in \mathbb{N}$. For any $\eta \in (0, \infty)$, these conditions can be shown to be satisfied for $\lambda \in [\frac{\eta+3}{\eta}, \infty)$. To visualize the structure of such a partition, suppose that $\lambda = 4$ and $\eta = 1$. Figure I below shows all the possible settlement offers that can be observed in this equilibrium, that is, it illustrates the half-open intervals that support equilibrium as defined above. As was mentioned in the last section, the distance between two settlement offers is seen to converge to zero with an immense speed, even though λ took the smallest value possible given $\eta = 1$.



Figure I: *An equilibrium partition.*

The equilibrium described above is just one of an infinite number of possible equilibria characterized by proposition 1. For any λ and η large enough, any equilibrium where $t_o^1 \in [\frac{1+\eta}{1+\eta\lambda}, 1)$ can be supported. It is obvious that this fact should reduce the possibility of predicting the path of play. But if there exists a subset of these equilibria that in some sense is more salient, this subset could work as a benchmark for further analysis. To make such an argument I am going to apply [Schelling, 1963]'s concept of *focal equilibria*. That is, equilibria that are likely to be observed due to, for instance, cultural or social norms.

To apply the concept of focal equilibria, it needs to be determined which of the equilibria characterized by proposition 1 that is preferred by each player. In general, for both the sender and the receiver the answer ex-post will be

different for different values of b , and there is not much to be gained from using the idea of salient equilibria.⁷ Ex-ante, however, the answer has a neat structure that is characterized by the players' expected utilities. Using the fact that $t_o^k = t_o^1 (\frac{1+\eta}{1+\eta\lambda})^{k-1}$ for any $k \in \mathbb{N}$, the infinite sum that constitutes the players' ex-ante expected utilities—as a function of t_o^1 —can be written as the utility from t_o^1 plus an alternating geometric series.

First, the receiver's ex-ante expected utility can be written as

$$\begin{aligned}\mathbb{E} U_r &= (1 - t_o^1)t_o^1 + (t_o^1 - t_o^2)t_o^2 + (t_o^2 - t_o^3)t_o^3 + \dots + \dots \\ &= t_o^1 - \left(t_o^1\right)^2 \sum_{k=0}^{\infty} (-1)^k \left(\frac{1+\eta}{1+\eta\lambda}\right)^k \\ &= t_o^1 - \left(t_o^1\right)^2 \left(\frac{1+\eta\lambda}{1+\eta\lambda+1+\eta}\right)\end{aligned}$$

This expression is a concave function in t_o^1 , with the unique maximizer in the domain $[\frac{1+\eta}{1+\eta\lambda}, 1)$ given by $\frac{2+\eta(\lambda+1)}{2(1+\eta\lambda)}$. This maximizer is always in the interior of $[\frac{1+\eta}{1+\eta\lambda}, 1)$. To see this, notice that the inequalities $1 > \frac{2+\eta(\lambda+1)}{2(1+\eta\lambda)} > \frac{1+\eta}{1+\eta\lambda}$ can be rewritten as $\lambda > 1 > 2 - \lambda$, which is always satisfied. That is, the receiver can in some sense be thought of as preferring a balanced equilibrium. The reason for this is as follows. The closer t_o^1 is to the lower bound, the more the receiver "loses" if it turns out that he has a favorable bargaining position (large b). On the other hand, if t_o^1 is getting too close to upper bound, the next interval is going to be very large since $t_o^2 = t_o^1 (\frac{1+\eta}{1+\eta\lambda})$. The receiver therefore prefers an equilibrium that takes this trade-off into account.

Now, the sender's ex-ante utility can be derived in the same way

$$\begin{aligned}\mathbb{E} U_s &= (1 - t_o^1)(1 - t_o^1) + (t_o^1 - t_o^2)(1 - t_o^2) + (t_o^2 - t_o^3)(1 - t_o^3) + \dots + \dots \\ &= 1 - t_o^1 + \left(t_o^1\right)^2 \sum_{k=0}^{\infty} (-1)^k \left(\frac{1+\eta}{1+\eta\lambda}\right)^k \\ &= 1 - t_o^1 + \left(t_o^1\right)^2 \left(\frac{1+\eta\lambda}{1+\eta\lambda+1+\eta}\right)\end{aligned}$$

This is a convex function in t_o^1 , with the unique maximizer in the domain $[\frac{1+\eta}{1+\eta\lambda}, 1)$ given by $t_o^1 = \frac{1+\eta}{1+\eta\lambda}$. That is, the sender always prefers the smallest t_o^1

⁷ In this context, ex-ante refers to the time *before* a player knows the bargaining power while ex-post refers to the time *after* a player knows the bargaining power.

possible. Notice that the global minimizer of the sender's ex-ante expected utility is actually the global maximizer of the receiver's ex-ante expected utility. In this way, the players' ex-ante preferences over the class of equilibria derived in proposition 1 can be seen to be in conflict. In order to visualize the structure of this conflict, figure II shows the ex-ante expected utility for the sender (at the top) and the receiver (at the bottom), for $\lambda = 4$ and $\eta = 1$.

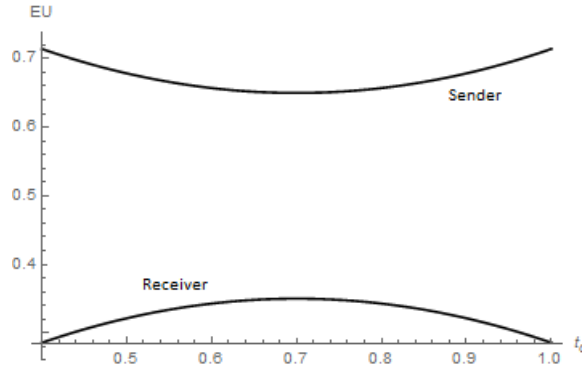


Figure II: *Ex-ante expected utility as a function of t_o^1 .*

As can be seen from this illustration, the dispute over which equilibrium that should be played is a zero-sum game. It follows that every $t_o^1 \in [\frac{1+\eta}{1+\eta\lambda}, 1)$ is ex-ante Pareto efficient, and that the Pareto efficient subsets of $[\frac{1+\eta}{1+\eta\lambda}, 1)$ do not identify any prominent equilibria. To be able to identify salient equilibria, one therefore has to resort to economic intuition. As was motivated in the introduction of the paper, the sender could for instance be an organization with a historical background and deep knowledge of the market structure. In this case the sender would seem to be in an advantageous position to control the path of play since she would be playing the game many times, while the receiver would only be playing it once. It therefore seems reasonable to presume that the sender has more authority than the receiver in the development of a focal point. Also, since the sender is the one who proposes the settlement offer, she again appears to be in control over the path of play.

The arguments in the example above are of course only conjectures, albeit not unreasonable ones. Which equilibrium that is played could easily depend on historical factors that have no general pattern, and are therefore outside the scope of this analysis. But since the sender holds the position of a Stackelberg

leader, it seems to me that the equilibrium where the sender maximizes her ex-ante utility, $t_o^1 = \frac{1+\eta}{1+\eta\lambda}$, would most likely be the focal point of the game. This presumption, however, builds on the assumption that the receiver's bargaining power follows the uniform distribution. The comparative statics considered here do therefore not necessarily cover any general pattern of the model, but has nonetheless provided a glimpse into the nature of the potential conflict between the sender and the receiver.

V. DISCUSSIONS

In this section I am going to examine the validity of the model and the subsequent conclusions that are drawn from the assumptions it builds on. The first complication is the nature of the receiver's formation of expectations over time. As was briefly mentioned in section II, the intertemporal structure that determines the reference point held by the receiver, at a given point in time, is not a trivial matter. The second complication is—as was illustrated in section IV—the high degree of loss aversion that may be needed to support an infinite partition equilibrium. Since it is standard in the literature to work with a loss aversion coefficient around two, it is imperative to consider the soundness and applicability of proposition 1. Finally I provide some thoughts on the learning process in which the receiver discovers his true bargaining power.

Expectations. The issue of how expectations behave over time is not a new one. In an application of their own model, [Kőszegi and Rabin, 2007] acknowledge the weakness of their theory in that it does not specify the lag with which new beliefs are incorporated into the reference point. In order to provide some structure on the revision of expectations, they propose two concepts: *Choice-acclimating personal equilibrium (CPE)* and *Unacclimating personal equilibrium (UPE)*. A CPE is when a person makes a committed decision long before the outcome occurs, that is, there is time to acclimate to the plan, i.e. the reference distribution adapts. An UPE is when a decision is made shortly before the outcome resulting from it occurs, that is, the reference distribution for evaluating the outcome is fixed by past expectations. Since their model is ambiguous as to the interpretation of "shortly before," these

concepts can be vague and largely determined by the author in the given context.

So how do these concepts fit into the context of this paper? When the receiver forms his expectations in period 0 he anticipates the decision he faces but cannot commit to a choice until period 1. If he makes his decision shortly before the outcome resulting from it, the beliefs determining the reference point are past and hence unchangeable. The receiver therefore maximizes his utility taking the reference point as given, so that he can rationally expect to follow the plan given his expectations. This is exactly the definition of an UPE. Since there can be many UPE—that is, multiple self-fulfilling expectations—the receiver would be expected to choose the best plan he knows he will follow through. Notice how this interpretation corresponds with idea of PE and PPE used in [Kőszegi and Rabin, 2006]. As stated by the authors:

"Our theory posits that preferences depend on lagged expectations, rather than on expectations contemporaneous with the time of consumption. This does not assume that beliefs are slow to adjust to new information or that people are unaware of the choices they have just made—but that preferences do not instantaneously change when beliefs do. When somebody finds out five minutes ahead of time that she will for sure not receive a long-expected 100, she presumably immediately adjusts her expectations to the new situation, but five minutes later she will still assess not getting the money as a loss."

But if the receiver makes his decision long before the outcome resulting from it, the concept of UPE would be hard to justify. In this case, the expectations relative to which a decision's outcome is evaluated are formed after—and therefore incorporate the implications of—the decision itself. So when the receiver makes his choice in period 1, this will determine his reference point by the time the outcome occurs. As noted by [Kőszegi and Rabin, 2007], UPE and CPE are not different theories of what outcomes people prefer, but are actually motivated by the same theory of preferences. Instead these concepts affect the revision of expectations by having it depend on whether the receiver can commit to his choice ahead of time or not. These observations make it clear that the interpretation formulated in section II makes use of both UPE and CPE. Having established this connection, the conclusions drawn from the previous paragraphs can be stated as follows.

When the receiver has to choose between settlement and bargain in period 1, the expectations he formed in period 0 are fixed. If he chooses to settle, he therefore evaluates the settlement outcome with respect to his period 0 expectations. Now, if he chooses to bargain, the ramifications are different. Since the bargaining outcome is not determined "shortly" after the decision of bargaining is made, the receiver has committed to the decision long before the outcome occurs. By the definition of CPE, the receiver therefore has time to acclimate his plans, that is, revising his expectations. With the receiver's expectations fully updated when the bargaining outcome is determined—in other words, with the receiver's preferences collapsed to the standard risk-neutral preferences—the bargaining outcome is given by $\langle b, 0 \rangle$, as was formally stated in equation (1).

But notice that the acclimation of the expectations takes place after the receiver has declined the settlement offer and chosen to bargain. When he faces the decision between settlement and bargaining in period 1, the expectations he formed in period 0 are still fixed. So when the receiver evaluates the payoff from bargaining in period 1, he knows that the outcome will be given by $\langle b, 0 \rangle$ if he chooses to bargain. But since his reference point is still fixed by the expectations he formed in period 0, he will compare the potential bargaining outcome with respect to these expectations. Hence, when the receiver decides whether to bargain, his period 0 expectations affect his choice, but if he actually chooses to bargain, his utility will not be affected by loss aversion at all.

Notice how this interpretation mimics the underlying intuition of time-inconsistent preferences. Even though the receiver knows that he could achieve a higher payoff by declining the settlement offer, such a strategy will not be optimal in period 1, since he is still burdened by his period 0 expectations. Along the same lines, a time-inconsistent individual maximizes his utility in a given period by making a choice that is going to be dominated ex-post. The shortcomings of the receiver's behaviour are of course contingent on the uncertainty he faces when forming his expectations. Conversely, the behavioural bias that fosters early gratification when an individual is time-inconsistent is built directly into the preferences. That belief-based preferences can generate time-inconsistent behaviour is not a new observation.

It was first pointed out in [Caplin and Leahy, 2001], and later scrutinized in [Kőszegi, 2010].

Given the chronicle presented in this subsection, and the fact that it is derived from established concepts that make the players' behaviour endogenous, the model of this paper seems to build on a solid foundation. However, as was also recognized, the exact timing that determines how these concepts are applied are built on ad hoc context-dependent narratives. This fact does not make the theory invalid, but it does depreciate the degree of confidence with which one can expound the theory in a general framework; a caveat that is present in any model that combines complex timing with expectation based loss aversion.

Loss aversion coefficient. As was apparent from the example presented in section IV, the receiver may need to be severely loss-averse for an infinite partition equilibrium to be supported. In fact, for any $\eta \in (0, \infty)$, the equilibria of section IV were shown to be supported for $\lambda \in [\frac{\eta+3}{\eta}, \infty)$. The question now is whether this solution set seems reasonable. That people in fact are loss-averse seems today to be a truism with a lot of experimental evidence to support it. But even though consilience is established among several independent experiments, there are still large variations in the data on the personal level. An obvious explanation for this is that the degree of loss aversion varies with the definition of loss aversion used. For an overview of the different elicitation procedures, and a formulation of parameter free estimation, see [Abdellaoui et al., 2007]. As the authors note:

"We found clear evidence of loss aversion both at the aggregate and at the individual level. The degree of loss aversion varied with the definition of loss aversion used."

Models that apply the concept of loss aversion typically assume a loss aversion coefficient around two. In this case, the weight the receiver attaches to the gain-loss part, η , should be at least three times the weight he attaches to intrinsic utility, for the solution set of section IV to support an infinite partition equilibrium. That the receiver should put such a high weight on the gain-loss part does not appear to conform with reality. And because theorem 1 hinges on its validity, the validity of theorem 1 itself should naturally be called into

question. However, as I will now argue, a loss aversion coefficient of two is not at all universal at the individual level.

To see that a loss aversion coefficient of two can be an imprecise assessment at the individual level, I review the top deciles of two different estimations from [Abdellaoui et al., 2007], that is, estimates using different elicitation methods. First, in the local method of [Köbberling and Wakker, 2005] they measure loss aversion at the reference point and therefore provide a natural measurement index for loss aversion to be manifested at the kink of the utility function. Using this method, the top decile is seen to have an average that is larger than ten. Second, in the global method of [Wakker and Tversky, 1993] they scan the specific utility domain under study for any loss aversion measurement. With this method the top decile has an average around four. In the light of these numbers it seems hard to dismiss the possibility that λ could take values around four—values that would be consistent with an η around one.

These observations are by no means an attempt to survey the literature of the estimation of loss aversion or even to provide a rudimentary overview. Instead my goal has been to convey the fact that very large degrees of loss aversion are actually observed empirically, even though the average over a group of people will stay close to the “standard value” of two. This should not be seen as strong empirical justification for the theory, but should instead provide support for the assumption that it could actually be true. Furthermore, the fact that experimental evidence is derived from subjects that—for example—have to give up a coffee mug, begs the question whether the observed degree of loss aversion would change for higher stakes. If an individual has to make a decision that could determine his/her career, it would seem plausible that losses could be felt really hard.

Learning process. Until now, the process in which the receiver discovers his true bargaining power has virtually been left unmentioned. However, the amount of information the receiver possesses before he makes his decision, and the certainty with which this information is obtained, can have significant consequences for the equilibrium dynamics. In this paper I have assumed a favorable environment for the receiver where he was able to acquire a complete picture of the bargaining relation. The reasons for this assumption

were twofold. First of all, it provided a simplified framework where the number of technical derivations were minimized. Second, it was motivated by the central inquiry put forward in the introduction: If the sender is able to obtain a payoff that exceeds the bargaining payoff—at the expense of the receiver—in the most favorable environment for the receiver, then it would seem reasonable to extrapolate the same conclusion to environments where the receiver has worse conditions.

Other kinds of learning processes could easily be incorporated into the model. When the receiver collects information in the market, he may not necessarily observe his true bargaining power with probability one; instead, he may observe his true bargaining power with probability $p \in (0, 1)$. Furthermore, the sender's settlement offer could be assumed to affect the receiver's learning process. The settlement offer would then get another strategic dimension that potentially could foster new economic insight, and could also help strengthen or weaken the conclusion of this paper. How such alterations would affect the equilibrium dynamics is, however, outside the scope of this discussion. But it seems intuitively appealing that the sender should be able to manipulate the receiver to her advantage when the receiver faces a less favorable environment.

VI. CONCLUSION

In economics the risk attitude of an individual tends to play an important role in how that individual behaves. In the bargaining literature, much attention has been paid to the influence of risk attitudes, and in particular risk aversion, on the outcomes assigned by specific bargaining solutions. In those studies, it is found that bargaining solutions often favor the less risk averse player; in other words, being constrained by a severe aversion to risk can limit an individual's ability to obtain a favorable bargaining outcome (see for example [Kihlstrom et al., 1981]). With loss aversion being increasingly seen as the primary explanation for individuals' distaste for risk, the present paper has analyzed how this risk-concept can affect asymmetric Nash bargaining.

The concept of loss aversion is not new to the bargaining literature, and it has been applied to the Nash bargaining solution and extensions in theoretical and experimental work (see [Shalev, 2002] and [Driesen et al., 2011]).

What these papers have in common is that loss aversion is built directly into the solution concept; both make use of [Shalev, 2000]’s work on loss aversion equilibrium.⁸ As with risk aversion, these studies find that increasing the loss aversion for a player leads to worse outcomes for that player in bargaining situations. However, more recent evidence lends support to the expectation-based model of reference-dependent preferences formulated by [Kőszegi and Rabin, 2006]. For example, in simple experiments a significant number of subjects stop working when they have earned the amount of money they had expected to earn (see [Abeler et al., 2011]), and they are also more likely to keep an item they had received if they had been expecting a lower probability of being able to exchange it (see [Ericson and Fuster, 2011]).

My contribution has been to incorporate this expectation-based model into a coherent framework with the asymmetric Nash bargaining solution; albeit in a way that differs from previous work. By acknowledging the strategic interactions that happen “outside” the bargaining solution, this paper has shown how a privately informed sender can affect the expectations of a loss-averse receiver in such a way that the receiver ends up being worse off than he would be in fair bargaining. This result lends support to the findings that loss-averse players in general are worse off. The justifications, however, were seen to be different. As was delineated in section V, the intertemporal structure of the receiver’s formation of expectations is a key factor in accomplishing the sender’s advantageous position. In other words, the fact that the sender is able to manipulate the receiver’s expectations before the actual bargaining process begins, is the foundation of the conclusions that are drawn from theorem 1.

These conclusions can help bring new insights into a number of politically relevant discussions. Labor unions, for example, whose main activity today centers on collective bargaining over wages, benefits, and working conditions, have often been criticized by economists. However, some of the positive reflections on the role of labor unions have been the increase in bargaining power they bring to their members. This argument can now be extended. Since individuals will often be uncertain of their bargaining position before they enter bargaining with a potential employer, an informed organization

⁸ [Shalev, 2000] was the first paper to formalize the idea that expectations determine the reference point and to specify a rule for deriving them endogenously in any environment.

would be able to obtain an unfair share. But if individuals group together in large risk-neutral groups—that is, unions—that are also more likely to be well informed, it could limit the possibility of exploitation. This would also support the use of transparency policies that can help protect the individual. By making workers' rights and data on the industry easily available, the government could potentially limit the degree to which workers are exploited. Applications can of course also be found outside labor economics, where many real life disputes from medical malpractice to consumer complaints are found to be resolved through settlement.

REFERENCES

- [Abdellaoui et al., 2007] Abdellaoui M., Bleichrodt H., and Paraschiv C. (2007): "Loss aversion under prospect theory: A parameter-free measurement," *Management Science*, 53: 1659–1674.
- [Abeler et al., 2011] Abeler J., Falk A., Götte L., and Huffman D. (2011): "Reference points and effort provision," *American Economic Review*, 101: 470–492.
- [Caplin and Leahy, 2001] Caplin A., and Leahy J. (2001): "Psychological expected utility and anticipatory feelings," *Quarterly Journal of Economics*, 116: 55–79.
- [Cho and Kreps, 1987] Cho, I., and Kreps D. (1987): "Signaling games and stable equilibria," *Quarterly Journal of Economics*, 102: 179–221.
- [Driesen et al., 2011] Driesen, B., Perea A., and Peters H. (2011): "The Kalai-Smorodinsky bargaining solution with loss aversion," *Mathematical Social Sciences*, 61(1): 58–64.
- [Ericson and Fuster, 2011] Ericson, K.M., and Fuster A. (2011): "Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments," *Quarterly Journal of Economics*, 126: 1879–1907
- [Kahneman et al., 1991] Kahneman, D., Knetsch L.J., and Tversky A. (1991): "Anomalies: The endowment effect, loss aversion, and status quo bias," *Journal of Economic Perspectives*, 5(1): 193–206.
- [Kahneman and Tversky, 1979] Kahneman, D., and Tversky, A. (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, XLVII 263–291.
- [Kalai, 1977] Kalai, E. (1977): "Nonsymmetric Nash Solutions and Replications of Two-Person Bargaining," *International Journal of Game Theory* 6: 129–33.
- [Kihlstrom et al., 1981] Kihlstrom, R.E., Roth A.E., and Schmeidler D. (1981): "Risk aversion and solutions to Nash's bargaining problem," *Game Theory and Mathematical Economics*, (O. Moeschlin and D. Pallaschke, eds.), North-Holland, Amsterdam.
- [Kőszegi, 2010] Kőszegi, B. (2010): "Utility from anticipation and personal equilibrium," *Economic Theory*, 44, 415–444.

- [Kőszegi and Rabin, 2006] Kőszegi, B., and Rabin M. (2006): "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121(4):1133-65.
- [Kőszegi and Rabin, 2007] Kőszegi, B., and Rabin M. (2007): "Reference-Dependent Risk Attitudes," *American Economic Review*, 97(4): 1047-73.
- [Köbberling and Wakker, 2005] Köbberling, V., and Wakker P.P. (2005): "An index of loss aversion," *Journal of Economic Theory*, 122: 119-131.
- [Leland and Pyle, 1977] Leland, H., and Pyle D. (1977): "Information asymmetries, financial structure and financial intermediaries," *Journal of Finance*, 32: 371-387.
- [Nash, 1950] Nash, J.F. (1950): "The Bargaining Problem," *Econometrica* 6: 18: 155-62.
- [Nash, 1953] Nash, J.F. (1953): "Two-Person Cooperative Games," *Econometrica* 21: 128-40.
- [Schelling, 1963] Schelling, T.C. (1963): "The Strategy of Conflict," *Oxford University Press*.
- [Shalev, 2002] Schelling, J. (2002): "Loss aversion and bargaining," *Theory and Decision* 52(3):201-232.
- [Shalev, 2000] Schelling, J. (2000): "Loss Aversion Equilibrium," *International Journal of Game Theory* XXIX 269-287.
- [Wakker and Tversky, 1993] Wakker, P.P., and Tversky A. (1993): "An axiomatization of cumulative prospect theory," *Journal of Risk Uncertainty*, 7: 147-176.