

**SAM 14 2011****ISSN: 0804-6824**

July 2011

**Discussion paper**

# **Fairness motivation in bargaining**

BY  
**Sigbjørn Birkeland d.y.**

This series consists of papers with limited circulation, intended to stimulate discussion.

# Fairness motivation in bargaining\*

Sigbjørn Birkeland d.y.<sup>†</sup>  
Department of Economics  
Norwegian School of Economics

July 2011

## Abstract

In this paper, we develop a model that captures the potential conflict between two individuals who follow different fairness principles in bargaining. This model is used to analyse the influence of fairness motivation on the possibility of reaching an agreement in bargaining, and to examine the properties of the agreement. We show that bargaining between two individuals who are strongly fairness motivated, but who disagree about what represents a fair division, ends in disagreement. This result contrasts the standard bargaining model with individuals who are only motivated by material self-interest, which always leads to agreement. Furthermore, by applying the Nash bargaining solution, we study the influence of fairness motivation on the bargaining outcome. A fairness motivated individual reaches an outcome that is closer to his fairness principle in bargaining against an individual who is only motivated by material self-interest.

---

\*I would very much like to thank Bertil Tungodden, Alexander Cappelen, James Konow, and Shachar Kariv for valuable comments and helpful discussions. Thanks also to participants at the 5th Nordic Conference on Behavioural and Experimental Economics at Aalto University.

<sup>†</sup>Address: Helleveien 30, 5045 Bergen, Norway. E-mail [sigbjorn.birkeland@nhh.no](mailto:sigbjorn.birkeland@nhh.no).

# 1 Introduction

An equal division of monetary rewards is a frequent outcome in many laboratory experiments in bargaining, and it is a common principle in many real life situations, for example, bequests to children (Camerer, 2003; Wilhelm, 1996). In other situations, for example, in bargaining over the output from production, experiments in economics and psychology have shown that many people follow a principle of proportionality, although a minority still prefer the equal division principle (Konow, 1996; Gächter and Riedl, 2005; Cappelen, Sørensen, and Tungodden, 2010). Fairness principles such as equality and proportionality may arise from moral or political philosophy or simply be accepted over time as a way of dealing with distributive issues.

Hirschman (1977) and Elster (1989) have pointed to the fact that bargaining between individuals who strongly believe in different fairness principles can easily lead to conflict. They both argue that material self-interest can moderate conflicts of fairness principles, in the words of Elster (1989):

The last case, norm conflict, is less likely to yield negotiated solutions. In norm-free bargaining, the only thing at stake is self-interest, a mild if mean-spirited passion. In norm conflict, the parties argue in terms of their honour, a notoriously strong passion capable of inspiring self-destructive and self-sacrificial behaviour. . . . Compromises are possible between opposing norms, if one or both parties pour some water in their wine and let self-interest override honour. (Elster, 1989, p. 244).

In this paper, we develop a model that captures the potential conflict between two individuals who follow different fairness principles in bargaining. An individual's preferences are represented by a utility function where he or she trades off material self-interest and deviations from a fairness principle. Bargaining experiments such as the ultimatum game show that people are willing to trade off monetary rewards to achieve a more fair outcome from bargaining (Camerer, 2003).

The model developed in this paper is a variation of the frequently used inequity aversion model, which assumes that bargainers agree on a principle of equal division (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Bruyn and Bolton, 2008). The model in this paper builds on Cappelen, Hole, Sørensen, and Tungodden (2007), which allows individuals to follow different fairness principles. This introduces a dimension of conflict between two bargainers who follow different fairness principles, in addition to the trade-off between material self-interest and fairness motivation.<sup>1</sup>

---

<sup>1</sup>It is acknowledged in many models that different fairness principles should be considered, but are left out for reasons of intractability. See, for example, Andreoni and Bernheim (2009)

Fairness motivation can influence both the possibility of reaching an agreement in bargaining and it can influence the properties of the agreement that is reached. The first part of the paper studies how fairness motivation influences the possibility of reaching an agreement. Proposition 2 formalizes the intuition of Elster (1989), that bargaining between two individuals who strongly believe in different fairness principles ends in conflict. This result shows the importance of considering a plurality of fairness principles to understand many bargaining problems. In contrast, Proposition 3 formalizes that if two bargainers follow the same fairness principle, it is always possible to reach an agreement.

The second part of the paper analyses the properties of the agreements that can be reached. We apply the Nash bargaining solution to bargaining situations between different types of individuals. We find that bargaining between an individual with strong fairness motivation and an individual motivated only by material self-interest, reaches an agreement that is closer to the fairness motivated individual's principle. If two bargainers who are motivated by fairness, but who disagree about what represents a fair division, reach an agreement, it will be a compromise between the two fair shares. We also show that in a bargaining situation where both individuals follow a fairness principle of strict equality, the Nash bargaining solution gives an equal division, and the trade-off between material self-interest and fairness motivation does not influence the solution.

Empirical studies have found that people follow a plurality of fairness principles in negotiations, and that this can explain bargaining impasses and how these are solved (Bazerman, 1985; Babcock, Loewenstein, Issacharoff, and Camerer, 1995; Babcock, Wang, and Loewenstein, 1996). Section 2 discusses data from a bargaining experiment that shows that it is also important to include a plurality of fairness principles to understand the properties of the agreements that are reached in bargaining. The agreements are from a bargaining experiment where participants have individually produced the endowment before they bargain over a division of the endowment. The two most common models for bargaining problems, material self-interest and preferences for equality, do not explain the experimental data in Section 2.

The model is presented in Section 3. Bargaining is then analysed in two steps in Section 4. First, the influence of fairness on the bargaining set is discussed without relying on a specific solution concept, and second, the Nash bargaining solution is applied to the problem. Section 5 contains some concluding remarks.

---

who discuss fairness norms other than the 50–50 norm: ‘If the players are asymmetric with respect to publicly observed inertia of merit, the fairness of an outcome might depend on the extent to which it departs from some other benchmark, such as  $x^F = 0.4$ . Provided the players agree on  $x^F$ , similar results would follow, except that the behavioural norm would correspond to the alternate benchmark. However, if players have different views of  $x^F$ , matters are more complex’ (footnote 12).

## 2 Experiment

The laboratory experiment reported in Birkeland (2011) illustrates the importance of including fairness principles other than equality in bargaining. This experiment consists of a production phase and a bargaining phase. First, participants produce individually an output by typing a text from a transcript on the computer, and they receive a monetary reward equal to each correct word typed, rounded off to the nearest 50 words, multiplied by a randomly assigned high or low price. Second, participants were randomly matched into pairs and instructed to bargain over the endowment, which in this experiment is the sum of the individual production values. The experiment used an alternating offer bargaining protocol with infinite horizon.<sup>2</sup>

The 112 bargaining outcomes are shown in Figure 1, where the share of the total production value to person one and to person two are on the axes. The left panel shows the outcomes from 15 situations where both bargainers have produced the same amount and have the same price, and the right panel shows the outcomes from 97 situations where there is a difference in either the amount produced or the price. All the points that are along the diagonal from the lower left corner to the upper right corner are equal splits of the production value.

The results show that in all of the situations where bargainers have produced the same amount (left panel), the bargaining outcome is an equal division (the circle indicates the 10 observations that are equal divisions of the initial production value).<sup>3</sup> This result is consistent with the standard solution for players motivated by material self-interest. Alternating offer bargaining between players who are motivated by material self-interest, and have equal discount factors, gives an almost equal split (Rubinstein, 1982).

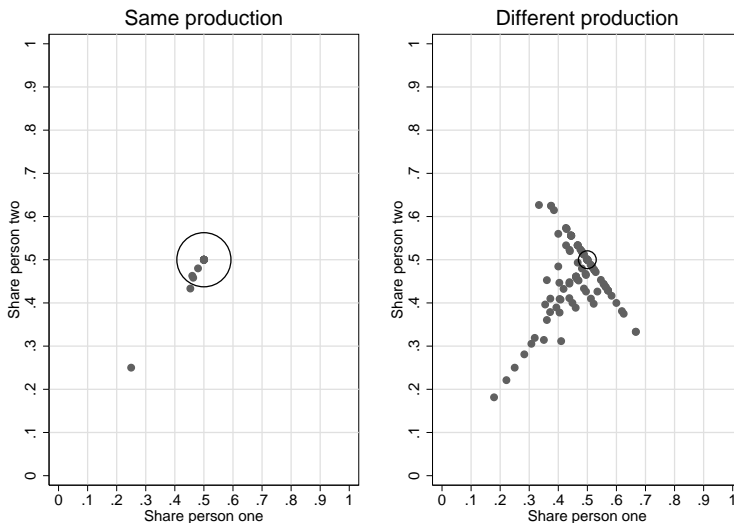
In the right panel, 49% of the outcomes are equal divisions that give the same amount of money to both participants. The number of equal divisions is significantly reduced when there are differences between the players from the production phase of the experiment. These observations are inconsistent with the prediction of the standard model where bargainers are only motivated by material self-interest. We show later in the paper that a model where bargainers are motivated by strict equality cannot explain these observations either. A likely

---

<sup>2</sup>The alternating offer protocol starts with one of the players being randomly assigned as the first mover who suggests an opening offer in the first round ( $t = 1$ ). Individual  $i$  proposes an amount of pay-off  $x_i$  for himself and  $Y - x_i$  for the other player in each round of bargaining. The second mover responds to the opening offer by either accepting it and the bargaining is closed without cost, or by giving a counter offer in a second round ( $t = 2$ ). The endowment shrinks in each round  $t$  by a discount factor  $\delta_i^t$ . An agreement is reached when one player accepts the offer from the other player. In the experiment discussed in this paper, both players were induced with an equal discount factor,  $\delta = 0.96$ , which is so high that there is an insignificant first-mover advantage.

<sup>3</sup>To accommodate rounding to the nearest NOK 5, all agreements within the 47.5–52.5 split range are characterized as equal splits.

Figure 1: Experimental bargaining results



*Note:* The left panel shows outcomes from 15 situations where bargainers have the same production value, and the right panel shows bargaining outcomes from 97 situations where there are differences between bargainers in terms of either the amount produced or the price. The circle indicates the number of observations that are exactly a 50–50 split of the initial production value.

explanation for this shift to a more unequal division of the production value is that bargainers are motivated by fairness principles that do not imply equal division in these situations. A post-experimental questionnaire confirmed this hypothesis: 96% of the participants supported principles that justify unequal division in these situations, whereas only 4% of the participants supported equal division. Thus, this experiment emphasizes the importance of allowing for fairness principles other than equal division in economic models of bargaining to understand better many bargaining problems.

### 3 Model

In this section, we describe the theoretical framework for the analysis, including the bargaining environment and a utility function that can accommodate bargainers who are motivated by different fairness principles. We consider a bargaining environment in which two players bargain over how to divide an endowment,  $Y$ . Players can agree on any pair  $x = (x_1, x_2)$  of shares of the endowment,  $x_i \in [0, 1]$ , such that the pair of shares is in the set  $X = \{x \mid x_1 + x_2 \leq 1\}$ , which is called the set of feasible agreements. In the following, we assume complete information, that is, the rules of the game and the utility functions of both players are common knowledge.

### 3.1 Fairness principle

An individual is assumed to have preferences that can be represented by a utility function where deviation from a fair share of the endowment reduces utility. The fair share of the endowment to individual  $i$ , according to his fairness principle  $k$ , is denoted as  $s^{k(i)} \in [0, 1]$ . We assume that the fairness principle gives a unique division of the endowment, which is the case for all the fairness principles discussed in this paper.<sup>4</sup>

**Assumption 1.** *For any endowment,  $Y$ , and fairness principle,  $k$ , there exists a unique fair division  $(s^{k(1)}, s^{k(2)})$ , such that  $s^{k(1)} + s^{k(2)} = 1$ .*

The following example illustrates how principles of fairness could be applied in a production context. Consider a case where the endowment,  $Y$ , is the sum of individual production values,  $y_i$ , which can be decomposed into the individual production of units,  $e_i$ , and price,  $p_i$ , such that  $y_i = e_i p_i$ . In a two-person case, let  $e_1 = 3$ ,  $e_2 = 1$ ,  $p_1 = \frac{1}{3}$ , and  $p_2 = 3$ , then the production values are  $y_1 = 1$  and  $y_2 = 3$ , and the endowment  $Y = 4$ . This could, for example, be bargaining between two executives about their share of a bonus in a corporation where one business area is exposed to the oil price and another business area is exposed to the aluminium price. In this context, there are three different distributive principles that are salient,  $k = E, L, P$ . The first fairness principle is an equal sharing of the monetary rewards, *strict equality*, which implies a fair share  $s^{E(i)} = \frac{1}{2}$ . The second fairness principle is a *laissez-faire* principle, where the individual production values determine the fair share to individual  $i$ ,  $s^{L(i)} = \frac{e_i p_i}{Y}$ . The third principle is *proportionality*, where the fair share to individual  $i$  is proportional to the level of the production of units,  $s^{P(i)} = \frac{e_i}{e_1 + e_2}$ , but where prices have no influence on the division.

In our example, implementation of these three fairness principles for two individuals gives the nine combinations of fair shares in Table 1. Each entry in the table shows the fair share that person one and person two claim according to their fairness principle, if that is only what they care about. The combinations of fairness principles can be divided into three categories: (i) both players follow the *same* fairness principle (diagonal elements), where by Assumption 1 the fair shares are always compatible; (ii) players one and two follow different fairness principles such that fair shares are incompatible,  $s^{k(1)} + s^{k(2)} > 1$ ; (iii) players one and two follow different fairness principles such that fair shares sum to less than one,  $s^{k(1)} + s^{k(2)} < 1$ . Category (ii) is a natural bargaining situation where there

---

<sup>4</sup>The relevant set of fairness principles must be specified for the context in the model is applied. Which fairness principle an individual follows in a particular context, may depend on his identity; for example, an individual may follow a different fairness principle if he is an employer or if he is an employee (Akerlof and Kranton, 2010). The formulation of a fairness principle as a fair share of the endowment excludes some possible fairness principles, for example, principles that are related to the size of the endowment.

Table 1: Combinations of fairness principles

	$s^E(2)$	$s^L(2)$	$s^P(2)$
$s^E(1)$	$(\frac{1}{2}, \frac{1}{2})$	$(\frac{1}{2}, \frac{3}{4})$	$(\frac{1}{2}, \frac{1}{4})$
$s^L(1)$	$(\frac{1}{4}, \frac{1}{2})$	$(\frac{1}{4}, \frac{3}{4})$	$(\frac{1}{4}, \frac{1}{4})$
$s^P(1)$	$(\frac{3}{4}, \frac{1}{2})$	$(\frac{3}{4}, \frac{3}{4})$	$(\frac{3}{4}, \frac{1}{4})$

*Note:* The table shows combinations of the fairness principles  $E, L, P$  for person one and person two for parameters  $e_1 = 3, e_2 = 1, p_1 = \frac{1}{3}, p_2 = 3$ . Each entry shows the fair shares that person one and person two claim according to their principles if they only care about fairness.

is conflict of interest. Category (iii) is less important in bargaining and will not be discussed in the following analysis.

In this numerical example, different fairness principles give different fair shares, but this is not necessarily the case in all situations. Different principles can also give the same fair shares in some situations; for example, an equal division follows both from the fairness principle of strict equality, and from the principle of proportionality if both individuals have produced the same number of units.

### 3.2 Utility function

We assume that the utility function is additively separable for individual  $i$  in his own share of the endowment,  $x_i$ , and the cost of deviating from the fair share,  $x_i - s^{k(i)}$ . The endowment is assumed to be non-negative,  $Y \geq 0$ .

**Assumption 2.** *Individual  $i$ 's preferences can be represented by the utility function:*

$$u_i(x_i Y, s^{k(i)} Y) = (x_i - \beta_i (x_i - s^{k(i)})^2) Y.$$

The utility loss from deviating from the fair share is squared, which implies that the utility loss from deviation to the better or the worse is symmetric, and that the utility loss increases exponentially with the distance. The weight individual  $i$  has on not deviating from his fair share is given by the parameter  $\beta_i$ , which is assumed to be non-negative,  $\beta_i \geq 0$ . If  $\beta_i = 0$ , the model is reduced to material self-interest, a utility function that is linear in  $x_i$ . The parameter  $\beta_i$  captures tension between an individual's motivation to follow a fairness principle and that of material self-interest.<sup>5</sup> This utility function is continuous, twice

<sup>5</sup>The functional form of the utility function follows the two-person case of Bolton and Ockenfels (2000), Lopomo and Ok (2001), Cappelen et al. (2007), and Bruyn and Bolton (2008). The parameter  $\beta_i$  captures the trade-off that is represented by the fraction  $\frac{b_i}{2a_i}$  in Bolton and Ockenfels (2000). The utility function in this paper differs from Bolton and Ockenfels (2000)



differentiable and concave in  $x_i$ . The utility function attains its inner maximum when:

$$x_i^* = \frac{1}{2\beta_i} + s^{k(i)}.$$

The interior solution to an individual,  $x_i^*$ , is not defined for  $\beta_i = 0$ , and it is independent of the endowment  $Y$ . The utility function is strictly increasing in the interval  $0 \leq x_i \leq s^{k(i)}$ . Only for small values of  $\beta_i$ , that is,  $\beta_i < \frac{1}{2(1-s^{k(i)})}$ , is the utility function strictly increasing in the entire interval  $0 \leq x_i \leq 1$ . For high values of  $\beta_i$ , the importance of obtaining the fair share outweighs the utility of obtaining a larger share of the endowment, see also Bolton and Ockenfels (2000). In the case that  $\beta_i \rightarrow \infty$ , the interior solution approaches the fair share,  $x_i^* \rightarrow s^{k(i)}$ .

This model allows for different types of players: a material self-interested player, a strongly fairness motivated player, as well as an intermediate type that trades off material self-interest and fairness motivation. It also introduces differences regarding which fairness principle players follow.

The effects of changing the parameters of the weights assigned to fairness,  $\beta_i$ , and the fair share,  $s^{k(i)}$ , are illustrated in Figure 2. We observe that all three utility functions illustrated in the figure attain negative values for a small  $x$ , but only the utility function with a high  $\beta_i$  attains negative values for a large  $x$ . Shifting the fair share,  $s^{k(i)}$ , to a higher value raises the point where the utility function is zero. We also observe that the utility function with a high  $\beta_i$  envelops the utility function with the same fair share,  $s^{k(i)}$ , but a lower  $\beta_i$ , because an increase in fairness motivation reduces the utility from deviating from the fair share.

In bargaining, people evaluate their utility from possible agreements against the utility from the situation where no agreement is reached. We assume that the disagreement utility is zero.

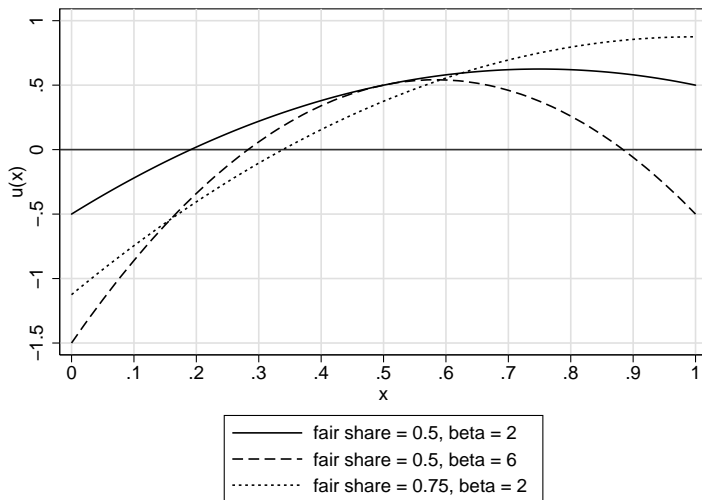
**Assumption 3.** *Individual  $i$ 's utility from disagreement is zero,  $u_i^d = 0$ .*

Assumption 3 follows directly from the specification of the utility function in an economic environment where the endowment is zero in disagreement,  $Y = 0$ , which is the case in many bargaining experiments. This assumption implies,

---

in that it allows for fair shares other than  $s^{k(i)} = \frac{1}{2}$ . In the model of Lopomo and Ok (2001), a bargainer's utility depends on both his absolute gain and the relative share he gets compared with the average share. The model of Lopomo and Ok (2001) allows for uncertainty about the weight that the other players have on the deviation from the average share. The utility function in Cappelen et al. (2007) allows for different fairness principles, but the principles are not defined in shares of the endowment. The utility function used in Bruyn and Bolton (2008) is linear above an equal division. If you introduce players with different fairness principles in a model with asymmetries in valuing deviations from a fair share, e.g., the Bruyn and Bolton (2008) model, this could easily result in a non-convex Pareto frontier of the bargaining set, which could give multiple Nash bargaining solutions.

Figure 2: The utility function



*Note:* The utility function given in Assumption 2 for different parameter values,  $Y = 1$ .

however, that an individual's fairness consideration does not apply to the disagreement outcome.

### 3.3 Reservation points

An individual's reservation point is defined as the share that makes an individual indifferent between accepting an offer or choosing disagreement, which gives zero utility. As illustrated in Figure 2, the present model allows for more than one reservation point. The lower reservation point,  $x_i^L$  in the interval  $0 \leq x_i^L \leq s^{k(i)}$ , where  $u_i(x_i Y, s^{k(i)} Y) = 0$  is given by:

$$x_i^L = \frac{1 + 2\beta_i s^{k(i)} - \sqrt{1 + 4\beta_i s^{k(i)}}}{2\beta_i}.$$

The reservation point is influenced both by the fairness principle and the weight that an individual attaches to following the fairness principle. For a fairness motivated individual,  $\beta_i > 0$ , an offer below the lower reservation point would be considered too unfair, and it is therefore rejected, although the offer represents a positive share of the endowment. In contrast, the utility function for an individual who is only motivated by material self-interest,  $\beta_i = 0$ , always has a lower reservation point at zero, and he would not reject a positive share of the endowment.

An important property of the lower reservation point is that when  $\beta_i$  increases the reservation point,  $x_i^L$ , approaches the fair share:

$$\lim_{\beta_i \rightarrow \infty} x_i^L = s^{k(i)}.$$

We focus on the lower reservation because it is used later in the paper to analyse the bargaining set. For large values of  $\beta_i$ , there is also an upper reservation point,  $x_i^H$ , in the interval  $s^{k(i)} \leq x_i^H \leq 1$ . The upper reservation point will correspond to situations where a player, for example, is offered the whole gain, and he rejects this as unfair even though it benefits him.

**Proposition 1.** Reservation points. *For any fair share  $s^{k(i)}$ , there is a value  $\hat{\beta}_i$  such that for any  $0 < \beta_i < \hat{\beta}_i$  there exists a unique reservation point,  $x_i^L$ , and for any  $\beta_i \geq \hat{\beta}_i$  there exist two reservation points,  $x_i^L$ , and  $x_i^H$ .*

*Proof.* See Appendix A. □

## 4 Bargaining solutions

In this section, we use cooperative bargaining theory to analyse the outcome from bilateral bargaining with fairness motivated individuals. A cooperative bargaining solution is not based on a specific bargaining process. It is assumed that an agreement from bargaining is binding and enforceable through a legal system outside of the model. Cooperative bargaining theory builds on a utility representation of the feasible agreements, which is the convex utility set:<sup>6</sup>

$$\mathcal{U} = \{(u_1(x_1Y, s^{k(1)}Y), u_2(x_2Y, s^{k(2)}Y)) : x \in X\}.$$

First, following standard analysis we require that the agreement is at the frontier of the utility set, which implies that the agreement is in the set  $Z = \{x \mid x_1 + x_2 = 1\}$ .<sup>7</sup> Second, we require that the agreement gives both players at least as much utility as they can get without an agreement, which from Assumption 3 implies that the agreement must give both players at least zero utility. These two requirements ensure that the agreement is in the *bargaining set*,  $\mathcal{B}$ .<sup>8</sup>

**Assumption 4.** *An agreement is in the bargaining set:*

$$\mathcal{B}(\mathcal{U}) = \{(u_1(x_1Y, s^{k(1)}Y), u_2(x_2Y, s^{k(2)}Y)) \geq (0, 0) : x \in Z\}.$$

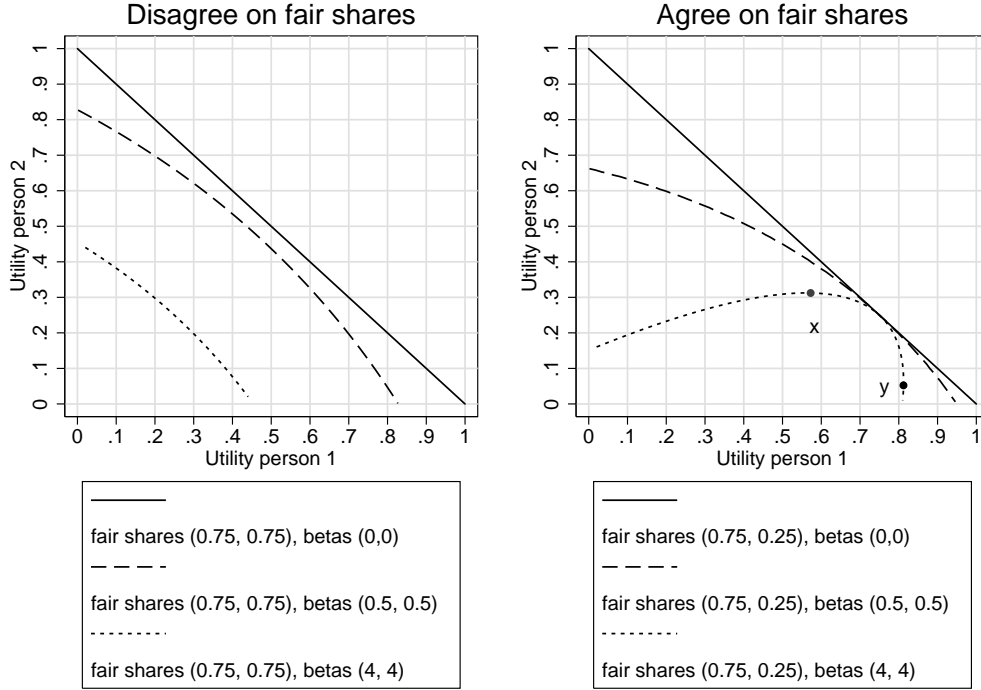
Bargaining sets for two individuals are illustrated in Figure 3, where the left panel shows individuals who disagree on what is a fair division, and the right panel shows individuals who agree on what is a fair division. Each line represents a bargaining set for different parameter values for person one and person two. The two end-points of a bargaining set are defined where person one gets maximum utility given that person two gets enough utility to accept the agreement, which must be at his lower reservation point,  $(u_1^{max}, u_2(x_2^L) = 0)$ , and similarly where person two gets maximum utility given that person one gets enough utility to accept the agreement  $(u_1(x_1^L) = 0, u_2^{max})$ .

<sup>6</sup>Two strictly increasing concave functions give a convex combination (Binmore, 2007).

<sup>7</sup>Roth (1979) discusses the standard assumptions of cooperative bargaining theory.

<sup>8</sup>Here,  $\geq$  is defined coordinatewise, that is,  $(x_1, x_2) \geq (y_1, y_2)$  iff  $x_i \geq y_i$  for each  $i = 1, 2$ .

Figure 3: Bargaining sets



*Note:* The left panel shows bargaining sets where players have incompatible fair shares, and the right panel shows bargaining sets where they agree on fair shares. The lines represent different levels of  $(\beta_1, \beta_2)$ . The endowment  $Y = 1$ . Points marked  $x$  and  $y$  are the end-points of the Pareto frontier for the dotted line.

We observe that the bargaining set for fairness motivated individuals is typically smaller than the bargaining set for players who are more motivated by material self-interest. The bargaining set shrinks when  $\beta_i$  increases because an individual gets disutility from deviating from the fair share, and he consequently increases his reservation point. A smaller bargaining set means that there are fewer possible agreements that can be realized. It is often argued that the frequency of disagreements in bargaining is higher when there are fewer possible agreements that can be realized. The relationship between the size of the bargaining set and the efficiency of bargaining is, however, unclear. Crawford (1982), for example, develops a bargaining model where disagreements are reduced when the size of the bargaining set shrinks.

The left panel in Figure 3 represents bargaining sets where players have incompatible fair shares. In this case, the bargaining set moves towards the origin for bargainers who are more fairness motivated, and at some point, the bargaining set is empty, which occurs when reservation points are incompatible, i.e., the reservation points combined constitute more than the endowment. Consequently,

players prefer to disagree. This shows that strong preferences for conflicting fairness principles make it impossible to reach an agreement.

**Proposition 2.** *Principled disagreement. If the fair shares are incompatible,  $s^{k(1)} + s^{k(2)} > 1$ , then there exists a  $(\hat{\beta}_1, \hat{\beta}_2)$  such that for any  $(\beta_1 \geq \hat{\beta}_1, \beta_2 \geq \hat{\beta}_2)$ , the only feasible solution is the disagreement outcome.*

*Proof.* See Appendix A. □

The right panel in Figure 3 represents bargaining sets where both players follow the same fairness principle. We observe from Figure 3 that for players who are more fairness motivated, the bargaining set shrinks and envelopes the point that both players consider a fair division. The bargaining set shrinks further for players who are strongly motivated by the fairness principle, and in the limit, the bargaining set only contains the fair solution. Proposition 3 formalizes the point that, for bargainers who follow the same fairness principle, it is always possible to reach an agreement.

**Proposition 3.** *Principled agreement. If two individuals follow the same fairness principle,  $k$ , then there always exists a non-empty bargaining set,  $\mathcal{B}$ . Increases in  $\beta_1$  and  $\beta_2$  give a shrinking bargaining set, and in the limit, it collapses to a single point, which represents the fair division  $(s^{k(1)}, s^{k(2)})$ .*

*Proof.* See Appendix A. □

The reservation point of a player can drop below the other player's utility maximizing offer, and still give the first player more utility than in disagreement. This occurs when the two bargainers agree on the fairness principle, and at least one of them is strongly fairness motivated. The bargaining set marked with a dotted line in the right panel in Figure 3 has two points marked  $x$  and  $y$ . The line connecting these two points is the Pareto frontier.<sup>9</sup> A high  $\beta_i$  changes the curvature of the frontier of the bargaining set such that the line segment up to the point marked  $x$  represents a Pareto improvement for player one, and the line segment up to the point marked  $y$  represents a Pareto improvement for player two. Pareto optimality is a requirement for the Nash bargaining solution discussed in the next section.

---

<sup>9</sup>The point marked  $x$  in Figure 3 is defined as the maximum utility that player two can achieve, given that no further Pareto improvement for player one is possible,  $(\tilde{u}_1, u_2^{max})$  where  $\tilde{u}_1 > u_1(x^L) = 0$ , and the point marked  $y$  is the maximum utility that player one can achieve, given that no further Pareto improvement for player two is possible,  $(u_1^{max}, \tilde{u}_2)$ , where  $\tilde{u}_2 > u_2(x^L) = 0$ .

## 4.1 Nash bargaining solution

A commonly used concept for finding a unique outcome in the bargaining set is the Nash bargaining solution (Nash, 1950).<sup>10</sup> The Nash bargaining solution is the maximum of the product of the utility minus the utility of disagreement,  $u_i^d$ :

$$\max(u_1 - u_1^d)(u_2 - u_2^d).$$

The analytical solution to the Nash bargaining solution for the model developed in this paper is derived in Appendix B. The non-linearity of the utility function makes the analytical solution difficult to interpret, but the effect of changing the parameters can easily be interpreted by studying numerical computations.

Figure 4 shows the Nash bargaining solution for different combinations of players. Each point on the four panels shows the share that player one receives,  $x_1$ , at different levels of the fairness weight,  $\beta_1$ .<sup>11</sup> The four panels show matching of player one against different types of player two. The standard solution for two players only motivated by material self-interest is an equal division of the monetary gain, which is the starting point in the upper left panel. This panel shows that if player one has a higher weight on following his fairness principle, the bargaining solution gives a share that is closer to his fair share, which in this example is  $s^{k(1)} = \frac{3}{4}$ . A fairness motivated player who takes a principled stand in bargaining will achieve a solution that is closer to his fairness principle if he bargains against a player who is only motivated by material self-interest.

In the upper right panel both players are fairness motivated, and they agree on the fair division ( $s^{k(1)} = \frac{3}{4}, s^{k(2)} = \frac{1}{4}$ ). We observe that all the bargaining solutions are close to the fair division. An increase in the trade-off between self-interest and fairness motivation,  $\beta_1$ , has an insignificant effect on the bargaining solution. In line with Proposition 2, sufficiently high weights on following the same fairness principle give the fair division.

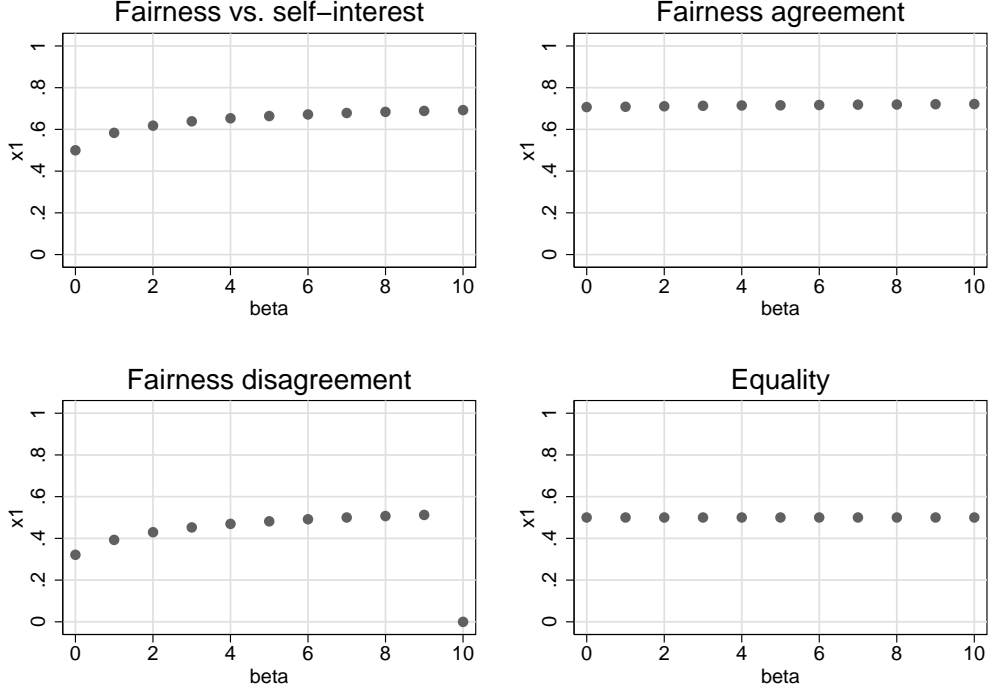
In the lower left panel, player two is also a fairness motivated player, but in this example player two disagrees with player one about the fairness principle. Both players believe that it would be fair if they get three-quarters of the endowment; thus, they both have the same fair share ( $s^{k(1)} = \frac{3}{4}, s^{k(2)} = \frac{3}{4}$ ). From the lower left panel, we see that at low levels of  $\beta_1$ , the bargaining solution is a division that is close to player two's fair share. At high levels of  $\beta_1$ , the bargaining solution is a compromise solution between the fair shares, which in this example is an equal

---

<sup>10</sup>Nash (1950) proves that this is the only solution that fulfils four reasonable axioms: (i) the solution should be independent of affine transformations of the utility function; (ii) the solution should be independent of irrelevant alternatives; (iii) the solution should treat players symmetrically; and (iv) the solution should be Pareto optimal. An introduction to bargaining theory and the Nash axioms is found in Roth (1979) and Binmore (2007).

<sup>11</sup>Bruyn and Bolton (2008) and Cappelen et al. (2007) estimate the average weight that an individual has on following his fairness principle. The average weight, converted to a value comparable to  $\beta_i$ , is 6.0 for a three-round bargaining game in Bruyn and Bolton (2008), and 7.7 for a dictator game in Cappelen et al. (2007).

Figure 4: Nash bargaining solution



*Note:* The graph shows the Nash bargaining solution for the share that player one receives,  $x_1$ , for different levels of the fairness weight,  $\beta_1$ . Player one has a fairness principle  $s^{k(1)} = \frac{3}{4}$  in the first three panels, and  $s^{k(1)} = \frac{1}{2}$  in the lower right panel. The parameters for player two are: ‘Fairness vs. self-interest’:  $\beta_2 = 0$ ; ‘Fairness agreement’:  $\beta_2 = 7$ ,  $s^{k(2)} = \frac{1}{4}$ ; ‘Fairness disagreement’:  $\beta_2 = 7$ ,  $s^{k(2)} = \frac{3}{4}$ ; ‘Equality’:  $\beta_2 = 7$ ,  $s^{k(2)} = \frac{1}{2}$ . The endowment  $Y = 1$ .

division. Importantly, in line with Proposition 3, two players who disagree about what is a fair share cannot reach an agreement if their fairness motivation is too strong. At the level of  $\beta_1 = 10$  in the lower left panel, the bargain solution is the disagreement outcome of zero.

The lower right panel shows the case where both players are fairness motivated and both players agree on a principle of strict equality. We can see that, in this case, the trade-off between self-interest and fairness motivation,  $\beta_1$ , does not influence the solution. This last result follows from the property of symmetric treatment of players in the Nash bargaining solution.

**Proposition 4.** *Fairness weight impotency. If bargainers follow the fairness principle of strict equality,  $s^{k(1)} = s^{k(2)} = \frac{1}{2}$ , then the Nash bargaining solution is  $(x_1^N = \frac{1}{2}, x_2^N = \frac{1}{2})$  for any  $\beta_1, \beta_2$ .*

*Proof.* See Appendix A. □

## 4.2 The generalized Nash bargaining solution

There is a version of the Nash bargaining solution that allows for bargaining power,  $\alpha_i$ , to influence the solution:

$$\max(u_1 - u_1^d)^{\alpha_1}(u_2 - u_2^d)^{\alpha_2}.$$

In the standard case where individuals are only motivated by material self-interest, the individual with more bargaining power gets a larger share of the endowment than a player with less bargaining power. Similarly, in bargaining between two equally fairness motivated individuals who disagree about fair shares, the individual with more bargaining power gets closer to his fair share. This argument also works the other way, a more fairness motivated individual gets closer to his fair share for a given distribution of bargaining power. Fairness motivation can therefore counterbalance the influence of unfavourable bargaining power.

Moreover, the same outcome that follows from bargaining between two equally strong fairness motivated individuals who agree on the fairness principle, may also be the result of bargaining between self-interested individuals who have bargaining power distributed in the same proportion as the fair shares. An interpretation of this result is that there are two ways to achieve a fair outcome in bargaining, through agreement about fairness principles or regulation of bargaining power.

However, in one case where the relative bargaining power is distributed in the exact same proportion as the fair shares that follow from a fairness principle, fairness motivation does not influence the generalized Nash bargaining solution. This result is similar to Proposition 4 where individuals are motivated by the fairness principle of strict equality and they have equal bargaining power. Numerical computation gives support to the following conjecture (see Appendix A).

**Conjecture 1.** Generalized fairness weight impotency. *If  $s^{k(1)} + s^{k(2)} = 1$ ,  $\alpha_1 = s^{k(1)}$  and  $\alpha_2 = s^{k(2)}$ , then the generalized Nash bargaining solution is  $(x_1^N = s^{k(1)}, x_2^N = s^{k(2)})$ , for any  $\beta_1, \beta_2$ .*

## 5 Concluding remarks

Individuals who are only motivated by material self-interest are always able to make a compromise and find an agreement, provided that the monetary reward from agreement is higher than from the disagreement outcome. For fairness motivated individuals, the outcome from bargaining will depend both on the principle they follow, and the trade-off they make between following the principle and material self-interest. First, people who are motivated by the same fairness principle have a non-empty bargaining set and it is always possible to reach an agreement, and if they have a high weight on following the principle, they will agree on the fair outcome. Second, if people disagree about what is a fair share, it may be impossible to reach an agreement, particularly if they have a high



weight on following their principles. Disagreement can easily be the outcome from bargaining between players that insist on different fairness principles.

The Nash bargaining solution shows that bargaining between an individual with strong fairness motivation and an individual only motivated by material self-interest reaches an outcome that is closer to the fairness motivated individual's principle. In bargaining between two individuals motivated by fairness, and who disagree about what represents a fair division, the Nash bargaining solution gives an outcome that is a compromise between the fair shares. If individuals follow the commonly assumed fairness principle of strict equality, the trade-off between following the fairness principle and material self-interest does not influence the outcome. The generalized Nash bargaining solution shows that a strongly fairness motivated individual can balance the higher bargaining power of another individual.

This research could be extended both theoretically and empirically. An interesting theoretical extension is to incorporate characteristics of the disagreement outcome into individuals' fairness principles. Another important issue for further research is how social preference models influence the efficiency of bargaining.

Finally, I would like to point to several empirical hypotheses for fairness motivation that can be derived from the analysis in this paper. First, material self-interest may be more predominant in societies where there is a great deal of plurality of fairness principles among people, and conversely, in a more homogeneous society, people may be more fairness motivated. In societies with a great deal of heterogeneity, it is important to reach compromises in transactions. Thus, an environment that fosters material self-interest may perform better than one that fosters fairness motivation. Second, the analysis in this paper shows that fairness motivation could develop among groups in societies where the bargaining power is to their disadvantage. The mobilization of fairness motivation can neutralize the imbalance of bargaining power. The development of strong fairness principles among unions in wage negotiations could be an example of this. Third, in societies with strong groups that are motivated by different fairness principles, the analysis shows that there can be more conflicts, for example, in societies where employers and employees strongly believe in different principles of wage setting.

## References

- Akerlof, George A. and Rachel E. Kranton (2010). *Identity Economics*, Princeton University Press.
- Andreoni, James and Douglas Bernheim (2009). "Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects", *Econometrica*, 77: 1607–1636.

- Babcock, Linda, George Loewenstein, Samuel Issacharoff, and Colin Camerer (1995). “Biased judgements of fairness in bargaining”, *American Economic Review*, 85(5): 1337–1343.
- Babcock, Linda, Xianghong Wang, and Georg Loewenstein (1996). “Choosing the wrong pond: Social comparisons in negotiations that reflect a self-serving bias”, *The Quarterly Journal of Economics*, CXI(1): 1–19.
- Bazerman, Max H. (1985). “Norms of distributive justice in interest arbitration”, *Industrial and Labor Relations Review*, 38(4): 558–570.
- Binmore, Ken (2007). *Playing for Real*, Oxford University Press.
- Birkeland, Sigbjørn (2011). “Negotiation under possible third party settlement”, Norwegian School of Economics and Business Administration, discussion paper.
- Bolton, Gary E. and Axel Ockenfels (2000). “Erc: A theory of equity, reciprocity and competition”, *American Economic Review*, 90: 166–193.
- Bruyn, Arnaud De and Gary E. Bolton (2008). “Estimating the influence of fairness on bargaining behavior”, *Management Science*, 54(10): 1774–1791.
- Camerer, Colin (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press.
- Cappelen, Alexander W., Astrid D. Hole, Erik Ø. Sørensen, and Bertil Tungodden (2007). “The pluralism of fairness ideals: An experimental approach”, *American Economic Review*, 97(3): 818–827.
- Cappelen, Alexander W., Erik Ø. Sørensen, and Bertil Tungodden (2010). “Responsibility for what? fairness and individual responsibility”, *European Economic Review*, 54(3): 429–441.
- Crawford, Vincent P. (1982). “A theory of disagreement in bargaining”, *Econometrica*, 50: 607–38.
- Elster, Jon (1989). *The Cement of Society*, Cambridge University Press.
- Fehr, Ernst and Klaus M. Schmidt (1999). “A theory of fairness, competition and cooperation”, *The Quarterly Journal of Economics*, 114(3): 917–868.
- Gächter, Simon and Arno Riedl (2005). “Moral property rights in bargaining with infeasible claims”, *Management Science*, 51(2): 249–263.
- Hirschman, Albert O. (1977). *The Passions and the Interests : Political Arguments for Capitalism before Its Triumph*, Princeton University Press.

- Konow, James (1996). “A positive theory of economic fairness”, *Journal of Economic Behaviour and Organisation*, 31: 13–35.
- Lopomo, Giuseppe and Efe A. Ok (2001). “Bargaining, interdependence, and the rationality of fair division”, *RAND Journal of Economics*, 32: 263–283.
- Nash, John F. (1950). “The bargaining problem”, *Econometrica*, 18: 155–62.
- Roth, Alvin E. (1979). *Axiomatic Models of Bargaining*, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag.
- Rubinstein, Ariel (1982). “Perfect equilibrium in a bargaining model”, *Econometrica*, 50: 97–109.
- Wilhelm, Mark O. (1996). “Bequest behavior and the effect of heirs’ earnings: Testing the altruistic model of bequests”, *The American Economic Review*, 86: 874–892.

## Appendix A Proof of propositions

### Proof of Proposition 1

We want to show that for any fair share,  $s^{k(i)}$ , there is a value  $\hat{\beta}_i$  such that for any  $0 < \beta_i < \hat{\beta}_i$  there exists a unique reservation point,  $x_i^L$ , and for any  $\beta_i \geq \hat{\beta}_i$  there exist two reservation points,  $x_i^L$ , and  $x_i^H$ . The utility function in Assumption 1:

$$u_i(x_i Y, s^{k(i)} Y) = (x_i - \beta_i(x_i - s^{k(i)})^2) Y,$$

is a quadratic equation:

$$a_1 x^2 + a_2 x + a_3 = 0,$$

where the coefficients are reduced to:

$$\begin{aligned} a_1 &= -\beta_i, \\ a_2 &= (1 + 2\beta_i s^{k(i)}), \\ a_3 &= -\beta_i (s^{k(i)})^2. \end{aligned}$$

The discriminant,  $a_2^2 - 4a_1a_3 = 1 + 4\beta_i s^{k(i)}$ , is positive and hence the utility function has two real, distinct roots. The quadratic formula gives the two solutions:

$$x_i^L = \frac{1 + 2\beta_i s^{k(i)} - \sqrt{1 + 4\beta_i s^{k(i)}}}{2\beta_i}, \quad x_i^H = \frac{1 + 2\beta_i s^{k(i)} + \sqrt{1 + 4\beta_i s^{k(i)}}}{2\beta_i}.$$

These solutions are not defined for  $\beta_i = 0$ . By definition, a fair share,  $s^{k(i)}$ , can have values in the interval  $0 \leq s^{k(i)} \leq 1$ . We see that if  $s^{k(i)} = 0$ , then  $x_i^L = 0$ . Differentiate  $x_i^L$  with respect to  $\beta_i$ :

$$\frac{dx_i^L}{d\beta_i} = \frac{1 + 2\beta_i s^{k(i)} - \sqrt{1 + 4\beta_i s^{k(i)}}}{2\beta_i^2 \sqrt{1 + 4\beta_i s^{k(i)}}}.$$

For the numerator to be positive,  $1 + 2\beta_i s^{k(i)} > \sqrt{1 + 4\beta_i s^{k(i)}}$ . By squaring both sides of the inequality we see that the numerator is always positive for  $s^{k(i)} > 0$ . Hence,  $\frac{dx_i^L}{d\beta_i} > 0$ , and  $x_i^L$  is strictly increasing in  $\beta_i$  for  $s^{k(i)} > 0$ . We know from Section 3.3 that  $\lim_{\beta_i \rightarrow \infty} x_i^L = s^{k(i)}$ . Thus, there always exists a lower reservation point,  $x_i^L$ , which attains values in the interval  $0 \leq x_i^L \leq s^{k(i)}$ .

We then consider the upper reservation point,  $x_i^H$ . We can see that if  $s^{k(i)} = 1$ , then  $x_i^H > 1$ , which is outside of the domain of the utility function for argument  $x_i$ . Differentiate  $x_i^H$  with respect to  $\beta_i$ :

$$\frac{dx_i^H}{d\beta_i} = \frac{-1 - 2\beta_i s^{k(i)} - \sqrt{1 + 4\beta_i s^{k(i)}}}{2\beta_i^2 \sqrt{1 + 4\beta_i s^{k(i)}}}.$$

We see that  $x_i^H$  is strictly decreasing in  $\beta_i$ , since  $\frac{dx_i^H}{d\beta_i} < 0$ . If  $s^{k(i)} < 1$ , then  $x_i^H$  may attain values in the interval  $0 < x_i^H \leq 1$ , depending on the relationship between  $s^{k(i)}$  and  $\beta_i$ . Define  $\hat{\beta}_i$  such that  $x_i^H = 1$ , which gives:

$$\hat{\beta}_i = \frac{1}{(1 - s^{k(i)})^2}.$$

Any  $\beta_i \geq \hat{\beta}_i$  will give an upper reservation point in the interval  $0 < x_i^H \leq 1$ . Hence, for  $\beta_i$  in the interval  $0 < \beta_i < \hat{\beta}_i$ , there exists a unique reservation point,  $x_i^L$ , and for  $\beta_i \geq \hat{\beta}_i$  there exist two reservation points,  $x_i^L$ , and  $x_i^H$ .

## Proof of Proposition 2

We want to show that if the fair shares are incompatible,  $s^{k(1)} + s^{k(2)} > 1$ , then there exists a  $(\hat{\beta}_1, \hat{\beta}_2)$  such that for any  $(\beta_1 \geq \hat{\beta}_1, \beta_2 \geq \hat{\beta}_2)$ , the only feasible solution is the disagreement outcome. Note that by assumption, if  $s^{k(1)} + s^{k(2)} > 1$ , then  $s^{k(1)} > 0$  and  $s^{k(2)} > 0$ .

1. From Section 4 we know that  $\mathcal{B}$  is empty if  $x_1^L + x_2^L > 1$ .
2. From Proposition 1 we know that there exists a lower reservation point,  $x_i^L$ , which is strictly increasing in  $\beta_i$  for  $s^{k(i)} > 0$ .
3. Define  $\varepsilon < |\frac{1 - s^{k(1)} - s^{k(2)}}{2}|$ . Find  $(\hat{\beta}_1, \hat{\beta}_2)$  such that  $x_1^L = s^{k(1)} - \varepsilon$  and  $x_2^L = s^{k(2)} - \varepsilon$ .

4. It follows from step 2 and step 3 that  $x_1^L + x_2^L > 1$  for any  $(\beta_1 \geq \hat{\beta}_1, \beta_2 \geq \hat{\beta}_2)$ . Hence, from step 1 it then follows that  $\mathcal{B}$  is empty, and the only feasible solution is the disagreement outcome.

### Proof of Proposition 3

Consider any combination of  $(\beta_1, \beta_2)$ . We want to show that  $\mathcal{B}(\beta_1, \beta_2)$  is non-empty, i.e.  $x_1^L + x_2^L \leq 1$ , if two individuals who follow the same fairness principle,  $k$ .

1. From Assumption 1 we know that if two individuals follow the same fairness principle,  $k$ , then  $s^{k(1)} + s^{k(2)} = 1$ .
2. From Proposition 1 we know that there exists a lower reservation point,  $x_i^L$ , which is monotonically increasing in  $\beta_i$  in the interval  $0 \leq x_i^L \leq s^{k(i)}$ . For a given  $(\beta_1, \beta_2)$ , we find  $(x_1^L, x_2^L)$  by using the formula in Proposition 1.
3. By taking into account step 1, it follows from step 2 that  $x_1^L + x_2^L \leq 1$ , and hence  $\mathcal{B}(\beta_1, \beta_2)$  is non-empty. This completes the proof of the first part of Proposition 3.
4. We also want to show that an increase in  $\beta_1$  and  $\beta_2$  in the limit collapses to a single point, which represents the fair division  $(s^{k(1)}, s^{k(2)})$ . If  $(\beta_1, \beta_2) \rightarrow \infty$ , it follows from Proposition 1 that  $(x_1^L, x_2^L) \rightarrow (s^{k(1)}, s^{k(2)})$ . By Assumption 1,  $(s^{k(1)}, s^{k(2)})$  is a unique element. Hence, in the limit  $\mathcal{B}$  only contains one element, which is the utility representation of  $(s^{k(1)}, s^{k(2)})$ . This completes the proof of the second part of Proposition 3.

### Proof of Proposition 4

Consider any combination of  $\beta_1 \geq 0$  and  $\beta_2 \geq 0$ . We want to show that if bargainers follow the fairness principle of *strict equality*, then the Nash bargaining solution is  $(x_1^N = \frac{1}{2}, x_2^N = \frac{1}{2})$ .

1. Consider the case where  $\beta_1 = \beta_2 = 0$ . The model is then reduced to a standard utility function, and it follows straightforwardly that the solution is  $(x_1^N = \frac{1}{2}, x_2^N = \frac{1}{2})$ .
2. Consider the case where  $\beta_1 > 0$  and  $\beta_2 > 0$ . By assumption, the parameter values are  $s^{k(1)} = s^{k(2)} = \frac{1}{2}$  in the Nash bargaining solution.
3. Differentiate the Nash bargaining solution (as stated in Appendix B) with respect to  $\beta_1$  and  $\beta_2$ . If you evaluate these two expressions for any  $\beta_1$  and  $\beta_2$ , the outcome is zero. Hence, changes in  $\beta_1$  and  $\beta_2$  do not influence the Nash bargaining solution.

## Examples of Conjecture 1

Table 2: Asymmetric bargaining power

		$\beta_2$		
		1	5	10
$\beta_1$	1	(0.75, 0.25)	(0.75, 0.25)	(0.75, 0.25)
	5	(0.75, 0.25)	(0.75, 0.25)	(0.75, 0.25)
	10	(0.75, 0.25)	(0.75, 0.25)	(0.75, 0.25)

*Note:* Table 2 shows a numerical computation of the generalized Nash bargaining solution for person one and person two for parameters  $\alpha_1 = s^{k(1)} = 0.75$ , and  $\alpha_2 = s^{k(2)} = 0.25$ .

Table 3: Symmetric bargaining power

		$\beta_2$		
		1	5	10
$\beta_1$	1	(0.5, 0.5)	(0.5, 0.5)	(0.5, 0.5)
	5	(0.5, 0.5)	(0.5, 0.5)	(0.5, 0.5)
	10	(0.5, 0.5)	(0.5, 0.5)	(0.5, 0.5)

*Note:* Table 3 shows a numerical computation of the generalized Nash bargaining solution for person one and person two for parameters  $\alpha_1 = s^{k(1)} = 0.5$ , and  $\alpha_2 = s^{k(2)} = 0.5$ .

## Appendix B Analytical solution

The Nash bargaining solution can be found by solving the optimization problem:

$$\begin{aligned} \max \quad & (u_1 - u_1^d)(u_2 - u_2^d) \\ \text{s.t.} \quad & x_1 + x_2 = 1. \end{aligned}$$

From Section 3.2 we have:

$$\begin{aligned} u_i(x_i Y, s^{k(i)} Y) &= (x_i - \beta_i(x_i - s^{k(i)})^2)Y, \\ u_i^d &= 0. \end{aligned}$$

By substituting the constraint into the objective function, the optimization problem can be written as:

$$\max f(x_1) = \left( (x_1 - \beta_1(x_1 - s_1)^2)Y \right) \left( (1 - x_1 - \beta_2(1 - x_1 - s_2)^2)Y \right).$$

Differentiating  $f$  with respect to  $x_1$  gives a cubic equation (the subscript on  $x$  is suppressed):

$$a_1x^3 + a_2x^2 + a_3x + a_4 = 0,$$

where the coefficients are:

$$\begin{aligned} a_1 &= 4\beta_1\beta_2Y^2, \\ a_2 &= (3\beta_1 - 3\beta_2 - 6\beta_1\beta_2 - 6s_1\beta_1\beta_2 + 6s_2\beta_1\beta_2)Y^2, \\ a_3 &= (-2 - 2\beta_1 - 4s_1\beta_1 + 4\beta_2 - 4s_2\beta_2 + 2\beta_1\beta_2 + 8s_1\beta_1\beta_2 + 2s_1^2\beta_1\beta_2 \\ &\quad - 4s_2\beta_1\beta_2 - 8s_1s_2\beta_1\beta_2 + 2s_2^2\beta_1\beta_2)Y^2, \\ a_4 &= (1 + 2s_1\beta_1 + s_1^2\beta_1 - \beta_2 + 2s_2\beta_2 - s_2^2\beta_2 - 2s_1\beta_1\beta_2 - 2s_1^2\beta_1\beta_2 \\ &\quad + 4s_1s_2\beta_1\beta_2 + 2s_1^2s_2\beta_1\beta_2 - 2s_1s_2^2\beta_1\beta_2)Y^2. \end{aligned}$$

Define the following relationships:

$$\begin{aligned} Q &\equiv \frac{a_3}{3a_1} - \left(\frac{a_2}{3a_1}\right)^2, \\ R &\equiv \frac{a_3a_2}{6a_1^2} - \frac{a_4}{2a_1} - \left(\frac{a_2}{3a_1}\right)^3, \\ D &\equiv Q^3 + R^2, \\ S &\equiv \left(R + \sqrt{D}\right)^{\frac{1}{3}}, \\ T &\equiv \left(R - \sqrt{D}\right)^{\frac{1}{3}}. \end{aligned}$$

$D$  is the discriminant that determines the nature of the roots of the equation. If  $D > 0$ , there is one real root and two conjugate complex roots; if  $D = 0$ , there are real roots of which at least two are equal; if  $D < 0$ , there are three distinct real roots. In this model,  $D$  is negative and there are three distinct real roots. Cardano's formulae for the roots are as follows:

$$\begin{aligned} root_1 &= -\frac{a_2}{3a_1} + (S + T), \\ root_2 &= -\frac{a_2}{3a_1} - \frac{1}{2}(S + T) + \frac{1}{2}i\sqrt{3}(S - T), \\ root_3 &= -\frac{a_2}{3a_1} - \frac{1}{2}(S + T) - \frac{1}{2}i\sqrt{3}(S - T), \end{aligned}$$

where  $i = \sqrt{-1}$ . It turns out for this model that  $root_2 < root_3 < root_1$ . The optimal solution is  $x_1^N = root_3$ , and the Nash bargaining solution is  $(x_1^N, 1 - x_1^N)$ . The solution is only defined for  $\beta_1 > 0$  and  $\beta_2 > 0$ . To make sure the solution is in the bargaining set  $\mathcal{B}(\mathcal{U})$ , check that  $(u_1(x_1^N), u_2(1 - x_1^N)) > (0, 0)$ .



# NHH

---

**Norges  
Handelshøyskole**

Norwegian School of Economics  
and Business Administration

NHH  
Helleveien 30  
NO-5045 Bergen  
Norway

Tlf/Tel: +47 55 95 90 00  
Faks/Fax: +47 55 95 91 00  
[nhh.postmottak@nhh.no](mailto:nhh.postmottak@nhh.no)  
[www.nhh.no](http://www.nhh.no)