Bargaining with random implementation: an experimental study

Nejat Anbarci
School of Accounting, Economics and Finance
Deakin University
70 Elgar Road
Burwood VIC 3125, Australia
nejat.anbarci@deakin.edu.au

Nick Feltovich*
Department of Economics
Monash University
Clayton VIC 3800, Australia
nicholas.feltovich@monash.edu

July 23, 2012

Abstract

We use a laboratory experiment to study bargaining with *random implementation*. We modify the standard Nash demand game so that incompatible demands do not necessarily lead to the disagreement outcome. Rather, with exogenous probability q, one bargainer receives his/her demand, with the other getting the remainder. We use an asymmetric bargaining set (favouring one bargainer) and disagreement payoffs of zero, and we vary q over several values.

Our results mostly support game theory's directional predictions. As with conventional arbitration, we observe a strong *chilling effect* on bargaining for q near one: extreme demands and low agreement rates. Increasing q reinforces the game's built–in asymmetry – giving the favoured player an increasingly large share of payoffs – but also raising efficiency. These effects are non–uniform: over sizable ranges, increases in q have minimal effect, but for some q, small additional increases lead to sharp changes in results.

Journal of Economic Literature classifications: C78, D74, C72.

Keywords: Nash demand game, random implementation, chilling effect, equilibrium selection, arbitration, experiment.

*Corresponding author. Financial support from Deakin University is gratefully acknowledged. Some of this research took place while Feltovich was at University of Aberdeen. We thank two anonymous referees, John Boyd III, Ehud Kalai, Hervé Moulin, and Joe Swierzbinski for helpful suggestions and comments.

1 Introduction

Many economic transactions involve a decentralised element, with the price (and perhaps other attributes) set by a single buyer and a single seller, each with some degree of market power. Well–known examples of at least partly decentralised markets include those for houses, new cars and used cars in most countries, as well as labour markets in many professions. In such a market, associated with any potential transaction is a relation–specific surplus for the parties involved: for example, if a painting is worth \$50,000 to its current owner and \$80,000 to a potential buyer, then a surplus of \$30,000 is available to the two parties. The fundamental role of bargaining in determining how surpluses are divided (and indeed, whether they are even realised) in decentralised markets has long been recognised, with work in this area going back at least to the late nineteenth century (Edgeworth, 1881). However, until the 1950s, bilateral bargaining situations were deemed by economists to lack a clear predicted outcome.[1] The only quasi–prediction was that the division of the surplus would depend on the two parties' relative bargaining power.

Nash (1950) approached this indeterminacy problem by proposing a set of four axioms that the outcome of bargaining ought to follow, and proving that together, these axioms entail a unique solution to any bargaining situation that satisfies a few weak conditions, giving rise to the *Nash bargaining solution*.[2] Nash (1953) followed up this axiomatic approach by introducing a very simple non–cooperative game, now known as the *Nash Demand Game* (which we abbreviate as NDG). In the simplest version of the NDG, there is a fixed sum of money (a "cake") available to the two bargainers, and each simultaneously makes a single, irrevocable demand. If the demands are compatible (their sum does not exceed the size of the cake), then there is "agreement", and each bargainer receives the amount he/she demanded. If not, then a predetermined "disagreement" outcome is imposed. By introducing the NDG, which was meant to capture the key aspects of real bargaining, Nash established a new research agenda, now called the "Nash program" (Binmore, 1998). This program uses non–cooperative game theory to provide a foundation for axiomatic (cooperative) bargaining solution concepts like the Nash solution.

The simplicity of the Nash Demand Game is a great virtue, but as a model of real bargaining it has two major disadvantages: one theoretical and one practical. Its theoretical disadvantage is that most versions of the NDG have a large number of Nash equilibria. In particular, every efficient, individually rational division of the surplus corresponds to a Nash equilibrium. (There are also inefficient Nash equilibria.) This multiplicity of equilibria – and resulting lack of predictive power – clearly limits the usefulness of the NDG for analysing real bargaining.

The practical shortcoming of the NDG concerns the implication of incompatible demands. In this case, the disagreement outcome is imposed, resulting in a severe punishment to the bargainers, with no chance of avoiding it (e.g., through renegotiation). The fact that failure to agree immediately leads to irrevocable disagreement – irrespective of how close to being compatible the two demands were – flies in the face of most people's intuitive understanding of how bargaining works. Despite this seeming deficiency, some researchers have defended the NDG as at least capturing the most important features of real bargaining. Binmore (2007) points out that when bargainers can commit to demands, but neither has the ability to commit before the other, the NDG is the limiting case where both bargainers "rush to get a take–it–or–leave–it demand on the table first" (p. 496), resulting in simultaneous irrevocable demands. Moreover, Skyrms (1996) argues that in modelling the bargaining process, "[o]ne might imagine some initial haggling...but in the end each of us has a bottom line" (p. 4); focussing

[1]For example, as Roth (1979) points out, von Neumann and Morgenstern's (1944) bargaining solution coincides with the set of all efficient outcomes that both bargainers prefer to disagreement.

[2]Formally, a two–person bargaining problem is described by a pair (S, d) where $S \subset \mathbf{R}^2$ is the *set of feasible agreements* with a *disagreement point* $d = (d_1, d_2) \in S$ being the allocation that results if no agreement is reached. Nash's solution requires only that S is compact and convex, and that it contains some (x_1, x_2) with $x_1 > d_1$ and $x_2 > d_2$ (that is, gains from agreement are available); these conditions will be satisfied for the bargaining problems considered in this paper.

on these bottom lines results in the NDG. However, experimental evidence suggests that there are indeed systematic differences in behaviour between the NDG and less structured bargaining settings – both in the likelihood of reaching agreement and in how the resulting surplus is divided (Feltovich and Swierzbinski, 2011) – suggesting that some important features of real bargaining are lost by modelling it with the NDG. In particular, subjects in NDG experiments tend to leave more money "on the table", hedging against the game's strategic uncertainty by reducing demands, compared to less structured settings.

Nash (1953) himself provided the first attempt to rectify these problems with the NDG, using a "smoothing" approach. Under smoothing, incompatible pairs of demands do not necessarily lead to zero payoffs; rather, the probability of a pair of demands being accepted decreases continuously from one (at the boundary of the original bargaining set) to zero based on a smoothing function.[3] Clearly, such a modification treats incompatible demands differently according to how close to being compatible they were. Less obviously, when an appropriate smoothing function is used, it results in the set of Nash equilibria shrinking to a unique equilibrium corresponding to the Nash bargaining solution. Although this smoothing attempt was the first to provide non–cooperative foundations for the Nash solution, it has typically not been deemed reasonable by game theorists since that time.[4]

A more recent attempt is that by Anbarci and Boyd (2011), whose "probabilistic simultaneous procedure" (p. 18) modified the NDG so that bargainers submitting incompatible demands do not necessarily receive their disagreement payoffs; instead, this happens only with exogenous probability $1 - q$, where $q \in [0, 1]$. With the remaining probability q (conditional on incompatible demands), a fair coin toss determines which of the two bargainers receives his/her demand, with the remainder going to the other bargainer. Anbarci and Boyd's game can be thought of as tacking a (stylised) element of arbitration onto the standard NDG. If demands made in the first stage are incompatible, with probability q the bargainers move to a second stage where they undergo final–offer arbitration, and where these final offers are just the bargainers' demands from the first stage.[5] This mechanism differs from many models of final–offer arbitration in that the arbitrator does not have a preferred outcome with which to compare the demands to arrive at a decision, and simply chooses either of the demands with equal probability, thus providing incentives in line with conventional arbitration.[6]

In our theoretical analysis (as in Anbarci and Boyd's) and our experiment, the second stage of the game – the draw to determine what happens in case of incompatible first–stage demands – is automated. As there is no strategic arbitrator in our setup, but rather just random choice of one of the bargainers' demands to be implemented, we refer to this class of games as "Nash demand games *with random implementation*". The parameter q (the probability of random implementation in the case of disagreement) admits several interpretations. In an abstract vein, it serves to relate standard Nash bargaining without random implementation (the case where $q = 0$) to bargaining with certain random implementation in case of disagreement ($q = 1$), and by analogy, to highlight a connection between the theoretical implications of bargaining with and without arbitration. In a more concrete vein, one could consider imposition of the disagreement outcome with probability $1 - q$ as reflecting a possibility that subsequent external events render a first–stage disagreement permanent (e.g., by making arbitration impossible).

Anbarci and Boyd (2011) were primarily concerned with the set of outcomes in which bargainers agree and leave no

[3]See Binmore et al. (1993) for an experiment using a "smoothed" bargaining set.

[4]For example, Luce and Raiffa (1957) write that "Nash offers an ingenious and mathematically sound argument for [resolving the indeterminacy problem], but we fail to see why it is relevant" (p. 141) and go on to call smoothing "completely artificial" (p. 142). Schelling (1960) was more sympathetic, but even so, stated that smoothing was "in no sense logically necessary" (p. 283) and that while it provided a way of selecting one of the multiplicity of equilibria, the same argument "equally supports any other procedure that produces a candidate for election among the infinitely many potential [solutions]" (p. 284).

[5]Since the outcome imposed in the second stage is binding on the bargainers, the mechanism is more similar to arbitration than to *mediation*, where the second–stage result is not binding on the bargainers and thus must be agreeable to both of them.

[6]Chatterjee's (1981) theoretical analysis of arbitration mentioned the version of this mechanism with $q = 1$ in an aside as an example of final–offer arbitration, but didn't consider it further. See Section 2.1 for additional discussion of how this mechanism relates to real bargaining and arbitration.

money on the table (what we will call "agreement equilibria", where the random implementation stage is not reached), and in particular how this set relates to the predictions of some of the well–known axiomatic bargaining solutions. However, their game, which we will often abbreviate as NDG(q), has an additional property that makes it interesting from a theoretical standpoint: as q increases, the set of Nash equilibria changes in a manner that is monotonic in a sense, but not uniform. The specifics of this will be discussed in Section 2, but intuitively, for some ranges of q, fairly large changes in q will have little or no effect on the set of equilibria, while for other ranges of q, the set of equilibria will be extremely sensitive to small changes in q.

The purpose of this paper is to put the theoretical implications of Anbarci and Boyd's (2011) game to the test, with the use of a human–subjects experiment. To our knowledge, there has been no previous experimental study of this game, though of course both bargaining and arbitration have immense literatures.[7] Our study begins with an underlying bargaining environment that is biased, in that one bargainer has a favourable position relative to the other. Such asymmetry has two advantages from an experimental–design standpoint. First, it adds an element of genuine strategic uncertainty to the game, since in asymmetric bargaining settings, there is no single obvious focal outcome (in contrast to symmetric bargaining games, which are overwhelmingly likely to lead to equal splits). Second, it opens the possibility of examining whether varying q affects the extent to which the favoured player is able (or willing) to exploit this favourable position. Our experiment is designed to allow such an examination (among other things), utilising several values of q ranging from zero to one.

Our experimental results give positive, but not unequivocal, support to standard game theory. As is often the case in experiments, we find that the theory performs fairly poorly when faced with severe tests based on equilibrium point predictions; these are typically not seen in the experimental data. However, qualitative implications of the theory, based on the directions of effects on behaviour resulting from varying q, usually parallel what is seen in the experiment. In particular, compared to lower values, higher values of q are more often associated with a "chilling effect" on bargaining (as is often seen in models of conventional arbitration), with substantially lower agreement rates (e.g., less than 25% when $q = 1$, as compared to over 90% when $q = 0$) resulting from extreme demands. This effect tends to reinforce the inherent bias of the bargaining environment, increasing the share of payoffs going to the favoured player at the expense of the unfavoured player. Specifically, the ratio between favoured– and unfavoured–player payoffs when q is low is close to that of the most equitable efficient division of the cake (the *lexicographic egalitarian* solution; see Chun, 1989), while when q is high, this ratio is comparable to that of the more uneven *Kalai–Smorodinsky* (1975) outcome. This result ought to give players diametrically opposed preferences over q, but we also find that efficiency tends to increase with q (in contrast to the theoretical prediction), so that higher values of q sometimes lead to Pareto improvements over lower values. Finally, we find that consistent with the theory, increases in q have non–uniform effects. For a fairly large range of q, increases in q have minimal effect on bargaining behaviour and outcomes. From there, even a small additional increase in q can lead to sharp changes in results, while further increases have little effect beyond this.

2 Theoretical background

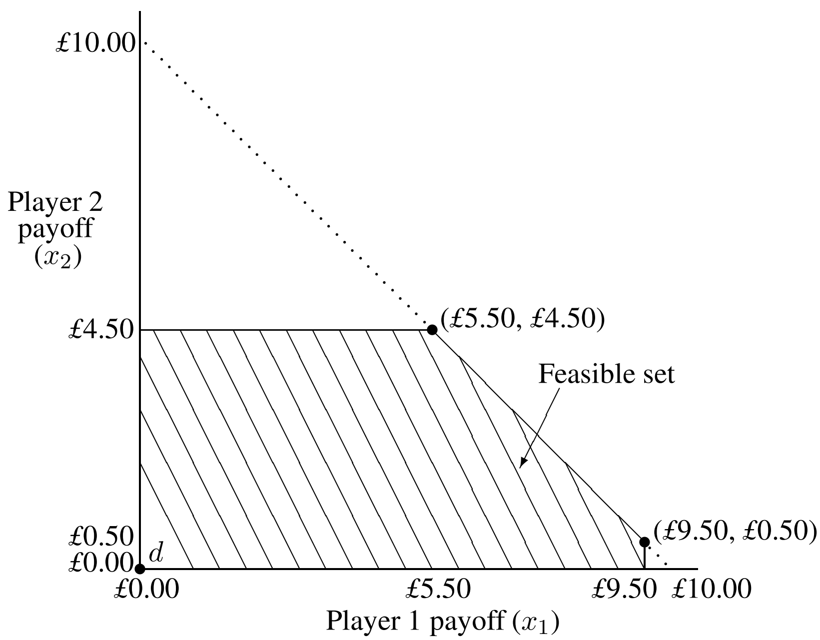
The game used in the experiment is an adaptation of the Nash Demand Game (Nash, 1953), which we abbreviate NDG. In the version we use, there is a fixed surplus ("cake") of £10 available to be divided by two bargainers. The bargainers make simultaneous demands; Player 1 (the favoured player) can choose any demand x_1 between zero and $\overline{x}_1$ = £9.50, while Player 2 (the unfavoured player) can choose any demand x_2 between zero and $\overline{x}_2$ = £4.50. If the demands are compatible (total at most the cake size of £10), there is "agreement", and each bargainer receives the amount demanded, with any remainder left "on the table". In this basic game, if the demands are incompatible ("disagreement"), both bargainers receive zero. In

[7] See Section 2.3 for a brief review of relevant previous work.

either case, the game ends with no opportunity for further negotiation. All rules of the game, including the limits placed on demands, are assumed to be common knowledge between the players.

Under the assumption that bargainers' game–theoretic payoffs are affine functions of their monetary payments, the bargaining problem can be depicted in Figure 1. As noted already, setting the players' maximum allowable demands to

Figure 1: The bargaining problem – feasible set and disagreement outcome d



$\overline{x}_1 = 9.50$ and $\overline{x}_2 = 4.50$ makes the bargaining problem asymmetric, decreasing the likelihood that the players will agree on a single obvious focal point.[8] Indeed, the most common focal point in bargaining games, a 50–50 split of the £10 cake, is impossible in our game, and while other equal–payoff outcomes are possible (e.g., each player receives £4.50), they are neither equilibria nor efficient.

As is common in NDGs, this basic game has a large number of Nash equilibria. There are efficient pure–strategy equilibria in which Player 1 demands k and Player 2 demands $10 - k$, for $k \in [10 - \overline{x}_2, \overline{x}_1]$, as well as inefficient mixed–strategy equilibria. Ruling out any of these equilibria requires the use of additional assumptions; however, we note that both Harsanyi and Selten's (1988) risk dominance and their general equilibrium selection procedure select the most equitable of the efficient equilibria (5.5, 4.5).[9] This outcome is also implied by some of the well–known axiomatic bargaining solutions for the corresponding unstructured bargaining game, such as the Nash (1950) and lexicographic egalitarian (Chun, 1989) solutions (though not the Kalai–Smorodinsky (1975) solution, which selects an agreement with Player 1 getting £$6.78\frac{4}{7}$ of the £10).[10] Additionally, while Rawls's (1971) "difference principle" did not deal with bargaining specifically, arguments in that spirit would select (5.5, 4.5) as the fairest efficient division of the cake (as it maximises the payoff of the worse–off

[8] A typical result in experiments involving symmetric bargaining games is that they tend to yield a high frequency of agreements on 50–50 splits of the cake. Introducing any kind of asymmetry substantially lowers both agreement frequencies and equal splits, even when such play is still consistent with equilibrium, possibly due to players' self–serving views of fairness (Babcock et al., 1995, Roth and Murnighan, 1982). See Nydegger and Owen (1975) and Roth and Malouf (1979) for experimental comparisons of symmetric and asymmetric bargaining games.

[9] Risk dominance formalises the intuitive notion that when players have little information about the choices others will make, they will prefer strategies that are (in some sense) less risky. In the simplest case of a symmetric 2x2 game with strategic complementarities and two strict Nash equilibria (s, s) and (t, t), (s, s) is risk dominant if the threshold probability of the opponent choosing s at which s becomes a best response is lower than the corresponding threshold probability for t. Harsanyi and Selten's (1988) text extends this intuition for general non–cooperative games.

[10] The lexicographic egalitarian solution differs from the strict egalitarian solution in tolerating increases in inequity as long as they don't actually

player). Given that so many cooperative and non–cooperative methods select the (5.5, 4.5) outcome, we might expect it to serve as an appealing focal point.

2.1 The Nash demand game with random implementation

From the basic NDG, we make one modification: in case of incompatible demands, the players might nonetheless avoid receiving their disagreement payoffs of zero. Specifically, with (common knowledge) probability $q \in [0, 1]$, play goes to a second stage, with one of the players randomly selected – both equally likely – to receive the amount he/she demanded, with the other player receiving the remainder of the cake. With probability $1 - q$, both players receive zero, as before. We will use the notation NDG(q) to refer to a generic version of this game, so that NDG(0) is equivalent to the basic Nash demand game, and NDG(1.0) is the game where the second stage occurs with certainty (though there is still uncertainty regarding which of the two demands will be implemented) in the event of incompatible demands.

Obviously, our NDG(q) is much simpler than actual situations involving bargaining and arbitration. However, enough real–world settings have the flavour of this game to make it worthy of study. The mechanism's restriction to implementing one of the two demands – with no compromises possible – may seem severe, but it is found in real final–offer arbitration (as used, for example, in North America by Major League Baseball when salary negotiations between teams and players break down, as well as for labour disputes in several US states involving public–sector workers), though as already mentioned, the lack of a bliss point for the arbitrator in our mechanism means that it presents incentives more along the lines of conventional arbitration. The lack of a bliss point itself might seem strange, but this can be rationalised as a reduced form of an incomplete–information game where bargainers do not know the arbitrator's preferred outcome – either because the arbitrator is drawn after demands are chosen from a heterogeneous population of potential arbitrators, or simply because the preferences of a particular arbitrator are unknown.[11] Incomplete information about the arbitrator's preferred outcome is a very common assumption in the study of arbitration, both in theoretical analyses (Chatterjee, 1981; Brams and Merrill, 1983; Wittman, 1986), and in experimental designs (Ashenfelter et al., 1992). In the limiting case, as information about the arbitrator's preferred outcome becomes extremely diffuse (e.g., the distribution from which the bliss point is drawn is uniform and contains all feasible splits of the cake), traditional final–offer arbitration will yield the same outcome as our mechanism (conditional on arbitration occurring), with the arbitrator choosing either demand with equal probability.

Other real–life bargaining situations resemble our game in that there is genuine uncertainty about whether a second–stage outcome would be imposed (i.e., the case with q strictly between 0 and 1). For example, in some cases of bargaining breakdown between firms, arbitration is neither automatic nor automatically ruled out. Ryan (1999) notes in the context of international commercial disputes:

> "Chapter 2 of Title 9 of the United States Code contains the New York Convention and the enabling legislation by which it was ratified by the United States in 1970 [...] The Convention contemplates a limited inquiry by courts when considering a motion to compel arbitration: 1. Is there an agreement in writing to arbitrate the dispute? 2. Does the agreement provide for arbitration in the territory of a Convention signatory? 3. Does the agreement to arbitrate arise out of a commercial legal relationship? 4. Is a party to the agreement not an American citizen or does the commercial relationship have some reasonable relation with one or more Foreign States?"

harm the worse–off player. For the bargaining set we use, the inefficient outcome (4.5, 4.5) is the strict egalitarian solution; weak Pareto improvements from here are possible, but benefit Player 1 only and are thus ruled out by a strict egalitarian. By contrast, a lexicographic egalitarian would allow any (weak or strong) Pareto improvements, and would thus select (5.5, 4.5) for our bargaining set.

[11] This might be because the arbitrator is new to the job, or because the arbitrator's beliefs about what constitutes a "fair" settlement are highly situation–specific, so that past decisions offer little information about future ones.

The third and fourth of these criteria, relying on the interpretations of "commercial legal relationship" and "reasonable relation", may in some cases be ambiguous enough to permit real uncertainty regarding whether arbitration will be compelled – along with the usual uncertainty about which side would prevail in that case.

Two recent labour disputes in Australia also show some resemblance to our mechanism. In November 2011, the Australian federal government intervened in the Qantas industrial dispute to end the company lockout, siding with the unions. In a speech defending the government's position, the Minister for Infrastructure and Transport, Anthony Albanese, noted that "Fair Work Australia [the national workplace relations tribunal] went through 16 hours of hearings to come up with a decision", implying that multiple outcomes were possible: the decision could have gone in either direction, and taking no action was also a possibility.[12] The genuine uncertainty regarding the government's response was summarised by Transport Workers Union national secretary Tony Sheldon: "The government has stepped in, it's the first to my knowledge in the history of this country and of course it means that new laws, new approaches need to be instigated..."[13] By contrast, in February 2012, Fair Work Australia abruptly ordered the Australian Nurses Federation (ANF) to end its programme of nurses' and midwives' rolling work stoppages during an industrial dispute in the state of Victoria, with the threat of large fines and up to 12 months' imprisonment if they refused.[14] This intervention was particularly surprising since the federal Minister for Workplace Relations, Bill Shorten, had just urged the Victoria state government to seek private arbitration to resolve the dispute. Thus, within a few months the same industrial–relations arbiter – under the same Labor government – intervened in union–management disputes, fully siding with unions in one case and with management in the other, and in both cases the outcome was surprising to a non–trivial extent.

2.2 Theoretical predictions

As noted in the introduction, this modification to the game can be interpreted as (with probability q) adding a second stage consisting of a stylised version of final–offer arbitration, where the final offers are simply the players' demands from the first stage, and the arbitrator chooses either with equal probability. From a theoretical standpoint, this "random implementation" rule differs in an important way from those in typical models of final–offer arbitration, where the likelihood of a player's offer being the one chosen is negatively related to its distance from some bliss point held by the arbitrator, thus punishing extreme demands to some extent. Our random implementation rule is not sensitive to the specific demands chosen by the players. Indeed, the incentives it introduces can be similar to those of conventional arbitration; the random implementation procedure chooses an outcome with expected value midway between the bargainers' demands, as a simple model of conventional arbitration might choose the outcome whose actual value is midway between the demands. As a result, when q is close to one, predicted behaviour is the typical "chilling effect" seen in models of conventional arbitration, with both players choosing the maximum demand they are able to make.[15] As long as q is strictly less than one, this equilibrium is inefficient, and indeed, Pareto dominated by some non–equilibrium strategy profiles that lead to agreement. Thus, when $q = 0$, NDG(q) has the usual problem of multiplicity of Nash equilibria, while when $q = 1$, the only equilibrium involves the chilling effect just described. However, for some $q \in (0, 1)$, the game's theoretical prediction can be precise as well as avoiding the most extreme demands, as the analysis below will show.

Clearly, for any allowable demand x_{3-i} by the opponent ($i = 1, 2$), Player i's best response will be either the highest

[12] http://www.minister.infrastructure.gov.au/aa/speeches/2011/AS37_2011.aspx

[13] http://www.smh.com.au/travel/fair-work-ends-qantas-industrial-dispute-20111031-1mqsq.html#ixzz1qS0s0Upn

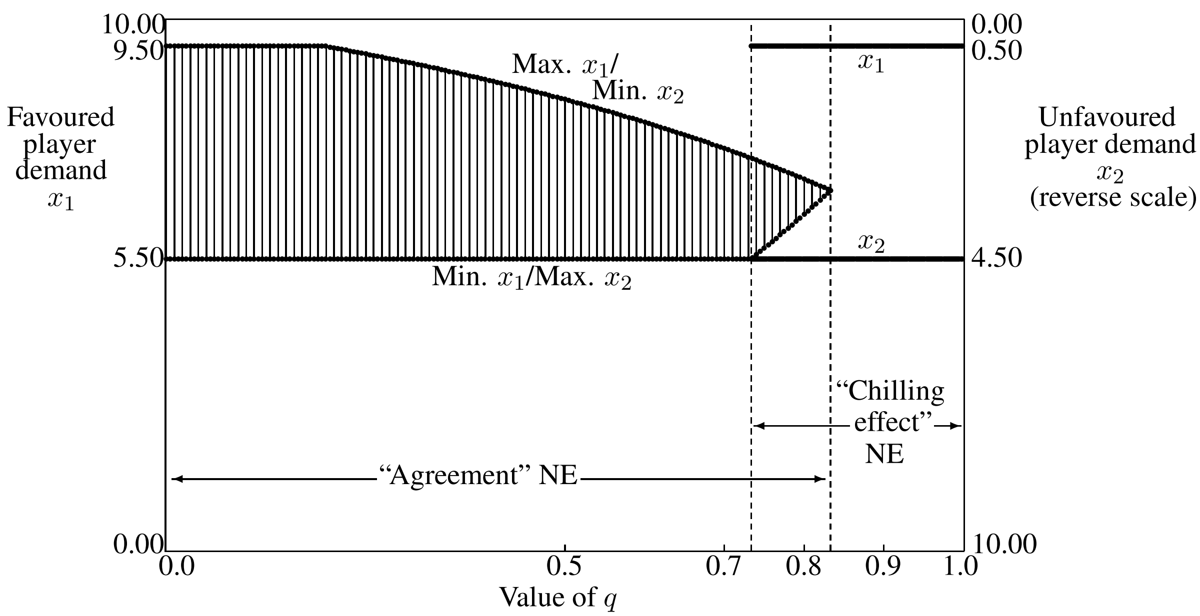
[14] http://www.abc.net.au/news/2012-02-24/fair-work-australia-rules-on-nurses/3851088/?site=melbourne and http://www.wsws.org/articles/2012/feb2012/nurs-f25.shtml.

[15] The chilling effect that conventional arbitration, and to a lesser extent final–offer arbitration, can have on bargaining is well covered in the theoretical and empirical literature. See, for example, Feuille (1975).

compatible demand $x_i = 10 - x_{3-i}$ or the maximum possible demand $x_i = \overline{x}_i$. Choosing $\overline{x}_i$ is best if either (a) x_{3-i} is low enough that agreement is reached even when $\overline{x}_i$ is chosen, or (b) q is high enough that the $\frac{q}{2}$ probability of being selected to get $\overline{x}_i$ is worth the risk of disagreement (noting that the player still gets $10 - x_{3-i}$ in the event that the opponent's demand is chosen to be implemented). This leads to the possibility of two types of pure–strategy Nash equilibrium. The first type, which corresponds to the efficient equilibria of the standard NDG, has $x_1 + x_2 = 10$, and agreement occurs with probability one. The second type of equilibrium has both players choosing their maximum allowable demands: $(x_1, x_2) = (\overline{x}_1, \overline{x}_2)$, so expected payoffs are $\frac{q}{2}[10 + \overline{x}_1 - \overline{x}_2]$ and $\frac{q}{2}[10 - \overline{x}_1 + \overline{x}_2]$ respectively. We will sometimes refer to these two types of equilibrium as "agreement equilibria" and "chilling effect equilibria" respectively.

Which of these types of equilibria exists depends on q; as q increases, a choice of the maximum allowable demand becomes more profitable relative to lower demands that might lead to agreement; more precisely, a demand of $\overline{x}_i$ strictly dominates demands below $\frac{q}{2-q}\overline{x}_i$. So, as a function of q, the favoured player's minimum demand in an agreement equilibrium is $Max\left\{9.5\frac{q}{2-q},\ 5.5\right\}$, while her maximum demand in an agreement equilibrium is 9.5 for all q. Similarly, the unfavoured player's minimum demand in an agreement equilibrium is $Max\left\{4.5\frac{q}{2-q},\ 0.5\right\}$, while his maximum demand in an agreement equilibrium is 4.5 for all q. As a result, as q increases, the continuum of agreement equilibria seen in the basic NDG shrinks,

Figure 2: Theoretical predictions – Nash equilibrium demands of favoured players (x_1) and unfavoured players (x_2), contingent upon q



becoming a single point when $q = \frac{5}{6} \approx 0.833$, with the favoured player receiving $\frac{19}{28} \approx 0.679$ of the cake.[16] Beyond this value of q, only the "chilling effect" equilibrium remains; the favoured player demands 9.5 and the unfavoured player demands 4.5, implying expected payoffs of $7.5q$ and $2.5q$ respectively. Chilling effect equilibria also exist for some lower values of q: specifically, whenever $q \geq \frac{11}{15} \approx 0.733$. Except when $q = 1$, these chilling effect equilibria are inefficient, and indeed are payoff dominated by outcomes in which the players make compatible demands. For $q \in \left[\frac{11}{15}, \frac{5}{6}\right]$, these latter outcomes may be equilibria themselves, but for higher q, they are necessarily non–equilibrium outcomes (in which case the game has some characteristics of the prisoners' dilemma).

Figure 2 shows the correspondence between the value of q and the set of Nash equilibria for NDG(q). As the figure shows,

[16]The unique Nash equilibrium outcome for this value of q coincides with the Kalai–Smorodinsky (1975) solution for the corresponding unstructured bargaining problem. As Anbarci and Boyd (2011) show, this is true for any bargaining set satisfying minimal properties.

there is an element of monotonicity to the correspondence: for $q_1 < q_2$, the set of agreement equilibria of NDG(q_1) contains the corresponding set for NDG(q_2), and once we reach a value of q for which a chilling–effect equilibrium exists, it continues to exist for all higher q. However, the correspondence is not uniform. For $q \in [0, 0.2]$, the set of equilibria is exactly the same as it is for the basic Nash demand game, and for q from 0.2 all the way up to $\frac{11}{15}$, the set of agreement equilibria shrinks at a fairly constant rate, as relatively asymmetric agreements become worse for the unfavoured player than demanding his maximum; in particular, the risk–dominant equilibrium (5.50, 4.50) of the basic NDG continues to be an equilibrium of NDG(q) for this range of q. At $q = \frac{11}{15}$, there is a sudden change, as chilling–effect equilibria appear and, beyond this point, the outcome (5.50, 4.50) is no longer an equilibrium. For $q \in \left(\frac{11}{15}, \frac{5}{6}\right)$, the set of agreement equilibria shrinks quickly, as demanding the maximum becomes more attractive for the favoured player than the most equitable agreements, and for the unfavoured player, more attractive than the most inequitable agreements. At $q = \frac{5}{6}$, there is another sudden change, as agreement equilibria cease to exist beyond this point; only the chilling–effect equilibria are left.

2.3 Related literature

A full review of the experimental literature on bargaining is well beyond the scope of this paper; the interested reader can find a survey of early work in Roth (1995), and some of the more recent work in Camerer (2003, pp. 151–198). Our examination of the effect of varying the probability of random implementation places our work into the strand of the bargaining literature focussing on the connection between structural aspects of the bargaining situation and the predicted frequency of agreement. This topic has received much attention, both theoretical (Axelrod, 1967; Crawford, 1982; Myerson, 1984) and experimental (Malouf and Roth, 1981; Coursey, 1982; Rapoport and Sundali, 1996). However, we know of no previous experimental investigation into Anbarci and Boyd's (2011) mechanism specifically.

As we have already noted, while our NDG(q) game is at best a highly stylised model of bargaining with arbitration, it does resemble real bargaining with arbitration in some ways, including the incentives it provides to bargainers; indeed, Chatterjee (1981) briefly mentioned our $q = 1$ case as an example of final–offer arbitration that yields a result equivalent to that of conventional arbitration. Surveys of experiments involving bargaining with arbitration can be found in Kuhn (2009) and Charness and Kuhn (2010). Our random–implementation rule, on its own, is similar to the "random–dictator" rule used in one of the treatments in Forsythe, Kennan and Sopher's (1991) experimental study of incomplete–information bargaining. This treatment had no first–stage bargaining round; the two players simply submitted proposals for dividing the cake, and one of the proposals was randomly drawn and implemented if feasible. They found substantial evidence for a chilling effect (see, e.g., their Table 8), as is predicted by theory. There is also a literature on random implementation of decisions in contexts other than bargaining; see, for example, Zizzo (2003).[17]

2.4 Treatments and hypotheses

In the experiment, we use a total of six versions of NDG(q), with $q = 0$, 0.5, 0.7, 0.8, 0.9 and 1.0. Figure 2 shows that for the first three values of q, only agreement equilibria exist; for the last two, only a chilling effect equilibrium exists; and for $q = 0.8$, both kinds of equilibrium exist. Some equilibrium features of the NDG(q) with these particular values of q are summarised in Table 1. Here, efficiency is defined as the sum of the bargainers' payoffs divided by the size of the cake, thus

[17]Indeed, random implementation is increasingly used as a cost–reducing methodological tool in economics experiments. In experiments involving asymmetric games, subjects' roles might be assigned only at the end of the game, and their strategies can be elicited contingent on being chosen for each possible role. For example, a dictator–game experiment might have all subjects submitting the division of the cake they would choose if they were dictator, even though only half of the subjects will end up being chosen as dictator (see, e.g., Cabrales et al., 2010). This and other "strategy elicitation" (sometimes called "strategy method") techniques aren't without their drawbacks (see Brandts and Charness (2011) for a survey of the effect of these versus direct–response methods), but they clearly lower the cost–per–observation of gathering data.

taking on values between zero and one; efficiency is one in any agreement equilibrium and is equal to q in any chilling–effect equilibrium.

Table 1: Theoretical predictions for the versions of NDG(q) used in the experiment, to 2 decimal places

Value of q	Nash Equil.	Equilibrium demands (£) Favoured player	Unfavoured player	Equilibrium payoffs (£) Favoured player	Unfavoured player	Agreement frequency	Efficiency
0	All	[5.5, 9.5]	[0.5, 4.5]	[5.5, 9.5]	[0.5, 4.5]	1	1
	Risk dom.	5.5	4.5	5.5	4.5	1	1
0.5	All	[5.5, 8.5]	[1.5, 4.5]	[5.5, 8.5]	[1.5, 4.5]	1	1
	Risk dom.	5.5	4.5	5.5	4.5	1	1
0.7	All	[5.5, 7.58]	[2.42, 4.5]	[5.5, 7.58]	[2.42, 4.5]	1	1
	Risk dom.	5.5	4.5	5.5	4.5	1	1
0.8	All	[6.33, 7], 9.5	[3, 3.67], 4.5	6, [6.33, 7]	2, [3, 3.67]	0, 1	0.8, 1
	Risk dom.	9.5	4.5	6	2	0	0.8
0.9	(Unique)	9.5	4.5	6.75	2.25	0	0.9
1.0	(Unique)	9.5	4.5	7.5	2.5	0	1

As noted already, in the cases of $q = 0$, 0.5, 0.7 and 0.8, a continuum of Nash equilibria exist, making sharp predictions difficult without the use of additional assumptions. In order to overcome this difficulty, we also consider the implications arising from additionally imposing the selection criterion of *risk dominance* (Harsanyi and Selten, 1988); in all four of these versions of NDG(q), this leads to a unique prediction. For $q = 0$, 0.5 and 0.7, risk dominance implies the most equitable efficient outcome, with agreement occurring with probability one, and with demands (and thus payoffs) of £5.50 and £4.50 by the favoured and unfavoured player respectively. For $q = 0.8$, although there are agreement equilibria, the risk dominant outcome is the chilling effect equilibrium, where both players choose their maximum demands and thus disagree with probability one; notably, this outcome is payoff dominated by all of the agreement equilibria.[18] NDG(0.9) and NDG(1.0) have only the chilling effect equilibrium, so the equilibrium prediction for these games is precise.

Comparison of these predictions, based on Nash equilibrium and in some cases risk dominance, for the various values of q in the experiment gives us the following hypotheses, which will structure our analysis of the experimental results.

Hypothesis 1 *(within p = 0, 0.5, 0.7) There are no systematic differences across the NDG(0), NDG(0.5) and NDG(0.7) games.*

Hypothesis 2 *(within p = 0.8, 0.9, 1.0) There are no systematic differences in favoured–player demands, unfavoured–player demands or agreement frequencies across the NDG(0.8), NDG(0.9) and NDG(1.0) games.*

Hypothesis 3 *(within p = 0.8, 0.9, 1.0) Within the NDG(0.8), NDG(0.9) and NDG(1.0) games, favoured– and unfavoured–player payoffs and efficiency increase as q increases.*

Hypothesis 4 *(between p = 0, 0.5, 0.7 and p = 0.8, 0.9, 1.0) There are no systematic differences in unfavoured–player demands between the NDG(0.8), NDG(0.9) and NDG(1.0) games and the NDG(0), NDG(0.5) and NDG(0.7) games.*

[18] In the case of NDG(0.8), risk dominance and Harsanyi and Selten's (1988) selection criterion have different implications. The risk dominant equilibrium is the chilling effect equilibrium, but this is payoff–dominated by all of the agreement equilibria. Since the Harsanyi–Selten criterion gives priority to payoff dominance over risk dominance, it will select any agreement equilibrium over the chilling effect equilibrium. Amongst the agreement equilibria, players' interests are perfectly opposed, so payoff dominance carries no implication once the chilling effect equilibrium is eliminated. Thus, the $\left(6\frac{1}{3}, 3\frac{2}{3}\right)$ equilibrium – which risk–dominates all of the other agreement equilibria – is selected.

Hypothesis 5 *(between p = 0, 0.5, 0.7 and p = 0.8, 0.9, 1.0) Favoured–player demands and payoffs are higher, unfavoured–player payoffs and agreement frequencies are lower, and efficiency is weakly lower, in the NDG(0.8), NDG(0.9) and NDG(1.0) games than in the NDG(0), NDG(0.5) and NDG(0.7) games.*

3 Experimental design and procedures

Our experimental design varies the game both within– and between–subjects. All subjects became familiarised with the experimental setting by playing ten rounds of the basic NDG (i.e., $q = 0$). Then, all subjects play an additional thirty rounds of NDG(q) with one of the strictly positive values of q (0.5, 0.7, 0.8, 0.9 or 1.0). The value of q in the second part of a given session was the same in all rounds and for all subjects in that session. Subjects also remained in the same role (favoured or unfavoured player) in all rounds, but they were randomly re–matched in each round to a subject in the opposite role.

The experimental sessions took place at the Scottish Experimental Economics Laboratory (SEEL) at the University of Aberdeen, between autumn 2010 and spring 2011. Subjects were primarily undergraduate students from University of Aberdeen, and were recruited using the ORSEE system (Greiner, 2004) from a database of people expressing interest in participating in economics experiments. No one took part in this experiment more than once.

At the beginning of a session, subjects were seated in a single room and given written instructions for the first ten rounds.[19] These instructions stated that the experiment would be made up of two parts and that the second half would comprise thirty additional rounds, but additional details of the second half were not announced until after the first half had ended. The instructions were also read aloud to the subjects, in an attempt to make the rules of the game common knowledge. Then, the first round of play began. After the tenth round was completed, each subject was given a copy of the instructions for rounds 11–40. These instructions were also read aloud, after which time round 11 was played.

The experiment was run on networked personal computers, and was programmed using the z–Tree experiment software package (Fischbacher, 2007). Subjects were asked not to communicate with other subjects except via the computer program. No identifying information was given about opponents (in an attempt to minimise incentives for coordination across rounds, reputation, and other repeated game effects). Also, in order to minimise demand effects, we referred to a subject's opponent as "the other player" or "the player paired with you" which, while sometimes cumbersome–sounding, avoids the negative framing of "opponent" or the positive framing of "partner".

Each round of the experiment began with subjects being prompted to choose their demands (called "claims" in the experiment). Demands were restricted to be whole–number multiples of £0.01, between zero and the subject's maximum allowable demand (£4.50 or £9.50); both own and opponent maximum allowable demands were displayed on the computer screen at this time. After all subjects had entered their demands, the round ended and they received feedback. In rounds 1–10 (when subjects played the basic NDG), feedback comprised the subject's own demand, the opponent demand, whether agreement was reached, own payoff and opponent payoff. In rounds 11–40, feedback included all of these and in the case of disagreement, whether the subject, the opponent or neither was chosen to have his/her demand implemented. After viewing these results and clicking a button to continue, the next round began.

At the end of the fortieth round, the experimental session ended and subjects were paid, privately and individually. For each subject, one round from rounds 1–10 and three from rounds 11–40 were randomly chosen, and the subject was paid the sum of his/her earnings in those rounds, to the penny. Subjects' total earnings ranged from £4.50 to £34.00, and averaged approximately £17.85, for a session that typically lasted about 60 minutes.

[19] Sample instructions and screenshots are shown in the Appendix. The remaining sets of instructions, as well as other experimental materials and the raw data from the experiment, are available from the corresponding author upon request.

4 Experimental results

A total of 264 subjects participated in the experiment (see Table 2), in 18 sessions. We will begin the discussion of experi-

Table 2: Treatment information

Value of q	Rounds	Groups	Subjects
0	1–10	30	264
0.5	11–40	6	64
0.7	11–40	6	50
0.8	11–40	6	48
0.9	11–40	6	56
1.0	11–40	6	46

mental results with an analysis of the aggregate data. Our hypotheses concerning these aggregates will be tested with the use of conservative non–parametric statistical tests. The unit we will use for these tests is the "group", which we define to be a (not necessarily proper) subset of an experimental session, made up of equal numbers of favoured and unfavoured subjects, closed with respect to interaction. That is, favoured subjects in a particular group interacted only with unfavoured subjects in the same group. Thus, data from any group can be considered statistically independent of data from any other group. Each experimental session, depending on its size, comprised one or more groups; as Table 2 shows, the 18 experimental sessions comprised a total of 30 groups.

4.1 Treatment aggregates

Some summary data are presented in Table 3 and Figures 3–5. For each version of NDG(q), Table 3 shows six treatment–wide statistics – the mean demands of both types of player (favoured and unfavoured), the mean payoffs of both types, the frequency of agreement, and mean efficiency (i.e., the sum of payoffs, normalised to a zero–one scale) – along with their corresponding standard deviations. The figures show these statistics at both the aggregate level and disaggregated by group, to give an impression of the level of heterogeneity within treatments.

Figure 3: Favoured– and unfavoured–player demands, as fraction of cake (aggregate and by group)

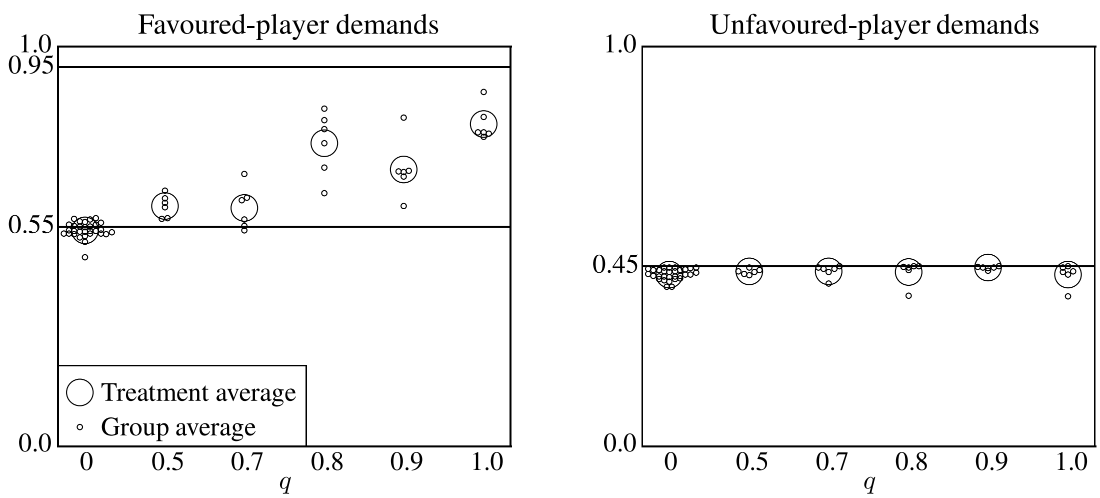
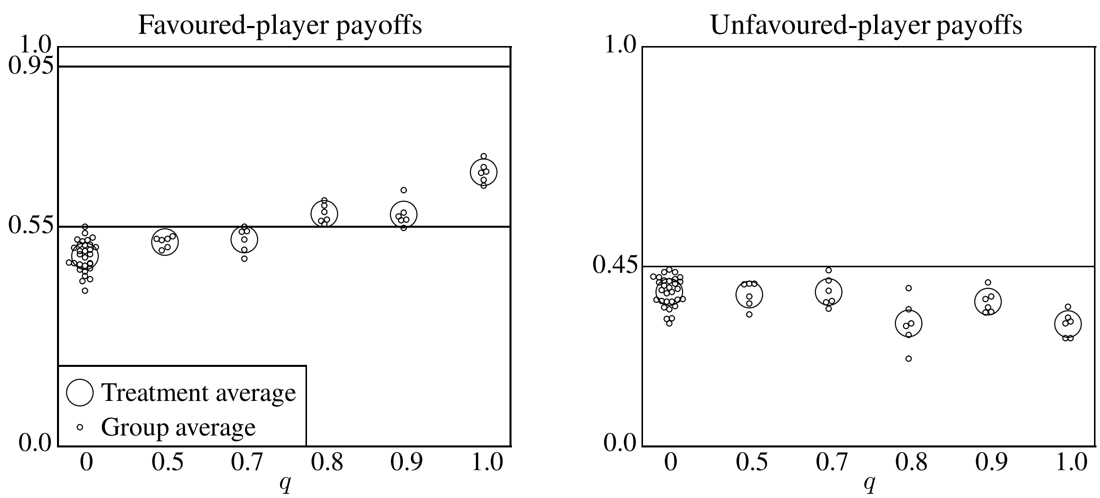


Table 3: Aggregate results – treatment means, standard deviations (in parentheses) and significance test results

Value of q	0	0.5	0.7	0.8	0.9	1.0
Favoured player demand (£):	5.41^{a}	6.04^{b}	5.95^{b}	7.60^{cd}	6.88^{c}	8.06^{d}
	(0.73)	(1.30)	(1.42)	(1.92)	(1.80)	(1.78)
Unfavoured player demand (£):	4.29^{a}	4.36^{b}	4.37^{bc}	4.37^{abc}	4.47^{c}	4.29^{abc}
	(0.50)	(0.44)	(0.51)	(0.50)	(0.21)	(0.73)
Favoured player payoff (£):	4.76^{a}	5.10^{b}	5.19^{b}	5.80^{c}	5.78^{c}	6.88^{d}
	(1.68)	(1.87)	(1.74)	(2.58)	(1.94)	(1.82)
Unfavoured player payoff (£):	3.85^{c}	3.77^{bc}	3.86^{bc}	3.05^{a}	3.62^{b}	3.02^{a}
	(1.37)	(1.49)	(1.40)	(1.90)	(1.58)	(1.79)
Agreement frequency (%):	90.1^{d}	79.0^{c}	76.4^{c}	39.7^{ab}	52.3^{b}	24.2^{a}
	(29.9)	(40.8)	(42.5)	(49.0)	(50.0)	(42.9)
Efficiency (%):	86.2^{a}	88.6^{b}	90.5^{abc}	88.5^{ab}	94.0^{c}	99.0^{d}
	(29.5)	(29.8)	(26.1)	(31.0)	(22.5)	(4.2)

Note: Within each statistic, entries with no superscripts in common are significantly different at the 5% level (group–level data, see text for additional details); superscripts earlier in the alphabet correspond to significantly lower values.

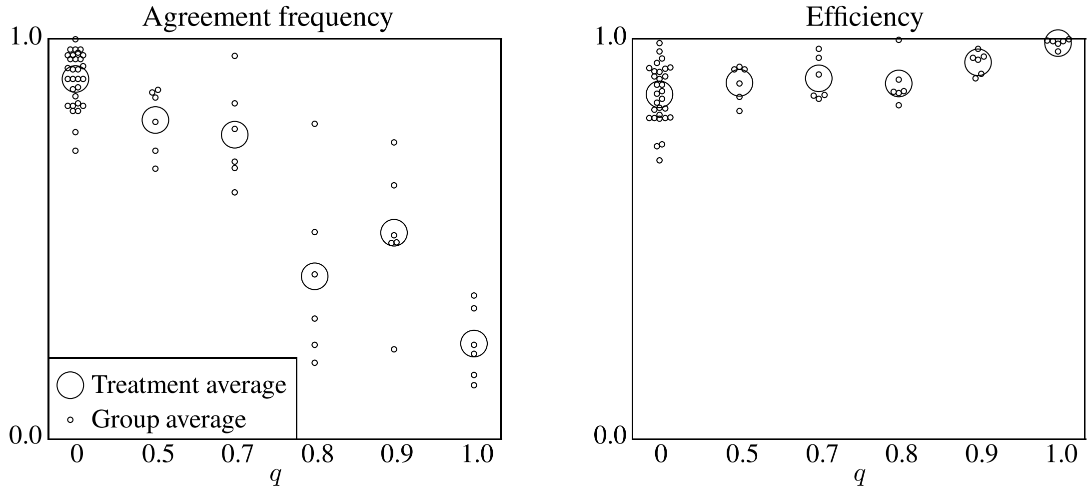
Figure 4: Favoured– and unfavoured–player payoffs, as fraction of cake (aggregate and by group)



Also shown in Table 3 are significance results from nonparametric tests of differences across individual cells for each of these five variables, using group–level data, and with significance defined as a p–value of 0.05 or less.[20] These significance results are displayed as superscript letters to the table entries; for a given statistic, entries sharing a superscript are not

[20]For pairwise comparisons between NDG(0.5), NDG(0.7), NDG(0.8), NDG(0.9) and NDG(1.0), we use the robust rank–order test, since no group faces more than one of these games, making them independent samples. For comparisons between any of these games and NDG(0), we use the Wilcoxon signed–ranks test for matched samples, with data from a particular positive–q NDG group (e.g., NDG(0.9)) compared only to the subset of NDG(0) groups that subsequently played that particular version of NDG(q) – reducing any variability that might have been due to session or group effects. See Siegel and Castellan (1988) for descriptions of the nonparametric statistical tests used in this paper. Some critical values for the robust rank–order tests are from Feltovich (2005).

Figure 5: Agreement frequency and efficiency (aggregate and by group)



significantly different, while entries with letters earlier in the alphabet correspond to significantly lower values (e.g., a statistic with a b superscript is significantly higher than one with an a superscript, but neither is significantly different from one with an ab superscript.)

We begin our discussion of the aggregate data by remarking on a few salient features that emerge from Table 3 and the three figures. First, the data show a fair amount of variation across treatments, which we attribute to treatment effects, and which in some cases is particularly striking. As an example, consider the sharp decrease in agreement frequencies as one moves from NDG(0) to NDG(1.0): from over 90% to less than one–quarter.[21] As a second example, note that as q increases, favoured–player payoffs rise substantially, while unfavoured–player payoffs tend to fall. Consequently, the ratio of favoured– to unfavoured–player payoffs for low values of q (e.g., roughly 1.24 when $q = 0$) is comparable to $\frac{11}{9} \approx 1.22$, the common prediction of the Nash and lexicographic egalitarian bargaining solutions applied to the corresponding unstructured–bargaining setting, but for high values of q this ratio (e.g., approximately 2.28 when $q = 1$) is closer to $\frac{19}{9} \approx 2.11$, the ratio implied by the less equitable Kalai–Smorodinsky solution.[22]

Second, we also see heterogeneity within each cell, and in some cases this is also substantial. In particular, the range of mean favoured–player demands across groups in each of the NDG(0.8) and NDG(0.9) cells is roughly 20 percentage points (i.e., £2), and the range of agreement frequencies is over 50 percentage points. This heterogeneity suggests that group effects (analogous to session effects, but at the level of the individual group) exist in our data. The non–parametric tests we report here cannot accommodate group effects, but we will try to control for these in the regressions of Section 4.3.

Third, separate from the success or failure of our hypotheses, all of which are of a qualitative nature (involving the direction of an effect, but not its size), the point predictions of Nash equilibrium (see Table 1) have, at best, equivocal success in characterising subject behaviour. In some cases, these point predictions fare well; for example, average unfavoured–player demands in all cells are close to the equilibrium point prediction of £4.50, and the ranges of predicted values for favoured–player demands and both favoured and unfavoured types' payoffs in NDG(0), NDG(0.5), NDG(0.7) and NDG(0.8) usually

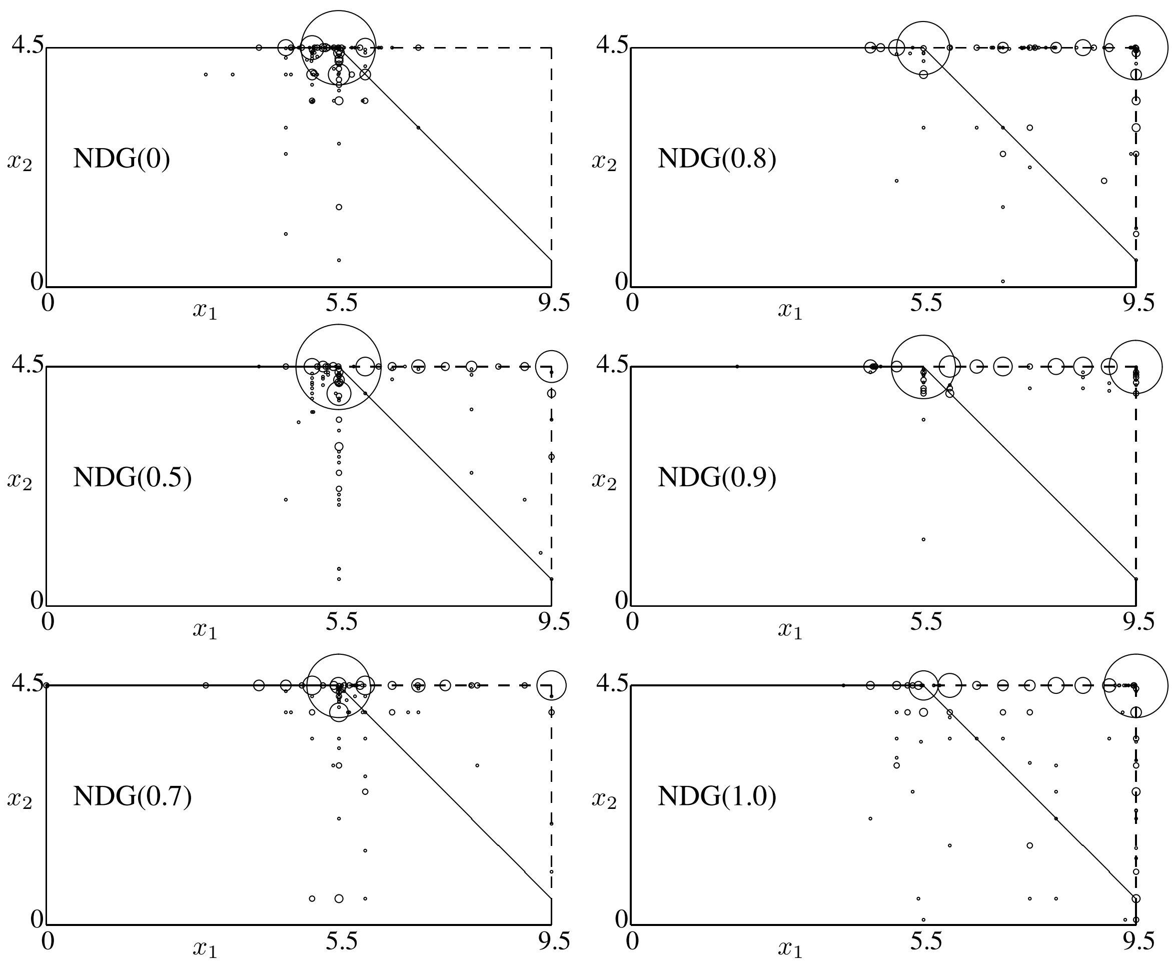
[21] Unsurprisingly, as agreements become less common, the chilling effect is observed more often. The frequency of (9.50, 4.50) outcomes tends to rise with q, from 0.1% of all outcomes when $q = 0$ and 8.2% and 8.8% for $q = 0.5$ and $q = 0.7$ (where there is no chilling–effect equilibrium) to 38.3% when $q = 0.8$, 22.7% when $q = 0.9$ and 60.2% when $q = 1$.

[22] Non–parametric Kruskall–Wallis tests (which differ from the robust rank–order test in that they detect differences across any number of independent samples, rather than between two samples) reject the null hypothesis of no difference across the positive–q treatments in favoured–player demands, favoured– and unfavoured–player payoffs, agreement frequencies, and efficiency ($p < 0.01$ in all cases), though no significant differences are found in unfavoured–player demands ($p \approx 0.24$).

contain the corresponding observed value from the experiment. However, these three statistics in NDG(0.9) and NDG(1.0), and agreement frequencies in all cells except NDG(0.8), are far away from their corresponding point predictions. Efficiency is well below the theoretical prediction in NDG(0), NDG(0.5) and NDG(0.7), above it in NDG(0.8) and NDG(0.9), and roughly equal to it in NDG(1.0).

In order to shed some light on the mixed success of equilibrium point predictions, Figure 6 displays the outcomes from each individual pair of bargainers in all cells, from all rounds after the first five using that value of q (i.e., from rounds 6–10 for NDG(0) and rounds 16–40 for the other games). Each outcome is completely characterised by the corresponding favoured– and unfavoured–player demands, shown on the horizontal and vertical axis respectively. For each pair of demands, the figure shows a circle with area proportional to the number of times that pair of demands occurred in that cell.

Figure 6: All pairs of favoured– and unfavoured–player demands (in £), rounds 6–10 for NDG(0) and rounds 16–40 for other games *(area of circle is proportional to number of observations)*



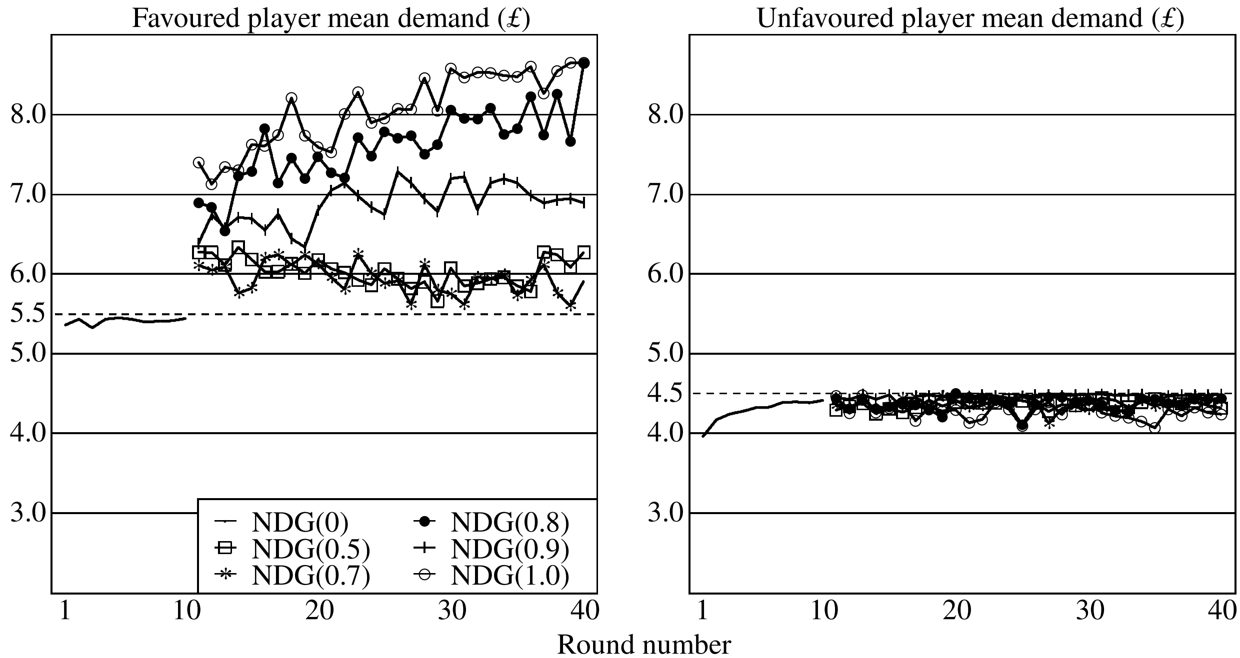
The figure confirms what we saw previously in Table 3 and Figures 3–5: there is typically substantial heterogeneity within cells. Also, coordination on any division of the cake other than (5.5, 4.5) is extremely rare, despite the continuum of agreement equilibria for some values of q. In NDG(0), the agreement equilibrium (5.5, 4.5) is by far the most frequent outcome, while in the other cells, (5.5, 4.5) is one of two modes, along with the chilling–effect outcome (9.5, 4.5). This is true in NDG(0.5) and NDG(0.7), where there exist agreement equilibria but no chilling–effect equilibrium; in NDG(0.9) and NDG(1.0), where there is a chilling–effect equilibrium but no agreement equilibria; and in NDG(0.8), where both types of equilibrium exist, but (5.5, 4.5) is not an equilibrium. The substantial amount of non–equilibrium play in all cells except

NDG(0) provides an explanation for the mixed success of theoretical point predictions to characterise play at the aggregate level, along with showing that this ability is also limited at the individual level.

4.2 Evolution of play over time

Some more information about subject behaviour comes from Figures 7, 8 and 9. Figure 7 shows the round–by–round time paths of favoured and unfavoured players' demands for each game, and Figures 8 and 9 do the same for favoured and unfavoured players' payoffs, agreement frequencies, and efficiency. The figures show that over the first ten rounds (when

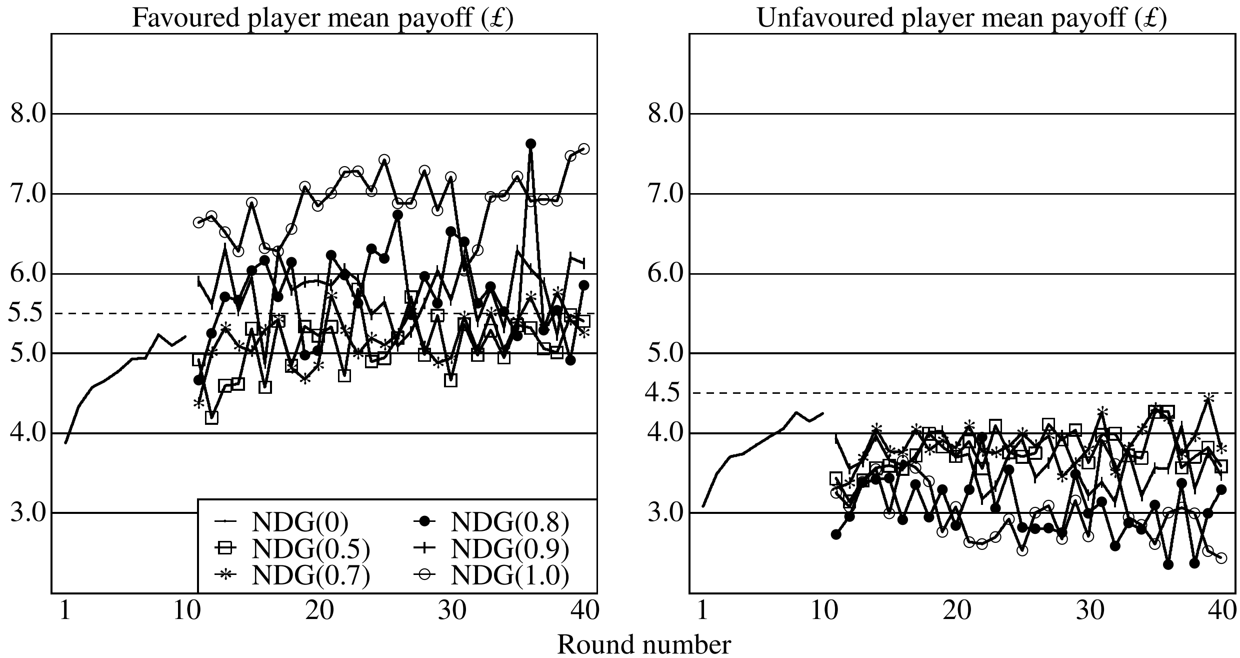
Figure 7: Favoured and unfavoured player demands – time paths



all subjects play NDG(0)), behaviour converges approximately to the 55–45 split implied by risk dominance and other equilibrium selection criteria. Average favoured–player demands start and remain near £5.50, and average unfavoured–player demands, while beginning below £4.00, rise over time toward £4.50. Neither side's demands quite reach these asymptotes, however, as some subjects continue to hedge against the strategic uncertainty of the game by reducing their demands. Despite average favoured–player demands staying roughly constant over time and unfavoured–player demands rising, agreement frequencies also rise, due to decreases in the variance in both types' demands – also leading to increases in both types' average payoffs, toward (£5.50, £4.50).

The left panel of Figure 7 shows that when a positive probability of random implementation is implemented (between rounds 10 and 11), mean favoured–player demands immediately rise by £0.70–1.95, depending on the treatment (an increase on the order of 25%, on average), as subjects in this role apparently attempt to take advantage of their perceived increase in bargaining power. In NDG(0.5) and NDG(0.7), this increase is followed by a slow decline, reflecting the continued existence of (5.5, 4.5) as an equilibrium, and suggesting its robustness as a focal point (as evidenced by the right panel of this figure, where unfavoured–player demands are, at least on average, nearly unaffected by introducing a positive q). By the last ten rounds, favoured–player demands in these two games fall back to between £5.50–6.00 – nearly where they were in NDG(0) – though there are signs of some sort of endgame effect in the last few rounds of NDG(0.5). By contrast, favoured–player demands in NDG(0.8), NDG(0.9) and NDG(1.0) continue rising over time (though more slowly in NDG(0.9) than in the other two cells), averaging more than £7 over the last twenty rounds, and in the case of NDG(1.0), more than £8. While these

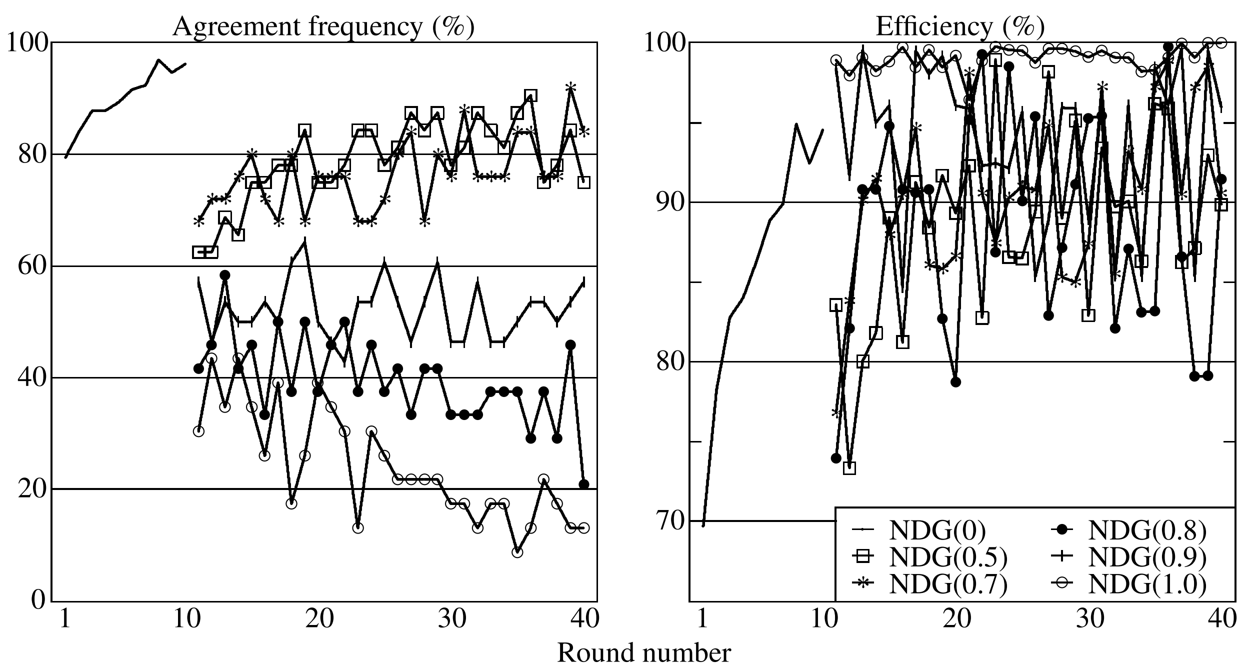
Figure 8: Favoured and unfavoured player payoffs – time paths



amounts are still far from the Nash equilibrium point prediction of £9.50, the movement away from the non–equilibrium (5.5, 4.5) outcome is clear.

Similar conclusions can be drawn from the time paths of agreement frequencies, shown in Figure 9. Here, we see

Figure 9: Agreement frequency and efficiency – time paths



that from near–certain agreement in the last few rounds of NDG(0), there is a precipitous drop, to between 30% and 70% depending on q, when the possibility of random implementation is introduced. These drops reflect (as noted above) initial attempts by favoured players in each of these treatments to take advantage of their perceived improvement in bargaining position by raising demands, while unfavoured players do not appreciably decrease their own demands. As the sessions progress, the agreement frequencies further diverge across cells. In NDG(0.5) and NDG(0.7), favoured players are largely unsuccessful in their attempts to exploit their (seeming) improved bargaining power, and agreement frequencies slowly rebound as favoured players lower their demands back toward £5.50. This does not happen in the treatments with larger q. In NDG(0.9), agreement frequencies stay at roughly one–half; in NDG(0.8), they continue declining to approximately 30%; and in NDG(1.0) they fall even further, averaging below 20% over the last twenty rounds. As with favoured–player demands, while these levels are far from the equilibrium point prediction (zero agreement frequency), they are substantially below the corresponding frequencies in the NDG(0), NDG(0.5) and NDG(0.7) cells, consistent with the qualitative prediction of the theory.

In contrast, the time paths of efficiency shed little additional light compared with what could be seen in the aggregate data. Efficiency increases over time in NDG(0), as agreements become more frequent. From round 11 on, round–to–round noise appears to overwhelm differences across treatments, with the exception that efficiency in NDG(1.0) is consistently higher than in the other treatments, and close to 1 in all rounds.

4.3 Parametric statistics

We continue our examination of the experimental data by using parametric statistical methods, in order to disentangle the effects of the multiple factors that could influence outcomes in our NDG(q) games, and thus address our hypotheses. We consider five dependent variables: favoured–player demands, favoured– and unfavoured–player payoffs (all as fractions of the cake), agreement frequencies, and efficiency. (Since there is no interesting variation across treatments in unfavoured–player demands, we leave that statistic out of the rest of our analysis to save space.) For demands and payoffs, we estimate Tobit models on the pair–level data with endpoints zero and the maximum allowable demand; for efficiency, we use Tobits with zero and one as the endpoints; and for agreement frequencies, we use probits. For each dependent variable, there are two sets of explanatory variables: a restricted set which does not allow for differences between the NDG(0.5) and NDG(0.7) cells or amongst the NDG(0.8), NDG(0.9) and NDG(1.0) cells, and an unrestricted set which does. Thus, right–hand–side variables in the restricted set included an indicator for $q \in \{0.8, 0.9, 1.0\}$, called NDG(8–10). By contrast, the unrestricted set had indicators for individual cells (values of q), labelled NDG(0.5), NDG(0.7),... NDG(1.0). All regressions included the products of these cell variables with the round number (to control for time–dependence of the treatment effects), the round number itself, and a constant term.

Finally, due to the within–cell heterogeneity seen in Figures 3–5, we sought to control for possible group effects. Rather than simply including group fixed effects (which would partly mask treatment effects, due to our between–subjects variation across positive values of q), we defined a new variable, "round 1–10 average", equal to the group mean value of the left–hand–side variable over rounds 1–10. By definition, this variable will not change within a group (across either subjects or rounds), but can vary across groups. Since we use the data from rounds 1–10 to calculate this variable, we restrict the data set to rounds 11–40, meaning that we have dropped the NDG(0) observations. Hence, our odd–numbered models, which use the restricted set of right–hand–side variables, use $q \in \{0.5, 0.7\}$ as the baseline, while the even–numbered models, which use the unrestricted set, use NDG(0.5) as the baseline.

All of the models were estimated using Stata (version 11.2), and incorporated individual–subject random effects and bootstrapped standard errors. Table 4 presents coefficient estimates, standard errors, and log likelihoods. Also shown are the estimated marginal effects of the treatment variables – taking into account interaction effects – with the round number set to 25 (the midpoint of the second half of a session) and the "round 1–10 average" variable set to its mean. Additionally, each model's performance according to the Bayesian Information Criterion (Schwarz, 1978) and the Akaike (1974) Information

Criterion are shown.[23]

Table 4: Regression results for rounds 11–40 – coefficient estimates, bootstrapped standard errors, and marginal effects

Dep. var.:	Demand (favoured)		Payoff (favoured)		Payoff (unfavoured)		Agreement		Efficiency	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Coefficients:										
constant	0.550	0.238	0.423^{***}	0.476^{***}	0.324^{**}	0.327^{*}	–2.573	–2.015	–0.252	0.587
	(0.405)	(0.558)	(0.086)	(0.101)	(0.158)	(0.174)	(2.020)	(2.132)	(0.709)	(0.744)
round	–0.001	–0.001	0.002^{***}	0.002^{***}	0.007^{***}	0.006^{**}	0.027^{**}	0.029^{*}	0.021^{***}	0.020^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.013)	(0.016)	(0.005)	(0.006)
round 1–10	0.141	0.701	0.130	–0.007	0.708^{*}	0.687	3.898^{*}	3.382	1.868^{**}	0.917
average	(0.743)	(1.002)	(0.019)	(0.203)	(0.400)	(0.453)	(2.188)	(2.414)	(0.814)	(0.863)
NDG(0.7)		0.018		0.013		0.008		–0.252		–0.083
		(0.047)		(0.022)		(0.074)		(0.398)		(0.179)
NDG(0.8)		0.107^{*}		0.096^{***}		–0.060		-0.899^{*}		0.354^{*}
		(0.057)		(0.028)		(0.067)		(0.532)		(0.205)
NDG(0.9)		0.070		0.106^{***}		0.060		-0.913^{*}		0.685^{***}
		(0.056)		(0.026)		(0.067)		(0.521)		(0.197)
NDG(1.0)		0.187^{***}		0.213^{***}		–0.071		-1.447^{***}		1.032^{***}
		(0.059)		(0.051)		(0.065)		(0.350)		(0.270)
NDG(8–10)	0.110^{***}		0.130^{***}		–0.023		-0.964^{***}		0.702^{***}	
	(0.033)		(0.019)		(0.053)		(0.295)		(0.161)	
Average marginal effects										
NDG(0.7)		0.013		0.012		0.015		–0.351		–0.029
		(0.042)		(0.012)		(0.064)		(0.422)		(0.159)
NDG(0.8)		0.220^{***}		0.087^{***}		-0.179^{***}		-1.672^{***}		0.162
		(0.050)		(0.018)		(0.050)		(0.471)		(0.165)
NDG(0.9)		0.118^{**}		0.083^{***}		–0.048		-1.365^{***}		0.471^{***}
		(0.050)		(0.017)		(0.052)		(0.518)		(0.165)
NDG(1.0)		0.331^{***}		0.222^{***}		-0.216^{***}		-2.522^{***}		1.042^{***}
		(0.069)		(0.030)		(0.048)		(0.375)		(0.343)
NDG(8–10)	0.205^{***}		0.122^{***}		-0.151^{***}		-1.668^{***}		0.509^{***}	
	(0.033)		(0.014)		(0.041)		(0.248)		(0.128)	
Hypothesis tests of equal average marginal effects:										
NDG(0.8), NDG(0.9), NDG(1.0)		$p \approx 0.011$		$p < 0.001$		$p < 0.001$		$p \approx 0.017$		$p \approx 0.026$
Interactions with round number?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3960	3960	3960	3960	3960	3960	3960	3960	3960	3960
$\|ln(L)\|$	657.22	688.52	931.56	905.49	2307.02	2300.46	1579.58	1566.84	2179.48	2168.80
BIC	**1347.58**	1435.03	1896.26	**1868.97**	**4647.18**	4658.91	3192.30	**3191.67**	**4392.10**	4395.59
AIC	**1322.44**	1391.04	1871.12	**1824.98**	4622.04	**4614.92**	3167.16	**3147.68**	4366.96	**4351.60**

* (**,***): *Coefficient significantly different from zero at the 5% (1%, 0.1%) level.*

We note first, based on the BIC and AIC values, that neither restricted nor unrestricted models are systematically better

[23]The Bayesian and Akaike information criteria function like standard likelihood–ratio tests in that they reward goodness–of–fit but punish free parameters, but can be used even when the competing models are not nested, as is the case here. It is possible for these two criteria to select different models, as the BIC punishes additional free parameters more severely than the AIC does. Both criteria are commonly used, so we report both here.

than the other. For favoured–player demands, the restricted model performs better according to both criteria (lower values are better), while both criteria prefer the unrestricted model for favoured–player payoffs and agreement frequencies, and the criteria disagree in the cases of unfavoured–player payoffs and efficiency. Since neither version of the model is consistently better than the other, we will generally report results that are common to both versions as much as possible.

Moving to the effects of our treatments, we see that the average marginal effect of the NDG(0.7) variable is insignificant in four of the five models in which it appears, and only significant at the 10% level in the fifth, consistent with our Hypothesis 1 (similar behaviour across NDG(0), NDG(0.5) and NDG(0.7)). However, we note the danger involved in making positive conclusions based on failure to reject null hypotheses. Also, because we had to drop the NDG(0) data in order to use round 1–10 averages as a right–hand–side variable, the resulting regressions are only able to detect differences between NDG(0.5) and NDG(0.7), rather than the differences between these games and NDG(0) that were typically seen in the aggregate data.

For the remaining cells (NDG(0.8), NDG(0.9) and NDG(1.0)), these regressions show some systematic differences. In all five regressions in which the NDG(0.8)–NDG(1.0) indicators appear separately, we can reject the null hypothesis of equal average marginal effects. Looking at the three marginals individually indicates that there are sharp differences in behaviour between NDG(1.0) and the other two cells, and usually comparatively smaller differences between NDG(0.8) and NDG(0.9).[24] The differences in favoured–player demands and agreement frequencies are at odds with our Hypothesis 2 (similar levels in these variables across NDG(0.8)–NDG(1.0)), and the differences in unfavoured–player payoffs are nearly opposite of our Hypothesis 3 (increasing payoffs for both players and efficiency as q increases). On the other hand, the increases with q in favoured–player payoffs and efficiency are broadly consistent with this hypothesis.

Comparisons between the NDG(0.5) and NDG(0.7) cells and the NDG(0.8), NDG(0.9) and NDG(1.0) cells are captured easily by the marginal effect of the NDG(8–10) variable. In all five models, this variable is significant at the 1% level (indeed, at the 0.1% level). The sign of its coefficient is in the direction predicted by our Hypothesis 5 (higher demands and payoffs for the favoured player, lower unfavoured payoffs and agreement frequencies, and weakly lower efficiency) for all of the variables except efficiency, where the increased marginal effect for high values of q is the opposite of that stated by this hypothesis.

4.4 Summary of results

In this section, we first collect together the main results of the experiment, and where applicable, we compare them to the hypotheses we listed in Section 2.4. Then, we attempt to put the results into perspective.

Result 1 *There are no economically relevant differences in unfavoured–player demands across games.*

Table 3 shows that across all of the positive–q versions of NDG(q), the only statistically significant difference in unfavoured–player demands is between NDG(0.5) and NDG(0.9), and the actual difference in their means is £0.11. While unfavoured–player demands are significantly lower in NDG(0) than in several of the other cells, the size of the difference is again small (at most £0.18). Indeed, the largest difference between any two NDG(q) is £0.18, between NDG(0.9) and either NDG(1.0) or NDG(0). The lack of any meaningful difference across cells is consistent with our Hypothesis 4.

Result 2 *We find no evidence of treatment effects between NDG(0.5) and NDG(0.7), but substantial differences exist between these games and NDG(0). Favoured–player demands and payoffs are higher, and agreement frequencies are lower, in NDG(0.5) and NDG(0.7) than in NDG(0). Efficiency is higher in NDG(0.5) than NDG(0).*

[24] Pairwise tests between NDG(0.8) and NDG(0.9) reject the null of equal marginal effects in only two of the five models (unfavoured player payoffs and efficiency), both at the 5% level but not at the 1% level. This suggests that a large portion of the apparent differences seen in the aggregate data (e.g., Figures 3–5) between these two games disappears when individual heterogeneity is controlled for via the "round 1–10" variables.

The differences between NDG(0) and both NDG(0.5) and NDG(0.7), seen in the non–parametric test results in Table 3, are inconsistent with our Hypothesis 1, which predicted no differences. The lack of differences between NDG(0.5) and NDG(0.7) is suggested by the insignificant (and usually very small) coefficients and marginal effects for the NDG(0.7) variable in our regressions (Table 4), and is consistent with Hypothesis 1, though as always, we should point out that failure to reject a null hypothesis of no difference is only weak evidence that there actually is no difference, and one should be careful in drawing conclusions based on such results.

Result 3 *We find significant variation across the NDG(0.8), NDG(0.9) and NDG(1.0) games, in all statistics except for unfavoured–player demands. Efficiency increases from NDG(0.8) to NDG(0.9) and thence to NDG(1.0), while the other variables do not change systematically with q.*

This result is apparent from the tests of equal average marginal effects in Table 4 (though fewer differences are seen in the more conservative non–parametric tests in able 3). The differences across these three cells in favoured–player demands and agreements are inconsistent with Hypothesis 2. The monotonic change in efficiency with q supports Hypothesis 3, while the lack of such changes in favoured– and unfavoured–player payoffs is inconsistent with this hypothesis.[25] As noted earlier, the largest differences within these cells are between NDG(1.0) and the other two.

Result 4 *Favoured–player demands and payoffs and efficiency are higher, and unfavoured–player payoffs and agreement frequencies are lower, in the NDG(0.8), NDG(0.9) and NDG(1.0) games than in the NDG(0), NDG(0.5) and NDG(0.7) games.*

These differences are seen in both the non–parametric tests and the regression results (in particular, the coefficient and marginal effect of NDG(0.8–1.0)). All are consistent with our Hypothesis 5 except for the efficiency result, which is nearly the opposite of the (weak) decrease hypothesised there.

Taken together, our results can be characterised in terms of three stylised facts. First, the tendency of favoured–player demands to increase, and agreement frequencies to decrease, from low to high values of q (Result 4) is consistent with a move away from agreement equilibria and toward chilling–effect equilibria. Second, the tendency of favoured–player payoffs to rise with q (Results 2 and 4), coupled with the less systematic effect on unfavoured–player payoffs (the increase in Result 4, and the lack of significant differences in Results 2 and 3), is consistent with a change in the distribution of relative payoffs, away from the unfavoured player and toward the favoured player, along with a general increase in payoff efficiency (also seen in Result 4). Third, the large differences seen between the games with $q \leq 0.7$ and those with $q \geq 0.8$ (Result 4), along with the smaller and usually insignificant differences between NDG(0.5) and NDG(0.7) (Result 2), and the typical lack of systematic differences across NDG(0.8)–NDG(1.0) once group effects are controlled for (Result 3), suggest that behaviour changes relatively little as q rises up to a value of 0.7, and beyond a value of 0.8, but there are sharp differences over a fairly small interval (between 0.7 and 0.8).

5 Discussion

The Nash demand game (NDG) has long been used as a model of how two–party bargaining occurs. The game has two disadvantageous features, however. First, almost all variations of the game have a large number of Nash equilibria, lessening its value as a source of predictions. Second, only two outcomes are possible: immediate agreement or immediate disagreement,

[25]The effect on efficiency is also seen in a non–parametric Jonckheere test, which is similar to the Kruskall–Wallis test, but with a directional alternative hypothesis. The increase in efficiency from NDG(0.8) to NDG(0.9) to NDG(1.0) is highly significant ($J = 96$, $p < 0.005$).

with no opportunity for renegotiation. Anbarci and Boyd (2011) proposed a modification of the standard NDG, under which incompatible demands – rather than leading to certain disagreement – only lead to disagreement with probability $1 - q$. With the remaining probability q, there is *random implementation*, with one of the bargainers randomly chosen to receive his/her demand and the other receiving the remainder.

Our paper is an experimental examination of this "NDG with random implementation". We begin with a bargaining setting in which one player is favoured relative to the other, and we vary the random implementation parameter q. We show theoretically that as q increases, the set of "agreement equilibria" (Nash equilibria in which agreement is reached) weakly shrinks. Beyond a threshold value of q, there also exists a "chilling–effect" equilibrium with both players demanding their maximum amount (and thus not reaching agreement); beyond an even higher higher value of q, this is the only Nash equilibrium.

As is often the case, our experimental data provide only weak support for standard game theory's point predictions. However, we find fairly strong support for the theory's directional predictions, whether we use conservative non–parametric statistical tests or standard regression techniques. Firstly, we find stark decreases in agreement frequencies as q increases, consistent with the rising prominence of the chilling–effect equilibrium in comparison with the set of agreement equilibria.

Secondly, we find that raising q tends to reinforce the asymmetry of the underlying bargaining setting, increasingly benefiting the favoured player at the expense of the unfavoured player. On average, we find that the ratio of favoured– to unfavoured–player payoffs increases with q, from being roughly comparable to the ratio implied by the lexicographic egalitarian outcome (the most equitable efficient outcome) at the lowest values of q to being about the same as that implied by the more unequal Kalai–Smorodinsky outcome at the highest values of q. On the other hand, efficiency also tends to increase with q (a result which is only partly consistent with the theoretical prediction), which tempers the overall effect on unfavoured players. For some increases in q, unfavoured–player payoffs fall, while for others, they rise, so that in some cases there are Pareto improvements associated with increasing q.

While we acknowledge that our random–implementation mechanism is, at best, a stylised model of real bargaining and arbitration, the two main directional effects of increasing q (the distributional effect and the efficiency effect) raise the possibility that some version of this mechanism could arise. To the extent that Pareto improvements are possible, there are obvious incentives for bargainers to agree to be bound by it. Alternatively, to the extent that the mechanism reinforces the bargaining power the favoured–player already had, and given the likelihood that favoured players may have further advantages outside the bargaining process itself (e.g., more resources for political activities), there is a clear incentive for them to lobby policy makers to adopt such mechanisms, even in cases where unfavoured players would suffer as a result, and even where measures of social benefit like efficiency or agreement frequency would also suffer. This distributional aspect of random implementation – and in particular, the possibility that bargainers may have almost diametrically opposed preferences over the size of q – may be an important topic for further research.

Finally, consistent with the theory, we find that increases in q yield non–uniform effects on observed bargaining outcomes in general. For some fairly substantial ranges within the unit interval – from 0 to 0.7, and again from 0.8 to 1.0, changes to q are observed to have minimal effect on either the likelihood of agreement or the distribution of the surplus given that agreement is reached. But a small change to q, moving from $q = 0.7$ to $q = 0.8$, has a sharp impact on results, consistent with the implication of risk dominant Nash equilibrium.

References

Akaike, H. (1974), "A new look at the statistical model identification", *IEEE Transactions on Automatic Control* 19, pp. 716-723.

Anbarci, N. and J.H. Boyd III (2011), "Nash demand game and the Kalai–Smorodinsky solution", *Games and Economic Behavior* 71, pp. 14–22.

Ashenfelter, O., J. Currie, H.S. Farber and M. Spiegel (1992), "An experimental comparison of dispute rates in alternative arbitration systems", *Econometrica* 60, pp. 1407–1433.

Axelrod, R. (1967), "Conflict of interest: an axiomatic approach", *Journal of Conflict Resolution* 11, pp. 87–99.

Babcock, L., G. Loewenstein, S. Issacharoff, and C. Camerer (1995), "Biased judgments of fairness in bargaining", *American Economic Review* 85, pp. 1337–1343.

Binmore, K. (2007), *Playing for Real: a Text on Game Theory*, Oxford University Press, Oxford.

Binmore, K., J. Swierzbinski, S. Hsu and C. Proulx (1993), "Focal points and bargaining", *International Journal of Game Theory* 22, pp. 381–409.

Brandts, J. and G. Charness (2011), "The strategy versus the direct–response method: a first survey of experimental comparisons", *Experimental Economics* 14, pp. 375–398.

Brams, S.J. and S. Merrill, III (1983), "Equilibrium strategies for final–offer arbitration: there is no median convergence", *Management Science* 29, pp. 927–941.

Cabrales, A., R. Miniaci, M. Piovesan, and G. Ponti (2010), "Social preferences and strategic uncertainty: an experiment on markets and contracts", *American Economic Review* 100, pp. 2261–2278.

Camerer, C.F. (2003), *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press, Princeton, NJ.

Charness, G. and Kuhn, M.A. (2011), "Lab labor: what can labor economists learn from the lab?" In *Handbook of Labor Economics, v. 4a*, O. Ashenfelter and D. Card, eds., North–Holland, Amsterdam, pp. 229–330.

Chatterjee, K. (1981), "Comparison of arbitration procedures: models with complete and incomplete information", *IEEE Transactions on Systems, Man, and Cybernetics* SMC–11, pp. 101–109.

Chun, Y. (1989), "Lexicographic egalitarian solution and uncertainty in the disagreement point", *Zeitschrift für Operations Research* 33, pp. 259–266.

Coursey, D.L. (1982), "Bilateral bargaining, Pareto optimality, and the empirical frequency of impasse", *Journal of Economic Behavior & Organization* 3, pp. 243–259.

Crawford, V.P. (1982), "A theory of disagreement in bargaining", *Econometrica* 50, pp. 607–638.

Edgeworth, F.Y. (1881), *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*, Kegan Paul, London.

Feltovich, N. (2005), “Critical values for the robust rank–order test”, *Communications in Statistics—Simulation and Computation* 34, pp. 525–547.

Feltovich, N. and J. Swierzbinski (2011), “The role of strategic uncertainty in games: an experimental study of cheap talk and unstructured bargaining in the Nash demand game”, *European Economic Review* 55, pp. 554–574.

Feuille, P. (1975), “Final offer arbitration and the chilling effect”, *Industrial Relations* 14, pp. 302–310.

Forsythe, R., J. Kennan, and B. Sopher (1991), “An experimental analysis of strikes in bargaining games with one–sided private information”, *American Economic Review* 81, pp. 253–278.

Greiner, B. (2004), “An online recruitment system for economic experiments”, in Kremer, K. and V. Macho eds., *Forschung und wissenschaftliches Rechnen*. GWDG Bericht 63. Göttingen: Gesellschaft für Wissenschaftliche Datenverarbeitung, pp. 79–93.

Harsayni, J.C. and R. Selten (1988), *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, MA.

Kalai, E. and M. Smorodinsky (1975), “Other solutions to Nash’s bargaining problem”, *Econometrica* 43, pp. 513–518.

Kuhn, M.A. (2009), “To settle or not to settle: a review of the literature on arbitration in the laboratory”, working paper, University of California, San Diego.

Luce, R.D. and H. Raiffa (1957), *Games and Decisions*, Wiley, New York.

Malouf, M.W.K. and A.E. Roth (1981), “Disagreement in bargaining: an experimental study”, *Journal of Conflict Resolution* 25, pp. 329–348.

Myerson, R.B. (1984), “Two–person bargaining problems with incomplete information”, *Econometrica* 52, pp. 461–488.

Nash, J. (1950), “The bargaining problem”, *Econometrica* 18, pp. 155–162.

Nash, J. (1953), “Two–person cooperative games”, *Econometrica* 21, pp. 128–140.

Nydegger, R.V. and G. Owen (1975), “Two person bargaining: an experimental test of the Nash axioms”, *International Journal of Game theory* 3, pp. 239–249.

Raiffa, H. (1953), “Arbitration schemes for generalized two–person games”, in H.W. Kuhn and A.W. Tucker, eds., *Contributions to the Theory of Games. II: Annals of Mathematics Studies*, Princeton University Press, Princeton, NJ, pp. 361–387.

Rapoport, A. and J.A. Sundali (1996), “Ultimatums in two–person bargaining with one–sided uncertainty: offer games”, *International Journal of Game theory* 25, pp. 475–494.

Rawls, J. (1971), *A Theory of Justice*, Harvard University Press, Cambridge, MA.

Roth, A.E. (1979), *Axiomatic Models of Bargaining*, Springer Verlag.

Roth, A.E. (1995), “Bargaining experiments”, in *Handbook of Experimental Economics*, J.H. Kagel and A.E. Roth, eds., Princeton University Press, Princeton, NJ, pp. 253–348.

Roth, A.E. and M.W.K. Malouf (1979), "Game–theoretic models and the role of information in bargaining," *Psychological Review* 86, pp. 574–594.

Roth, A.E. and K.J. Murnighan (1982), "The role of information in bargaining: an experimental study," *Econometrica* 50, pp. 1123–1142.

Ryan, R. (1999), "Enforcement of foreign arbitration agreements and awards in insurance coverage disputes", *McCullough, Campbell & Lane LLP Publications*, http://www.mcandl.com/arbitration.html, accessed 26.10.2011.

Schelling, T.C. (1960), *The Strategy of Conflict*, Harvard University Press, Cambridge, MA.

Schwarz, G. (1978), "Estimating the dimension of a model", *Annals of Statistics* 6, pp. 461–464.

Siegel, S. and N.J. Castellan, Jr. (1988), *Nonparametric Statistics for the Behavioral Sciences*, McGraw–Hill, New York.

Skyrms, B. (1996), *Evolution of the Social Contract*, Cambridge University Press, Cambridge.

von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.

Wittman, D. (1986), "Final–offer arbitration", *Management Science* 32, pp. 1551–1561.

Zizzo, D.J. (2003), "Money burning and rank egalitarianism with random dictators", *Economics Letters* 81, pp. 263–266.

A Sample instructions and screenshots

Below is the text of instructions from our cell with q = 0 in rounds 1-10 (given out before round 1) and q = 0.8 in rounds 11-40 (given out between rounds 10 and 11), followed by two sample screenshots (the decision screen and the end-of-round feedback screen). The other sets of instructions are the same except for where the value of q is mentioned; these are available from the corresponding author upon request.

Instructions: first part of experiment

You are about to participate in a decision making experiment. Please read these instructions carefully, as the money you earn may depend on how well you understand them. If you have a question at any time, please feel free to ask the experimenter. We ask that you not talk with the other participants during the experiment.

This experiment consists of ***two*** parts. These instructions are for the first part, which will be made up of ***10 rounds***. The second part will be made up of ***30 rounds***; you will receive instructions for the second part after this part has ended. Each round in this part consists of a simple computerised bargaining game. At the beginning of a round, ***you are randomly paired with another participant***, with whom you play the game. You will not be told the identity of the person you are paired with in any round, nor will they be told yours – even after the session ends.

The bargaining game is as follows. You and the person paired with you bargain over a ***£10.00*** prize. You and the other person make simultaneous claims for shares of this prize. The claims must be multiples of £0.01, and cannot be less than zero. ***The maximum allowable claim is different for the two people in a pair: for one, the maximum is £9.50, and for the other, it is £4.50.*** Your maximum allowable claim will be the same in all rounds.

Your profit in a round depends on the claims made by you and the person paired with you:
- If your claims add up to ***the amount of the prize or less***, your profit equals your claim, and the other person's profit equals his/her claim.
- If your claims add up to ***more than*** the amount of the prize, both you and the other person receive a profit of zero.

Sequence of Play: The sequence of play in a round is as follows.
(1) The computer randomly pairs up the participants. Your computer screen will display your maximum allowable claim and that of the other person.
(2) You choose a claim for your share of the £10.00 prize. The other person chooses a claim for his/her share of the prize. Your claim can be any multiple of 0.01, between zero and your maximum allowable claim (inclusive). Both of you choose your claim before being informed of the other's claim.
(3) The round ends. You receive the following information: your own claim, the claim made by the person paired with you, your own profit for the round, the profit of the person paired with you.

After this, you go on to the next round.

Payments: At the end of the experimental session, ***one round*** from this part will be chosen randomly for each participant. You will be paid the total of your profits in this round. In addition, there will be opportunities for payments in the second part of the session. Payments are made privately and in cash at the end of the session.

Instructions: second part of experiment

The procedure in this part of the experiment is similar to that in the first part. You will play a computerised bargaining game for ***30 additional rounds***. The participant paired with you will still be chosen randomly in every round, and the amount you bargain over and your maximum allowable claims will be the same as before.

The difference from the first part of the experiment is that ***you might not receive a zero profit if your claim and the other person's claim add up to more than £10.00***. Specifically, the computer randomly determines what happens in this case.
- There is now a ***40%*** (***8 out of 20***) chance that ***you receive the amount you claimed***, and the person paired with you receives the remainder (£10 minus the amount you claimed).
- There is now a ***40%*** (***8 out of 20***) chance that the person paired with you receives the amount he/she claimed, and ***you receive the remainder*** (£10 minus the amount he/she claimed).
- There is now a ***20%*** (***4 out of 20***) chance that ***both of you receive zero***.

As before, if your claims add up to the amount of the prize or less, you receive your claim, and the other person receives his/her claim.

Payments: At the end of the experimental session, ***three rounds*** from this part will be chosen randomly for each participant. You will be paid the total of your profits in those three rounds. Your earnings from this part of the experiment will be added to your earnings from the previous part.

Decision screen:

Round

2 of 2

Remaining time [sec]: 39

History of your past outcomes:

Round	Amount bargained over (£)	Your claim (£)	Other person's claim (£)	Agreement?	Claim implemented	Your profit (£)	Other person's profit (£)
1	10.00	9.50	4.50	No	Neither	0.00	0.00

This is the beginning of Round 2. You have been randomly matched to another person for this round. You and this person are bargaining over **£10.00**. Your claim must be a multiple of 0.01, between zero and **£9.50** inclusive. The other person's claim must be a multiple of 0.01, between zero and **£4.50** inclusive.

If you reach an agreement (your claims total less than or equal to **£10.00**), you and the other person will each receive the amounts you claimed.

If you do not reach an agreement (your claims total more than £10.00), then there are three possibilities.

There is a **45%** chance that your claim is imposed. In that case, you will receive the amount you claimed, and the other person will receive the remainder (10.00 minus the amount you claimed).

There is a **45%** chance that the other person's claim is imposed. In that case, he/she will receive the amount he/she claimed, and you will receive the remainder (10.00 minus the amount the other person claimed).

There is a **10%** chance that bargaining breaks down. In that case, you will receive **£0.00** and the other person will receive **£0.00**.

Please choose your claim for this round. £

OK

Feedback screen:

Round

2 of 2

History of your past outcomes:

Round	Amount bargained over (£)	Your claim (£)	Other person's claim (£)	Agreement?	Claim implemented	Your profit (£)	Other person's profit (£)
1	10.00	9.50	4.50	No	Neither	0.00	0.00
2	10.00	9.50	4.50	No	Other person's	5.50	4.50

THIS ROUND'S RESULTS:

Your claim was **£9.50.**

The other person's claim was **£4.50.**

Your combined claims were **MORE THAN** the amount you were bargaining over.

The other person's claim was chosen to be implemented.

Your payoff is **£5.50**.

The other person's payoff is **£4.50**.

OK