THE UNIVERSITY *of York*

Discussion Papers in Economics

No. 2006/18

What Price Compromise? Testing a Possibly Surprising Implication of Nash Bargaining Theory

by

John Bone, John D Hey and John Suckling

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

What Price Compromise? Testing a Possibly Surprising Implication of Nash Bargaining Theory

John Bone — Department of Economics, University of York, UK

John D Hey — Department of Economics, University of York, UK and Facolta' Di Economia, LUISS, Roma, Italy

John Suckling — Department of Economics, University of York, UK

August 2006

Abstract

This paper identifies, and tests experimentally, a prediction of Nash Bargaining Theory that may appear counterintuitive. The context is a simple bargaining problem in which two players have to agree a choice from three alternatives. One alternative favours one player and a second favours the other. The third is an apparently reasonable compromise, but is in fact precluded as an agreed choice by the axioms of Nash Bargaining Theory. Experimental results show that agreement on this third alternative occurs rather often. So the axiomatic Nash theory is not well-supported by our evidence. Our subjects' behaviour could be interpreted as the paying of an irrationally (according to the Nash theory) high price in order to reach a compromise agreement.

Keywords: Experiments, Nash Bargaining Theory

JEL codes: C78, C92

Corresponding Author: John Bone
Department of Economics,
University of York,
York, YO10 5DD
UK

jdb1@york.ac.uk

1. INTRODUCTION

This paper reports on an experimental investigation based on the following decision problem. There is a (desirable) prize that may be allocated to either of two individuals, J and K. The allocation is to be decided randomly, by drawing a coloured ball from an opaque bag. The prize goes to J if the drawn ball is yellow, and to K if blue. If the drawn ball is red, however, then they each receive nothing. The problem for J and K is that they have to agree which one of three bags is to be used for the draw, with contents as shown in Table 1. If they fail to do so then by default no draw occurs and they each receive nothing, an outcome we denote as *z*.

If they are both self-interested then J and K have opposing preferences, *ex ante*, over the three bags. Bag C, being the middle-ranked alternative for both partners, represents a compromise. However, this compromise comes at a price, in the form of the red ball. Five or more red balls in Bag C, yellows and blues being unchanged throughout, would be too high a price, in that both partners would be better off, *ex ante*, with either of Bags A and B (and thus, of course, C would no longer be a compromise). We can say that, in that case, agreement on Bag C would be *collectively irrational* for the two partners.[1] But what if, as here, Bag C contains only one red ball? Is this an irrationally high price to pay for the compromise?

It would be so if J and K could (bindingly) agree to toss a coin to decide between Bags A and B, since this would give them each, *ex ante*, a 50% chance of winning the prize. It would similarly be so if they could agree to share the prize afterwards, either by direct division or indirectly through side-payments or further randomisation, since each of Bags A and B delivers with certainty the prize thereby to be shared. But suppose that neither type of

[1] This has to be distinguished from individual rationality. If agreement (and thus the avoidance of *z*) requires unanimous individual assent, then each partner assenting to C is a Nash strategic equilibrium, even with five or more red balls.

agreement is possible, so that the only options for J and K are just as initially described. Might they then reasonably agree to choose the costly compromise in the form of Bag C? Our conjecture was that reasonable, self-interested people probably would do so. However, as we will show, such an agreement would violate the axioms of Nash Bargaining Theory, according to which just one red ball is indeed too high a price to pay.

The following section demonstrates this proposition. In section 3 we describe an experiment based on this decision problem, the results of which are discussed in section 4.

2. COLLECTIVE RATIONALITY IN THE COSTLY COMPROMISE PROBLEM

Define a *Costly Compromise Problem* (CCP) where, for given parameters $\pi \in (\frac{1}{2}, 1)$ and $\gamma \in (1-\pi, \frac{1}{2})$, and a given prospective prize P, individuals J and K have to agree a choice from three alternative "bags" (or equivalent devices) as defined in Table 2. Thus Table 1 represents a CCP with $\pi = 0.75$ and $\gamma = 0.4$. The general definition makes no assumption about the correlation between J winning and K winning. Section 1 describes a CCP in which these events are mutually exclusive. But we can allow for a variation in which for each partner an independent draw, with replacement, is to be made from the agreed bag, with the implied possibility of them each winning P.

Let $\hat{\pi} \equiv \sqrt{\pi(1-\pi)}$, where necessarily $\hat{\pi} \in (1-\pi, \frac{1}{2})$. In Table 1, for example, we have $\hat{\pi} \approx 0.43$. Our central proposition is that if $\gamma < \hat{\pi}$, as in Table 1, then Nash Bargaining Theory rules out Bag C as an agreed choice in the CCP.

We demonstrate this in two ways. The first, and simpler, is in terms of the *Nash Product*. In abstract, a two-person bargaining problem comprises a set of available alternatives X, each of which is weakly preferred by both bargaining partners to some given default outcome *d*. Nash's axioms together require the existence of some $\rho \in (0,1)$ such that the agreed choice $x \in \mathrm{X}$ maximises the Nash Product:

$$\mathrm{N}(x) = [u_j(x)]^{\rho}[u_k(x)]^{1-\rho} \quad (1)$$

where $u_j(x)$ and $u_k(x)$ are the two partners' individual vNM utilities,[2] each normalised to zero at *d*. (The further inclusion of a "symmetry" axiom would require the agreed choice to maximise (1) specifically for $\rho = ½$.)

The CCP is a two-person bargaining problem with $d = z$ and X={A,B,C}, taking as given the prospective prize P and the probability parameters $\{\pi, \gamma\}$. Here vNM utilities may be normalised to:

$$u_j(z) = 0 \qquad u_j(\mathrm{A}) = \pi \qquad u_j(\mathrm{B}) = 1-\pi \qquad u_j(\mathrm{C}) = \gamma$$

$$u_k(z) = 0 \qquad u_k(\mathrm{A}) = 1-\pi \qquad u_k(\mathrm{B}) = \pi \qquad u_k(\mathrm{C}) = \gamma$$

so that, for any given ρ:

$$\mathrm{N(A)} = \pi^{\rho}(1-\pi)^{1-\rho} \qquad \mathrm{N(B)} = (1-\pi)^{\rho}\pi^{1-\rho} \qquad \mathrm{N(C)} = \gamma \quad (2)$$

Evidently $\mathrm{N(C)} \geq \mathrm{N(A)}$ and $\mathrm{N(C)} \geq \mathrm{N(B)}$ together imply that $\gamma^2 \geq \pi(1-\pi)$. Given $\gamma < \hat{\pi}$, therefore, there is no value of ρ for which Bag C maximises the Nash Product. This completes the first demonstration.

Collective rationality, as characterised by the Nash axioms, is implicit in the maximization of (1).[3] The second demonstration makes this more explicit by appealing directly to Nash-type axioms. For any bargaining problem, given the partners and their individual preferences, define the *eligible subset* $\Gamma(\mathrm{X}, d) \subseteq \mathrm{X}$. This is analogous to the set of most-preferred alternatives in the case of individual choice,[4] with the following three axioms

---

[2] Actually the Nash theory does not require that individual utilities are vNM, in the sense of representing preferences under uncertainty, but only that they have (at least) the same degree of cardinality as vNM. In the present context, however, vNM is a natural interpretation.

[3] The seminal reference is Nash (1950). An exposition of the Nash theory can be found in Osborne and Rubinstein (1990, ch.2).

[4] For an individual we would expect the eligible subset not to depend on *d*, the default

thus being interpretable as requirements of rationality:

Non-emptiness $\Gamma(\mathrm{X}, d)$ is non-empty.

Efficiency $x \notin \Gamma(\mathrm{X}, d)$ if there exists some $y \in \mathrm{X}$ strictly preferred to *x* by each partner.

Consistency $x \notin \Gamma(\mathrm{X}, d)$ if there exists some $y \in \mathrm{X}$ and $\mathrm{Y} \supset \mathrm{X}$ such that $x \notin \Gamma(\mathrm{Y}, d)$ and $y \in \Gamma(\mathrm{Y}, d)$.

Now consider a hypothetical variant on the CCP in which J and K have to agree not only a choice of bag but also, at the same time, a choice of prize from Q and R, over which they have opposing preferences. Specifically, given π we can hypothesise prizes Q and R such that:

$$[\mathrm{Q}, \hat{\pi}] \sim_j [\mathrm{R}, \pi] \qquad \text{and} \qquad [\mathrm{R}, \hat{\pi}] \sim_k [\mathrm{Q}, \pi] \tag{3}$$

where $[\mathrm{P}, \varphi]$ denotes the prospect of winning P with probability φ, otherwise winning nothing. If J and K each have preferences conforming to Expected Utility (EU) theory, then (3) implies additionally that:

$$[\mathrm{Q}, 1-\pi] \sim_j [\mathrm{R}, \hat{\pi}] \qquad \text{and} \qquad [\mathrm{R}, 1-\pi] \sim_k [\mathrm{Q}, \hat{\pi}] \tag{4}$$

since the two probabilities in each prospect-pair have a common ratio throughout.

We are here considering, in effect, a two-person bargaining problem with $d = z$ and $\mathrm{X}=\{\mathrm{A_Q}, \mathrm{B_Q}, \mathrm{C_Q}, \mathrm{A_R}, \mathrm{B_R}, \mathrm{C_R}\}$, the subscripts indicating the variable prize. Now, if $\gamma < \hat{\pi}$ then for both partners:

$$[\mathrm{Q}, \hat{\pi}] \succ_{j,k} [\mathrm{Q}, \gamma] \qquad \text{and} \qquad [\mathrm{R}, \hat{\pi}] \succ_{j,k} [\mathrm{R}, \gamma] \tag{5}$$

It follows from (3), (4) and (5) that $\mathrm{C_Q}$ and $\mathrm{C_R}$ are both Pareto-dominated, the former by $\mathrm{A_R}$

---

outcome in the event of failure to choose. However in any well-defined individual choice problem there must always be such a default, even if only tacit.

and the latter by B_Q. Efficiency therefore requires that Bag C is ineligible here, whatever the agreed choice of prize, if $\gamma < \hat{\pi}$.

Next decompose this hypothetical problem into its two constituent CCPs. In each of these J and K have to agree a choice from {A,B,C} with the same $\{\pi, \gamma\}$ values as in the composite problem, but with the prize in each case being fixed, respectively, as Q and R. In the composite problem, given that C_Q and C_R are both ineligible, Non-emptiness requires that at least one of $\{A_Q, B_Q, A_R, B_R\}$ is eligible. Consistency then implies that in the constituent problems either $C_Q \notin \{A_Q, B_Q, C_Q\}$ or $C_R \notin \{A_R, B_R, C_R\}$. Thus, if $\gamma < \hat{\pi}$ then C is ineligible in at least one of these constituent CCPs.

So for any $\{\pi, \gamma\}$ values such that $\gamma < \hat{\pi}$ there exists at least one, albeit hypothetical, CCP in which the agreed choice of Bag C is precluded by the above three axioms of collective rationality. Now add a fourth, more context-specific, axiom:

Prize-Independence In any CCP, the eligibility of each bag depends only on the values of $\{\pi, \gamma\}$.

It then follows that Bag C is ineligible in any CCP with $\gamma < \hat{\pi}$, irrespective of the prize at stake. This completes the second demonstration.

In the next section we describe an experiment designed to test this specific implication of the Nash theory, i.e., that Bag C cannot be the agreed choice in a CCP if $\gamma < \hat{\pi}$. As is evident from the axiomatic argument, in effect this is a joint test of a number of assumptions. It may be thought that such an exercise is superfluous, given the prior existence of adverse experimental evidence on some of these assumptions individually. Most obviously, the assumption of EU preferences, by which we derived (4) from (3), is contradicted by a long record of experimental findings.[5] However, for our purposes this assumption is unnecessarily

[5] Camerer (1995) provides an excellent survey.

strong. The axiomatic argument would still go through if (4) was weakened to:

$$[\mathrm{Q}, 1-\pi] \succsim_j [\mathrm{R}, \hat{\pi}] \qquad \text{and} \qquad [\mathrm{R}, 1-\pi] \succsim_k [\mathrm{Q}, \hat{\pi}]$$

which, in conjunction with (3), would be consistent not only with EU but also with the *common-ratio* violation of EU, regularly observed under experimental conditions, whereby an individual's preference for the riskier prospect (i.e., that with the preferred prize but lower probability of winning) over the safer is inversely related to the overall probability levels.

Similarly, consider the axiom of Prize-Independence. This is implied by the Nash theory, as can be seen in the Nash Products (2) which, like the vNM utility values on which they are based, are independent of the prize at stake. Indeed, these utility values would be the same even if there were different prizes in prospect for each partner. Thus in our definition of a CCP we can allow the more general possibility that $\mathrm{P} = [\mathrm{P}_\mathrm{J}, \mathrm{P}_\mathrm{K}]$ is a pair of prizes, with P_J awarded to J in the event that J wins, and likewise P_K to K. Both demonstrations of the ineligibility of Bag C, given $\gamma < \hat{\pi}$, go through without amendment. However, there is prior evidence to suggest that Prize-Independence would be violated, empirically, for some CCPs of this type. In a series of experiments by Roth and various associates,[6] a monetary prize was allocated by lottery to one of two partners, who had to agree in advance how to divide a given total of lottery tickets between themselves. In treatments where it was common knowledge that the partners faced different prizes, compensating unequal divisions of tickets were regularly agreed, with the effect of equalizing expected monetary values. Equivalent behaviour in a CCP would be for J and K to agree a choice of Bag A given common knowledge of $\mathrm{P} = [\pounds\, y/\pi, \pounds\, y/(1-\pi)]$, but Bag B given $\mathrm{P} = [\pounds\, y/(1-\pi), \pounds\, y/\pi]$, each partner having an expected return of £*y* throughout. The behaviour of Roth's subjects could be

[6] Two key references are Roth and Malouf (1979) and Roth and Murnighan (1982). For full references and an overview, see pp. 40-49 of Roth (1995a)

interpreted in terms of fairness or of focal points. Either way, it casts doubt on the predictive power of the Nash theory, and in particular Prize-Independence.

To avoid our axiomatic argument, and thus our experiment, being vitiated by Roth's findings we can simply restrict our definition of a CCP. One possibility is to require, for a CCP, that if $P_J \neq P_K$ then the prizes are not common knowledge. Another is to require (common knowledge) that $P = [P_J, P_K]$ is envy-free, i.e., that neither partner strictly prefers the other's prize to their own. The form of CCP used in our experiment satisfies the first of these conditions and can be interpreted as also satisfying the second.

Finally, consider our assumption that individuals are self-interested. Not only through the above experiments of Roth *et al*, but also within a large and growing body of subsequent experimental literature, it has become a commonplace that considerations of fairness are essential to the understanding of bargaining agreements. Even in the Ultimatum Game, for example, there is substantial accumulated experimental evidence that proposers do not fully exploit their obvious strategic advantage.[7] One currently favoured account is that in such situations individuals are motivated by other-regarding preferences, rather than solely by self-interest. Various structural forms for such preferences have been posited and tested.[8] Now, the inclusion of other-regarding preferences in the Nash Product (1) might be expected to increase the relative value of compromise alternatives within X. It might further be supposed that, in the CCP, the agreed choice of Bag C could thus be rationalised within the Nash theory after all. However this supposition would be false. Assume that J winning and K winning are mutually exclusive, and that J and K have identical vNM preferences under which the utility value of one's self winning is 1, of the other winning is $\alpha \in [0,1]$, and of no-one winning is 0. Thus:

---

[7] A seminal reference is Guth *et al* (1982). Roth (1995b) provides a survey and discussion.

[8] See for example Cox *et al* (2006).

$$u_j(\mathrm{A}) = \pi + \alpha(1-\pi) \qquad u_j(\mathrm{B}) = 1-\pi+\alpha\pi \qquad u_j(\mathrm{C}) = \gamma(1+\alpha)$$

$$u_k(\mathrm{A}) = 1-\pi+\alpha\pi \qquad u_k(\mathrm{B}) = \pi+\alpha(1-\pi) \qquad u_k(\mathrm{C}) = \gamma(1+\alpha)$$

from which it may be verified that $\mathrm{N(C)} \geq \mathrm{N(A)}$ and $\mathrm{N(C)} \geq \mathrm{N(B)}$ together imply that :

$$\gamma^2 \ \geq \ \frac{[\pi+\alpha(1-\pi)][1-\pi+\alpha\pi]}{(1+\alpha)^2}$$

If $\alpha > 0$ then the right hand side of this condition is strictly greater than $\pi(1-\pi)$, so that for the Nash value to be maximised by Bag C requires a higher value of γ, i.e. fewer red balls, than if $\alpha = 0$. In this sense, therefore, altruism makes the agreed choice of Bag C actually *less* rationalisable within the Nash theory, rather than more so. This result may appear counter-intuitive, but is easily understood by considering the case $\alpha = 1$ where each player cares only that someone, rather than no-one, wins the prize. Here, both J and K are indifferent between Bags A and B, each of which delivers the prize to someone for sure, and they both prefer either of these to Bag C with its attendant possibility of no-one winning.

In summary, therefore, the hypothesis that $\gamma < \hat{\pi}$ makes Bag C ineligible in the CCP follows from a relatively weak version of Nash Bargaining Theory, the testing of which does not appear to have been pre-empted by already existing empirical evidence.

3. THE EXPERIMENT

In designing the experiment a central concern was to prevent collusive agreements of the type described in Section 1. Thus we required our subjects to negotiate anonymously and via computer, and with minimal exploitable information about the value of the prize in prospect, which could differ between partners.

The experiment was conducted in the **EXEC** laboratory at the University of York, the subjects being undergraduate and postgraduate students. There were four experimental treatments, as will shortly be described. For each treatment there were two separate sessions, each lasting around 45 minutes and employing a group of sixteen new subjects seated at individual computer terminals. Apart from oral instructions, pre-recorded and played back to the whole group, the session was carried out in silence with subjects communicating only with or via the computer. The principal off-screen instructions are presented in Appendix A, and can usefully be read now.

In the main part ("Part 2") of the experiment the sixteen subjects were randomly and anonymously paired. Each pair had to agree a choice of bag. This process was repeated in each of three further rounds, with re-matching of partners in such a way as to avoid cross-contamination by previous matches. In each round, the two partners negotiated within a structured protocol of alternating offers. The partner randomly designated as J (i.e., Yellow) opened by proposing one of the three bags, optionally accompanying the proposal with a brief message. K could either accept this proposal, thus ending the negotiations with agreement, or reject it. Rejection would trigger a computerised randomiser to determine whether the process would end at that stage in disagreement or could continue, with K making a counterproposal. Negotiations continued in this way until either a proposal was accepted, or the randomiser ended the process in disagreement. Appendix B shows a representative screenshot, in this case for a proposer about to compose a message to accompany the proposal of Bag A. The randomiser took the form of an onscreen spinning wheel, visible simultaneously to both partners, containing two sectors: green for continuation and red for termination. Immediately prior to being paired-up for the first round, individual subjects were given dummy screens so that they could practice making proposals and responses. In particular they were invited to spin the randomiser wheel as many times as they wanted, the aim being to give them confidence that it was genuine (which it was).

The practical purpose of a continuation probability $\theta < 1$ was to keep negotiations to a manageable length. But for self-interested and rational subjects θ is an important parameter.

A sufficiently low value of θ produces, in effect, an Ultimatum Game. If $1-\pi > \theta\pi$ then K will accept J's proposal of Bag A, since the best that K could achieve otherwise is agreement on Bag B at the next stage, which occurs only with probability θ. So, given common knowledge of self-interest and rationality (CKSR), if $\theta < (1-\pi)/\pi$ then the unique (subgame-perfect Nash) equilibrium is agreement on Bag A.

By contrast, a sufficiently high value of θ produces a negotiation game in which agreement on any of the three Bags is sustainable as an equilibrium, given CKSR. Suppose that K's strategy is always to propose, and to accept only, Bag B. Then J can do no better than likewise to propose or accept B at any stage. It follows in turn that K's strategy of accepting only B, and thus in particular of rejecting C, is rational if $\gamma < \theta\pi$. So given this condition there is an equilibrium in which Bag B is the agreed choice irrespective of who makes the first proposal. But, by symmetric reasoning, this same condition implies a second equilibrium in which both partners always propose/accept Bag A. Strategically this resembles a one-off Chicken Game, having one equilibrium in which J defers to K, and another in which K defers to J. With a sufficiently high θ there is also an equilibrium in which Bag C is the agreed choice irrespective of who makes the first proposal. Suppose that K's strategy is always to propose or accept C and to reject A. Given this, it is rational for J likewise always to propose or accept C. It is additionally rational for J always to reject B if $1-\pi < \theta\gamma$. The two strategies here are symmetric, so this same condition rationalises K's strategy, given J's. Overall, therefore, if $\theta > \max[\gamma/\pi, (1-\pi)/\gamma]$ then for each of the three bags there exists (given CKSR) an equilibrium supporting agreement on that bag, and in each case irrespective of who makes the first proposal. If $\gamma < \hat{\pi}$ then the relevant condition is $\theta > (1-\pi)/\gamma$; if $\gamma > \hat{\pi}$ then it is $\theta > \gamma/\pi$. [9]

[9] This reveals a connection between the Nash and strategic theories of bargaining. Given

As discussed in Section 2, our test of the Nash theory does not assume self-interest. However, in case subjects were self-interested it was important for a clean test that the bargaining protocol did not preclude agreement on any of the three bags as a strategic equilibrium. So our principal treatment had $\gamma < \hat{\pi}$ and $\theta > (1-\pi)/\gamma$, the tested hypothesis being no C-agreements in these circumstances. But this framework also provided the opportunity for a symmetric test of CKSR strategic equilibrium: if $\gamma \geq \hat{\pi}$ then agreement on any of the three bags is consistent with the Nash theory (without the symmetry axiom), but if in addition $\theta < (1-\pi)/\pi$ then the unique CKSR strategic equilibrium is agreement on Bag A.

These considerations suggested a 2×2 structure of treatments, which Table 3 categorises in terms of parameters γ and θ. Treatment 1 is a control treatment; here any agreement is consistent with both the Nash theory and CKSR strategic equilibrium. In Treatment 2 any agreement is likewise consistent with CKSR equilibrium, but C-agreements are precluded by Nash. In Treatment 3 any agreement is consistent with Nash but only A-agreements are consistent with CKSR equilibrium. In Treatment 4 C-agreements are precluded by Nash, and only A-agreements are consistent with CKSR equilibrium. This treatment is theoretically gratuitous, and we must admit to including it mainly for aesthetic completeness. But, as it happened, it did provide some useful additional data.

In implementing these treatments we used bags in which the number of yellow and blue balls was double that shown in Table 1. Thus, as in Table 1, each treatment had $\pi = 0.75$ and $\hat{\pi} \approx 0.43$. This doubling made possible Treatments 1 and 3, for which there was just one red

---

$\gamma < \hat{\pi}$, which in the Nash theory precludes agreement on Bag C, there is no value of θ at which agreement on C is the unique strategic equilibrium, whereas there are values of θ at which there is equilibrium agreement on each of A and B but not on C. For an extensive discussion of the relationship between these two theories of bargaining see Osborne and Rubinstein (1990, ch4).

ball in Bag C, corresponding to $\gamma \approx 0.44$. For Treatments 2 and 4 there were two red balls, thus corresponding to $\gamma = 0.4$ just as in Table 1. The continuation probabilities were $\theta = 0.85$ for Treatments 1 and 2, and $\theta = 0.15$ for Treatments 3 and 4, these values being comfortably within the required constraints indicated in Table 3.

To inhibit *ex post* sharing agreements, negotiation over the agreed choice of bag was carried out with both partners unaware not only of each others' prospective prize but also of their own. We also wanted to make it difficult for partners to estimate expected values for these prizes, including from any prior communication with subjects from previous sessions. To this end we preceded the main part of the experiment with an individual decision problem ("Part 1") in which each of the sixteen subjects privately nominated one of seven virtual boxes, labelled A-G. Each box contained £30 to be divided equally between all those subjects nominating it. The subject's individual dividend from this process then became his (or her) prospective monetary prize in Part 2. There were two practice rounds, after each of which the distribution of nominations across the seven boxes was displayed on all screens, with the corresponding hypothetical dividend being shown individually and privately for each subject. The third round of nominations was for real. Each subject could of course infer his own dividend from the distribution of nominations in this round, and that information was indeed disclosed, simultaneously to all subjects, but not until the end of Part 2 when the negotiations had all been completed. After this, each subject was called separately, in turn, into the adjacent office, where he first drew a numbered ball from a bag to determine which of the four rounds in Part 2 was actually to count and be played out for real. If the indicated round was one in which he had failed to reach agreement with his partner, then he left with nothing. Otherwise he next drew a ball from the bag he had agreed in that round. If this ball matched his designated colour in that round, then he was paid his dividend, the value of which was known only to himself and the experimenter; otherwise he left with nothing.

4. THE RESULTS

For each of the four treatments there were 32 subjects and 64 instances of a CCP (two sessions, each comprising four rounds, each matching eight pairs of subjects). The outcome of each CCP was either agreement on one of the bags {A,B,C} or else disagreement (*z*). By comparison with Table 3, which indicates the theoretical predictions, Figure 1 charts for each treatment the observed frequency of each outcome, aggregating over rounds and sessions.

The strong test of Nash Bargaining Theory is its implication of no C-agreements in Treatment 2. The theory clearly fails this test. There were 46 such agreements, comprising 72% of all outcomes. A weaker test comparing Treatment 2 with Treatment 1 offers no mitigation, there being more, rather than fewer, C-agreements in Treatment 2. Furthermore, these C-agreements were less contentious than those in Treatment 1, in that proportionately more of them (63% compared with 50%) were reached immediately, on the opening proposal by Partner J (Yellow).

At the individual level there is similarly little evidence in support of the Nash theory. Every one of the 32 individual subjects in Treatment 2 was party to at least one C-agreement over the four rounds. And there were very few messages, at any stage, expressing aversion to Bag C. Of the 57 proposals for Bag A or B throughout Treatment 2, 44 were accompanied by messages. Only five of these clearly referred, either directly or indirectly, to the red balls in Bag C. These five messages came from three different subjects, all in Session 2. One was Subject 5 whose pairing in Round 4 with Subject 15 produced the following brief negotiation:

<u>Treatment 2 Session 2 Round 4</u>					<u>Pair 8</u>
5	**A**	no reds which are a waste for us	15	**C**	we wont agree on the other 2 – reds are a fair price to pay for equality !
5	**accept C**				

In this respect there was little difference between the two treatments. Throughout Treatment 1 there were 79 proposals for Bag A or B; 66 of these were accompanied by messages and, of these, only six expressed aversion to Bag C.[10]

Our secondary aim was to investigate the extent of self-interested rational behaviour. The strong test of CKSR strategic equilibrium is its implication of solely A-agreements in Treatment 3. The theory clearly fails this test: there were 12 A-agreements, comprising only 19% of all outcomes. However, even this small number of A-agreements was significantly more than in Treatment 1, as a proportion both of outcomes (chi-squared $p = 0.013$) and of agreements $(p = 0.006)$. So the weaker test provides some support. Besides, in itself the strong test confirms only the absence of CKSR. In Treatment 3 a self-interested rational (SR) subject will (SR1) accept whatever is proposed, and (SR2) propose the best for herself that her partner will accept. If she knows her partner to be SR, and therefore following SR1, then SR2 prescribes proposing her most-preferred bag. But if instead she believes it sufficiently likely that her partner would for some reason reject this, then proposing it would not be rational. So any number of C-agreements, or even B-agreements, here would be consistent with all individual subjects being self-interested and rational, but some of them being doubtful of this fact.

Nevertheless, one outcome here not consistent with strategic rationality (either self-interested or altruistic) is disagreement, since this requires the rejection of a proposal and thus a violation of SR1. There were 16 disagreements in Treatment 3, from which we can infer that not all our subjects were SR. More precisely, an upper bound to the number of SR subjects is given by those who satisfied SR1 throughout, by not rejecting any proposal in any of the four rounds. There were 16 (out of 32) such subjects in Treatment 3, two of whom

[10] A full transcript is available at: http://www-users.york.ac.uk/~jdb1/3johns/mk4tscript.pdf

happened to be designated J (Yellow) in every round and as such received no proposals at all.

It is not clear simply from the outcomes what our non-SR subjects were up to. But the general predominance of C-agreements across all treatments suggests that at least some of them may have been motivated by the pursuit of compromise rather than self-interest. So a natural question is how many subjects systematically behaved in this way, throughout all four rounds of their session. Define as a *Strategic Compromiser* (SC) a subject who always proposes Bag C and always accepts Bag C. This is of course consistent with SR1. A stronger type (SC+) additionally never accepts his or her least-preferred bag (which in this treatment was usually Bag A), thus potentially violating SR1. In Treatment 3 there were eleven subjects meeting the criteria for SC, of whom nine also did so for SC+. Of these SC+ subjects, only one was so by default, never having received a proposal other than for Bag C. So there is evidence of systematic compromise-seeking, even to the individual's own disadvantage, by around one third of our subjects in Treatment 3.

Table 4 extends this analysis to all treatments. It also includes a count of subjects who were instead systematically avoiding compromise, whom we might expect to find especially in Treatments 2 and 4. Specifically, we define as a *Strategic Non-compromiser* (SN) a subject who never proposes Bag C, and a correspondingly stronger type (SN+) who additionally never accepts Bag C. Table 4 also records how many subjects satisfied the criteria for only one of SR, SC and SN. Note in particular that SR implies no specific prescriptions in Treatments 1 and 2, so that no subject can be ruled out as being SR here.

Treatment 4, as Figure 1 shows, was unique in producing more A-agreements than C-agreements. An obvious question is whether this is evidence in favour of the Nash theory, which implies no C-agreements here, or of self-interested strategic rationality which, given common knowledge, implies solely A-agreements. With regard to the latter, a comparison with Treatment 2 is similar to the parallel comparison already made between Treatments 3 and 1. We can confidently (chi-squared $p < 0.001$) reject the hypothesis that the overall four-outcome frequency is drawn from the same population as that of Treatment 2, there being significantly more A-agreements in Treatment 4 as a proportion both of outcomes and of agreements ($p < 0.001$ in either case). To this extent Treatment 4 provides additional weak evidence for the presence of self-interested, rational subjects. Table 4 records that there could be up to twenty such subjects here, although only four of these could only have been SR, from even the limited range of strategy types considered here.

However, albeit less confidently $(p = 0.040)$, we can also reject the hypothesis that the four-outcome frequency in Treatment 4 is drawn from the same population as that in Treatment 3, there being significantly fewer C-agreements in Treatment 4 as a proportion both of outcomes $(p = 0.033)$ and of agreements $(p = 0.005)$. Table 5 provides some additional detail, recording the opening proposals and responses in each round for each of Treatments 3 and 4. The ratio of opening A-proposals to opening C-proposals, over all four rounds together, was substantially (1.6 compared with 0.7) and significantly $(p = 0.034)$ higher in Treatment 4 than in Treatment 3. The difference was particularly marked (3.7 compared with 0.5, and $p = 0.014$) in Round 1; in fact through the four rounds the number of opening A-proposals fell in Treatment 4 and rose in Treatment 3. Over all four rounds, the acceptance rate of A-proposals was higher in Treatment 4 than in Treatment 3 (66% compared with 46%), but not significantly so $(p = 0.127)$.

So the additional red ball appears to have had some independent effect here, and in particular on the behaviour of proposers. One possibility is that subjects were more averse to Bag C as such, in accordance with the Nash theory. As Table 4 records, the number of SN subjects is at its highest in Treatment 4, and is almost double the number in Treatment 3. But of these eighteen SN subjects only five were SN+, and each of these was so by default, in that they never received a proposal of Bag C. In fact only one subject in Treatment 4 actually rejected Bag C at any stage (compared with three such subjects in Treatment 3), and this subject twice accepted Bag C in other rounds.

Table 6 provides some partial additional evidence. It lists all the opening A-proposals and C-proposals in Round 1 of Treatments 3 and 4, with individuals identified by session:subject numbers. Alongside each proposal is shown the accompanying message, if any (three of which have been abbreviated for this table, as indicated by "…"). Also shown is that subject's possible strategy type(s) as classified on the basis of proposals and responses over the whole session. The messages accompanying C-proposals are all rather similar, generally appealing in some way to equality and/or fairness. Those accompanying A-proposals are less uniform. We interpret those in Treatment 3 as pointing out the strategic situation to the receiver. In Treatment 4 there are also a couple of messages of that type (from 1:16 and 2:4) but most messages accompanying A-proposals here appear relatively vacuous, in simply stating the nature of the proposal and/or urging or requesting its acceptance. This pattern is broadly repeated in subsequent rounds, with vacuous messages accompanying A-proposals generally outnumbering strategic ones. But noticeably absent are any messages explicitly expressing aversion to Bag C. Indeed, of the 37 messages throughout Treatment 4 accompanying proposals of Bag A or Bag B, none at all referred to the red balls in Bag C or to the absence of them in Bags A or B.

So it is difficult to draw any firm conclusions about what accounts for the relatively low number of C-Agreements and high number of A-agreements in Treatment 4. But in our view the evidence does not favour the Nash theory. Instead it appears to be more consistent with a simpler explanation – that, in the absence of CKSR, the additional red ball caused SR subjects to be more optimistic, especially at the outset of the session, that their opening A-proposals would be accepted. To that extent Treatment 4 provides additional indication that a considerable number of our subjects were rationally pursuing their own self interest.

5. CONCLUSIONS

Nash Bargaining Theory rests on axioms of collective rationality. We investigated an implication of these axioms that seems to be counterintuitive: bargaining pairs should reject what appears to be a reasonable compromise. In our experimental test we found instead that the majority of bargaining pairs reached agreement on this compromise. So the Nash theory fails our test. We also extended our experimental framework to incorporate a symmetric test of self-interested individual rationality. We found clear evidence that not all individuals are self-interested and strategically rational, and evidence of the systematic pursuit of compromise by some individuals, even to their own disadvantage. However this latter behaviour was by no means universal or even predominant, and there is suggestive evidence of the rational pursuit of self-interest by many, if not most, of our subjects.

REFERENCES

CAMERER, COLIN (1995), "Individual Decision Making", chapter 8 in Kagel and Roth (1995).

COX, JAMES, DANIEL FRIEDMAN and STEVEN GJERSTAD (2006), "A tractable model of reciprocity and fairness", forthcoming in *Games and Economic Behavior*.

GUTH, WERNER, R. SCHMITTBERGER and B. SCHWARZ (1982), "An experimental analysis of ultimatum bargaining," *Journal of Economic Behavior and Organization*, 3, 367-88.

KAGEL, JOHN H. and ALVIN E. ROTH (1995), *Handbook of Experimental Economics*, Princeton University Press.

NASH, JOHN F (1950), "The Bargaining Problem", *Econometrica*, 18, 155-162

OSBORNE, MARTIN J. and ARIEL RUBINSTEIN (1990), *Bargaining and Markets*, San Diego, Academic Press.

ROTH, ALVIN E. (1995a), "Introduction to Experimental Economics", chapter 1 in Kagel and Roth (1995).

______ (1995b), "Bargaining Experiments", chapter 4 in Kagel and Roth (1995).

ROTH, ALVIN E. and MICHAEL W. K. MALOUF (1979), "Game-Theoretic Models and the Role of Information in Bargaining", *Psychological Review*, 86, 574-94.

ROTH, ALVIN E. and J. KEITH MURNIGHAN (1982), "The Role of Information in Bargaining: An Experimental Study", *Econometrica*, 50, 1123-42.

TABLE 1

THE THREE BAGS

	Bag A	Bag B	Bag C
yellow balls	3	1	2
blue balls	1	3	2
red balls	0	0	1

TABLE 2

THE GENERAL COSTLY COMPROMISE PROBLEM

	Bag A	Bag B	Bag C
Prob[J wins]	π	$1-\pi$	γ
Prob[K wins]	$1-\pi$	π	γ

TABLE 3

THE FOUR TREATMENTS IN THE EXPERIMENT

	$\gamma > \hat{\pi}$	$\gamma < \hat{\pi}$
$\theta > \max\left[\frac{\gamma}{\pi}, \frac{1-\pi}{\gamma}\right]$	**Treatment 1** Nash: any bag CKSR: any bag	**Treatment 2** Nash: not Bag C CKSR: any bag
$\theta < \frac{1-\pi}{\pi}$	**Treatment 3** Nash: any bag CKSR: Bag A	**Treatment 4** Nash: not Bag C CKSR: Bag A

TABLE 4

POSSIBLE STRATEGY TYPES (OUT OF 32 SUBJECTS IN EACH TREATMENT)

	consistent with						consistent only with		
	SR	SC	SC+	SN	SN+	none	SR	SC	SN
Treatment 1	32	8	8	8	1	0	16	0	0
Treatment 2	32	9	9	4	0	0	18	0	0
Treatment 3	16	11	9	10	2	5	7	7	3
Treatment 4	20	9	4	18	5	4	4	3	3

TABLE 5

OPENING PROPOSALS AND RESPONSES IN TREATMENTS 3 AND 4

		Treatment 3			**Treatment 4**		
		A	B	C	A	B	C
R1	proposed	5	1	10	11	2	3
	accepted	3	1	9	6	2	3
R2	proposed	6	1	9	10	1	5
	accepted	3	1	8	7	0	5
R3	proposed	6	0	10	8	2	6
	accepted	2	-	9	4	2	5
R4	proposed	9	0	7	6	2	8
	accepted	4	-	6	6	2	8
all	**proposed**	**26**	**2**	**36**	**35**	**7**	**22**
	accepted	**12**	**2**	**32**	**23**	**6**	**21**

TABLE 6

OPENING A-PROPOSALS AND C-PROPOSALS IN ROUND 1 OF TREATMENTS 3 AND 4

Treatment 3

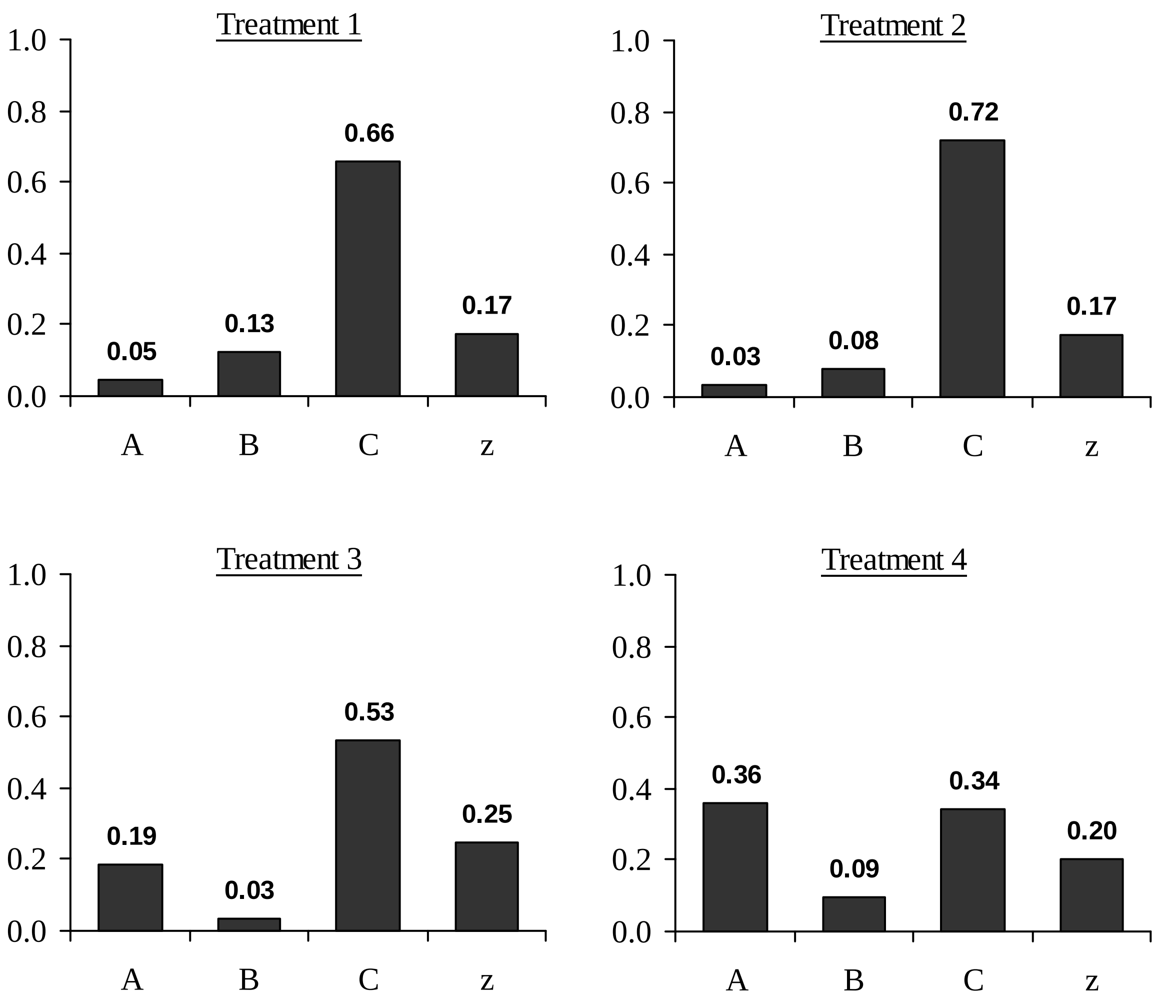
	ID	message	type
A	1:2	-	SN
A	1:8	This is gambling, consider it	SR
A	1:9	next you will be yellow, think!	SR SN
A	1:13	-	SR SN
A	2:14	You know it makes sense!	SR
C	1:6	-	SR SC
C	1:7	just to have equal chance	-
C	1:10	Are you agree?	SR
C	1:11	this gives us more chances to agree	SR
C	2:3	-	SR
C	2:9	… we both have equal chances of winning ...	SR
C	2:11	There is an even chance of either of us winning	SC+
C	2:12	because … we'd have an equal chance	SR SC
C	2:13	-	-
C	2:16	I think we should pick this bag becaus it's fair	SC+

Treatment 4

	ID	message	type
A	1:4	2 more yellow instead of the 2 blue?	SR
A	1:5	I think this is the best chance we get	-
A	1:6	The prob is higher for me	SR SN+
A	1:8	4y 2b- acceptable?	-
A	1:9	looks promising	-
A	1:12	Is bag a ok by you	SR SN
A	1:16	go on, makes sense	SR SN+
A	2:3	-	SR SN+
A	2:4	if it goes again - this will be reject	SR
A	2:5	-	SN
A	2:7	pick bag a	SR SN
C	1:1	this gives us both the fairest chance	SR
C	2:10	This gives us both a fair chance of winning …	SR SC
C	2:12	we've both got a better chance this way	SR

FIGURE 1

OUTCOME FREQUENCIES BY TREATMENT (*n*=64 IN EACH TREATMENT)



APPENDIX A

INSTRUCTIONS

Initial (pre-recorded) oral instructions prior to Part 1

Thank you for participating in this experiment. We hope that you will enjoy it. If you have a mobile phone with you, please check now that it is switched off. [*pause*]

The experiment requires you to make a few simple decisions which, together with a random factor, will determine the amount you are paid at the end of the session.

There are sixteen participants in this session, all facing the same decisions and receiving the same instructions. Beside your terminal you have an envelope, some blank paper, and a pen. Please do not open the envelope until instructed to do so. The pen and paper are provided should you wish to keep a record of your decisions, although it is not necessary to do this. Please leave the pen here at the end of the session.

The session is in two parts. Decisions in Part 1 will determine an amount of money which we will call your dividend. This amount may vary from one individual to another.

However, whether or not you receive your dividend will depend on Part 2, where you will have to agree some decisions with other participants. We will give you further details on this at the start of Part 2.

You will receive instructions both orally, like this, and also on the computer screen. In addition, at all times there will be an information bar at the bottom of the screen. This will remind you what action needs to be taken at that time.

You will have opportunities to ask questions should the instructions not be clear to you. Otherwise, however, you must remain silent throughout the session. At various times you may have to wait for other participants to complete their decisions. If so, please be patient.

Before we proceed to Part 1, are there any questions? [*pause*]

Please click the Start button now. Read the onscreen information and then wait for further instructions. [*pause*]

Your task in Part 1 is simply to choose one of the seven boxes. Each box contains £30, to be shared equally among the participants choosing that box. There will be three rounds. The first two are for practice only, and will not count. But the third round is for real, and will determine your dividend.

There will be no further oral instructions until Part 1 is completed. Are there any questions? [*pause*]

Please make your first practice selection now and then follow onscreen instructions until Part 1 is complete.

(Pre-recorded) oral instructions prior to Part 2

Part 1 is now complete. Your dividend has been computed, but will not be revealed to you until the end of the session.

We will now proceed to Part 2, which consists of four rounds. In each round the computer will pair you, at random, with another participant. It will designate one of you as Yellow and the other as Blue. The pair of you have to agree a decision, which will be explained shortly. You will then be assigned a new partner for the next round, and so on.

Thus, after four rounds, you will have agreed four decisions, each with a different partner. However, only one of these four agreements will actually count for you.

At the end of the session, each participant will be paid individually in private, in the adjoining office. So no other participant will know what payment you receive, unless you yourself choose to reveal it to them afterwards.

Your payment will be determined as follows. Firstly you will draw a number from 1 to 4, from this bag. This will select which of the four rounds in Part 2 is to count for you. Your color, either Yellow or Blue, will be as designated in your selected round. Then you will draw a ball from this bag, which will contain some yellow and blue balls, and possibly some red balls. If you draw your designated color, then you will be paid your dividend. Otherwise you will be paid nothing.

We have not yet told you how many balls of each color will be in your bag. In fact, this is the decision you have to agree with your partner. The contents of your bag will be as agreed by you and your partner in your selected round.

The envelope contains a summary of the information so far. Please open it now and read the summary. [*pause*]

You may consult the summary again at any time during Part 2.

In each round you will communicate with your partner only via the computer. Instructions for doing this will appear on your screen. Are there any questions? [*pause*]

Please click the Continue button now. The next few screens give you further details on Part 2, and enable you to practice communicating with your partner. Please note that for the purpose of these practice screens you will be communicating with yourself, as if you were your own partner.

Please read and follow the instructions, continuing through the practice screens in your own time. [*pause*]

Are there any questions? [*pause*]

Then please begin Part 2 now.

Written summary information, provided prior to Part 2

Part 1

Your dividend is determined. It will be revealed to you after Part 2.

Part 2

Round 1	The computer randomly assigns you a partner, and designates one of you as Yellow and the other as Blue. You and your partner agree the contents of the bag.
Round 2	The computer randomly assigns you a new partner, and designates one of you as Yellow and the other as Blue. You and your partner agree the contents of the bag.
Round 3	The computer randomly assigns you a new partner, and designates one of you as Yellow and the other as Blue. You and your partner agree the contents of the bag.
Round 4	The computer randomly assigns you a new partner, and designates one of you as Yellow and the other as Blue. You and your partner agree the contents of the bag.

Payment

You are paid individually and privately in the office, as follows....

You select one round (1-4) at random. Your color (Yellow or Blue) is as designated in that round, and the contents of your bag are as agreed with your partner in that round.

You draw a ball from your bag. If it is your designated color, then you are paid your dividend. Otherwise you are paid nothing.

APPENDIX B

SCREENSHOT FROM EXPERIMENT