Theory Dec. (2015) 79:463–478
DOI 10.1007/s11238-014-9477-5

The Nash solution is more utilitarian than egalitarian

Shiran Rachmilevitch

Published online: 23 November 2014

Abstract I state and prove formal versions of the claim that the Nash (Econometrica 18: 155–162, 1950) bargaining solution creates a compromise between egalitarianism and utilitarianism, but that this compromise is "biased": the Nash solution puts more emphasis on utilitarianism than it puts on egalitarianism. I also extend the bargaining model by assuming that utility can be transferred between the players at some cost (the transferable and non-transferable utility models are polar cases of this more general one, corresponding to the cases where the transfer cost is zero and infinity, respectively); I use the extended model to better understand the connections between egalitarianism and utilitarianism.

Keywords Bargaining · Egalitarianism · Nash solution · Utilitarianism

1 Introduction

Nash's (1950) *bargaining problem* is described in the utility space: two players face a compact, convex, and comprehensive set $S \subset \mathbb{R}^2_+$ of available utility allocations, from which they need to choose one; if they agree on $x \in S$, then each player i receives the utility payoff x_i, but if they fail to reach an agreement both get zero.[1,2] The *Nash solution* (1950) to the problem S, denoted as $N(S)$, is the maximizer of $x_1 \cdot x_2$ over

[1] It is assumed that $\mathbf{0} \equiv (0, 0) \in S$ for every bargaining problem S, so zero payoffs are always feasible; also, it is assumed that there is an $x \in S$ with $x > \mathbf{0}$, so cooperation is worthwhile (uRv means that $u_i R v_i$ for both i, for each $R \in \{>, \geq\}$; $u \gneq v$ means that $u \geq v$ and $u \neq v$).

[2] Comprehensiveness means that $\{y \in \mathbb{R}^2_+ : y \leq x\} \subset S$ for all $x \in S$.

S. Rachmilevitch (✉)
Department of Economics, University of Haifa, Mount Carmel, 31905 Haifa, Israel
e-mail: shiranrach@gmail.com; shiranrach@econ.haifa.ac.il

$x \in S$. In general, a *bargaining solution* is a *selection*—a function that assigns a unique point of S for every such S.

Two other known bargaining solutions are the *egalitarian solution*, E, and the *utilitarian solution*, U. Like the Nash solution, they are related to certain maximizations. The utilitarian solution selects an agreement under which the sum of the players' utilities is maximized. The egalitarian solution selects the weakly efficient agreement under which both players receive identical payoffs; this agreement maximizes $\min\{x_1, x_2\}$ over $x \in S$.[3,4] In this paper, I study various connections between the Nash solution and the aforementioned ones. As we will see, there are several non-trivial connections.

It is known that the Nash solution exhibits attractive fairness properties. For example, Brock (1979) showed that this solution is actually egalitarian, but not in the sense of recommending equal payoffs, but in the sense of recommending payoff ratios that are equal to the marginal rate of utility substitution between the players. An additional fairness property of the Nash solution is that it guarantees for each player at least one half of his maximum possible utility. This *midpoint domination* property was originally formulated by Sobel (1981); when combined with other standard conditions, it characterizes the Nash solution (Moulin 1983; de Clippel 2007). Another fairness-related result is by Mariotti (1999), who characterized the Nash solution on the basis of *Suppes-Sen dominance,*[5] a criterion that requires that if x is the selected agreement and (a, b) is some other feasible agreement, then both $(a, b) > x$ and $(b, a) > x$ are false. This requirement captures both fairness and efficiency restrictions, since it stems from a combination of impartiality and the Pareto principle: if, for example, $(b, a) > x$ for some feasible utility pair (a, b), then since x is Pareto inferior to (b, a) and impartiality implies that (b, a) and (a, b) are indistinguishable from an ethical point of view, x should not be selected.[6] Further works on the Nash solution in the context of fairness and distributive justice include Binmore (1989, 1991, 2005), Sacconi (2010) and Trockel (2005).

In addition to exhibiting these fairness properties, the Nash solution also creates the following compromise between egalitarianism and utilitarianism (Fig. 1):

Proposition 1 *For any bargaining problem S, the point $N(S)$ lies on S's Pareto boundary in between $E(S)$ and $U(S)$.*[7]

Hereafter, I will refer to the bounds that are described in Proposition 1 as the *E-U bounds*.

Proposition 1 brings about the following questions:

[3] E was axiomatized for the first time by Kalai (1977). U, in general, is multi-valued; in this paper, I will only consider problems for which it is single-valued.

[4] For problems S whose Pareto frontier is strictly concave, $U(S)$ is the unique maximizer of the utility sum over S and $E(S)$ is the unique maximizer of $\min\{x_1, x_2\}$ over $x \in S$ (in the following Section, I formally introduce an important class of such problems—*smooth* bargaining problems). Compromising on precision just a tiny bit, I will sometimes refer to $\sum_i x_i$ and $\min\{x_1, x_2\}$ as the utilitarian and egalitarian objectives, respectively.

[5] Suppes (1966), Sen (1970).

[6] Related results have been obtained by Anbarci and Sun (2011). See also Mariotti (2000).

[7] Later in this paper, I will prove two generalizations of Proposition 1—Propositions 6 and 7 below. Hence, for brevity, a proof of Proposition 1 is not provided.

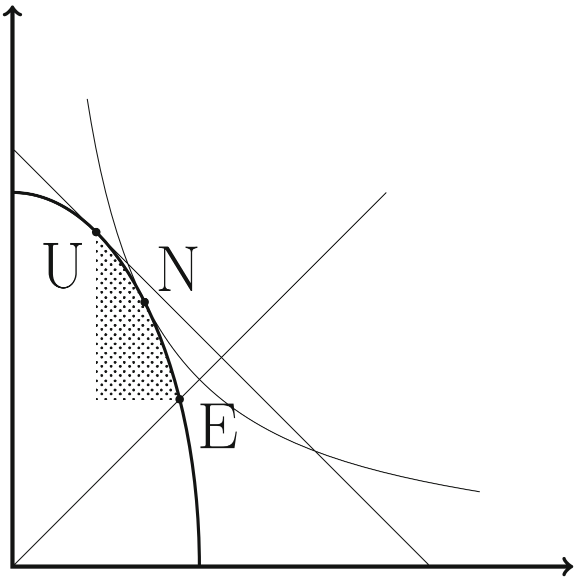

Fig. 1 Illustration of Proposition 1

- (I) Which side of the compromise between egalitarianism and utilitarianism, if any, is "favored" by the Nash solution?
- (II) What is the meaning of "being between E and U"?

In light of the extensive work emphasizing the fairness aspects of the Nash solution, one might suspect that, in the abovementioned compromise, it is egalitarianism that gets the upper hand. I suggest otherwise—I will highlight the utilitarian aspects of the Nash solution and show that, in a certain sense, it is "more utilitarian than egalitarian."

This sense takes on the following form. First, Proposition 2 shows that whenever N coincides with E, this common point is also utilitarian. Hence, the Nash solution is utilitarian "more frequently" than it is egalitarian. Proposition 3 gives sufficient conditions for the converse implication. Next is Proposition 4, which shows that for a normalized problem S,[8] the Nash solution point for S is not only between $U(S)$ and $E(S)$, but, moreover, is between $U(S)$ and $A(S)$, where $A(S)$ is the maximizer of $\frac{1}{2}(x_1 + x_2) + \frac{1}{2}\min\{x_1, x_2\}$ over $x \in S$—that is, the maximizer of the average of the utilitarian and egalitarian objectives.

Proposition 5 is a generalization of Proposition 4: it shows that for any problem S, normalized or not, the Nash solution point for S lies between $RU(S)$ and $NA(S)$, where RU is the *relative utilitarian solution* (Dhilon and Mertens 1999, Pivato 2009, Segal 2000 and Sobel 2001) and NA, the "normalized average" solution, is defined as follows: for every S, the point $NA(S)$ maximizes $\frac{1}{2}\left(\frac{x_1}{a_1(S)} + \frac{x_2}{a_2(S)}\right) + \frac{1}{2}\min\left\{\frac{x_1}{a_1(S)}, \frac{x_2}{a_2(S)}\right\}$ over $x \in S$, where $a_i(S) \equiv \max\{x_i : x \in S\}$.

More than thirty years ago, Cao (1982) presented a result that is similar to Proposition 1: he proved that for every problem S, the point $N(S)$ lies between the points

[8] A bargaining problem is normalized if for each player the minimum and maximum utilities are 0 and 1.

that are selected by the *Kalai-Smorodinsky solution* (Kalai and Smorodinsky 1975) and the relative utilitarian solution.[9] On the class of normalized problems, his result and Proposition 1 amount to the same statement.[10] In addition, Proposition 5 tightens Cao's (1982) bounds.

Next is Proposition 6 which shows that for any $\rho \in [-\infty, 1]$, the corresponding *constant elasticity solution* (CES) for S—the maximizer of $[x_1^\rho + x_2^\rho]^{1/\rho}$ over $x \in S$—lies between $U(S)$ and $E(S)$; the fact that the Nash solution adheres to these bounds is a particular instance of this more general result, since the Nash solution corresponds to the limit $\rho \to 0$. Therefore, Proposition 1 is a corollary of Proposition 6.[11]

An additional generalization of Proposition 1 is obtained by considering the weighted versions of N, U, and E. I prove, in Proposition 7, that an appropriately weighted Nash solution lies "between" the weighted utilitarian and egalitarian solutions. Proposition 1 is a corollary of Proposition 7.

In Section 5, I turn to question (II) from above: namely, to the meaning of "falling between E and U." I prove that it is equivalent to egalitarianism if utility can be transferred between the players at some cost, which is convex in the amount of the transfer. The models of transferable and non-transferable utility are particular instances of this more general model—they correspond to the cases where the cost function is the identity function and where it is identically infinity, respectively.

The remainder of the paper is organized as follows. In Sect. 2, I prove results that formalize the idea that the Nash solution is "at least as utilitarian as it is egalitarian." They are formal versions of some well-known facts about the geometry of the Nash solution, but, to the best of my knowledge, they have not been previously published. Section 3 is dedicated to the various bounds. In Sect. 4, I consider the non-symmetric generalization of Proposition 1. In Sect. 5, I prove that "falling between E and U" is equivalent to egalitarianism if utility can be transferred between the players at some (convex in the amount of the transfer) cost. Section 6 contains a discussion about the interpretation of the results, and Sect. 7 provides a conclusion.

The analysis in this paper is only for 2-person bargaining. The reconciliation of utilitarianism and egalitarianism with more than two players is beyond the scope of the present paper.

2 Egalitarianism implies utilitarianism, but not vice versa

Let $\mathcal{B}$ denote the collection of bargaining problems S such that $U(S)$ is single-valued.

Proposition 2 *Let $S \in \mathcal{B}$ be such that $N(S) = E(S)$. Then $N(S) = U(S)$.*

Proof Let S be a problem and assume by contradiction that $N(S) = E(S) = (e, e) \neq U(S) = (x, y)$. Note that $x + y > 2e$. For each $\lambda \in [0, 1]$, consider $\lambda(e, e) + (1 -$

[9] Cao refers to the relative utilitarian solution as the *modified Thomson solution* and to the Kalai-Smorodinsky as the *Raiffa solution*. I will introduce these solutions formally in Sect. 3.

[10] Both results are related to the fact that the Nash solution is the only solution that jointly satisfies the egalitarian and utilitarian objectives for some rescaling of the individual utilities (Harsanyi 1959, Shapley 1969).

[11] CES solutions have been studied by Sobel (2001), Bertsimas et al. (2012), and Haake and Qin (2013).

$\lambda)(x, y) = (\lambda e + (1-\lambda)x, \lambda e + (1-\lambda)y)$, and let g denote the associated Nash product:

$$g(\lambda) \equiv \lambda^2 e^2 + \lambda(1-\lambda)e(x+y) + (1-\lambda)^2 xy.$$

Note that $g'(\lambda)|_{\lambda=1} \geq 0$.[12] However, $g'(\lambda)|_{\lambda=1} = 2e^2 - e(x+y)$, so $g'(\lambda)|_{\lambda=1} \geq 0$ implies $2e \geq x + y > 2e$, a contradiction. □

The "converse" of Proposition 1 is not true: the fact that N coincides with U on a certain problem does not imply that it also coincides with E on that problem. To see this, look at $S^* = \text{conv}\{\mathbf{0}, (1,0), (0,1), \left(\frac{3}{4}, \frac{3}{4}\right), (1, 0.7)\}$, for which we have $N(S^*) = U(S^*) = (1, 0.7) \neq (\frac{3}{4}, \frac{3}{4}) = E(S^*)$.

There is, however, some inelegance in S^*: its boundary has a "kink," meaning that the "price" of one person's utility in terms of his partner's utility exhibits discontinuities. When this is ruled out, a "converse" of Proposition 1 obtains. Formally, let $\mathcal{B}^*$ be the collection of *smooth problems*—those $S \in \mathcal{B}$ whose frontier assumes the form $\{(x, f(x)) : x \in [0, M]\}$, where $M > 0$ is some number and f is some strictly concave, strictly decreasing differentiable function, with $f(0) > 0$ and $f(M) = 0$.

Proposition 3 *Let $S \in \mathcal{B}^*$ be such that $N(S) = U(S)$. Then $N(S) = E(S)$.*

Proof Let $S \in \mathcal{B}^*$ be such that $N(S) = U(S)$. Let f be the function describing S's boundary. The point $U(S) \equiv (x, f(x))$ satisfies the first-order condition $-f'(x) = 1$. By the first-order condition associated with N, $-f'(x) = \frac{f(x)}{x}$. Therefore $N(S) = E(S)$. □

So, even though the Nash solution's utilitarianism is generally more common (or more frequent) than its egalitarianism, on the restricted domain $\mathcal{B}^*$ the two coincide. The normative content of E and U relies on the assumption that utilities are interpersonally comparable (otherwise, for example, their summation has no normative meaning). There is, additionally, an alternative sense in which the Nash solution is more utilitarian than egalitarian; this sense, which is described in the following Section, requires not only that the comparison of utilities be meaningful, but also that these utilities be measured on the same scale.

3 Tighter bounds

One of the simplest ways to create a compromise between egalitarianism and utilitarianism is to consider the average of their respective objectives: that is, to maximize $\frac{1}{2}(x_1 + x_2) + \frac{1}{2}\min\{x_1, x_2\}$ over $x \in S$. Let A denote this bargaining solution.[13]

[12] Locally, the value of the Nash product does not decrease, as one gets closer to the Nash solution point.

[13] Alvarez-Cuadrado and van Long (2009) consider a maximization of a convex combination of utilitarian and egalitarian objectives in the context of intergenerational equity (their objectives are defined on infinite utility streams).

Since $\frac{1}{2}(x_1 + x_2) + \frac{1}{2}\min\{x_1, x_2\}$ is continuous, it has a maximum on every bargaining problem. However, since the level curves ("indifference curves") of this expression are linear on both sides of the 45° line, the aforementioned maximum may be obtained at multiple points. Let $\tilde{\mathcal{B}} \equiv \{S \in \mathcal{B} : A(S) \text{ is single-valued}\}$. Obviously, $\mathcal{B}^* \subset \tilde{\mathcal{B}}$. When attention is restricted to *normalized problems* in $\tilde{\mathcal{B}}$—$S \in \tilde{\mathcal{B}}$ such that $a_i(S) \equiv \max\{x_i : x \in S\} = 1$ for each i—the Nash solution is not only between E and U (this is true on the entire $\mathcal{B}$), but, moreover, it is also between A and U—it is "closer to utilitarianism than to egalitarianism." I will prove this result for the normalized problems in $\mathcal{B}^*$, but it will be easy to see that it holds for any normalized problem in $\tilde{\mathcal{B}}$; i.e., the boundary does not have to be smooth.

The following lemma will be useful.

Lemma 1 *Let $S \in \mathcal{B}^*$ be a normalized problem and let $(x, y) = A(S)$. Then $x \in (0, 1)$.*

Proof Make the aforementioned assumptions, and assume by contradiction that $x \in \{0, 1\}$. Let f be the function describing S's boundary and let $(e, e) \equiv E(S)$.

Case 1 $x = 1$. In this case, A maximizes $R(t) \equiv \frac{1}{2}[t + f(t)] + \frac{1}{2}f(t) = \frac{1}{2}t + f(t)$ over $(e, 1]$. Since $x = 1$, $R'(1) \geq 0$, or $\frac{1}{2} \geq -f'(1)$, which is impossible for a normalized problem.

Case 2 $x = 0$. In this case, A maximizes $L(t) \equiv \frac{1}{2}[t + f(t)] + \frac{1}{2}t = t + \frac{1}{2}f(t)$ over $[0, e)$. Since $x = 0$, $L'(0) \leq 0$, or $1 + \frac{1}{2}f'(0) \leq 0$, or $2 \leq -f'(0)$; again, this is impossible for a normalized problem. □

Proposition 4 *Let $S \in \mathcal{B}^*$ be a normalized problem. Then the following is true for each $i \in \{1, 2\}$:*

$$min\{A_i(S), U_i(S)\} \leq N_i(S) \leq max\{A_i(S), U_i(S)\}.$$

That is, the Nash solution, N, is "between" A and U.

Proof Let $S \in \mathcal{B}^*$ be a normalized problem. Let f be the function describing its boundary, let $(x, f(x)) \equiv A(S)$ and let $(e, e) \equiv E(S)$. If $(x, f(x)) = E(S)$, then we are done, since N is always between E and U. Suppose then that $(x, f(x)) \neq E(S)$.

Case 1 $x > f(x)$. In this case, A maximizes the objective function $R(t) \equiv \frac{1}{2}(t + f(t)) + \frac{1}{2}f(t) = \frac{1}{2}t + f(t)$ over $t \in (e, 1]$. $R'(t) = \frac{1}{2} + f'(t)$. By Lemma 1, the solution is interior, and therefore $R'(x) = 0$, or $f'(x) = -\frac{1}{2}$; so $A(S)$ is to the left of $U(S)$. If $N(S) \equiv (n, f(n))$ is to the left of $A(S)$, then $-f'(n) < \frac{1}{2}$. By the first-order condition associated with N, $-f'(n) = \frac{f(n)}{n}$, and so we would obtain $n > 2f(n)$; this, however, is impossible in a normalized problem, because $f(n) \geq \frac{1}{2}$.[14]

Case 2 $x < f(x)$. In this case, A maximizes the objective function $L(t) \equiv \frac{1}{2}(t + f(t)) + \frac{1}{2}t = t + \frac{1}{2}f(t)$ over $[0, e)$. $L'(t) = 1 + \frac{1}{2}f'(t)$. Here $L'(x) = 0$, or

[14] This is due to midpoint domination (Sobel 1981).

$f'(x) = -2$, so $A(S)$ is to the right of $U(S)$. If $N(S)$ is to the right of $A(S)$, then $\frac{f(n)}{n} > 2$, which contradicts midpoint domination. □

As mentioned in the text above, it is not important for Proposition 4 that S has a smooth boundary. It is important, however, that each player has the same maximum utility in S. Otherwise, it may be the case that the utilitarian objective becomes overwhelmingly more important than the egalitarian one, and, consequently, the maximizations of A and U coincide; in such a case, N clearly fails the betweenness condition. For example, in the problem $S = \text{conv}\{\mathbf{0}, (1, 0), (0, k)\}$ with $k > 2$, it is the case that $U(S) = A(S) = (0, k) \neq \left(\frac{1}{2}, \frac{k}{2}\right) = N(S)$.

However, with suitable normalizations, an analog (in fact, a generalization) of Proposition 4 is obtained for $\tilde{\mathcal{B}}$. In order to state it, some additional definitions are needed. Recall that the *relative utilitarian solution*, RU, is defined as follows: for each problem S, it selects the point that maximizes the (normalized) sum $\frac{x_1}{a_1(S)} + \frac{x_2}{a_2(S)}$ over $x \in S$.[15] "Relative egalitarianism" is analogous to "relative utilitarianism," and is expressed by the *Kalai-Smorodinsky solution* (Kalai and Smorodinsky 1975), which assigns to every S the point $\lambda a(S) \equiv KS(S)$, where λ is the maximum possible. Similar to the egalitarian solution, the Kalai-Smorodinsky solution maximizes the minimum of the normalized utilities, namely $\min\left\{\frac{x_1}{a_1(S)}, \frac{x_2}{a_2(S)}\right\}$. On the basis of the normalized versions of utilitarianism and egalitarianism, we can define, analogously to the definition of A, the solution NA as the one that picks for every S the maximizer of $\frac{1}{2}\left(\frac{x_1}{a_1(S)} + \frac{x_2}{a_2(S)}\right) + \frac{1}{2}\min\left\{\frac{x_1}{a_1(S)}, \frac{x_2}{a_2(S)}\right\}$ over $x \in S$.

Like RU and KS, the solution NA can be operationalized in three steps: first, the bargaining problem, say S, is normalized, which means that player i's utility is multiplied by $\frac{1}{a_i(S)}$; next, the relevant solution (A in the case of NA, U/E in the cases of RU/KS) is applied to the normalized problem, resulting in the choice of a point of the normalized problem, call it v; finally, the solution point of the original problem is $(v_1 a_1(S), v_2 a_2(S))$. The solutions NA, RU, and KS, as well as N, are *scale invariant*.[16] I will utilize this fact in the proof of the following result.

Proposition 5 *Let $S \in \mathcal{B}^*$. Then the following is true for each $i \in \{1, 2\}$:*

$$\min\{NA_i(S), RU_i(S)\} \leq N_i(S) \leq \max\{NA_i(S), RU_i(S)\}.$$

That is, the Nash solution, N, is "between" NA and RU. In other words, the Nash solution is "closer to relative utilitarianism than to relative egalitarianism."

Proof Let $S \in \mathcal{B}^*$. Let $V \equiv l \circ S$, where $l_i(t) \equiv \frac{t}{a_i(S)}$. By Proposition 4,

$$\min\{A_i(V), U_i(V)\} \leq N_i(V) \leq \max\{A_i(V), U_i(V)\}. \tag{1}$$

[15] Like U, the solution RU is also, in principle, multi-valued. For simplicity, I assume that the problems under consideration in this paper are such that it is single-valued (this is the case, for example, on the domain $\mathcal{B}^*$).

[16] A solution, μ, is scale invariant if $\mu(l \circ S) = l \circ \mu(S)$ for every S and every pair of positive linear transformations $l = (l_1, l_2)$. A positive linear transformation is also called a *rescaling*.

Applying l to (1) we get

$$\min\{NA_i(S), RU_i(S)\} \leq N_i(S) \leq \max\{NA_i(S), RU_i(S)\},$$

as desired. □

I end this Section by showing that the E-U bounds apply not only to the Nash solution, but also to any CES solution. Given $\rho \in [-\infty, 1]$, denote the corresponding such solution by μ^ρ; namely, $\mu^\rho(S) \equiv \text{argmax}_{x\in S}[x_1^\rho + x_2^\rho]^{1/\rho}$.

Proposition 6 *Let $\rho \in [-\infty, 1]$ and $S \in \mathcal{B}$. Then the following is true for each $i \in \{1, 2\}$:*

$$\min\{E_i(S), U_i(S)\} \leq \mu_i^\rho(S) \leq \max\{E_i(S), U_i(S)\}.$$

That is, any CES solution is "between" E and U.

Proof Let $\rho \in (-\infty, 1]$ and $S \in \mathcal{B}$ (there is nothing to prove if $\rho = -\infty$, as in this case the CES solution is E). Wlog, suppose that $U(S)$ is weakly to the left of $E(S)$.[17] Also, since the three solutions in question are *continuous*,[18] we can assume that $S \in \mathcal{B}^*$. The first-order condition associated with μ^ρ is

$$\left(\frac{f(x)}{x}\right)^{1-\rho} = -f'(x).$$

If this point is to the left of $U(S)$, then $-f'(x) < 1$, which implies $f(x) < x$, in contradiction to the fact that $(x, f(x))$ is to the left of $E(S)$. If, on the other hand, $(x, f(x))$ is to the right of $E(S)$, then $f(x) < x$, which implies that—$f'(x) < 1$, which implies that $(x, f(x))$ is to the left of $U(S)$—again, a contradiction. □

4 Asymmetry

Given the weights $(p, 1-p) > \mathbf{0}$, the weighted utilitarian solution maximizes the sum $px_1 + (1-p)x_2$ over $x \in S$, the weighted egalitarian solution is given by $(pe, (1-p)e)$, where e is the maximal number such that the latter expression belongs to S, and the weighted Nash solution maximizes the product $x_1^p x_2^{(1-p)}$ over $x \in S$. I will denote these solutions by U^p, E^p, and N^p, respectively.

I restrict my attention to $\mathcal{B}^*$; this guarantees that $U^p(S)$ is single-valued, for any $p \in (0, 1)$. Let $\theta \equiv \frac{p}{1-p}$. Note that $E^p(S)$ takes the form $(\theta y, y)$, N^p maximizes the product $x_1^\theta x_2$ over $x \in S$, and $U^p(S)$ maximizes $\theta x_1 + x_2$ over $x \in S$.

[17] This assumption is wlog, since each $\mu \in \{E, U, \mu^\rho\}$ is an *anonymous* solution; a solution μ is *anonymous* if for each S it is true that $\pi \circ \mu(S) = \mu(\pi \circ S)$, where $\pi(a, b) \equiv (b, a)$.

[18] A solution μ is *continuous* if $\mu(S_n) \to \mu(S)$, provided that $\{S_n\}$ converges to S in the Hausdorff topology.

Consider the following function $h\colon (0, 1) \to (0, 1)$:

$$h(p) \equiv \frac{p^2}{2p^2 - 2p + 1}.$$

A solution that makes its selection from "in between" $U^p(S)$ and $E^p(S)$, for every S, will be called *p-EU robust*.

Proposition 7 *Let $p \in (0, 1)$. The solution $N^{h(p)}$ is the unique scale invariant solution that is p-EU robust.*

In the proof of the proposition, I will make use of the following lemma.

Lemma 2 *Let $p \in (0, 1)$ and $S \in \mathcal{B}^*$. Then there is a rescaling of S, $T = \lambda \circ S$, such that $E^p(T) = U^p(T)$.*

Proof Let $p \in (0, 1)$ and $\theta = \frac{p}{1-p}$. Let $S \in \mathcal{B}^*$ and let f be the function, defined on $[0, M]$, which describes its boundary. Since both U^p and E^p are *homogeneous*—$\mu(cS) = c\mu(S)$ for every S, $c > 0$, and $\mu \in \{U^p, E^p\}$—it suffices to consider rescaling of one player's utility. Wlog, I will consider rescaling of player 2's utility by $\lambda > 0$. With $E^p(T) = (a, \lambda f(a))$ for some $a \in [0, M]$, the required equalities are $\theta\lambda f(a) = a$ and $\lambda f'(a) = -\theta$. That is, it is sufficient to find an $a \in [0, A]$ such that $\frac{a}{f(a)\theta} = \frac{-\theta}{f'(a)}$, or $\psi(a) \equiv \frac{-af'(a)}{f(a)} = \theta^2$. There exists a unique such a because the function ψ is strictly increasing and satisfies $\psi(0) = 0$ and $\psi(a) \to \infty$ as $a \to M$. □

Proof of Proposition 7 Let $p \in (0, 1)$. It is well-known that every weighted Nash solution satisfies scale invariance; hence, $N^{h(p)}$ satisfies it; let us verify that it is also p-EU robust. Let $S \in \mathcal{B}^*$. Let f be the smooth function describing S's boundary. Let $\theta = \frac{p}{1-p}$ and $\beta = \frac{h(p)}{1-h(p)}$. Note that $\beta = \theta^2$. Assume by contradiction that $N^{h(p)}(S)$ is not in between $U^p(S)$ and $E^p(S)$. Note that there are two possibilities: either $U_1^p(S) \leq E_1^p(S)$ or $U_1^p(S) > E_1^p(S)$. Suppose that $U_1^p(S) \leq E_1^p(S)$. There are two further possibilities: $N_1^{h(p)}(S) < U_1^p(S)$ or $N_1^{h(p)}(S) > E_1^p(S)$. Consider first the case $N_1^{h(p)}(S) < U_1^p(S)$. Letting $a \equiv N_1^{h(p)}(S)$, we see that the tangency condition associated with $N^{h(p)}$ is $\beta\frac{f(a)}{a} = -f'(a)$. Since $N_1^{h(p)}(S) < U_1^p(S)$, $-f'(a) < \theta$; therefore, $\beta\frac{f(a)}{a} < \theta$. Combining this with $\beta = \theta^2$, we obtain $\frac{f(a)}{a} < \frac{1}{\theta} = \frac{1-p}{p}$, in contradiction to $N_1^{h(p)}(S) < E_1^p(S)$. Next, consider $N_1^{h(p)}(S) > E_1^p(S)$. Here we have $\frac{f(a)}{a} < \frac{1}{\theta}$, and therefore $-f'(a) = \beta\frac{f(a)}{a} < \beta\frac{1}{\theta} = \theta$, in contradiction to $N_1^{h(p)}(S) > U_1^p(S)$. The possibility $U_1^p(S) > E_1^p(S)$ is ruled out on the basis of similar arguments.

I now turn to uniqueness. Let μ be a scale invariant p-EU robust solution. Let $S \in \mathcal{B}^*$. By Lemma 2, S can be rescaled such that the resulting problem, call it T, satisfies $E^p(T) = U^p(T) = N^{h(p)}(T)$.[19] By scale invariance, $\mu(S) = N^{h(p)}(S)$. □

[19] The last equality here is due to the fact that we just proved that $N^{h(p)}$ is p-EU robust.

Note that Proposition 1 is a corollary of Proposition 7 because $h\left(\frac{1}{2}\right) = \frac{1}{2}$. It is interesting to note that $\left(p - \frac{1}{2}\right)(h(p) - p) > 0$ for all $p \neq \frac{1}{2}$: in bridging the gap between (weighted) utilitarianism and egalitarianism, the stronger player is favored in terms of having augmented weight in the Nash product.

5 Costly transferable utility

Recall that when utility is 1-to-1 transferable between the players, utilitarianism coincides with Pareto efficiency. However, since 1-to-1 utility transfers are not always feasible, it would be fruitful to examine a richer set of circumstances—when transfers are possible, but only in a constrained, costly manner.

Consider, then, the following scenario: we fix a point on the Pareto frontier, say (x, y), and then, without changing the *bargaining outcome* (i.e., in addition to, or "on top of" the bargaining resolution (x, y)) we seek to increase player i's utility by t units; typically, this will entail the need to decrease j's utility. Denote the minimal required decrease by $c^{ij}(t)$. Thus, c^{ij} is a *cost function*, which specifies the cost of transferring utility from j to i. The case of 1-to-1 transferable utility corresponds to $c^{12}(t) = c^{21}(t) \equiv t$, and the case of non-transferable utility corresponds to $c^{12} = c^{21} \equiv \infty$.

The case $t < c^{ij}(t) < \infty$ has not received (to the best of my knowledge) treatment within the bargaining framework, even though it is particularly relevant from an applied point of view. For instance, think of a university administrator who needs to allocate a single job vacancy between two departments, i and j. If he gives it to i, then it may be the case that he will be able to subsequently implement policies that are more favorable to department j than to department i. This, in turn, can be viewed as an implicit transfer from i to j. Constructing such transfers, however, is not always easy, and certainly not free. These transfers are not 1-to-1.

From here on, I will assume that the players are symmetric in regards to the costs of utility transfers, namely that $c^{12} = c^{21} \equiv c$. I will consider cost functions c that are differentiable, convex, and satisfy $c(0) = 0$ and $c'(0) \geq 1$. The last condition says that it is easiest to make transfers in the transferable utility world. Let $\mathcal{C}$ be the set of these cost functions. For convenience, I will consider the domain $\mathcal{B}^*$ from here on.

Given S, consider the following two-step process: (i) choosing a point $(x, y) \in S$, and (ii) transferring utility between the players, where the cost of transfers is given by some $c \in \mathcal{C}$. Now consider egalitarianism in this setting: namely, the task of generating, via the aforementioned two-step process, a point in the utility space of the form (e, e), on which Pareto improvements are impossible. Call this the **general egalitarian problem**. Assuming that it is solved via some non-zero transfer between the players, there are two possibilities: a positive transfer from player 1 to player 2, or a positive transfer from player 2 to player 1. Letting f denote the boundary function of S, we see that the optimization problems corresponding to the two cases are as follows.

Case 1: Transfer from 1 to 2: Given that we start from the point $(x, f(x))$, where $x > f(x)$, and we increase player 2's utility by t, the post-transfer utilities are $(x -$

$c(t)$, $f(x)+t)$. The optimal t maximizes $f(x)+t$ subject to the constraint $x-c(t) \geq f(x)+t$. It is easy to check that there is a unique $t = t(x)$ that solves this problem. The optimization problem that corresponds to Case 1 is therefore as follows:

$$\max_{x:x>f(x)} \qquad f(x)+t(x). \tag{2}$$

Since this optimization is on an open set, a maximum may not exist. In this case—abusing terminology a little—the value of the problem, call it $V_{12}(S)$, is declared to be zero. Therefore,

$$V_{12}(S) \equiv \begin{cases} f(x^*)+t(x^*) & \text{if (2) has a solution and } x^* \text{is such a solution} \\ 0 & \text{otherwise} \end{cases}$$

Case 2: Transfer from 2 to 1: Given that we start from the point $(x, f(x))$, where $x < f(x)$, and we increase player 1's utility by t, the post-transfer utilities are $(x+t, f(x)-c(t))$. The optimal t maximizes $x+t$ subject to the constraint $x+t \leq f(x)-c(t)$. Similar to Case 1, it is easy to check that there is a unique solution, $\tilde{t} = \tilde{t}(x)$. The optimization problem that corresponds to Case 2, therefore, is as follows:

$$\max_{x:x<f(x)} \qquad x+\tilde{t}(x). \tag{3}$$

Similar to Case 1, define

$$V_{21}(S) \equiv \begin{cases} f(x^*)+\tilde{t}(x^*) & \text{if (3) has a solution and } x^* \text{is such a solution} \\ 0 & \text{otherwise} \end{cases}$$

The value corresponding to the problem S is

$$V(S) = \max\{V_{12}(S), e(S), V_{21}(S)\},$$

where $(e(S), e(S)) = E(S)$. Thus, $V(S)$ is the value of the general egalitarian problem. In order to highlight its solution's dependence on the particular cost function, denote the solution to the first stage of the two-stage process—namely, the point in S on which we improve by transfers—by $GE(S|c)$.

Proposition 8 *Let $S \in \mathcal{B}^*$ and $c \in \mathcal{C}$. Then the following is true for each $i \in \{1, 2\}$:*

$$min\{E_i(S), U_i(S)\} \leq GE_i(S|c) \leq max\{E_i(S), U_i(S)\}.$$

That is, any general egalitarian point is "between" E and U.

Proof Assume by contradiction that there are such S and c for which this assertion is false. Let f be the boundary function of S. Wlog, suppose that $U(S)$ is to the left of $E(S)$. There are two possibilities: the general egalitarian solution for this problem involves a transfer from 1 to 2 or from 2 to 1.

Case 1 The transfer is from 1 to 2.[20] In this case, the value of the objective is $f(x)+t$, where $x = GE_1(S|c)$ and t is the transfer. I will now show that if we start from $e \equiv E(S)$ instead of from x, we can find another transfer that keeps player 2 just indifferent relatively to the original situation and makes player 1 strongly better off. Let s be the transfer which is uniquely defined by

$$f(e) + s = f(x) + t. \tag{4}$$

I argue that $e - c(s) > x - c(t)$. To see this, assume by contradiction that

$$e - c(s) \leq x - c(t). \tag{5}$$

Adding up equations (4) and (5), we obtain $e + f(e) + s - c(s) \leq x + f(x) + t - c(t)$. Since our assumption implies $e + f(e) > x + f(x)$, it follows that $s - c(s) < t - c(t)$. Since the function $\psi(a) \equiv a - c(a)$ is decreasing, $s > t$. Since $s > t$, (4) implies that $f(e) < f(x)$, which implies $x < e$—a contradiction.

Case 2 The transfer is from player 2 to player 1, and so $x = GE_1(S|c) < U_1(S)$. As in Case 1, I will construct a transfer, s, that makes 2 indifferent but strictly improves 1's welfare. As opposed to Case 1, where our starting point for constructing the Pareto improvement was $E(S) = (e, f(e))$, here it will be $U(S)$; let $(y, f(y)) = U(S)$. The condition that defines s is $f(y) - c(s) = f(x) - c(t)$, and we would like to prove that $y + s > x + t$. Assume by contradiction that $y + s \leq x + t$. Adding this equation to $y - c(s) = f(x) - c(t)$, we obtain $y + f(y) + s - c(s) \leq f(x) + x + t - c(t)$, and therefore $s - c(s) \leq t - c(t)$, which implies $s > t$. Since $y + s \leq x + t$, it follows that $y < x$—a contradiction. □

The following is the "converse" of Proposition 8; together, the two propositions establish an equivalence relation between "$E - U$ betweenness" and general egalitarianism.

Proposition 9 *Let $S \in \mathcal{B}^*$ and let x be a point in S's Pareto boundary such that the following is true for each $i \in \{1, 2\}$:*

$$min\{E_i(S), U_i(S)\} \leq x_i \leq max\{E_i(S), U_i(S)\}.$$

Then there is a $c \in \mathcal{C}$ such that $x = GE(S|c)$.

Proof Let S be a smooth problem, such that $U(S)$ is to the left of $E(S)$ and let $(x, f(x))$ be some point in between them. In the general egalitarian solution that corresponds to this case, the egalitarian objective assumes the value $x+t(x)$. The associated first-order condition is

$$1 + t'(x) = 0. \tag{6}$$

[20] This means that the solution point is to the right of $E(S)$.

Since $t + c(t) - f(x) + x = 0$, it follows from the Implicit Function Theorem that $\frac{\partial t}{\partial x} = -\frac{1-f'}{1+c'}$. Combining this with (6), we obtain

$$c'(t(x)) = -f'(x). \tag{7}$$

I will now construct a function $c \in \mathcal{C}$ that satisfies (6) and

$$t + c(t) = f(x) - x \equiv \Delta. \tag{8}$$

Given $a > 0$, let $c_a(t) = t + at^2$. Note that $c_a \in \mathcal{C}$ for all $a > 0$, and hence it is enough to prove that there is an $a > 0$ such that c_a satisfies (7) and (8). Equation (8) becomes $2t + at^2 = \Delta$, which is solved by $t = -1 + \sqrt{1 + a\Delta}$. Then (7) becomes $1+2a[-1+\sqrt{1 + a\Delta}] = -f'(x)$. Since x is to the right of $U(S)$, at $a = 0$, we obtain the strict inequality $1 < -f'(x)$. However, the LHS diverges to infinity as $a \to \infty$, and hence there is an a such that the equation is satisfied. □

6 Discussion

Two types of results have been presented in the paper: in Sect. 2–4, the (geometrical) relationships among the three solutions $\{E, U, N\}$ are analyzed, and in Sect. 5, the idea of costly utility transfers is studied. I will now discuss these two parts separately.

The results of Sect. 2–4 show, plain and simple, that N is "between" E and U. This fact can be understood in light of the philosophical interpretations of these solutions, as follows. The Nash and the egalitarian solutions are *contractarian*: the former is a solution proposed by a theory of fair bargains, and the latter expresses the Rawlsian view of how collective choices should be made in a hypothetical *original position*—a thought experiment which is contractarian in its very nature (Rawls (1971)). In contrast, the utilitarian solution rests upon the following, non-contractarian foundation, which is due to Harsanyi.[21] The Harsanyian view describes an out-of-society standpoint, from which an impartial but sympathetic (to the members of society) observer seeks a desirable social alternative. Thus, in contrast to N and E, both of which corresponding to philosophies that concern the *agents' point of view*, the philosophy associated with U takes a stand that is external to society.

One can therefore interpret the results of Sect. 2–4 to imply that a contractarian who takes N to be the right bargaining solution can reassure a utilitarian that it is OK to embrace the former's theory: despite being built on different philosophical premises, the two views often lead to the same social choice. Moreover, in case of disagreement between U and N, the "distance" between them is bounded; in particular, it can be no greater (in a certain formal sense) than the "distance" between N and E. This is far from trivial because "qualified egalitarianism" is built into the Nash solution (in the form of Nash's symmetry axiom), whereas there is nothing utilitarian in the axiomatic foundations of this solution.[22]

[21] See Fleurbaey et al. (2008) for illuminating discussions on the subject.

[22] I am grateful to a thorough referee for offering this interpretation.

That N is "more utilitarian than egalitarian" has alternative formal manifestations. For example, consider the maximization of the weighted average $\mu^\alpha(x) \equiv \alpha\left[\frac{x_1}{a_1(S)} + \frac{x_2}{a_2(S)}\right] + (1-\alpha)\min\left\{\frac{x_1}{a_1(S)}, \frac{x_2}{a_2(S)}\right\}$ over $x \in S$, where $\alpha \in [0, 1]$ is an arbitrary weight.[23] Given a problem S, and a point on its boundary which lies between $KS(S)$ and $RU(S)$, say x, we can ask the following question: for what α (if any) does the point x maximize μ^α over S? It can be shown that for every S, the α that corresponds to $N(S)$ belongs to $\left[\frac{1}{2}, 1\right]$. Conversely, for any $\alpha \in \left[\frac{1}{2}, 1\right]$, there is a problem S such that $N(S)$ maximizes μ^α over S.[24]

Let us now turn to the analysis of Sect. 5. In that Section, I introduced the general egalitarian problem. The main upshot of the analysis in that Section is that it rationalizes (under the assumption of costly utility transfers) the E-U bounds. The shortcoming of this analysis is that it does not assign any particular importance or significance to the Nash solution—it is only one of many solutions adhering to these bounds. In the remainder of the current Section, I will discuss some feature that are common (albeit indirectly) to the aforementioned analysis and a classic work by Shapley (1969). This discussion's goal is to clarify the meaning of the costly transfers idea, and to indicate possible directions for future research.

Given a problem S, Shapley assumed that there are some weights, $(p, 1-p)$, that describe the "correct" rates of exchange between the individual utilities. These weights are allowed to depend on the problem, so $p = p(S)$. Given a point on S's frontier, x, if we could operationalize utility transfers in accordance with these correct weights, then we would face a TU game, namely $H(x, p) \equiv \{u \in \mathbb{R}^2_+ : pu_1 + (1-p)u_2 \leq px_1 + (1-p)x_2\}$. Now, suppose that we have a solution concept for TU games, and suppose, moreover, that this solution assigns to any such game H the midpoint of its frontier, call it $m(H)$ (as was mentioned in the proof of Proposition 4, the Nash solution satisfies this property; for TU games, this property is also satisfied by many other well-known solutions; see Kalai and Kalai (2013)). Could we apply our TU-solution to the original (NTU) game S? We certainly could, if we were fortunate enough to discover that $m(H(x, p(S))) \in S$. Shapley proposed a solution method for bargaining problems that guarantees that $m(H(x, p(S))) \in S$.

Specifically, he proved that for every problem S, there is a $p = p(S)$ such that $m(H(x^*, p)) \in S$, where the point x^* is egalitarian with respect to p, which means that $(1-p)x_1^* = px_2^*$. Call this approach *endogenous weighted egalitarianism*. The cherry on Shapley's sundae is that x^* is the Nash solution point for S.

Both Shapley's work and the analysis from Sect. 5 involve a form of egalitarianism as an ethical guideline for the resolution of the bargaining problem—endogenous weighted egalitarianism in the former, and general egalitarianism in the latter. In both cases, utility transfers play a role in a somewhat non-standard way: in the analysis from Sect. 5, transfers take place, but in a non-linear way; the transfers in Shapley's work are linear, but they are not carried out.

The major difference between the two works is that Shapley's pins down a unique solution, whereas the costly transfers model only pins down a range from which

[23] $\alpha = \frac{1}{2}$ corresponds to the solution NA.

[24] For the sake of brevity, I omit the proof. It is available upon request.

the solution makes its selection. Whether and how the costly transfers model can be enriched so as to narrow this range remains a task for future research.

7 Conclusion

I have studied in detail the geometric relationships among three important bargaining solutions: the Nash solution N, the utilitarian solution U, and the egalitarian solution E. I have described a formal sense in which N creates a compromise between U and E; this compromise is "biased," in the sense that N puts more emphasis on utilitarianism than on egalitarianism.

I have also introduced the idea of costly utility transfers and applied it to the bargaining model. It was shown that in the model with costly transfers, the central property from the aforementioned geometric analysis—"falling between E and U"—is equivalent to general egalitarianism.

Acknowledgments Insightful and informative reports from several anonymous referees are gratefully acknowledged. The comments of the participants in the T.S. Kim Memorial Seminar at Seoul National University have also contributed significantly to the paper; I am grateful to the seminar participants, and, in particular, to Youngsub Chun and Biung-Ghi Ju.

References

Alvarez-Cuadrado, F., & Van Long, N. (2009). A mixed Bentham-Rawls criterion for intergenerational equity: theory and implications. *Journal of Environmental Economics and Management*, *58*, 154–168.

Anbarci, N., & Sun, C. J. (2011). Distributive justice and the Nash bargaining solution. *Social Choice and Welfare*, *37*, 453–470.

Bertsimas, D., Farias, V., & Trichakis, N. (2012). On the efficiency-fairness trade-off. *Management Science*, *58*, 2234–2250.

Binmore, K. (1989). Social contract I: Harsanyi and Rawls. *The Economic Journal*, *99*, 84–102.

Binmore, K. (1991). Game theory and the social contract. In R. Selten (Ed.), *Game equilibrium models II, methods, morals, markets*. Berlin: Springer Verlag.

Binmore, K. (2005). *Natural justice*. Oxford: Oxford University Press.

Brock, H. W. (1979). *A game theoretic account of social justice*., Game theory, social choice and ethics Netherlands: Springer, pp 239–265.

Cao, X. (1982). Performance functions and bargaining solution. In *Proceedings of the 21st IEEE Conference on Decision and Control*, Orlando, Dec 1982.

de Clippel, G. (2007). An axiomatization of the Nash bargaining solution. *Social Choice and Welfare*, *29*, 201–210.

Dhilon, A., & Mertens, J.-F. (1999). Relative utilitarianism. *Econometrica*, *67*, 471–498.

Fleurbaey, M., Salles, M., & Weymark, J. A. (Eds.). (2008). *Justice, political liberalism, and utilitarianism: themes from Harsanyi and Rawls*. Cambridge: Cambridge University Press.

Haake, C-J, & Qin, C-Z. (2013). A unification of solutions to the bargaining problem, working paper.

Harsanyi, J. C. (1959). A bargaining model for the cooperative n-person games. In A. W. Tucker & R. D. Luce (Eds.), *Contributions to the theory of Games IV* (pp. 325–355). Princeton: Princeton University Press.

Kalai, E. (1977). Proportional solutions to bargaining situations: interpersonal utility comparisons. *Econometrica*, *45*, 1623–1630.

Kalai, A., & Kalai, E. (2013). Cooperation in two person games, revisited. *Quarterly Journal of Economics*, *128*, 917–966.

Kalai, E., & Smorodinsky, M. (1975). Other solutions to Nash's bargaining problem. *Econometrica*, *43*, 513–518.

Mariotti, M. (1999). Fair bargains: distributive justice and Nash bargaining theory. *Review of Economic Studies*, *66*, 733–741.
Mariotti, M. (2000). An ethical interpretation of the Nash choice rule. *Theory and Decision*, *49*, 151–157.
Moulin, H. (1983). Le choix social utilitariste. Ecole Polytechnique Discussion Paper.
Nash, J. F. (1950). The bargaining problem. *Econometrica*, *18*, 155–162.
Pivato, M. (2009). Twofold optimality of the relative utilitarian bargaining solution. *Social Choice and Welfare*, *32*, 79–92.
Rawls, J. (1971). *A theory of justice, Belknap*. Cambridge: Harvard University Press.
Sacconi, L. (2010). A Rawlsian view of CSR and the game theory of its implementation (Part II): fairness and equilibrium. In L. Sacconi, M. Blair, R. E. Freeman, & A. Vercelli (Eds.), *Corporate social responsibility and corporate governance: the contribution of economic theory and related disciplines*. Basingstoke: Palgrave Macmillan.
Segal, U. (2000). Let's agree that all dictatorships are equally bad. *Journal of Political Economy*, *108*, 569–589.
Sen, A. (1970). *Collective choice and social welfare*. San Francisco: Holden-Day.
Shapley, L. S. (1969). Utility comparison and the theory of games. *La Décision: Agrégation et Dynamique des Ordres de Préférence* (pp. 251–263). Paris: Editions du CNRS.
Sobel, J. (1981). Distortion of utilities and the bargaining problem. *Econometrica*, *49*, 597–619.
Sobel, J. (2001). Manipulation of preferences and relative utilitarianism. *Games and Economic Behavior*, *37*, 196–215.
Suppes, P. (1966). Some formal models of grading principles. *Synthese*, *6*, 284–306.
Trockel, W. (2005). In what sense is the Nash solution fair? In H. Schmidt & S. Traub (Eds.), *Advances in public economics: Utility, choice, and welfare*. Heidelberg: Springer.