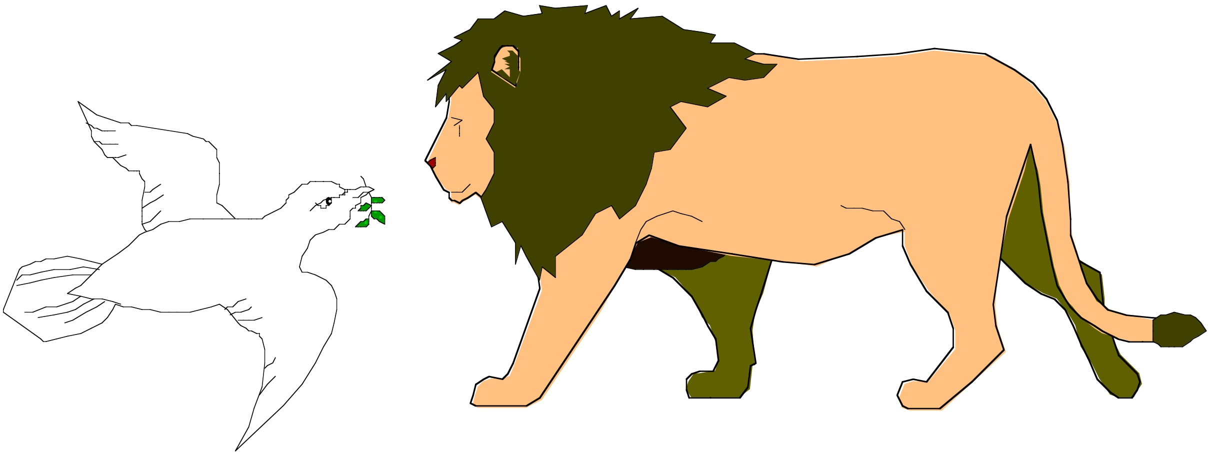
Why do lions get the lion’s share? A Hobbesian theory of agreements

Joan Esteban
Institut d'Anàlisi Econòmica, CSIC

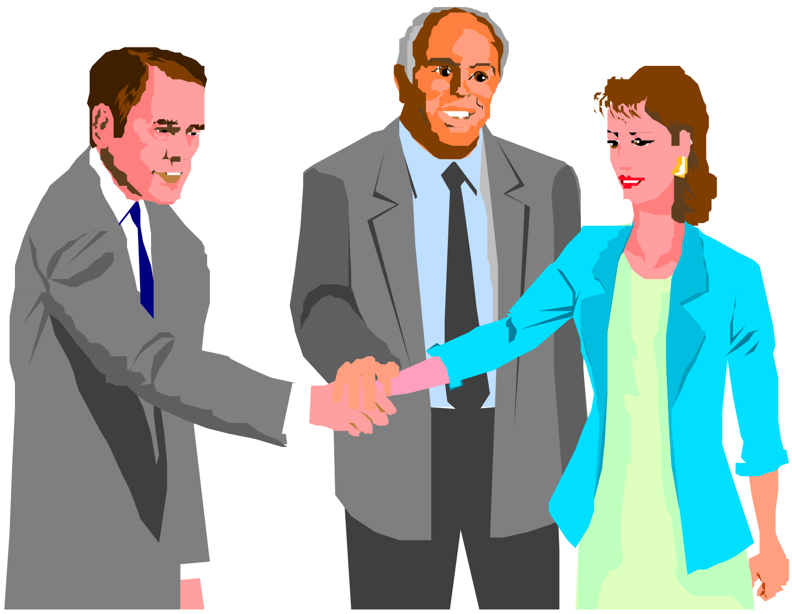
József Sákovics
University of Edinburgh and Institut d'Anàlisi Econòmica, CSIC

October 11, 1999

Why do lions get the lion’s share?



A Hobbesian theory of agreements



A disagreement theory of bargaining

Let Σ denote the set of compact subsets of the utility space, $\Re_+^N$.

There are N players who wish to reach an agreement in $S^0 \in \Sigma$.

<u>*Disagreement function*,</u> *D(.):* assigns a disagreement point, $d \in \Re_+^N$, to every $S \in \Sigma$. .

If the set of alternatives considered were *S*, the outcome of disagreement would be $d = D(S)$.

This mapping is to be interpreted as shorthand for the solution to an underlying *disagreement game*.

A <u>bargaining problem in the shadow of conflict</u> (BPSC) is completely described by the pair $(S^0, D(.))$.

Three assumptions on *D(.):*

Assumption 1 *D is continuous in the Hausdorff topology: if a sequence of elements of Σ converges to S in the Hausdorff topology then the corresponding sequence of disagreement points converges to D(S).*

Assumption 2 *There exists at least one agreement which weakly Pareto dominates the disagreement outcome: for all $S \in \Sigma$, there exists $z \in S$ such that $D(S) \leq z$.*

Assumption 3 *Unless S is a singleton, the disagreement outcome is strictly preferred to her worst possible agreement by at least one player: for all $S \in \Sigma$, there exists $z \in S$ such that $z_i < D_i(S)$ for at least some i, i=1,...,N.*

Let $\mathcal{B}$ denote the set of all BPSCs, with $S^0 \in \Sigma$ and $D(.)$ satisfying Assumptions 1 and 2.

A bargaining solution for BPSCs is a mapping $f\colon \mathcal{B} \to \Sigma$, satisfying $f(S^0,D(.)) \subseteq S^0$.

That is, the solution selects a subset of the alternatives as acceptable.

The Hobbes solution

Let $S_X = \{s \in S \mid s \geq x\}$.

That is, S_X is the subset of S which weakly Pareto dominates x.

Note that if $(S^0, D(.)) \in B$, then $(S^0_{D(S^0)}, D(.)) \in B$.

AXIOM:

Independence of Individually Irrational Alternatives

$$f(S,D(.)) = f(S_{D(S)},D(.)) \text{ for all } (S,D(.)) \in \mathcal{B}.$$

Any solution has to be (weakly) individually rational with respect to the disagreement point. If we eliminate from S all points that have no chance of becoming an agreement, the solution should be the same.

Compare with Independence of Irrelevant Alternatives.

Definition 1 A bargaining solution to BSCPs yields the Hobbesian set of agreements if and only if it satisfies IIIA.

We denote this Hobbesian set of agreements by $H(S^0,D(.))$.

Proposition 1 *Let Assumptions 1, 2 and 3 hold. Then, the Hobbesian (set of) agreement(s) is a singleton and it is Pareto efficient.*

Note that efficiency is a derived, not assumed, property of the solution.

<u>DISCUSSION</u>

The role of assumptions on *D*:

1. <u>*D(S) is a point*</u>: This implicitly assumes that the underlying non-cooperative game has a unique equilibrium. Then we can redefine the IIIA axiom in terms of the minimum utilities over all equilibria.

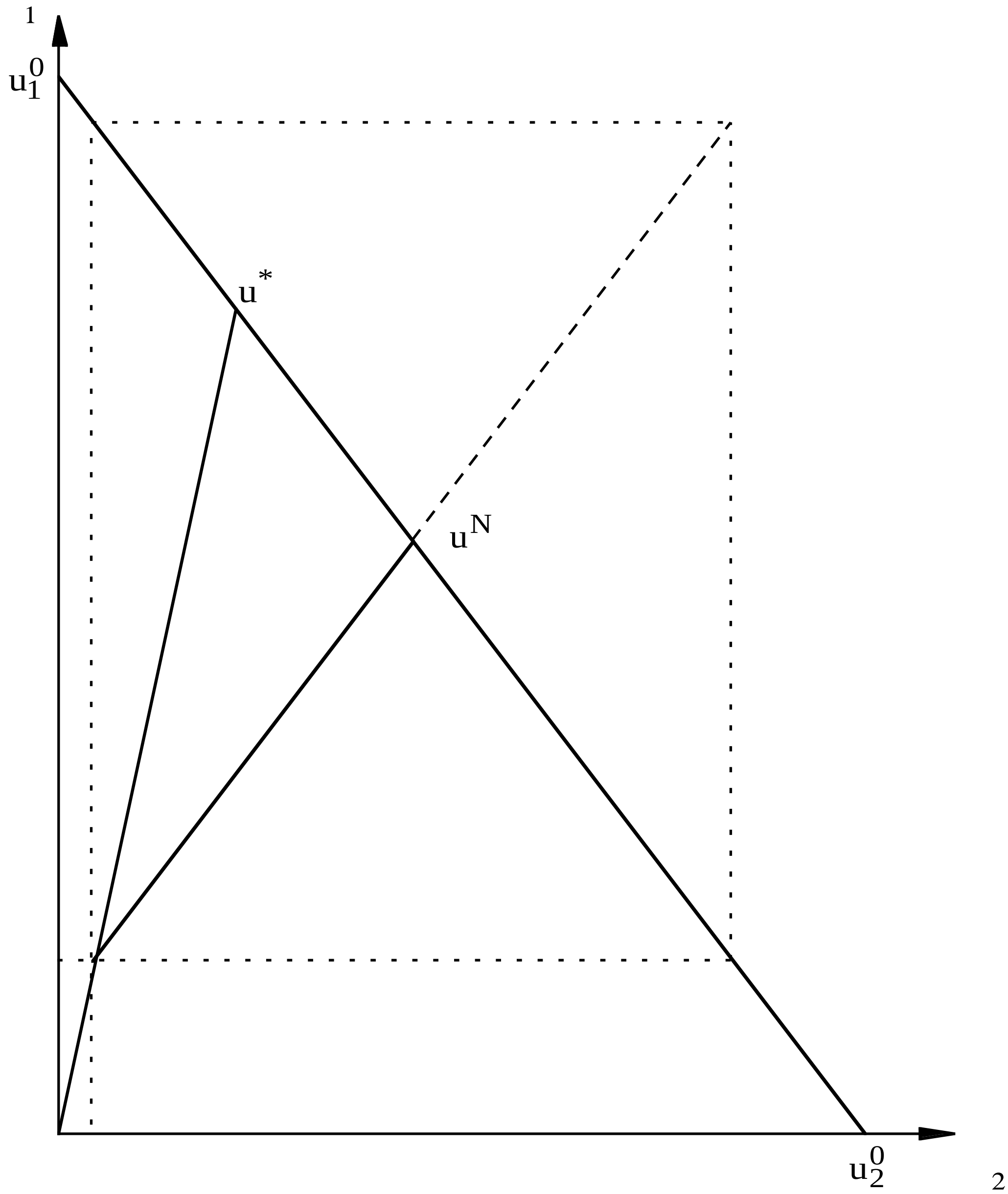
2. <u>There exists *z* ∈*S* such that *D(S)*≤*z*.</u> This assumption just plays the role of making agreements socially desirable.

3. <u>*D(S) is strictly preferred to the worst point in S by at least one player*</u>. This assumption is critical to obtaining that the solution is a singleton.

The role of the IIIA axiom: Describing the bargaining problem with a disagreement function allows a recursive application of the axiom.

The Hobbes solution compared with Nash's: example

1. Suppose that the Pareto frontier of *S* is linear.
2. Assume *D(a+bS)=a+bD(S)*.



Go to the motivating example and see how the difference varies with *d*.

REMARKS

A. The Hobbes solution yields a cooperative agreement that is fully driven by the underlying non-cooperative game that motivated the need for it.

B. The agreement predicated by the Hobbes solution critically depends on the specific disagreement function corresponding to each particular problem. This theory can be termed “positive” because it substantially reduces the “normative” content.

C. The Hobbes solution can be conceived as the limit of a process in which players try to narrow down the area of dissent. With complete information we obtain that players will always agree to fully eliminate dissent. With asymmetric information it might not be possible to completely eliminate disagreement.

Hobbes' as an asymmetric Nash solution

Recall that the asymmetric Nash solution (see Harsányi and Selten, 1972) results from the constrained maximization of a social welfare function where the individual welfare weights are supposed to embody the differential (bargaining) power of the players: $W(x,d)=\prod_{i=1}^{N}(x_i-d_i)^{\gamma_i}$. We shall now discuss the relationship between the vector γ and the power of the parties as embodied in the disagreement function.

Since the Hobbes solution selects a unique point on the Pareto frontier, it can obviously be interpreted as an asymmetric Nash solution. As it is well known, this solution can be characterized as the point on the Pareto frontier, where the pair-wise elasticity of this frontier is equal to the corresponding ratio of the bargaining weights. In general, one needs to calculate the Hobbes solution, in order to derive the associated bargaining weights. However, restricting attention to a relevant subset of disagreement functions, these weights can be directly given.

Let us, then make the simplifying assumption that the disagreement function satisfies *proportionality*, i.e. $D(\lambda S) = \lambda D(S)$ for all $\lambda > 0$.[1] Note that this scenario is equivalent in "richness" to the one analyzed by Rubinstein (1982), in the sense that in both models at each step of the process, the pie remaining in dispute decreases at some given

[1] This condition is somewhat weaker than homogeneity. It is easy to show that the endogenous contest model of Esteban and Ray (1999) mentioned earlier, satisfies this assumption whenever the Pareto frontier of the bargaining set is linear.

proportion.[2] As our next proposition shows, in this setting the Hobbes solution is very simple and intuitive: the players distribute utilities (efficiently) in the same proportion as the disagreement function does. Consequently, the pair-wise ratio of bargaining weights corresponds to the elasticity of the Pareto frontier at the point where the utilities are distributed in the same proportion as in the disagreement point.

Proposition 3 *Let the disagreement function be proportional. Then, the Hobbes solution satisfies* $\frac{f_i^H(S,D)}{D_i(S)} = \frac{f_j^H(S,D)}{D_j(S)}$ *for i, j = 1,2,…, N.*

Proof: Recalling the proof of Proposition 1, we only need to prove that $\frac{d_i^t}{d_j^t} = \frac{d_i^{t+1}}{d_j^{t+1}}$. Note that $d^{t+1} = d^t + D(S^t)$, by definition. Therefore,

$$\frac{d_i^{t+1}}{d_j^{t+1}} = \frac{d_i^t + D_i(S^t)}{d_j^t + D_j(S^t)} = \frac{d_i^t\left(1 + \frac{D_j(S^t)}{d_j^t}\right)}{d_j^t\left(1 + \frac{D_j(S^t)}{d_j^t}\right)} = \frac{d_i^t}{d_j^t} \; ,$$

where the second equality follows by the hypothesis, $\frac{d_i^t}{d_j^t} = \frac{D_i(S^t)}{D_j(S^t)}$. Q.E.D.

In Rubinstein's alternating-offer bargaining model the unique subgame-perfect equilibrium yields an agreement as a function of the discount factors and the selection of the first mover. As the time between offers shrinks to zero this solution converges to the same outcome as the asymmetric Nash solution –with bargaining weights[3] $\gamma_1 = \log\delta_2$ and

---

[2] Recall, however, that Rubinstein also assumes that there are only two players and the Pareto frontier of the bargaining set is linear (with slope –1).

[3] See Binmore (1987a,b) and Binmore et al. (1986). Wilson (2000), has obtained the same result in a model with a mediator who makes random proposals.

$\gamma_2 = \log \delta_1$ – independently of the identity of the first mover. Assuming that the Pareto frontier is the unit simplex, like Rubinstein does, we can prove a similar result for the Hobbes solution, without having to resort to taking limits. That is, the Hobbes solution will exactly coincide with the asymmetric Nash solution, while Rubinstein's does so only in an approximate sense.

Proposition 4 *When the Pareto frontier is the unit simplex and the disagreement function is proportional, the bargaining weights corresponding to the Hobbes solution are* $\gamma_i = D_i(S)$ *, i = 1,2.*

<u>Proof:</u> When the Pareto frontier is the unit simplex, the marginal rate of substitution is 1, everywhere. Consequently the elasticity of the Pareto frontier is equal at every point to the ratio of the utilities at that point. By Proposition 3, this ratio is equal to the ratio of the disagreement utilities. Q.E.D.

The following corollary is now immediate.

Corollary 1 *Under the Rubinstein assumptions (including the proportionality of the disagreement function), the Hobbes solution and the Rubinstein solution coincide if and only if* $\frac{D_1(S)}{D_2(S)} = \frac{\log \delta_2}{\log \delta_1}$.

When the disagreement function is not restricted to be proportional, our model still resembles somewhat a Rubinstein-like model, where the discount rates are not stationary (see Binmore, 1987b, for a detailed discussion of these games). Both models are still equivalent to some asymmetric

Nash solution. However, the bargaining weights –just as the actual solutions– are no longer easily computable. In terms of computability, the Hobbes solution has a significant advantage over the Rubinstein-like one: each step in the calculation of the Hobbes solution improves the precision of the current estimate, and this precision is known. In contrast, to calculate the subgame-perfect equilibrium of a Rubinstein-like game, one has to work backwards from the solution, trying to end up at the disagreement point. At no point in the process, can one have a precise idea about how good the approximation is.