The Pure Theory of Multilateral Obligations

Comment

by

CHRISTIAN EWERHART

Consider a general situation where decision-making by several parties occurs in an either simultaneous or dynamic way. With each party pursuing its individual objective, the presence of external effects may lead to distortions in incentives because private and social objectives may differ. In economic theory, the traditional doctrine has been to internalize those external effects by imposing a penalty on the injurer that equals the loss caused to one or several victims. This realigns private and social incentives and allows one to implement an efficient outcome.

SCHWEIZER [2005] takes an innovative perspective on this problem by pointing out that it suffices to ensure that those behaving efficiently are fully compensated for any potential loss. Once this *saddle-point property* has been assumed, one can show that the efficient outcome is a Nash equilibrium. Moreover, it turns out that all Nash equilibria are payoff-equivalent and therefore also efficient.

In this discussion, I will first briefly review, as a point of reference, the formal setup and the basic results, and then make a number of comments.

Review of the Formal Setup. The formal analysis considers a model with parties $i = 1, ..., I$, whose individual decisions $q_i \in Q_i$ result in a profile $q = (q_1, ..., q_I)$. Party i's pre-law payoff is denoted by $p_i(q)$. The legal framework is modeled as a system of bilateral damage payments $d_{ij}(q)$, to be made from party i to party j in the case of a joint decision q. The resulting post-law payoff for party i is denoted by

$$\pi_i(q) := p_i(q) + \sum_{j \neq i} \{d_{ji}(q) - d_{ij}(q)\} .$$

Call a profile q^* efficient if it maximizes the joint payoff

$$\pi^* := \sum_i p_i(q^*) = \sum_i \pi_i(q^*) .$$

The central result of SCHWEIZER [2005] can now be summarized as follows. Consider an efficient decision profile q^*, and assume that post-law payoff functions $\pi^i(\cdot)$ satisfy the following saddle-point property:

$$\pi_i(q_i^*, q_{-i}^*) \leq \pi_i(q_i^*, q_{-i}) \qquad \text{for all} \quad i \quad \text{and} \quad q_{-i} . \tag{1}$$

Journal of Institutional and Theoretical Economics
JITE 161 (2005), 1–5 © 2005 Mohr Siebeck – ISSN 0932-4569

Then q^* is a Nash equilibrium. To see why, note that if party i were to deviate from q^*, then the saddle-point property would ensure that any party other than i would receive weakly more than $\pi_i(q^*)$. However, as q^* already maximizes the parties' joint payoff, party i cannot gain from a deviation.

As mentioned above, the saddle-point property for one efficient profile q^* implies also that *all* Nash equilibria must be efficient (in fact, payoff-equivalent). As any subgame-perfect or renegotiation-proof equilibrium is a Nash equilibrium, efficiency holds obviously also for these refinements.

This second assertion follows by a similar two-step argument. Fix a party i. If $q^\#$ is a Nash equilibrium, then party i receives weakly more under the profile $q^\#$ than under the profile $\left(q_i^*, q_{-i}^\#\right)$. Moreover, by the saddle-point property for q^*, party i receives weakly more under the profile $\left(q_i^*, q_{-i}^\#\right)$ than under the profile q^*. Thus, all parties obtain under the profile $q^\#$ no less than under the profile q^*. As q^* maximizes joint payoff, this proves payoff equivalence and thus efficiency of $q^\#$.

Nontransferable Utility. My first comment concerns the general assumption in SCHWEIZER [2005] that utility must be transferable. Indeed, as we saw above, efficiency is defined in terms of the maximum of joint payoff, and compensation for damages is assumed to be purely monetary. However, in reality there may be compensations that are not monetary in nature, and there may also be losses that cannot easily be compensated for in monetary terms. As we are going to show now, the pure theory can be generalized without loss to situations where utility is nontransferable.

Let $U_i(q, a)$ denote party i's utility function, which may depend on the profile of decisions q and on a *compensatory allocation* a, which is taken from an abstract set A. The compensatory allocation is determined from the profile of individual decisions by the law according to some exogenous function $\alpha(q)$. One can define then party i's post-law utility as $\Pi_i(q) := U_i(q, \alpha(q))$. So far, we have only generalized the existing setup. Now, as usual, call q^* Pareto-efficient if there is no alternative profile q such that $\Pi_i(q) \geq \Pi_i\left(q^*\right)$ for all i, and such that the inequality is strict for at least one i. The definition of the saddle-point property (1) remains unchanged:

$$\Pi_i\left(q_i^*, q_{-i}^*\right) \leq \Pi_i\left(q_i^*, q_{-i}\right) \qquad \text{for all} \quad i \text{ and } q_{-i}\,. \tag{2}$$

Using these definitions, the following nontransferable utility analogue to Proposition 1 in SCHWEIZER [2005] can be derived essentially the same way as in the model with transferable utility.

PROPOSITION 1 *In the model with nontransferable utility, let q^* be Pareto-efficient, and assume that q^* satisfies the saddle-point property* (2). *Then q^* is a Nash equilibrium, and every Nash equilibrium is payoff-equivalent to q^*.*

PROOF Fix q^*, and consider a conjectural deviation q_i from q_i^* by party i. Then, if q^* satisfies the saddle-point property, all parties $j \neq i$ are at least as well off under the profile $\left(q_i, q_{-i}^*\right)$ as under q^*. If q^* is also Pareto-efficient, this implies that party i cannot strictly improve by the deviation, which proves the first assertion. Consider now an alternative profile $q^\#$, and fix i for the moment. If $q^\#$ is a Nash equilibrium, then

$$\Pi_i\left(q_i^\#, q_{-i}^\#\right) \geq \Pi_i\left(q_i^*, q_{-i}^\#\right). \tag{3}$$

If q^* satisfies the saddle-point property, it follows that

$$\Pi_i\left(q_i^*, q_{-i}^\#\right) \geq \Pi_i\left(q_i^*, q_{-i}^*\right). \tag{4}$$

Combining the last two inequalities yields that party i obtains at least the same payoff under $q^\#$ as under q^*. Since this is true for all i, and since q^* is assumed to be Pareto-efficient, we have proved the second assertion and thereby the proposition.

Q.E.D.

Limited Liability. Especially in a multiparty setting, an individual party may not always be able to provide sufficient funds to compensate all those injured by a harmful decision. As a consequence, it appears to me that restrictions to compensations such as those implied by limited capability or limited liability should be of central relevance for a pure theory of multilateral obligations. One approach would be to incorporate side constraints into the model, and to seek the constrained efficient decision profile that can be implemented in such a setup. As a matter of illustration, we will consider the case of limited liability in the original setup with purely monetary compensation.

Starting from a general setup given by nonnegative pre-law payoffs $p_i(\cdot)$, we will seek to determine the damage rule $d_{ij}(\cdot)$ that implements the highest possible joint payoff π^* in the society, taking account of limited liability of each individual party. This leads us to the following problem:

$$\max_{q^*, d_{ij}(\cdot)} \sum_i p_i\left(q^*\right) \tag{5}$$

$$\text{s.t.} \quad \text{(NE)} \quad \pi_i\left(q_i, q_{-i}^*\right) \leq \pi_i\left(q_i^*, q_{-i}^*\right) \quad \text{for all } i, q_i,$$

$$\text{(LL)} \quad \pi_i(q) \geq 0 \quad \text{for all } i, q.$$

How does the pure theory generalize to this type of setup? Before we solve the problem, we will argue with the help of an example that the saddle-point property may be too restrictive to analyze problems such as the one stated above.

Assume there are two parties $i = 1, 2$, with identical decision sets $Q_i = \{C, D\}$. Let pre-law payoffs be given by

$$p(C, C) = (2, 3), \qquad p(C, D) = (0, 1),$$
$$p(D, C) = (3, 0), \qquad p(D, D) = (1, 1).$$

The efficient decision profile is obviously $q^* = (C, C)$. However, q^* is not a Nash equilibrium in pre-law payoffs, because party 1 has an incentive to deviate to D. We will show now that in order to ensure that q^* satisfies the saddle-point property for some damage rule $d_{ij}(q)$, it is necessary that at least one party obtain a strictly negative post-law payoff for some profile of decisions. This example therefore illustrates that the sufficient condition captured by the saddle-point property may be too restrictive if the parties are protected by limited liability.

The assertion made in the previous paragraph is easily proved by considering the net transfer from party 2 to party 1, defined by

$$t(q) := d_{21}(q) - d_{12}(q)$$
$$= \pi_1(q) - p_1(q) = -\{\pi_2(q) - p_2(q)\}\,.$$

In this example, the saddle-point property with respect to q^* is tantamount to the conditions

$$t(C, D) \geq 2 + t(C, C) \quad \text{and} \quad -t(D, C) \geq 3 - t(C, C)\,.$$

Adding up yields

$$t(C, D) - t(D, C) \geq 5\,. \tag{6}$$

However, nonnegativity of post-law payoffs implies

$$t(C, D) \leq 1 \quad \text{and} \quad t(D, C) \geq -3\,,$$

which yields

$$t(C, D) - t(D, C) \leq 4\,,$$

in contradiction to (6). This shows that, as claimed above, the saddle-point property may be inconsistent with conditions of limited liability.

Nevertheless, there exists, in this example and more generally, a collection of very simple rules that solve the general implementation problem. Consider the following class of *take-and-distribute* damage rules, which "take" any potential surplus from a unilateral deviation and "distribute" the surplus subsequently among the nondeviators.

Formally, we specify that for all i and j, we have $d_{ij}(q^*) = 0$ and $d_{ij}(q) = 0$ whenever q differs from q^* in at least two entries. Thus, transfers are zero unless there is a unilateral deviation from q^*. Moreover, we require that $d_{ji}(q_i, q^*_{-i}) = 0$ for all i, j, and q_i, so that a unilateral deviator obtains no damage payments. Further, for any i and q_i such that $p_i(q_i, q^*_{-i}) \leq p_i(q^*)$, let $d_{ij}(q_i, q^*_{-i}) = 0$ for all j. This says that a unilateral deviator who is worse off than under q^* does not pay compensation. Finally, if for some i and q_i, we have $p_i(q_i, q^*_{-i}) > p_i(q^*)$, then we require the budget conditions $d_{ij}(q_i, q^*_{-i}) \geq 0$ for all j, and

$$\sum_{j \neq i} d_{ij}(q_i, q^*_{-i}) = p_i(q_i, q^*_{-i}) - p_i(q^*)\,. \tag{7}$$

This last requirement says that if a unilateral deviator makes a pre-law gain, then he pays nonnegative damages to the other parties that add up in total precisely to this gain.

It is clear that the described class of damage rules is nonempty. For example, any potential surplus could be divided in equal parts among the nondeviators, replacing (7) by

$$d_{ij}(q_i, q^*_{-i}) = \frac{1}{I-1}\{p_i(q_i, q^*_{-i}) - p_i(q^*)\}\,. \tag{8}$$

For the specific game considered above, this rule would imply $d_{12}(D, C) = 1$ and $d_{ij}(q) = 0$ otherwise. The following result now provides the general answer to the problem (5).

PROPOSITION 2 *Assume that pre-law payoffs are nonnegative, i.e., $p_i(q) \geq 0$ for all q and i. Let q^* be an efficient profile. Then any take-and-distribute damage rule implements q^* as a Nash equilibrium. Moreover, the post-law payoffs $\pi_i(q)$ are nonnegative for all q and i.*

PROOF By construction. *Q.E.D.*

This result shows that even under limited lability, the efficient profile of decisions can be implemented. However, as we have seen further above, this implementation is not feasible using a damage rule that satisfies the saddle-point property. Thus, while the saddle-point property has a number of desirable implications in the general and unrestricted model, it may be of limited value in a more specific setup.

Renegotiation. My final comment concerns the “renegotiation setup.” This variant of the model seems to allow parties, after having completed the first stage, to freely meet and negotiate about the equilibrium that is to be played in the second stage. To me, this raises the question why the parties would not agree on side payments in these renegotiations. This is a critical issue because, *e.g.*, if negotiations about side payments were allowed at the final stage, then the parties could implement any efficient decision profile by a multilateral contract that replicates suitable damage rules. One reason for the absence of contractual side payments could in principle be a missing legal environment. However, the law is an explicit and central element of the analysis. For these reasons, I am left somewhat uncomfortable about the motivation of the renegotiation setup. This, however, need not be a problem of this specific paper, given the well-known problems in building a consistent conceptual base for the theory of incomplete contracting.

References

SCHWEIZER, U. [2005], “The Pure Theory of Multilateral Obligations,” *Journal of Institutional and Theoretical Economics*, 161, ■–■.

Christian Ewerhart
Institute for Empirical Research in Economics
University of Zurich
Winterthurerstr. 30
8006 Zurich
Switzerland
E-mail:
christian.ewerhart@iew.unizh.ch