

John Nachbar  
Washington University in St. Louis  
December 11, 2019

# The Myerson-Satterthwaite Theorem<sup>1</sup>

## 1 Introduction.

The Myerson-Satterthwaite Theorem (MS) is an impossibility result for bilateral bargaining. The original cite is [Myerson and Satterthwaite \(1983\)](#). The proof here most closely follows [Krishna \(2009\)](#) and [Milgrom \(2004\)](#).

One person, the *seller*, owns one unit of an indivisible object. The seller values the object at  $c$ , which I refer to as the seller's *type*. The other person, the *buyer*, values the object at  $v$ , which I refer to as the buyer's *type*. Each knows her own type but not the other person's. More generally, one can interpret the environment as one in which two people are bargaining over a proposed change in the status quo, with the "buyer" gaining  $v$  from the proposed change and the "seller" losing  $c$ .

MS says that, subject to restrictions on the joint distribution of  $c$  and  $v$ , one cannot simultaneously have ex post efficiency, expected budget balance, and interim individual rationality. Ex post efficiency means that trade occurs when  $v > c$  (i.e., the buyer values the object more than the seller) but not when  $v < c$ . Expected budget balance means that while a third party (e.g., a judge) is allowed to provide a net subsidy for some type profiles  $(c, v)$  and collect a net tax for others, the third party's net transfer must be zero in expectation over  $(c, v)$ . A stronger requirement would be ex post budget balance, which requires zero net transfers for any  $(c, v)$ . Interim individual rationality means that the buyer and seller can opt out of the bargaining game after learning own type (e.g., after seeing the good to be traded).

Put differently, MS says that if one demands expected budget balance and interim individual rationality, then bargaining cannot be ex post efficient. This does not say that the outcome will be inefficient for *all* values of  $c$  and  $v$ , only that it will be inefficient for some values. For many standard bargaining models, the expected inefficiency can be large, on the order of a loss of 25% of expected surplus. One can construct games with equilibria that satisfy ex post efficiency and expected budget balance, and one can construct games with equilibria that satisfy ex post efficiency and interim

---

<sup>1</sup> This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License.

individual rationality. So it is the combination of all three criteria that causes a problem.

It is not surprising that bargaining is sometimes inefficient. Labor strikes, lawsuits, and even some wars are examples of bargaining failure. MS is important because it pinpoints information asymmetry as a robust cause of inefficiency. If instead the type profile  $(c, v)$  were known by the participants, then standard models of bargaining (e.g., Nash’s demand game; the Rubinstein-Stähl alternating offers game) typically have equilibria that are ex post efficient. There may also be inefficient equilibria, and in practice inefficiency might arise if either side miscalculates. But in such cases an arbitrator or other third party might be able to force efficient trade at a price of, say,  $(c + v)/2$ . But MS says that if  $c$  and  $v$  are private information, then ex post efficiency fails even with intervention by a third party.

The negative conclusion of MS is at least partly a small numbers problem. There are games, in particular auctions, for which the expected inefficiency, while still positive, becomes vanishingly small as the number of participants grows large. The classic cite is [Rustichini, Satterthwaite and Williams \(1994\)](#); for a more general treatment, see [Cripps and Swinkels \(2006\)](#)

## 2 Generalized Bargaining Games.

### 2.1 The information structure.

The seller’s type  $c$  (cost) and the buyer’s type  $v$  (value) are jointly distributed on  $\mathbb{R}_+^2$ . Let  $F$  denote the joint distribution and let  $\text{supp}(F)$  denote the support of  $F$ , meaning the smallest closed subset of  $\mathbb{R}_+^2$  that gets probability 1. Let  $C$  be the set of  $c$  that appear in  $\text{supp}(F)$ ,

$$C = \{c : \exists v \text{ s.t. } (c, v) \in \text{supp}(F)\}.$$

Similarly, define,

$$V = \{v : \exists c \text{ s.t. } (c, v) \in \text{supp}(F)\}.$$

- The Myerson-Satterthwaite Theorem (MS) assumes that  $F$  is independent, in which case  $\text{supp}(F) = C \times V$ . [McAfee and Reny \(1992\)](#) shows that MS fails if independence is relaxed. Section 5.1 provides an example.
- MS assumes that  $C$  and  $V$  are non-degenerate intervals. Section 5.2 shows by example that MS can fail if the interval assumption is dropped.

- MS assumes that the marginal distributions of  $c$  and  $v$  have densities; in particular, the probability of any given  $c$  or  $v$  is zero. This can be relaxed somewhat. See Section 5.3

Nature moves first and chooses  $(c, v)$ . The seller observes  $c$  but not  $v$ , and that the buyer observes  $v$  but not  $c$ . The joint distribution  $F$  is part of the description of the game; it is commonly known by both players.

Henceforth, hold this information structure fixed.

## 2.2 A generalized bargaining game form.

The seller, having observed  $c$ , chooses an action  $a_s \in A_s$ . The set  $A_s$  is non-empty but otherwise unrestricted. In particular,  $a_s$  could itself be a complicated extensive form strategy; for example,  $a_s$  could be a strategy for an infinite horizon, alternating offers bargaining game. Similarly, the buyer, having observed  $v$ , chooses an action  $a_b \in A_b$ .

A pure strategy for the seller in the incomplete information game (often called a Bayesian strategy) is then a function  $\rho_s : C \rightarrow A_s$ . A pure strategy for the buyer is a function  $\rho_b : V \rightarrow A_b$ .<sup>2</sup>

## 2.3 Utility.

The action profile  $(a_s, a_b)$  determines outcomes via the functions  $\pi : A_s \times A_b \rightarrow [0, 1]$ ,  $\tau_s : A_s \times A_b \rightarrow \mathbb{R}$ , and  $\tau_b : A_s \times A_b \rightarrow \mathbb{R}$ .

$\pi(a_s, a_b) = q$  is the probability that the buyer gets the object. In most bargaining games of interest,  $q$  is either 0 or 1.

$\tau_s(a_s, a_b) = t_s$  is the payment received by the seller ( $\tau$  or  $t$  for “monetary transfer”) while  $\tau_b(a_s, a_b) = t_b$  is the payment made by the buyer. I do *not* assume that  $t_s = t_b$ . If  $t_s \neq t_b$ , then some third party (the government) must make up the difference by taxing or subsidizing. This is why I refer to this game as a “generalized” bargaining game; a standard bargaining game would require  $t_s = t_b$  (which I am calling ex post budget balance).

Finally, utilities depend on types and on outcomes.<sup>3</sup> If the type profile is  $(c, v)$  and the outcome profile is  $(q, t_s, t_b)$ , then the expected utility to the seller (net of the cost  $c$ ) is

$$(1 - q)c + t_s - c = -qc + t_s,$$

---

<sup>2</sup>In most of my game theory notes, I use the notation  $s$  for a pure strategy. I am using  $\rho$  for a pure strategy here to avoid confusion with “s” for “seller”.

<sup>3</sup>In game theory, utilities are usually called payoffs; I use the term utility here to avoid confusion with the payments,  $t_s$  and  $t_b$ .

and the expected utility to the buyer is

$$qv - t_b.$$

Note that preferences are risk neutral.

*Remark 1.* The utility specification seems to rule out games in which a payment is made only when the object is transferred. This would be unfortunate, because many standard bargaining games have this feature. In fact, such games have not been ruled out. One merely has to reinterpret  $t_s$  and  $t_b$  as *expected* payments. For example, suppose that the seller receives the payment  $\hat{t}_s$  if the object is traded and a payment  $\tilde{t}_s$  (which could be 0) if it is not. In expectation, therefore, she receives  $q\hat{t}_s + (1 - q)\tilde{t}_s$ . Let  $t_s$  denote this expected payment. This trick works because of risk neutrality.  $\square$

## 2.4 Nash equilibrium.

A pure strategy profile  $(\rho_s, \rho_b)$  is a *Nash equilibrium (NE)* of the generalized bargaining game iff, for each  $c$ ,  $\rho_s(c)$  maximizes the seller's expected utility given  $\rho_b$  and, for each  $v$ ,  $\rho_b(v)$  maximizes the buyer's expected utility given  $\rho_s$ .<sup>4</sup> It is customary to call a NE in this sort of information setting a *Bayesian Nash Equilibrium*.<sup>5</sup>

## 2.5 Notation for some expectations.

Fix a strategy profile  $(\rho_s, \rho_b)$ . From the perspective of a seller with cost  $c$ , the probability of trade is given by,

$$Q_s(c) = \mathbb{E}_v[\pi(\rho_s(c), \rho_b(v))|c],$$

where the expectation is over  $v$  because the seller does not know the buyer's value. Similarly, from the perspective of a seller with cost  $c$ , the expected payment to the seller is,

$$T_s(c) = \mathbb{E}_v[\tau_s(\rho_s(c), \rho_b(v))|c].$$

---

<sup>4</sup>I can handle mixed strategy equilibria as well, but to avoid some non-essential issues, I do not do so.

<sup>5</sup>Strictly speaking, the equilibrium concept here is somewhat stronger than standard Nash equilibrium, because under standard Nash equilibrium,  $\rho_s$ , for example, is required to be optimal only *ex ante*: given  $\rho_b$ ,  $\rho_s$  maximizes the seller's expected utility, where the expectation is over both  $c$  and  $v$ . This is equivalent to requiring that  $\rho_s(c)$  maximize the seller's expected utility, where the expectation is over  $v$ , for every  $c$  in a subset of  $C$  of probability one. If the marginal distribution over  $C$  has a density, so that any given  $c \in C$  has probability zero, then  $\rho_s$  could satisfy the standard NE optimization condition even if  $\rho_s(c)$  is not optimal for some  $c$ .

Thus, given the strategy profile, a seller with cost  $c$  has an expected utility of,

$$U_s(c) = -Q_s(c)c + T_s(c).$$

The seller's cost influences her expected utility through three channels. There is a direct channel:  $c$  is the value of the object to the seller. There are also two indirect channels. First,  $c$  determines the seller's action  $a_s$  via the strategy  $\rho_s$ . Second, the expectations are conditional on  $c$ . For general  $F$ , it is possible that, for example, a seller with high cost infers that the buyer likewise has a high value, and this affects the seller's posterior over what action the buyer will take. However, MS assumes that  $F$  is independent, and hence assumes this second indirect channel away. Section 5.1 discusses the independence assumption further.

Similarly, for the buyer, define,

$$\begin{aligned} Q_b(v) &= \mathbb{E}_c[\pi(\rho_s(c), \rho_b(v))|v] \\ T_b(v) &= \mathbb{E}_c[\tau_b(\rho_s(c), \rho_b(v))|v] \\ U_b(v) &= Q_b(v)v - T_b(v). \end{aligned}$$

## 2.6 Efficiency, Budget Balance, and Individual Rationality.

**Definition 1.** A Nash equilibrium of a generalized bargaining game is ex post efficient iff, for any  $(c, v) \in \text{supp}(F)$  and associated  $(q, t_s, t_b)$ ,  $q = 1$  if  $v > c$  and  $q = 0$  if  $v < c$ .

**Definition 2.** A Nash equilibrium of a generalized bargaining game is ex post budget balanced iff, for any  $(c, v) \in \text{supp}(F)$  and associated  $(q, t_s, t_b)$ ,  $t_s = t_b$ . A Nash equilibrium of a generalized bargaining game is expected budget balanced iff the equilibrium expected subsidy,  $\mathbb{E}_{c,v}[T_s(c) - T_b(v)]$ , is zero.

**Definition 3.** A Nash equilibrium of a generalized bargaining game is interim individually rational iff, for any  $(c, v) \in \text{supp}(F)$ ,  $U_s(c) \geq 0$  and  $U_b(v) \geq 0$ .

*Remark 2.* A sufficient condition for individual rationality in Nash equilibrium is that the generalized bargaining game contains an action *quit*, available to either participant, with the property that if either quits, then the seller retains the object and there is no payment.  $\square$

### 3 The Myerson-Satterthwaite Theorem: Special Case.

The following result is the Myerson-Satterthwaite theorem for the special case in which  $C = V = [0, 1]$ . Section 4 provides a statement of a more general form of the theorem.

**Theorem 1** (Baby Myerson-Satterthwaite Theorem). *Suppose that  $F$  is independent,  $C = V = [0, 1]$ , and the marginal distributions have densities. Then there is no NE of any generalized bargaining game that is ex post efficient, expected budget balanced, and interim individually rational.*

The proof of Theorem 1 involves the following generalized bargaining game, which I call the VCG game in honor of Vickrey (1962), Clarke (1971), and Groves (1973), who considered games closely related to this one.

In the VCG game, the seller's action is to announce a type  $c^* \in [0, 1]$ , which could be different from the seller's true type, and the buyer's action is to announce a type  $v^* \in [0, 1]$ , which could be different from the buyer's true type. Trade occurs iff  $v^* \geq c^*$ . If there is trade then the buyer pays  $c^*$  and the seller receives  $v^*$ . If there is no trade then there are no payments.

There is a NE of the VCG game in which participants tell the truth:  $\rho_s(c) = c$  and  $\rho_b(v) = v$ . Indeed, truth telling is weakly dominant in this game.<sup>6</sup> It is immediate that this truth-telling NE is ex post efficient.<sup>7</sup> And it is also immediate that it is individually rational, since the buyer always gets either 0 (if there is no trade, which happens iff  $v < c$ ) or  $v - c \geq 0$  (if there is trade, which happens iff  $v \geq c$ ), and similarly for the seller.

But the VCG game isn't even close to being expected budget balanced. In fact, the subsidy equals  $v - c > 0$  whenever  $v > c$ , which occurs with positive probability, and the subsidy equals zero whenever  $v \leq c$ : the expected subsidy equals the expected surplus from efficient trade!

Remarkably, the truth-telling equilibrium of the VCG game is the *cheapest* way to get ex post efficiency and interim individual rationality.

---

<sup>6</sup>Briefly, consider the seller. Given  $c$ , suppose the seller reports  $c^* > c$ . The only circumstance in which lying in this way changes the seller's utility is when  $v^* \in (c, c^*]$ . In this case, there is trade if the seller tells the truth, in which case the seller gets a net utility of  $v^* - c > 0$ , while there is no trade if the seller reports  $c^*$ , in which case the seller gets a utility of 0. So lying yields a lower utility than telling the truth. The other cases are similar.

<sup>7</sup>There are other NE, not in weakly dominant strategies, that are not ex post efficient. For example, if  $c$  and  $v$  are both distributed on  $[0, 1]$ , then it is a NE of the VCG game for the buyer to announce  $v^* = 0$  and the seller to announce  $c^* = 1$  regardless of actual type; in this NE, there is never any trade.

**Theorem 2.** *Suppose that  $F$  is independent,  $C = V = [0, 1]$ , and the marginal distributions have densities. Then the truth-telling equilibrium of the VCG game has the smallest expected subsidy of any ex post efficient and interim individually rational NE of any generalized bargaining game.*

**Proof of Theorem 1.** Since the VCG game has a strictly positive expected subsidy, Theorem 2 implies that any NE of any generalized bargaining game is not expected budget balanced if it is ex post efficient and interim individually rational. ■

It remains to prove Theorem 2. The key fact driving Theorem 2 is the following.<sup>8</sup>

**Theorem 3** (The Integral Conditions). *Suppose that  $F$  is independent and that  $C = V = [0, 1]$ . Consider any strategy profile  $(\rho_s, \rho_b)$ . If  $\rho_s$  is optimal for the seller, given  $\rho_b$ , then,*

$$U_s(c) = U_s(1) + \int_c^1 Q_s(x) dx. \quad (1)$$

Similarly, if  $\rho_b$  is optimal for the buyer, given  $\rho_s$ , then,

$$U_b(v) = U_b(0) + \int_0^v Q_b(x) dx. \quad (2)$$

**Proof.** If  $\rho_s$  is optimal, then no deviation by the seller, say to  $\hat{\rho}_s$ , can be profitable. One particular deviation is for the seller, when her actual cost is  $c^*$ , to play as if her cost were  $c$ : she plays  $\rho_s(c) = a_s$  instead of  $\rho_s(c^*) = a^{*s}$ .

The probability of trade, from the perspective of a seller who has cost  $c^*$  but plays as if she has cost  $c$  is,

$$\tilde{Q}_s(c, c^*) = \mathbb{E}_v[\pi(\rho_s(c), \rho_b(v)) | c^*].$$

The assumption that  $F$  is independent implies,

$$\tilde{Q}_s(c, c^*) = Q_s(c).$$

Similarly, the expected payment to the seller, from the perspective of a seller who has cost  $c^*$  but plays as if she has cost  $c$ , is,

$$\tilde{T}_s(c, c^*) = \mathbb{E}_v[\tau_s(\rho_s(c), \rho_b(v)) | c^*].$$

---

<sup>8</sup>As a technical aside, note that Theorem 3 does not assume that the marginal distributions have densities.

Independence implies,

$$\tilde{T}_s(c, c^*) = T_s(c).$$

Therefore, the seller's expected utility, from the perspective of a seller who has cost  $c^*$  but plays as if she has cost  $c$ , is,

$$\tilde{U}_s(c, c^*) = -\tilde{Q}_s(c, c^*)c^* + \tilde{T}_s(c, c^*),$$

which, by independence, can be rewritten,

$$\tilde{U}_s(c, c^*) = -Q_s(c)c^* + T_s(c).$$

Note that  $U_s(c^*) = \tilde{U}_s(c^*, c^*)$ . In Section 5.1, I discuss what happens if independence is violated.

If  $\rho_s$  is optimal then, for any  $c$  and  $c^*$ , it must be that  $U_s(c^*) \geq \tilde{U}_s(c, c^*)$ , or,

$$U_s(c^*) = -Q_s(c^*)c^* + T_s(c^*) \geq -Q_s(c)c^* + T_s(c), \quad (3)$$

and conversely  $U_s(c) \geq \tilde{U}_s(c^*, c)$ , or,

$$U_s(c) = -Q_s(c)c + T_s(c) \geq -Q_s(c^*)c + T_s(c^*). \quad (4)$$

These inequalities are called *incentive compatibility* (IC) conditions for the seller. The seller's IC conditions imply,

$$\begin{aligned} U_s(c^*) - U_s(c) &\geq -Q_s(c)c^* + T_s(c) - U_s(c) \\ &= -Q_s(c)c^* + T_s(c) - [-Q_s(c)c + T_s(c)] \\ &= -Q_s(c)(c^* - c). \end{aligned}$$

Similarly,

$$\begin{aligned} U_s(c^*) - U_s(c) &\leq U_s(c^*) - [-Q_s(c^*)c + T_s(c^*)] \\ &= -Q_s(c^*)c^* + T_s(c^*) - [-Q_s(c^*)c + T_s(c^*)] \\ &= -Q_s(c^*)(c^* - c). \end{aligned}$$

Combining,

$$-Q_s(c)(c^* - c) \leq U_s(c^*) - U_s(c) \leq -Q_s(c^*)(c^* - c). \quad (5)$$

This implies that  $-Q_s$  is weakly increasing in  $c$ , hence  $Q_s$  is weakly decreasing. This should be intuitive: as  $c$  increases, it is increasingly unlikely that trade is worthwhile, so the probability of trade should decrease.

Since  $-Q_s$  is weakly increasing,  $-Q_s$  is (Riemann) integrable (Rudin (1976), Theorem 6.9); Inequality (5) implies that the integral satisfies,<sup>9</sup>

$$U_s(1) - U_s(c) = \int_c^1 -Q_s(x) dx. \quad (6)$$

Rearranging Equation (6) gives Equation (1). The argument for Equation (2) is similar and I omit it. ■

*Remark 3.* Assume for the moment that  $Q_s$  and  $T_s$  are differentiable. As in the proof of Theorem 3, the expected utility of a seller who has cost  $c^*$  but acts as if her cost were  $c$  is,

$$\tilde{U}_s(c, c^*) = -Q_s(c)c^* + T_s(c).$$

(Again, this notation assumes independence.) If  $\rho_s$  is indeed optimal, then  $\tilde{U}_s(c, c^*)$  is maximized when the seller with cost  $c^*$  actually acts as if her cost were  $c^*$  (so that the seller plays  $a^{*s} = \rho_s(c^*)$ ), in which case the expected utility is  $\tilde{U}_s(c^*, c^*) = U_s(c^*)$ .  $U_s$  is thus the value function for this maximization problem. By the Envelope Theorem,  $DU_s(c^*) = D_2\tilde{U}_s(c^*, c^*)$  (where  $D_2$  is the derivative with respect to the second argument), hence

$$DU_s(c^*) = -Q_s(c^*). \quad (7)$$

Integrating this gives Equation (6).

One cannot, in fact, simply assume that  $Q_s$  and  $T_s$  are differentiable, since they depend on the strategies, which are not required to be differentiable. Milgrom and Segal (2002) shows that, despite this issue, a generalized Envelope Theorem argument still applies and establishes Equation (6). □

Equation (1) and Equation (2) are often called the *Integral Conditions*. Substituting  $U_s(c) = -Q_s(c)c + T_s(c)$  and  $U_b(v) = Q_b(v)v - T_b(v)$  into the

---

<sup>9</sup>Explicitly, recall that a function is integrable iff its upper and lower integrals are equal (and finite). Consider any finite partition of  $[c, 1]$  determined by  $c = x_1 < \dots < x_K = 1$ . Consider any  $x_k < x_{k+1}$ . Since  $-Q_s$  is weakly increasing, the maximum value of  $-Q_s$  on this interval is  $-Q(x_{k+1})$ , so that the associated rectangle has area  $-Q_s(x_{k+1})(x_{k+1} - x_k)$ . By (5), this area is bounded below by  $U_s(x_{k+1}) - U_s(x_k)$ . Summing over such areas, all but two of the  $U_s(x_k)$  cancel, yielding a total lower bound of  $U_s(1) - U_s(c)$ . I get the same lower bound, namely  $U_s(1) - U_s(c)$ , for every partition. Hence the upper integral is bounded below by  $U_s(1) - U_s(c)$ . By an analogous argument, the lower integral is bounded above by  $U_s(1) - U_s(c)$ , which yields Equation (6).

Integral Conditions and rearranging yields,

$$T_s(c) = U_s(1) + Q_s(c)c + \int_c^1 Q_s(x) dx, \quad (8)$$

$$T_b(v) = -U_b(0) + Q_b(v)v - \int_0^v Q_b(x) dx. \quad (9)$$

These equations imply the following.

**Theorem 4.** *Suppose that  $F$  is independent and that  $C = V = [0, 1]$ . Consider any two generalized bargaining games and any two NE of these games. If  $Q_s$  and  $U_s(1)$  are the same in both NE, then  $T_s$  is the same in both NE. And if  $Q_b$  and  $U_b(0)$  are the same in both NE, then  $T_b$  is the same in both NE.*

**Proof.** Almost immediate from Equation 8 and Equation 9 since, in NE, both the buyer and seller are optimizing, and since, by Equation 8, the expected payment to the seller depends only on  $Q_s$  and  $U_s(1)$ , and similarly for the expected payment by the buyer. ■

*Remark 4.* Theorem 4 is a variant of the Revenue Equivalence Theorem, an important result in mechanism design theory. I have separate notes on Revenue Equivalence in the context of auctions. □

Ex post efficiency pins down  $Q_s$  and  $Q_b$ . If the equilibrium is ex post efficient then  $Q_s(c)$  equals the probability that  $v > c$  plus the probability that trade occurs if  $v = c$  times the probability that  $v = c$ , all conditional on  $c$ . But because of the assumption that the marginal distributions have densities, the probability that  $v = c$  is zero, so that  $Q_s(c)$  simply equals the probability that  $v > c$ , given  $c$ . Similarly,  $Q_b(v)$  equals the probability that  $v > c$ , given  $v$ . Thus, Theorem 4 implies that if two NE are ex post efficient then they have the same expected subsidy,  $\mathbb{E}_{c,v}[T_s(c) - T_b(v)]$ , iff they have the same  $U_s(1)$  and  $U_b(0)$ .

To complete the proof of Theorem 2, I need also to address the issue of interim individual rationality. In words, the following theorem says that a NE will be interim individually rational iff it is interim individually rational for the seller and buyer types who are in the worst position to trade.

**Theorem 5.** *Suppose that  $F$  is independent and that  $C = V = [0, 1]$ .  $(\rho_s, \rho_b)$  satisfies interim individual rationality iff  $U_s(1) \geq 0$  and  $U_b(0) \geq 0$ .*

**Proof.** If  $U_s(c) \geq 0$  for every  $c$ , then in particular  $U_s(1) \geq 0$ . So  $U_s(1) \geq 0$  is necessary for individual rationality for the seller. To see that it is sufficient, note that since  $Q_s(c) \geq 0$  for every  $c$ , (1) implies that if  $U_s(1) \geq 0$

then  $U_s(c) \geq 0$  for every  $c$ . And a similar argument holds for  $U_b$ . ■

I am now, finally, in a position to prove Theorem 2.

**Proof of Theorem 2.** Focus on ex post efficient NE of generalized bargaining games. The truth-telling equilibrium of the VCG game is one such equilibrium.

As already discussed, ex post efficiency implies that any such NE must have the same  $Q_s$  and  $Q_b$ . In view of equalities (8) and (9), this means that  $T_s$  and  $T_b$  are entirely determined by  $U_s(1)$  and  $U_b(0)$ . From (8) and (9),  $T_s$  is increasing in  $U_s(1)$  and  $T_b$  is decreasing in  $U_b(0)$ . Thus, if I want to minimize the net subsidy  $\mathbb{E}_{c,v}[T_s(c) - T_b(v)]$ , then I want to make both  $U_s(1)$  and  $U_b(0)$  as small as possible. By Theorem 5, interim individual rationality implies that the smallest possible value for either  $U_s(1)$  or  $U_b(0)$  is zero.

Putting this all together, and appealing to Theorem 4, if I can find a NE of a generalized bargaining game such that (a) the NE is ex post efficient, and (b)  $U_s(1) = U_b(0) = 0$ , then this NE has the smallest expected subsidy of any ex post efficient and interim individually rational NE of any generalized bargaining game. The truth-telling NE of the VCG game has these properties. ■

*Remark 5.* One sometimes sees MS stated in a weaker form: there does not exist any NE in *weakly dominant strategies* of any generalized bargaining game that is ex post efficient, expected budget balanced, and interim individually rational. The restriction to NE in weakly dominant strategies is actually without loss of generality, since the truth-telling NE of the VCG game is in weakly dominant strategies. A benefit of focusing on weak dominance, however, is that relatively easy proofs are available for MS in this special case; Jeff Ely provides a nice graphical illustration [here](#). □

## 4 The Myerson-Satterthwaite Theorem: General Case.

Suppose that  $F$  is independent, so that  $\text{supp}(F) = C \times V$ . Suppose further that  $C$  and  $V$  are non-degenerate intervals, with  $C = [\underline{c}, \bar{c}]$ ,  $V = [\underline{v}, \bar{v}]$ ,  $\bar{c} > \underline{c}$ , and  $\bar{v} > \underline{v}$ .

If the  $C$  and  $V$  intervals do not overlap then MS fails. In particular, if  $\bar{c} \leq \underline{v}$ , then it is always ex post efficient to trade. As a bargaining “game”, have the two players always trade at a price of  $(\bar{c} + \underline{v})/2$ . On the other hand,

if  $\underline{c} \geq \bar{v}$ , then it is always ex post efficient NOT to trade. As a bargaining “game”, have the two players never trade.

The general statement of MS, therefore, assumes that  $C$  and  $V$  overlap, meaning that  $\bar{c} > \underline{v}$  and  $\underline{c} < \bar{v}$ .

**Theorem 6** (The Myerson-Satterthwaite Theorem). *Suppose that  $F$  is independent,  $C$  and  $V$  are non-degenerate intervals, these intervals overlap, and that the marginal distributions have densities. Then there is no NE of any generalized bargaining game that is ex post efficient, expected budget balanced, and individually rational.*

**Proof.** Modify the VCG game so that if there is trade ( $v^* \geq c^*$ ) then the buyer pays  $\max\{c^*, \underline{v}\}$  (i.e., the buyer never pays less than  $\underline{v}$  if there is trade) and the seller receives  $\min\{v^*, \bar{c}\}$  (i.e., the seller never receives more than  $\bar{c}$  if there is trade).

In the modified VCG game, it is a Nash equilibrium (in weakly dominant strategies) to report one’s true type. In this equilibrium,  $U_s(\bar{c}) = 0$ : if  $\bar{v} < \bar{c}$  then there is no trade when  $c = \bar{c}$ ; otherwise, if  $\bar{v} \geq \bar{c}$ , then whenever  $v \geq \bar{c}$  (so that there is trade) and  $c = \bar{c}$ , the seller receives a payment of  $\bar{c}$ , for a utility of zero. By a similar argument,  $U_b(\underline{v}) = 0$ . Since  $U_s(\bar{c}) = U_b(\underline{v}) = 0$ , essentially the same arguments as before imply that interim individual rationality holds and that the expected subsidy by a third party is as small as possible.

The assumptions on  $C$  and  $V$  guarantee that  $\max\{\underline{c}, \underline{v}\} < \min\{\bar{c}, \bar{v}\}$ . There is thus an open subset of  $C \times V$  for which (a)  $c < v$  (there is trade), (b)  $v < \bar{c}$  (so that the seller receives  $v$ ) and (c)  $c > \underline{v}$  (so that the buyer pays  $c$ ). Thus, there is a positive probability that the third party pays a strictly positive subsidy, namely  $v - c > 0$ . Moreover, working through the remaining cases shows that the subsidy is never negative:

- if  $v \geq \bar{c}$  and  $c \leq \underline{v}$ , then (since, by assumption,  $\bar{c} > \underline{v}$ ),  $v > c$  (there is trade) and the subsidy is  $\bar{c} - \underline{v} > 0$ ;
- if  $v \geq \bar{c}$  and  $c > \underline{v}$  then (since  $c \leq \bar{c}$ )  $v \geq c$  (there is trade) and the subsidy is  $\bar{c} - c \geq 0$ ;
- if  $v < \bar{c}$  and  $c \leq \underline{v}$  then (since  $v \geq \underline{v}$ )  $v \geq c$  (there is trade) and the subsidy is  $v - \underline{v} \geq 0$ .

The overall expected subsidy is, therefore, strictly positive, which implies that expected budget balance fails. ■

## 5 Remarks on the Distribution Assumptions.

### 5.1 If the joint distribution is not independent.

If  $F$  is not independent then MS fails. In particular, McAfee and Reny (1992) shows that if independence does not hold then there is a generalized bargaining game that is ex post efficient, interim individually rational, and satisfies expected budget balance (but possibly not ex post budget balance).

As a simple example, suppose that  $c$  is uniformly distributed on  $[0, 1]$  and that  $v = \sqrt{c}$ . The support of  $F$  is then the graph of the square root function on  $[0, 1]$ . Consider the game in which the seller reports  $c^*$  (which may be different from the seller's true type), the buyer reports  $v^*$  (which may be different from the buyer's true type), there is trade iff  $v^* = \sqrt{c^*}$ , and if trade takes place then the seller sells to the buyer at the price  $(c^* + v^*)/2$ . There is an equilibrium in which, for any  $(c, v) \in \text{supp}(F)$ , both players report their true type. This equilibrium is ex post efficient, interim individually rational, and ex post budget balanced.

Intuitively, correlation allows a third party to use the buyer's behavior to infer information about the cost of the seller, and vice versa. In the example, this takes the form of having trade take place iff the seller and buyer agree, meaning that they report  $(c^*, v^*)$  such that  $v^* = \sqrt{c^*}$ . More generally, an ability to crosscheck increases the scope of what the equilibrium of a generalized bargaining game can achieve.

In terms of the formal argument, if  $F$  is not independent then it may not be true that, for example,  $\tilde{Q}_s(c, c^*) = Q_s(c^*)$ . This vitiates much of the analysis. And in the Envelope Theorem approach (Remark 3 in Section 3), if independence is dropped then Equation (7) becomes,

$$DU_s(c^*) = -Q_s(c^*) - D_2\tilde{Q}_s(c^*, c^*)c^* + D_2\tilde{T}_s(c^*, c^*),$$

which no longer implies the integral condition, Equation (6). In particular, note that the Envelope Theorem expression for  $DU_s$  now contains a term involving the expected payment,  $\tilde{T}_s$ .

### 5.2 If $C$ and $V$ are not non-degenerate intervals.

If  $C$  and  $V$  are intervals but one is degenerate then MS can fail. For example, suppose that  $C = [0, 1]$  but that  $V = \{1/2\}$  (so that  $\bar{v} = \underline{v} = 1/2$ ). (Since  $V$  is degenerate, this example also violates the marginal density assumption.) Consider the game in which the seller reports  $c^* \in C$  (which may be different from the seller's true type), there is trade iff  $c^* < 1/2$ , and if there is trade

then it is at a price of  $1/2$ . Then it is an equilibrium for both players to report their true type, and this equilibrium is ex post efficient, interim individually rational, and ex post budget balanced.

MS can also fail if  $C$  and  $V$  are not intervals. Suppose  $C = V = \{0, 1\}$ , and that the four possible outcomes are equally likely. (This example also violates the marginal density assumption.) Consider the game in which the seller reports  $c^* \in C$  (which may be different from the actual  $c$ ), the buyer reports  $v^* \in V$  (which may be different from the actual  $v$ ), there is trade iff  $(c^*, v^*) = (0, 1)$ , and if there is trade then it is at a price of  $1/2$ . Then it is an equilibrium for both players to report their true type, and this equilibrium is ex post efficient, interim individually rational, and ex post budget balanced.

In the first example, the suggested game is the modified VCG game used in the proof of Theorem 6. Here, however, because  $V$  is degenerate, there is zero probability that  $(c, v)$  falls into a range where a third party must provide a strictly positive subsidy.

In the second example, a partial intuition is that any equilibrium of any generalized bargaining game must satisfy the incentive compatibility (IC) conditions, Inequality 3 and Inequality 4. Very loosely, the more IC conditions there are, the more restricted is equilibrium behavior. If  $C$  and  $V$  are intervals, then the seller and buyer have, for each realized  $c$  or  $v$ , a continuum of IC conditions. In contrast, if  $C = V = \{0, 1\}$  then the seller and buyer have, for each realized  $c$  or  $v$ , only a single IC condition (e.g., it cannot be profitable for a  $c = 0$  seller to behave as if  $c = 1$ ). In terms of the formal argument, the integral conditions (Equations (6) and (2)) assume that  $Q_s$  and  $Q_b$  are defined on intervals.

### 5.3 If the marginal distributions on $C$ and $V$ do not have densities.

The density assumption is used to pin down  $Q_s$  and  $Q_b$ . Explicitly, if the marginal distribution of  $v$  has a density then ex post efficiency implies that  $Q_s(c)$  equals the probability that  $v > c$ , given  $c$ ; the case  $v = c$  (which is not pinned down by ex post efficiency) has probability zero and can be ignored. And a similar statement holds for  $Q_b$ .

A weaker sufficient condition for this argument to hold is that the probability that  $v = c$ , given either  $c$  or  $v$ , is zero. This condition could be met even if the marginal distributions do not admit densities, so long as any atoms are not in the same place for both the buyer and the seller.

And this issue can be side-stepped entirely, with no restriction on the marginal distributions (other than that their supports be non-degenerate

intervals) by strengthening ex post efficiency to require that trade occurs whenever  $v = c$ , since in this case, ex post efficiency implies that  $Q_s(c)$  must equal the probability that  $v \geq c$ , given  $c$ , and similarly for  $Q_b(v)$ . (The game discussed in the second example in Section 5.2 violates this stronger version of efficiency.)

## References

- Clarke, Edward. 1971. "Multipart Pricing of Public Goods." *Public Choice* 2:19–33.
- Cripps, Martin and Jeroen Swinkels. 2006. "Depth and Efficiency in Large Double Auctions." *Econometrica* 74(1):47–92.
- Groves, Theodore. 1973. "Incentives in Teams." *Econometrica* 41:617–631.
- Krishna, Vijay. 2009. *Auction Theory*. Second ed. Academic Press.
- McAfee, Preston and Philip Reny. 1992. "Correlated Information and Mechanism Design." *Econometrica* 60(2):395–421.
- Milgrom, Paul. 2004. *Putting Auction Theory To Work*. Cambridge University Press.
- Milgrom, Paul and Ilya Segal. 2002. "Envelope Theorems for Arbitrary Choice Sets." *Econometrica* 70(2):583–601.
- Myerson, Roger and Mark Satterthwaite. 1983. "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory* 28:265–81.
- Rubinstein, Ariel. 1982. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50:97–110.
- Rudin, Walter. 1976. *Principles of Mathematical Analysis*. Third ed. New York: McGraw-Hill.
- Rustichini, Aldo, Mark Satterthwaite and Steven Williams. 1994. "Convergence to Efficiency in a Simple Market With Incomplete Information." *Econometrica* 62:1041–63.
- Vickrey, William. 1962. Auctions and Bidding Games. In *Recent Advances in Game Theory*. Number 29 in "Princeton Conference Series" Princeton University Press pp. 15–27.