October 17, 2011

CORRIGENDUM TO "AGGREGATION AND LINEARITY IN THE PROVISION OF INTERTEMPORAL INCENTIVES"

ANDREW CLAUSEN [1]

This corrigendum concerns the following theorem of Holmstrom and Milgrom (1987),

THEOREM 4 *A strategy* $\{p^t(X^{t-1})\}$ *can be implemented if and only if for every date and history,* $p^t(X^{t-1}) \in P^0$ *(that is, if and only if each* $p^t(X^{t-1})$ *can be implemented in the single period problem). A sharing rule* $s(X^T)$ *implements* $\{p^t(X^{t-1})\}$ *with certain equivalent* w_0 *if and only if it can be written in the form*

$$s(X^T) = w_0 + \sum_{t=1}^{T} s_t(x^t; p^t(X^{t-1})), \tag{1}$$

where each $s_t(\cdot; p)$ *is a sharing rule that implements* p *with certain equivalent zero in the single-period problem.*

The *only if* part of the last sentence is mistaken. It implies that *only* sharing rules with the specified form of history independence can implement a strategy with certain equivalent w_0.[1] However, *every* sharing rule $s(X^T)$ implements some strategy $\{p^t(X^{t-1})\}$, and $\hat{s}(X^T) = s(X^T) - \bar{w}$ does so with certain equivalent w_0 for an appropriate choice of $\bar{w}$. In particular, the theorem asserts that sharing rules are additively separable across time whenever they implement a stationary strategy $p^t(X^{t-1}) = p^*$. Thus, the theorem incorrectly disqualifies non-stationary sharing rules of the form

$$s(X^2) = \begin{cases} s_a(x^1) + s_a(x^2) & \text{if } x^1 = 1, \\ s_a(x^1) + s_b(x^2) & \text{if } x^1 \neq 1, \end{cases}$$

from implementing a stationary strategy, even when s_a and s_b implement p^* with certain equivalent zero in the single-period problem. The mistake does not invalidate Holmstrom and Milgrom's main conclusion that all optimal sharing rules are linear in accounts. Bolton and Dewatripont (2005, pages 445-447) present a proof that does not apply Theorem 4.

A version of Theorem 4 can be proved that does not assert any history independence properties of the sharing rule. This theorem can be interpreted as a decomposition theorem, that any sharing rule can be uniquely decomposed into spot sharing rules that implement the appropriate action at each history:

[1] I am grateful to Xianting Hu, Steven Matthews, Guido Menzio, Mallesh Pai and Xi Weng for helpful comments.

[1] Holmstrom and Milgrom's Theorem 3 gives sufficient conditions for uniqueness of the sharing rule that implements p.

Definition (Holmstrom-Milgrom decomposition) $\{w_0, s_t(x^t; X^{t-1})\}$ *is a Holmstrom-Milgrom decomposition of the sharing rule* $s(X^T)$ *if*

$$(2) \qquad s(X^T) = w_0 + \sum_{t=1}^{T} s_t(x^t; X^{t-1}),$$

and $s(X^T)$ *implements a strategy* $\{p^t(X^{t-1})\}$ *if and only if each* $s_t(\cdot; X^{t-1})$ *implements* $p^t(X^{t-1})$ *with certain equivalent zero in the single-period problem.*

Holmstrom and Milgrom define

$$(3) \qquad V_t(X^t) = \max_{\{p^\tau(X^{\tau-1})\}_{\tau=t+1}^T} E\left\{u\left[s(X^T) - \sum_{\tau=t+1}^{T} c(p^\tau; \theta^\tau)\right] \middle| X^t, \{p^\tau\}\right\}$$

to be the agent's value excluding sunk costs at the end of period t. They write $w_t(X^t) = u^{-1}(V_t(X^t))$ as the corresponding certain equivalent. Note that $w_T(X^T) = s(X^T)$ and w_0 is the agent's certain equivalent of participating in the contract.

Theorem *Every sharing rule* $s(X^T)$ *has a unique Holmstrom-Milgrom decomposition,* $w_0 = u^{-1}(V_0)$ *and*

$$(4) \qquad s_t(x^t; X^{t-1}) = w_t(X^{t-1}, x^t) - w_{t-1}(X^{t-1}).$$

Proof: From $s(\cdot)$, $\{V_t(\cdot)\}$ and hence $\{w_t(\cdot)\}$ are defined. Then from the latter, define $s_t(\cdot)$ according to (4). We claim that $\{w_0, s_t(\cdot)\}$ is a Holmstrom-Milgrom decomposition. Summing (4) across all time periods gives

$$(5) \qquad \sum_{t=1}^{T} s_t(x^t; X^{t-1}) = \sum_{t=1}^{T} \left[w_t(X^{t-1}, x^t) - w_{t-1}(X^{t-1})\right] = w_T(X^T) - w_0 = s(X^T) - w_0$$

which is equivalent to (2). To study the decomposed sharing rules, it is helpful to construct a Bellman equation from (3) using the property $u(y + y') = -u(y)u(y')$ of CARA utility function $u(y) = -\exp(-ry)$:

$$(6) \qquad u(w_t(X^t)) = V_t(X^t) = \max_{\hat{p}} E\left\{-u(-c(\hat{p}, \theta^{t+1}))V_{t+1}(X^t, x^{t+1}) \middle| X^t, \hat{p}\right\}$$

$$(7) \qquad = \max_{\hat{p}} E\left\{u(w_{t+1}(X^t, x^{t+1}) - c(\hat{p}; \theta^{t+1})) \middle| X^t, \hat{p}\right\}.$$

This leads to the following principle of optimality: a sharing rule $s(X^T)$ implements a strategy $\{p^t(X^{t-1})\}$ if and only if at every history X^{t-1},

$$(8) \qquad p^t(X^{t-1}) \in \arg\max_{\hat{p}} E\left\{u(w_t(X^{t-1}, x^t) - c(\hat{p}; \theta^t)) | X^{t-1}, \hat{p}\right\}$$

which is equivalent to

$$p^t(X^{t-1}) \in \arg\max_{\hat{p}} E\left\{u(s_t(x^t, X^{t-1}) - c(\hat{p};\theta^t))|X^{t-1}, \hat{p}\right\}. \tag{9}$$

Thus, $s_t(x^t, X^{t-1})$ indeed implements $p^t(X^{t-1})$.
Substituting (4) into (7) gives

$$u(w_t(X^t)) = \max_{\hat{p}} E\left\{u(s_{t+1}(x^{t+1}; X^t) + w_t(X^t) - c(\hat{p};\theta^{t+1}))\big|X^t, \hat{p}\right\}$$

and dividing both sides by $-u(w_t(X^t))$ gives

$$u(0) = \max_{\hat{p}} E\left\{u(s_{t+1}(x^{t+1}; X^t) - c(\hat{p};\theta^{t+1}))\big|X^t, \hat{p}\right\}.$$

This establishes that $s_{t+1}(x^{t+1}; X^t)$ implements $p^{t+1}(X^t)$, with certain of equivalent of zero in the single-period problem. Therefore (4) defines a Holmstrom-Milgrom decomposition.

It remains to show that the decomposition is unique. Suppose some other $\{\hat{w}_0, \hat{s}_t(x^t; X^{t-1})\}$ implements each $p^t(X^{t-1})$ with certain equivalent zero in the single-period problem. This implies $s(X^T)$ implements $p^t(X^t)$ with certain equivalent $\hat{w}_0 = w_0$. It suffices to show that

$$w_t(X^t) = \sum_{\tau=1}^{t} \hat{s}_\tau(x^\tau; X^{\tau-1}) + w_0, \tag{10}$$

because taking the difference of this formula for t and $t-1$ gives (4).

This is trivial for $t = T$. Suppose the claim is true for $t+1$. Then substituting (10) for $t+1$ into (7) gives

$$\begin{aligned} V_t(X^t) &= \max_{\hat{p}} E\left\{u\left[\sum_{\tau=1}^{t+1} \hat{s}_\tau(x^\tau; X^{\tau-1}) + w_0 - c(\hat{p};\theta^{t+1})\right]\Bigg|X^t, \hat{p}\right\} \\ &= -u\left[\sum_{\tau=1}^{t} \hat{s}_\tau(x^\tau; X^{\tau-1}) + w_0\right] \max_{\hat{p}} E\left\{u\left[\hat{s}_{t+1}(x^{t+1}; X^t) - c(\hat{p};\theta^{t+1})\right]\big|X^t, \hat{p}\right\}. \end{aligned}$$

From the condition that $\hat{s}_{t+1}(\cdot; X^t)$ implements $p^{t+1}(X^t)$ with certain equivalent zero in the single-period problem, it follows that the final maximization problem obtains $u(0) = -1$, and the equation reduces to the induction hypothesis (10) for t. *Q.E.D.*

Holmstrom and Milgrom make several remarks on their theorem. It is possible to amend these with the decomposition theorem. Firstly, they present a "stochastic integral" formula in terms of "accounts". They define the account A_i^t to be the number of times in the first t periods that outcome i occurred, and stack these accounts into a vector $A^t = (A_1^t, \ldots, A_m^t)$. Similarly, the decomposed sharing rules $s_t(x^t; X^{t-1})$ may be stacked into a vector $s_t(X^{t-1})$. Their formula only requires a minor modification to support history dependence,

$$s(X^T) = \sum_{t=1}^{T} s_t(X^{t-1}) \cdot (A^t - A^{t-1}) + w_0,$$

where $s_t(X^{t-1})$ replaces $s_t(p^t(X^{t-1}))$.

Secondly, they remark that if the CARA utility assumption is dropped, then a weaker formula

$$s(X^T) = w_0 + \sum_{t=1}^{T} \left\{ s_t[x^t; p^t(X^{t-1}), w_{t-1}(X^{t-1})] - w_{t-1}(X^{t-1}) \right\},$$

applies, where $s_t(\cdot; p, w)$ implements p with certain equivalent w in the single-period problem.

REFERENCES

Bolton, P., and M. Dewatripont (2005): *Contract Theory*. MIT Press, Cambridge, Massachusetts.

Holmstrom, B., and P. Milgrom (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55(2), 303–328.