Wasteful sanctions, underperformance, and endogenous supervision

Online Appendix

By David A. Miller and Kareen Rozen*

This online appendix studies optimality within a special class of contracts: those with increasingly harsh marginal penalties, which we term "decreasing convex" contracts. Within this class, simple contracts with work-target strategies and kinked-linear sanctioning schemes are optimal.
JEL: C72, D03, D86
Keywords: Partnership, teams, moral hazard, monitoring, supervision, informal sanctions

Optimal convex contracts

In this online appendix we consider contracts that are symmetric with respect to task names and for which the amount of monitoring to be accomplished (denoted F) is public. In this case, the sanction depends on the number of failures f of inspection, where $f \in \{0, 1, \ldots, F\}$. Within this class, contracts which deliver increasingly large sanctions for larger numbers of inspection failures may be a focal class to consider. Such *decreasing convex (DC)* contracts satisfy the restriction $v(f) - v(f+1) \geq v(f-1) - v(f) \geq 0$. Convex contracts may be natural in settings where sanctions are imposed by third parties who are more inclined to exact sanctions if they perceive a consistent pattern of failures. Conversely, a non-convex contract may be particularly difficult to enforce via an affected third party, since it would require leniency on the margin for relatively large injuries. For arbitrary capacity M, we show that DC contracts optimally induce work target strategies. Furthermore, the optimal such contract forgives failures up to some threshold, and increases the sanction linearly thereafter.

Theorem 5. *For any M, work-target strategies with a kinked linear sanctioning scheme are optimal in the class of DC contracts.*

We first prove several lemmas. The first provides a sufficient condition on a one-parameter family of probability distributions for the expectation of a concave function to be concave in the parameter. Though it can be derived as a corollary of

* Miller: University of Michigan, Dept. of Economics, 611 Tappan St., Ann Arbor, MI 48109, econdm@umich.edu. Rozen: Yale, Dept. of Economics and the Cowles Foundation for Research in Economics, 30 Hillhouse Ave., New Haven, CT 06511, kareen.rozen@yale.edu.

a more general theorem of Susan Athey (2000), we provide a simple statement of the condition along with a direct proof. We say that a function $\psi : \{0,1,\ldots,R\} \to \mathbb{R}$ is *concave* if $\psi(r+1)-\psi(r) \le \psi(r)-\psi(r-1)$ for all $r = 1,\ldots,R-1$. A function $\phi : \mathcal{Z} \to \mathbb{R}$, where $\mathcal{Z} \subseteq \mathbb{R}$, is *double crossing* if there is a (possibly empty) convex set $A \subset \mathbb{R}$ such that $A \cap \mathcal{Z} = \big\{z \in \mathcal{Z} : \phi(z) < 0\big\}$.

Lemma 5 (Preservation of concavity)**.** *Let $\mathcal{R} = \{0,1,\ldots,R\}$, and let $\{q_z\}_{z\in\mathcal{Z}}$ be a collection of probability distributions on $\mathcal{R}$ parameterized by $z \in \mathcal{Z} = \{0,1,\ldots,Z\}$.*[1] *The function $\Psi(z) = \sum_{r=0}^{R} \psi(r) q_z(r)$ is concave if*

1) *There exists $k, c \in \mathbb{R}$, $k \neq 0$, such that $z = k\sum_{r=0}^{R} r q_z(r) + c$ for all $z \in \mathcal{Z}$;*
2) *$q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)$ for all $z = 1,\ldots,Z-1$, as a function of r, is double crossing;*
3) *$\psi : \{0,1,\ldots,R\} \to \mathbb{R}$ is concave.*

Proof. Since $z = k\sum_{r=0}^{R} r q_z(r) + c$, there exists $\hat{b} \in \mathbb{R}$ such that $\sum_{r=0}^{R}(mr + b)q_z(r) = \frac{m}{k}z + \hat{b} + c$ for any real m and b. Hence, for any m and b,

$$\text{(B1)} \qquad \sum_{r=0}^{R}(mr+b)\big(q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)\big) = \frac{m}{k}\big(z + 1 - 2z + z - 1\big) = 0,$$

for all $z = 1,\ldots,Z-1$. Therefore, for any m and b, the second difference of $\Psi(z)$ is

(B2)

$$\begin{aligned}\Psi(z+1) - 2\Psi(z) + \Psi(z-1) &= \sum_{r=0}^{R} \psi(r)\big(q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)\big)\\ &= \sum_{r=0}^{R}\big(\psi(r) - mr - b\big)\big(q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)\big).\end{aligned}$$

By assumption, $q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)$, as a function of r, is double crossing. Furthermore, since ψ is concave, we can choose m and b such that, wherever $\big(q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)\big)$ or $\frac{\partial^2}{\partial z^2} q_z(r)$ is nonzero, $\psi(r) - mr - b$ either has the opposite sign or is zero. From Eq. B2 we may conclude $\Psi(z)$ is concave. □

The next lemma says that the expected sanctioning scheme will be decreasing convex in the number of tasks completed.[2]

[1] A similar result holds if $z \in \mathcal{Z} = [0,1]$.

[2] Recall that $g(f,a) \equiv \sum_{k=f}^{F_i} \frac{\binom{p_i - a}{k}\binom{a}{F_i - k}}{\binom{p_i}{F_i}} \binom{k}{f} \gamma^f (1-\gamma)^{k-f}$.

Lemma 6. *If v is decreasing convex, then $h_v \equiv \sum_{f=0}^{F} v(f)g(f,\cdot)$ is decreasing convex.*

Proof. By letting $a \equiv |A|$, reversing the order of summation, and using fact that $\binom{k}{f} = 0$ when $k < f$, we can write $h_v(A)$ as follows:

$$\begin{aligned}
h_v(A) &= \sum_{f=0}^{F} g(f,a)v(f) \\
&= \sum_{f=0}^{F}\left(\sum_{k=0}^{F}\frac{\binom{p-a}{k}\binom{a}{F-k}}{\binom{p}{F}}\binom{k}{f}\gamma^f(1-\gamma)^{k-f}\right)v(f) \qquad \text{(B3)}\\
&= \sum_{k=0}^{F}\frac{\binom{p-a}{k}\binom{a}{F-k}}{\binom{p}{F}}\left(\sum_{f=0}^{F}\binom{k}{f}\gamma^f(1-\gamma)^{k-f}v(f)\right).
\end{aligned}$$

Therefore, the expectation is first with respect to the binomial, and then with respect to the hypergeometric. Applying Lemma 5 twice gives the result. First, note that the expectation of the binomial is γk, a linear function of k, while the expectation of the hypergeometric is $\frac{F}{p}(p-a)$, a linear function of a. Hence it suffices to show that the binomial second-difference in k is double-crossing in f (hence the inside expectation is decreasing convex in k) and the hypergeometric second-difference in a is double-crossing in k. To see this is true for the binomial, note that we may write the binomial second-difference in k as

$$\binom{k}{f}\gamma^f(1-\gamma)^{k-f}\left(\frac{(k+1)(1-\gamma)}{k+1-f} - 2 + \frac{k-f}{k(1-\gamma)}\right). \qquad \text{(B4)}$$

It can be shown that the term in parentheses is strictly convex in f and therefore double crossing in f, so the whole expression is double-crossing in f. To see this is true for the hypergeometric, note that we may write the hypergeometric second-difference in a as

$$\frac{\binom{p-a}{k}\binom{a}{F-k}}{\binom{p}{F}}\left(\frac{p-a-k}{p-a}\cdot\frac{a+1}{a+1-F+k} - 2 + \frac{p-a+1}{p-a+1-k}\cdot\frac{a-F+k}{a}\right). \qquad \text{(B5)}$$

It can be shown that the term in parentheses has either zero or two real roots.[3] If there are no real roots, then the term in parentheses is double-crossing in k (the region in which it is negative must be convex, but may be empty), and thus the whole expression is double-crossing in k. If there are two real roots, it can be shown that the derivative with respect to k is negative at the smaller root, and

[3] The term in parentheses does not account for the fact that the entire expression equals zero whenever $k > p-a$ or $F-k > a$. However, on the closure of these regions the second difference cannot be negative, and so these regions may be ignored.

thus both the term in parentheses and the whole expression are double-crossing in k. $\square$

Proof of Theorem 5. Fix any p, F, and λ. Suppose strategy s, with $p^* > 0$ the maximal number of tasks completed, is optimal. Consider the decreasing convex contract v that implements s at minimum cost. Because v is decreasing, MLRP (or FOSD in a) implies the expected sanction decreases in the number of completed tasks: $h(a) > h(a-1)$ for all a. By contradiction, suppose the downward constraint for p^* versus p^*-1 is slack: $h(p^*) - h(p^*-1) > c - b$. By Lemma 6 and monotonicity, for any $k > 1$, $h(p^*-k+1) - h(p^*-k) > c - b$. But then for any a with $s(a) = a$ and every $a' < a$, the downward constraint $h(a) - h(a') = \sum_{k=a'}^{a-1} h(k+1) - h(k) \geq (a-a')(c-b)$ is slack. Some constraint must bind at the optimum, else the strategy is implementable for free, so the downward constraint for p^* versus p^*-1 must bind. Again, each downward constraint is satisfied, and for any $a > p^*$, $h(a) - h(p^*) < (a-p^*)(c-b)$. So the strategy s has a work target of p^*.

Suppose we look for the optimal convex contract with p assigned tasks, F monitoring slots, and strategy s with work target p^*. By the above, the only binding incentive constraint is the downward constraint for completing p^* tasks. Since $v(0) = 0$, convexity implies monotonicity. The constraint $v(0) \geq 0$ does not bind,[4] so the cost minimization problem in primal form is

$$\tag{B6} \begin{aligned} &\max_{(-v)\geq\vec{0}} \sum_{f=0}^{F} \Big(-(-v(f)) \sum_{a=0}^{p} -g(f,a) t_s(a)\Big) \text{ subject to} \\ &\sum_{f=0}^{F} (-v(f)) \big(g(f,p^*) - g(f,p^*-1)\big) \leq -(c-b), \\ &2(-v(f)) - (-v(f+1)) - (-v(f-1)) \leq 0 \text{ for all } f = 1,\ldots,F-1, \end{aligned}$$

where $t_s(a) = \sum_{a'=a}^{p} \mathbb{I}\big(s(a') = a\big) \binom{p}{a'} \lambda^{a'} (1-\lambda)^{p-a'}$ is the probability of completing a tasks given strategy s. Let x be the Lagrange multiplier for the incentive compatibility constraint, z_f the multiplier for the convexity constraint $2(-v(f)) - (-v(f+1)) - (-v(f-1)) \leq 0$, and $\vec{z}$ the vector $(z_1, \ldots, z_{F-1})$. The constraint

[4] Although $v(0) \geq 0$ is satisfied with equality, the binding constraint on $v(0)$ is actually $v(0) \leq 0$.

set can be written $A^\top \cdot (-v(0), \ldots, -v(F))$, where, in sparse form,

$$A = \begin{pmatrix} g(0,p^*) - g(0,p^*-1) & -1 & & \\ \vdots & 2 & \ddots & \\ \vdots & -1 & \ddots & -1 \\ \vdots & & \ddots & 2 \\ g(F,p^*) - g(F,p^*-1) & & & -1 \end{pmatrix}. \tag{B7}$$

Let r be the vector of dual variables: $r = (x, z_1, \ldots, z_{F-1})$. The dual problem is

$$\min_{r \geq \vec{0}} (b-c)x \quad \text{s.t. } (Ar)_f \geq -\sum_{a=0}^{p} g(f,a)t_s(a) \text{ for all } f = 0, 1, \ldots, F, \tag{B8}$$

where $(Ar)_f$ is the (f)th component of $A \cdot r$; i.e.,

$$(Ar)_f = x\big(g(f,p^*) - g(f,p^*-1)\big) - z_{f-1} + 2z_f - z_{f+1}, \tag{B9}$$

where we define $z_0 \equiv 0$, $z_F \equiv 0$, and $z_{F+1} \equiv 0$. Let $\hat{f}$ be the smallest f with $v(f) < 0$. It must be that $v(f) < 0$ for all $f \geq \hat{f}$, so by duality, $(A \cdot r)_f \geq -\sum_{a=0}^{p} g(f,a)t_s(a)$ binds for all $f \geq \hat{f}$. Hence

$$x = \frac{\sum_{a=0}^{p} g(f,a)t_s(a) - z_{f-1} + 2z_f - z_{f+1}}{g(f,p^*-1) - g(f,p^*)} \text{ for all } f = \hat{f}, \ldots, F. \tag{B10}$$

In particular, this means that if $z_{F-1} = 0$ (implied for $\hat{f} = F$) then the optimal contract (which would have expected sanction $-x(c-b)$) has the same value as that derived in Lemma 4, completing the claim. Henceforth we assume $z_{F-1} > 0$. The sum of the z-terms over $(A \cdot r)_{-1}$ and $(Acdotr)_F$ is $-z_{F-1} + (2z_{F-1} - z_{F-2}) = z_{F-1} - z_{F-2}$. Note also the corresponding sum of z-terms over $F-2$, $F-1$, and F: $-z_{F-1} + (2z_{F-1} - z_{F-2}) + (-z_{F-3} + 2z_{F-2} - z_{F-1}) = z_{F-2} - z_{F-3}$. Iterating, the sum of the z-terms in $(A \cdot r)_f$ from any $\tilde{f} \geq \hat{f}$ to F is $z_{\tilde{f}} - z_{\tilde{f}-1}$. Summing the equalities in Eq. B10 thus yields a recursive system for $z_{\tilde{f}}$ for all $\tilde{f} = \hat{f}, \ldots, F$:

$$z_{\tilde{f}} = z_{\tilde{f}-1} - \sum_{f=\tilde{f}}^{F} \sum_{a=0}^{p} g(f,a)t_s(a) + x \sum_{f=\tilde{f}}^{F} \big(g(f,p^*-1) - g(f,p^*)\big). \tag{B11}$$

By definition, the convexity constraint is slack at $\hat{f}-1$, so $z_{\hat{f}-1} = 0$. By induction, for $f' = \hat{f}, \ldots, F$,

$$\text{(B12)} \quad z_{f'} = -\sum_{\tilde{f}=\hat{f}}^{f'} \sum_{f=\tilde{f}}^{F} \sum_{a=0}^{p} g(f,a)t_s(a) + x \sum_{\tilde{f}=\hat{f}}^{f'} \sum_{f=\tilde{f}}^{F} \big(g(f,p^* - 1) - g(f,p^*)\big).$$

Plugging Eq. B12 for $f' = F$ into the binding constraint $(Ar)_F \geq -\sum_{a=0}^{p} g(F,a)t_s(a)$ yields:

$$\text{(B13)} \quad x = \frac{\sum_{\tilde{f}=\hat{f}}^{F} \sum_{f=\tilde{f}}^{F} \sum_{a=0}^{p} g(f,a)t_s(a)}{\sum_{\tilde{f}=\hat{f}}^{F} \sum_{f=\tilde{f}}^{F} \big(g(f,p^*-1) - g(f,p^*)\big)}.$$

The expectation of a random variable X on $\{0,\ldots,n\}$, is $\sum_{j=1}^{n} j \Pr(X=j)$, which also equals $\sum_{j=1}^{n} \Pr(X \geq j)$. Since $\sum_{f=\tilde{f}}^{F} \sum_{a=0}^{p} g(f,a)t_s(a) = \Pr(f \geq \tilde{f})$, the numerator of Eq. B13 equals

$$\text{(B14)} \quad \sum_{\tilde{f}=\hat{f}}^{F} \sum_{f=\tilde{f}}^{F} \sum_{a=0}^{p} g(f,a)t_s(a) = \sum_{\tilde{f}=\hat{f}}^{F} \Pr(f \geq \tilde{f}) = \sum_{\tilde{f}=\hat{f}}^{F} (\tilde{f} - \hat{f} + 1) \Pr(f = \tilde{f})$$
$$= \sum_{\tilde{f}=1}^{F} (\tilde{f} - \hat{f} + 1)_+ \Pr(f = \tilde{f}) = \mathbb{E}\big((f - \hat{f} + 1)_+\big) \equiv \mathbb{E}\big(\phi(\hat{f})\big),$$

where $(y)_+ \equiv \max\{y, 0\}$ and ϕ is the random function $\phi(\hat{f}) \equiv (f - \hat{f} + 1)_+$. In words, $\phi(\hat{f})$ is the number of discovered unfulfilled tasks that exceed the threshold for sanctions $\hat{f}$. The denominator of Eq. B13 can be rewritten similarly, yielding

$$\text{(B15)} \quad x = \frac{\mathbb{E}\big(\phi(\hat{f})\big)}{\mathbb{E}\big(\phi(\hat{f}) \mid a = p^* - 1\big) - \mathbb{E}\big(\phi(\hat{f}) \mid a = p^*\big)}.$$

The minimized expected sanction is $\mathbb{E}\big(v(f)\big) = (b - c)x$, and is implemented by

$$v(f) = -\frac{(c-b)(f - \hat{f} + 1)_+}{\mathbb{E}\big(\phi(\hat{f}) \mid a = p^* - 1\big) - \mathbb{E}\big(\phi(\hat{f}) \mid a = p^*\big)} \text{ for all } f = 0, 1, \ldots, F. \quad \square$$

*

REFERENCES

Athey, Susan. 2000. “Characterizing properties of stochastic objective functions.” Working paper.