Draft: February 2016

Prison Work Programs in a Model of Deterrence

A. Mitchell Polinsky*

Abstract: This article considers the social desirability of prison work programs in a model in which the function of imprisonment is to deter crime. Two types of prison work programs are studied—voluntary ones and mandatory ones. A voluntary work program is socially beneficial: if prisoners are paid a wage that just compensates them for their disutility from work, the deterrent effect of the prison sentence is unaffected, but society obtains the product of the work program. But in the basic model considered, a mandatory work program is superior to a voluntary work program: if prisoners are forced to work without compensation, the deterrent effect of the prison sentence rises, allowing society to restore deterrence and save resources by reducing the probability of detection or the sentence length, and also to obtain greater output than under the optimal voluntary work program. In an extension of the basic model, however, in which prisoners vary in their disutility from work, a voluntary work program may be superior to a mandatory work program because prisoners with relatively high disutility from work can elect not to work.

Key words: deterrence; imprisonment; prison work programs; prison costs; prison labor

JEL codes: H23; K14; K42

---

* Stanford Law School and National Bureau of Economic Research. Research on this article was supported by the John M. Olin Program in Law and Economics at Stanford Law School. Helpful comments were provided by Steven Shavell and valuable research assistance by Paul Riskind.

1. Introduction

Many countries require prisoners to perform labor during the terms of their incarceration, and such practices often have been in place for centuries.[1] In the United States, for instance, prison work has been a significant feature of penal policy since the late eighteenth century.[2] Work programs of some kind exist in 88 percent of the nation's prisons, employing approximately 775,000 prisoners, and producing goods and services valued at an estimated $2 billion per year.[3] Prisoners perform myriad jobs, from farming, to making furniture and clothing, to maintaining the infrastructure of prisons themselves (painting prison buildings, caring for prison grounds, and so forth).

In this article I undertake an analysis of the social desirability of prison work programs within the conventional economic model of deterrence.[4] Work programs have the obvious benefit of generating valuable output, but they can have two conflicting effects on deterrence. On one hand, the disutility that prisoners bear from working increases the deterrent effect of any given prison term.[5] On the other hand, any compensation that prisoners receive for their work

---

[1] See England (1955), p. ix (listing 38 countries surveyed) and pp. 1-3 (describing scope of obligations to work and noting that many countries imposed such obligations by statute in the eighteenth and nineteenth centuries). See also Gibson (2011) on the history of prison labor in France (p. 1045), Holland and Germany (p. 1048), colonial Indochina (p. 1051-2), colonial Africa (pp. 1053, 1058), China (pp. 1054, 1058), and Japan (pp. 1055, 1058).

[2] See Barnes (1921) and American Correctional Association (1986, pp. 1-8).

[3] See Bureau of Justice Statistics (2008, p. 5, Table 6; p. 24, Appendix Table 16) for the number of work programs and participating inmates, respectively; and Correctional Industries Association (1999) for the value of output produced with prison labor.

[4] There is no discussion of prison work programs in four major surveys of the economics of crime and punishment — Garoupa (1997), Levitt and Miles (2007), Mookherjee (1997), and Polinsky and Shavell (2007). Nor is this topic addressed in the only survey of which I am aware of the economics of prisons, Avio (1998).

[5] The strength of this effect is diminished to the extent that prisoners enjoy some work; see comment (b) in Section 4 below.

decreases the deterrent effect of imprisonment. Thus, an assessment of the desirability of prison work programs must take into account both the direct benefits of prison output and the indirect effects on deterrence.[6]

I consider two types of prison work programs: voluntary ones in which inmates are offered opportunities to work, which they are free to accept or reject, and are paid a wage; and mandatory ones that require prisoners to work, and that might include some compensation.[7] I assess the social desirability of a prison work program compared to no work program; derive the socially optimal features of each type of prison work program (such as the optimal wage rate in a voluntary program); and compare the merits of a mandatory work program to a voluntary one.

I first analyze prison work programs in Section 2 in a benchmark model in which individuals are identical except for the potential gain they would obtain from committing a crime. In particular, were they to participate in a prison work program, it is assumed that they would each have the same productivity and would each suffer the same disutility from the effort to work.

A voluntary work program is superior to no work program in the benchmark case. To understand why, suppose that the wage in the voluntary program equals the disutility that prisoners would bear from working, and assume that prisoners would choose to work at this wage. A voluntary work program with this wage would leave the disutility of the prison

---

[6] I distinguish between a prison *work program* and a prison *vocational training program*. A work program, as I use that term, is for the purpose of producing valuable output within prisons given the present skills of prisoners, whereas a vocational training program is focused on developing prisoners' skills so as to better prepare them for productive work outside of prison after they complete their sentences. In practice, of course, there is not such a rigid distinction since prisoners may learn some useful skills in work programs and may produce some useful output in vocational training programs. But I will not in this paper be concerned with the enhancement of skills by prisoners or with their prospects for employment after serving their sentences (but see comment (d) in Section 4 below).

[7] In practice, mandatory work programs appear to predominate; see comment (e) in Section 4 below.

sentence—and thus deterrence—unchanged. But taxpayers are better off because they obtain the value of prisoners' output. Hence, a voluntary work program would be socially desirable.[8]

However, a mandatory work program can function even better than a voluntary work program. In essence, that is because prisoners are made worse off by being forced to work without pay. This results in an increase in deterrence, everything else equal. The level of welfare of prisoners, and hence deterrence, can be restored either by shortening the prison term or by lowering the probability of detection, either of which would improve the welfare of taxpayers by lowering public costs. Because taxpayers obtain this benefit in addition to the value of prisoners' output, while the level of welfare of prisoners is unaffected, a mandatory work program is superior to a voluntary one.[9]

In Section 3 I reconsider the merits of prison work programs under two extensions of the benchmark model. I first explain that if prisoners differ in terms of their productivity, the ranking of the policies is not affected. For essentially the same reasons as those discussed above, a mandatory work program is preferable to a voluntary one, which in turn is preferable to no work program.

However, if prisoners differ in terms of their disutility from work, the ranking of the policies might change. Although a voluntary prison work program would continue to be superior to no work program, a mandatory work program no longer is necessarily preferable to a voluntary work program. The potential advantage of a voluntary work program is that it allows

[8] I demonstrate that the optimal voluntary prison work program offers prisoners the opportunity to work at the first-best number of hours and pays them the lowest possible wage that induces them to work.

[9] I show that the optimal mandatory prison work program requires prisoners to engage in maximal work (exceeding the level of work under the voluntary program) and does not compensate them for this work.

prisoners with relatively high disutility from work to opt out of working. Everything else equal, this is socially desirable if their disutility from work exceeds the value of their output.

Section 4 concludes with several comments, including on the role of prison work programs when the purpose of imprisonment is incapacitation or rehabilitation.

2. Prison Work Programs in the Benchmark Model

In this section I begin by describing the standard model of deterrence through imprisonment, in which prison work programs are absent, and then analyze voluntary and mandatory work programs within this framework.

A. The Benchmark Model[10]

Risk-neutral individuals contemplate committing a harmful act in order to obtain a benefit that varies among them. If they commit the act, they are caught with some positive probability. They are assumed to have no wealth, so that the only sanction that can be imposed on them is a prison sentence.

Let

h = harm caused if the offense is committed; $h > 0$;

b = benefit to an individual from committing the offense; $b \geq 0$;

$v(b)$ = density of b among individuals; $v(b) > 0$ for all $b \geq 0$;

p = probability of catching an offender; $p > 0$;

$k(p)$ = cost to the state per unit time to maintain the probability p; $k'(p) > 0$; $k''(p) > 0$;

s = prison sentence for the offense; $s > 0$; and

c = cost to the state per unit time to imprison an individual.

[10] This model is well known; see, for example, Polinsky and Shavell (2007).

Sentences are measured in units of time corresponding to one dollar of disutility.[11]

An individual will commit an offense if his benefit b exceeds the expected disutility of the sentence ps.[12]

Social welfare equals the benefits that individuals obtain from committing harmful acts, less the harms they cause, less the disutility that they bear from imprisonment, and less the state's cost of imprisoning individuals and maintaining the probability of detection. Hence, social welfare in the absence of a prison work program is

$$\int_{ps}^{\infty}[b - h - ps(1 + c)]v(b)db - k(p). \tag{1}$$

The state's problem is to choose the probability of detection p and the sentence s to maximize (1). Let p_N^* and s_N^* be the optimal probability and sentence in the benchmark model (subscript "N" for "no" prison work program). For present purposes, it is not necessary to characterize these results.[13]

B. A Voluntary Prison Work Program

Now consider a voluntary work program in which prisoners choose whether to participate given the wage offered per unit time. Because of institutional constraints on the supervision of prisoners, I assume that they are not free to choose how many hours to work, only whether to participate in a work program chosen by the state with a specified number of hours of work.[14]

---

[11] Implicit in this construction is that the disutility of a sentence is proportional to its length. My results do not depend on this assumption, but it simplifies the analysis.

[12] There is no loss of generality in assuming that an individual will not commit the offense when he is indifferent.

[13] It will also be the case that I do not need to derive the optimal probability of detection and sentence under the voluntary and mandatory prison work programs.

[14] One could imagine a more general voluntary work program in which the state simply sets the wage and allows prisoners to work as many hours as they wish at this wage. This scheme is problematic, however, because

Let

$\bar{\lambda}$ = maximum fraction of time that a prisoner is capable of working per period; $\bar{\lambda} < 1$;[15]

λ_V = fraction of time prisoners work per period under a voluntary work program;

$0 \leq \lambda_V \leq \bar{\lambda}$;

$d(\lambda)$ = disutility of time worked per period; $d(0) = 0$; $d'(\lambda) > 0$; $d''(\lambda) > 0$;

n = output per period if full-time work (that is, if $\lambda = 1$); $n > 0$; and

w_V = wage per period if full-time work under a voluntary work program; $w_V \geq 0$.

For notational simplicity, I will in the present section drop the subscript on w (other than in the statements of the propositions). The disutility from working is in addition to the normal disutility of time in prison. Output is measured in units that correspond to one dollar of value. I make the realistic assumption that the value of prison work per period is less than the public cost of imprisonment per period, that is, $\bar{\lambda}n < c$.[16]

It will be useful to define the level of work that maximizes the net social value of work. Let

λ^* = first-best level of work per period,

---

prisoners are not free to come and go at different times. Often they are supervised by guards during their work and are transported to and from work as a group under guard. Hence, the most realistic characterization of a voluntary prison work program is one in which prisoners are free to accept or reject the opportunity to work a fixed number of hours at a specified wage.

[15] It is immaterial to my results what the magnitude of $\bar{\lambda}$ is, though it is certainly less than 1.

[16] Based on the evidence cited in the first paragraph of this article, the average value produced per prisoner in a work program is $10,000 per year, approximately one-third of the average public cost of maintaining a prisoner per year.

where λ^* maximizes the value of the output of work net of the disutility incurred to produce it, $\lambda n - d(\lambda)$. I assume that $\lambda^* > 0$, which implies that $n > d'(0)$, and that $\lambda^* < \bar{\lambda}$.[17] Thus, λ^* is an interior solution satisfying $d'(\lambda) = n$.

If a prisoner chooses to participate in the voluntary work program, he will earn $\lambda_V w$ and suffer disutility $d(\lambda_V)$ per period. Thus, assuming that he will participate if he is indifferent, he will participate if $\lambda_V w \geq d(\lambda_V)$ or, equivalently, if $w \geq d(\lambda_V)/\lambda_V$. Let

$\tilde{w}(\lambda_V)$ = minimum wage at and above which a prisoner would choose to participate in a voluntary work program and below which he would decline to participate,

where $\tilde{w}(\lambda_V) = d(\lambda_V)/\lambda_V$.

Assuming $w \geq \tilde{w}$, an individual will commit an offense if his benefit b exceeds the expected disutility of the sentence, which is $ps(1 + d(\lambda_V) - \lambda_V w)$. Relative to the situation in which there is no work program, he now suffers additional disutility from work $d(\lambda_V)$, but obtains wage income $\lambda_V w$.

Social welfare will reflect not only the factors that were included in the absence of a work program—the benefit from offenses less their harm, the disutility of imprisonment, and the public cost of imprisonment and enforcement—but also the disutility that prisoners bear from working and the value of their output. Accordingly, if $w \geq \tilde{w}$, social welfare under a voluntary prison work program is

$$\int_{ps(1 + d(\lambda_V) - \lambda_V w)}^{\infty} [b - h - ps(1 + c + d(\lambda_V) - \lambda_V n)]v(b)db - k(p). \qquad (2)$$

[17] That this is a reasonable assumption is suggested by the following observations. A prisoner could conceivably work fifteen hours or more per day. Given increasing marginal disutility from work, and the relatively low value of the output of prison work (trash removal along roads, manufacturing license plates, and so forth), it is implausible that the first-best level of work effort would be this high.

Note that because the wage is a transfer payment, it only affects social welfare through its effect on the level of deterrence (the lower bound of the integral).

The state's problem is to choose p, s, λ_V, and w to maximize (2), subject to the constraint that $\lambda_V \leq \bar{\lambda}$. The optimal values will be designated with an asterisk.

Proposition 1: A voluntary prison work program is superior to not having a prison work program.

Comment: Even though I assumed that some work is socially desirable—that is, $d'(0) < n$—it does not necessarily follow that a voluntary prison work program is superior to no work program. This is because, to motivate prisoners to work, they have to be paid a wage, which will reduce the deterrent effect of the prison sentence, everything else equal. But by paying prisoners a wage that just compensates them for their disutility from work, they can be induced to work without being deterred less, resulting in a social gain due to their productive output.

Proof: Set $p = p_N^*$, $s = s_N^*$. I will show that, given these values, there exists a voluntary prison work program that results in higher social welfare (2) than without a work program (1). For any $\lambda_V > 0$ under the voluntary work program, let the wage be $\tilde{w}(\lambda_V) = d(\lambda_V)/\lambda_V$. Then the level of deterrence will be the same as that without a work program since $ps(1 + d(\lambda_V) - \lambda_V \tilde{w}(\lambda_V)) = p_N^* s_N^*$. Thus, social welfare under the voluntary work program will be higher if the integrand in (2) exceeds that in (1), that is, if

$$b - h - ps(1 + c + d(\lambda_V) - \lambda_V n) > b - h - p_N^* s_N^* (1 + c), \tag{3}$$

which will hold if $\lambda_V n - d(\lambda_V) > 0$. Let $f(\lambda_V) = \lambda_V n - d(\lambda_V)$. Clearly, $f(0) = 0$. Moreover, $f'(0) = n - d'(0) > 0$, where the inequality follows from the assumption that $\lambda^* > 0$. Hence, there exists a $\lambda_V > 0$ in the neighborhood of 0 such that (3) holds. □

The next proposition characterizes the optimal voluntary prison work program.

Proposition 2: The optimal voluntary prison work program:

(a) offers prisoners the opportunity to work at the first-best level of effort, $\lambda_V^* = \lambda^*$; and

(b) employs the lowest possible wage that induces prisoners to work, which is less than the marginal value of the output of prison work, $w_V^* = \tilde{w}(\lambda^*) = d(\lambda^*)/\lambda^* < n$.

Comments: (a) It is not surprising that the optimal voluntary prison work program sets the level of work at the first-best level, since this maximizes the net social value of work.

(b) In a private labor market, in which workers can choose how much to work, they would need to be paid a wage equal to the marginal value of their output in order to induce them to choose the first-best level of work. But in a voluntary prison labor market, where the level of work is chosen by the state and can be set at the first-best level, prisoners only need to be paid enough to induce them to prefer working to not working. In other words, they only have to be paid enough to compensate them for their average disutility from work. Since their marginal disutility from work is increasing, their average disutility from work is less than their marginal disutility. This implies that, if prisoners are paid the lowest possible wage that induces them to work, they will be receiving less than the marginal value of their work.

(c) The reason that the optimal wage to employ in a voluntary prison labor program is the lowest possible wage that induces prisoners to work should by now be familiar. Any higher wage would reduce deterrence and require costly public expenditures on enforcement or imprisonment to offset this reduction.

Proof: I first establish that, for any voluntary work program $\lambda_V > 0$ in which prisoners are induced to work, the optimal wage equals $\tilde{w}(\lambda_V) = d(\lambda_V)/\lambda_V$. The optimal wage cannot be lower than this since prisoners would then choose to not work. Suppose the wage were higher, say $\hat{w} >$

$\tilde{w}(\lambda_V)$. Then the wage could be lowered to $\tilde{w}(\lambda_V)$, which would increase deterrence, and the probability of detection could be lowered to some $p^{\circ} < p$ so as to return deterrence to its original level: $p^{\circ}s(1 + d(\lambda_V) - \lambda_V\tilde{w}) = ps(1 + d(\lambda_V) - \lambda_V\hat{w})$.[18] This would have two beneficial effects on social welfare (2). First, the cost of maintaining the probability of detection would decline. Second, the expected sentence length ps would decline, which would raise social welfare if $(1 + c + d(\lambda_V) - \lambda_V n) > 0$; that this inequality holds follows from the assumption that $\bar{\lambda}n < c$. Hence, the optimal wage is $\tilde{w}(\lambda_V) = d(\lambda_V)/\lambda_V$.

Given the preceding result, the level of deterrence, $ps(1 + d(\lambda_V) - \lambda_V\tilde{w}(\lambda_V))$, equals ps regardless of λ_V. Thus, the optimal λ_V can be derived by maximizing the integrand of social welfare (2). The resulting first-order condition is $-ps(d'(\lambda_V) - n) = 0$, implying that $d'(\lambda_V) = n$; hence, $\lambda_V^* = \lambda^*$.

Since $d(0) = 0$, $d'(\lambda) > 0$, and $d''(\lambda) > 0$, it follows that $d'(\lambda) > d(\lambda)/\lambda$. Thus, $\tilde{w}(\lambda^*) = d(\lambda^*)/\lambda^* < d'(\lambda^*) = n$. □

C. A Mandatory Prison Work Program

Next consider a mandatory work program in which prisoners are told how much they have to work and might be compensated for this work. Let

λ_M = fraction of time prisoners are required to work per period under a mandatory work program; $0 \leq \lambda_M \leq \bar{\lambda}$; and

w_M = wage per period if full-time work under a mandatory work program; $w_M \geq 0$.

For notational simplicity, I will drop the subscript on w when there is no ambiguity.

[18] In this proof, as in subsequent proofs, it would be possible to obtain the same result by lowering the

Since the expected disutility of the sentence now is $ps(1 + d(\lambda_M) - \lambda_M w)$, social welfare under a mandatory prison work program is

$$\int_{ps(1 + d(\lambda_M) - \lambda_M w)}^{\infty} [b - h - ps(1 + c + d(\lambda_M) - \lambda_M n)]v(b)db - k(p). \quad (4)$$

The state's problem is to choose p, s, λ_M, and w to maximize (4), subject to the constraint that $\lambda_M \leq \overline{\lambda}$.

Proposition 3: A mandatory prison work program is superior to a voluntary prison work program.

Comment: A mandatory prison work program obviously can duplicate the level of work that would be achieved under a voluntary work program. The fundamental advantage of a mandatory work program stems from the fact that it can make prisoners work regardless of how little they are paid, whereas a voluntary program is constrained by having to offer a wage high enough to induce prisoners to work. Thus, a mandatory work program can make prisoners worse off than they would be under a voluntary work program, in which case the probability of detection or the length of the sentence (or both) can be lowered relative to that under a voluntary work program without lessening deterrence. The resulting savings in enforcement costs or prison costs makes a mandatory work program superior to a voluntary one.

Proof: Let p, s, λ, and w be the probability of detection, sentence, level of work offered, and wage under a voluntary prison work program, where $\lambda \in (0, \overline{\lambda}]$ and $w \geq \tilde{w}(\lambda)$. The resulting level of deterrence will be $ps(1 + d(\lambda) - \lambda w)$. Under the mandatory work program, set $s_M = s$, $\lambda_M = \lambda$, and choose a $w_M < w$. Then it will be possible to select a $p_M < p$ such that $p_M s(1 + d(\lambda_M) - \lambda_M w_M) = ps(1 + d(\lambda) - \lambda w)$, so that the levels of deterrence under the voluntary and mandatory

---

sentence length s rather than the probability p, or by lowering both to some extent. For simplicity in the proofs I just lower p, though in the comments following the proofs I refer to the possibility of lowering p or s (or both).

work programs will be the same. Social welfare under the mandatory program will exceed that under the voluntary program if

$$b - h - p_M s(1 + c + d(\lambda_M) - \lambda_M n) \geq b - h - ps(1 + c + d(\lambda) - \lambda n), \quad (5)$$

which follows from $\bar{\lambda} n \leq c$ (implying that the expression in parentheses is positive) and $p_M < p$. Moreover, the cost of enforcement will be lower under the mandatory work program, $k(p_M) < k(p)$. □

Proposition 4: The optimal mandatory prison work program:

(a) requires prisoners to work at the maximal level of effort, $\lambda_M^* = \bar{\lambda}$; and

(b) does not compensate prisoners for their work, $w_M^* = 0$.

Comments: (a) It is easy to see why it is optimal to not compensate prisoners for their work under a mandatory work program. Compensating them would improve their welfare in prison and thereby reduce deterrence, which would be socially costly to offset by raising the probability of detection or lengthening the prison sentence. Obviously, it is not necessary to compensate prisoners under a mandatory prison work program in order to get them to work.

(b) Parallel intuition explains why the optimal work requirement is maximal effort. By requiring prisoners to work as much as possible without compensation, the disutility of time in prison is maximized. This increases deterrence and allows the probability of detection or the sentence to be reduced, saving enforcement or sanctioning costs.

(c) It might be surprising that the optimal work requirement is maximal effort even if, as I assume, this level of effort exceeds the first-best level, $\bar{\lambda} > \lambda^*$. For example, how could it be that it is desirable to require maximal effort even if the corresponding level of disutility imposed on a prisoner, $d(\bar{\lambda})$, exceeds the value of the resulting output, $\bar{\lambda} n$? The explanation is that when the work requirement is raised, the expected disutility from imprisonment, $ps(1 + d(\lambda))$, can be kept

constant by lowering the probability of detection p or the sentence length s. With this adjustment, the aggregate welfare of offenders is unchanged, but society gains both through a reduction of the net cost of imprisonment per unit time and a reduction in the probability of detection or sentence length.

Proof: I first show that $w_M^* = 0$. Suppose otherwise and consider some $w_M > 0$; let p_M, s_M, and λ_M be the optimal probability, sentence, and work requirement under a mandatory work program given this w_M. The resulting level of deterrence is $p_M s_M(1 + d(\lambda_M) - \lambda_M w_M)$. If w_M is lowered to 0, deterrence will rise. Lower p_M to $\hat{p}_M$ so as to restore deterrence to its original level; hence $\hat{p}_M$ solves

$$\hat{p}_M s_M(1 + d(\lambda_M)) = p_M s_M(1 + d(\lambda_M) - \lambda_M w_M). \tag{6}$$

Lowering p_M to $\hat{p}_M$ raises social welfare (4) both because the expected net social costs of imprisonment, $ps(1 + c + d(\lambda) - \lambda n) > 0$, decline; and the cost of enforcement, $k(p)$, declines. Hence, $w_M > 0$ could not have been optimal.

That $\lambda_M^* = \bar{\lambda}$ will be demonstrated in a similar fashion. Consider some $\lambda_M < \bar{\lambda}$ and let p_M, s_M, and $w_M = 0$ be the optimal probability, sentence, and wage rate under a mandatory work program given this λ_M. The resulting level of deterrence is $p_M s_M(1 + d(\lambda_M))$. Raise λ_M to $\bar{\lambda}$ and lower p_M to $\hat{p}_M$ so as to maintain deterrence at its original level, where $\hat{p}_M$ solves

$$\hat{p}_M s_M(1 + d(\bar{\lambda})) = p_M s_M(1 + d(\lambda_M)). \tag{7}$$

This will have two beneficial effects on social welfare (4). First, the expected net social costs of imprisonment decline, that is

$$\hat{p}_M s_M(1 + c + d(\bar{\lambda}) - \bar{\lambda} n) < p_M s_M(1 + c + d(\lambda_M) - \lambda_M n). \tag{8}$$

To see that (8) holds, rewrite it using (7) and divide through by s_M to obtain

$$\hat{p}_M(c - \bar{\lambda}n) < p_M(c - \lambda_M n). \tag{9}$$

The result follows because $\hat{p}_M < p_M$ and $\lambda_M < \bar{\lambda}$. Second, lowering p_M to $\hat{p}_M$ reduces enforcement costs, $k(p)$. Thus, $\lambda_M < \bar{\lambda}$ could not have been optimal. □

3. Extensions

In this section I consider two generalizations of the benchmark model, first allowing prisoners to vary with respect to their productivity and then with respect to their disutility from imprisonment. In each case I discuss the ranking of the two types of prison work programs and the no work program.

A. Prisoners Differ in Productivity

For many types of prison work, there could be significant variations in the productivity of prisoners. Able-bodied younger prisoners may be able to collect much more trash along a highway than older prisoners. Some prisoners may be able to concentrate on complicated tasks better than others and consequently make fewer errors in operating prison machinery, such as machines that stamp out license plates.

Consider a generalization of the model described in Section 2 to allow for variations among prisoners in their productivity. Specifically, let

α = productivity coefficient; $\alpha \geq 0$; and

$z(\alpha)$ = density of α among prisoners; $z(\alpha) > 0$ for all $\alpha \geq 0$.

Thus, the productivity an α-type prisoner is αn. I assume that the mean of α is 1; that the densities of b (the benefit from committing the offense) and α are independent; and that the state only knows the distribution of α among prisoners, not its value for each one.

Voluntary prison work program: Given the state's information, the same work program and wage rate must be offered to all prisoners. Since a prisoner's decision whether to work only depends on these values and the disutility from work, which does not vary among prisoners in the present case, the behavior of prisoners will be the same as in the benchmark model. All will work if $w \geq \tilde{w}$ and none will work if $w < \tilde{w}$.

Assuming $w \geq \tilde{w}$, social welfare under a voluntary work program will now be

$$\int_0^{\infty} \{ \int_{ps(1 + d(\lambda_V) - \lambda_V w_V)}^{\infty} [b - h - ps(1 + c + d(\lambda_V) - \lambda_V \alpha n)] v(b) db \} z(\alpha) d\alpha - k(p). \tag{10}$$

Mandatory prison work program: Although the optimal wage under a mandatory prison work program was shown to be zero in the benchmark model, I will not preclude here the possibility that the wage rate is positive. Hence, social welfare under a mandatory prison work program is in the present case

$$\int_0^{\infty} \{ \int_{ps(1 + d(\lambda_M) - \lambda_M w_M)}^{\infty} [b - h - ps(1 + c + d(\lambda_M) - \lambda_M \alpha n)] v(b) db \} z(\alpha) d\alpha - k(p). \tag{11}$$

Proposition 5: If prisoners differ in their productivity, then the ranking of the policy alternatives is the same as that in the benchmark model:

(a) a voluntary prison work program is superior to not having a prison work program; and

(b) a mandatory prison work program is superior to a voluntary prison work program.

Comment: The ranking of the policy alternatives in the benchmark model is unaffected by variations in the productivity of prisoners since this variability does not affect the disutility borne by prisoners from working, their behavior, or their *total* output. Variation per se in worker productivity is immaterial.

Proof: The proof of part (a) is essentially identical to the proof of Proposition 1. All that needs to be recognized is that, since the mean of α is 1, (10) is the same as (2). Similarly, the proof of part (b) follows from the proof of Proposition 3 since (11) is the same as (4). □

B. Prisoners Differ in Disutility From Work

The disutility of work also could vary widely among prisoners. Working outdoors on a hot and sunny day picking up trash along a highway could be much more distasteful to individuals who grew up in a moderate climate than those who grew up, say, in the Southwest. Working in a noisy machine shop may be more stressful to some prisoners than to others.

To reflect such variations,[19] let

τ = disutility coefficient; $\tau \geq 0$; and

$r(\tau)$ = density of τ among prisoners; $r(\tau) > 0$ for all $\tau \geq 0$.

Thus, the disutility from work for a τ-type prisoner is $\tau d(\lambda)$. I assume that the mean of τ is 1, that the densities of b and τ are independent, and that the state can only observe the distribution of τ, not its value for a particular prisoner.

Voluntary prison work program: Since the work program and wage rate offered for work has to be the same for all prisoners, whether they will choose to work will vary among them. A τ-type prisoner will choose to work if $s(1 + \tau d(\lambda_V) - \lambda_V w_V) \leq s$ or $\lambda_V w_V - \tau d(\lambda_V) \geq 0$. Let

$\tilde{\tau}(\lambda_V, w_V)$ = critical disutility coefficient at and below which a prisoner will choose to work and above which would choose not to work,

where $\tilde{\tau}$ solves $\lambda_V w_V - \tau d(\lambda_V) = 0$, so $\tilde{\tau}(\lambda_V, w_V) = \lambda_V w_V / d(\lambda_V)$.

Social welfare under a voluntary work program will now be

[19] For simplicity, I consider one extension at a time, and therefore here assume that prisoners are equally productive.

$$\int_{0}^{\tilde{\tau}}\{\int_{ps(1+\tau d(\lambda_V)-\lambda_V w_V)}^{\infty}[b-h-ps(1+c+\tau d(\lambda_V)-\lambda_V n)]v(b)db\}r(\tau)d\tau + \int_{\tilde{\tau}}^{\infty}\{\int_{ps}^{\infty}[b-h-ps(1+c)]v(b)db\}r(\tau)d\tau - k(p). \quad (12)$$

The first term reflects the contribution to social welfare by prisoners who choose to work and the second term the contribution by those who decline to work.

Mandatory prison work program: Under the mandatory work program, variations in the disutility from work will affect the level of deterrence, but not whether a prisoner works.

Again allowing for the possibility of a positive wage, social welfare under a mandatory prison work program is in the present case

$$\int_{0}^{\infty}\{\int_{ps(1+\tau d(\lambda_M)-\lambda_M w_M)}^{\infty}[b-h-ps(1+c+\tau d(\lambda_M)-\lambda_M n)]v(b)db\}r(\tau)d\tau - k(p). \quad (13)$$

Proposition 6: If prisoners differ in their disutility from work, then

(a) a voluntary prison work program is superior to not having a prison work program; and

(b) a mandatory prison work program could be superior or inferior to a voluntary prison work program.

Comments: (a) The first part of the proposition can be explained as follows. Given a wage under a voluntary work program, prisoners with relatively low levels of disutility from work will choose to work. The marginal prisoner will be indifferent between working and not working, and therefore will be deterred to the same extent as he would be if there were no work program. But the inframarginal prisoners, whose levels of disutility from working are lower, will be strictly better off by participating in the work program and therefore will be deterred less. The optimal wage rate balances the social benefit from having more prisoners work, as the wage rate is raised, with the social detriment of more crime. Note, however, that starting from a wage

rate of zero, there are no inframarginal prisoners who will be deterred less if the wage rate is raised slightly; the only prisoners who pursue the opportunity to work will be those with negligible disutility from work and who will be indifferent between working and not working. Since there will be no first-order reduction in deterrence, but there will be a first-order enhancement of social welfare due to the productive output of these prisoners, it will always be desirable to employ a positive wage rate. In other words, it will always be better to have a voluntary prison work program than none at all.

(b) With respect to the second part of the proposition, first observe that if the variation among prisoners in the disutility from work is small, then the results tend towards those in the benchmark case in which a mandatory work program is preferable to a voluntary work program. Hence, the main result that needs to be explained here is why, if there is significant variation among prisoners in the disutility of work, a mandatory work program might be inferior to a voluntary one. In essence, the potential advantage of a voluntary work program is that it allows prisoners with relatively high disutility from work to opt out of working; were they forced to work, as they would be under a mandatory work program, their disutility from work could well exceed the value of their output. Put differently, a voluntary work program beneficially harnesses the private information of prisoners about their disutility from work in a way that a mandatory work program cannot.

(c) To illustrate this point more concretely, suppose that there are an equal number of two types of prisoners, those with low disutility from work and those with high disutility from work. Suppose, too, that the value of the output from work is midway between these two levels of disutility and is the same for both groups. Then, if the state requires all prisoners to work—it cannot observe their levels of disutility and distinguish between the two types—the aggregate

value of the output from prison work will be just offset by the aggregate disutility associated with that work. There would be no net social gain from prison work. In contrast, if the state were to use a voluntary work program and set the wage at a level that would just induce prisoners with low disutility from work to accept it, then there would be a net social gain from the work since the value of the output would exceed the low level of disutility.

(d) The preceding advantage of a voluntary work program needs to be balanced against its disadvantage with respect to achieving deterrence. As I have explained, a voluntary work program cannot increase deterrence because, if prisoners were made worse off as a result of it, they would elect not to participate in it. A mandatory work program, however, can increase deterrence by forcing prisoners to work without compensation. This advantage of a mandatory work program was what led the mandatory program to be superior to the voluntary program in the benchmark model. Hence, when prisoners vary in their disutility from work, which work program is preferable depends on a comparison of the disutility-sorting advantage of the voluntary program to the deterrence-enhancing advantage of the mandatory program. In the Appendix I provide an example in which this balancing favors a voluntary prison work program.

Proof: (a) Since a voluntary work program with $w = 0$ is equivalent to not having a work program, I will demonstrate the result of part (a) by evaluating the derivative of social welfare under a voluntary work program (12) with respect to w at $w = 0$, which is

$$[1 - V(ps)][ps\lambda_V^2 n/d(\lambda_V)]r(0) > 0, \tag{13}$$

where $V(.)$is the cumulative distribution function of b. Hence, $w_V^* > 0$, implying that a voluntary work program is superior to no work program.

(b) Proposition 3 established that if there is no variation among prisoners in the disutility from work, a mandatory work program is strictly superior to a voluntary one. Clearly, therefore,

if the variation among prisoners in the disutility from work is sufficiently small, a mandatory work program will remain superior to a voluntary work program. That a mandatory work program can be inferior to a voluntary one when the disutility from work varies among prisoners is demonstrated in the Appendix in an example. □

4. Concluding Remarks

I conclude with several comments.

(a) *What if the productivity of prisoners is under their control?* I assumed that the output of prisoners in work programs was fixed per unit time. In many circumstances, however, a prisoner's output might be influenced by his effort, such as how vigilant he is in picking up trash along a highway or how much he concentrates on the tasks required to make a pair of shoes.[20] When prisoner effort of this sort is unobservable by administrators of prison work programs, there is an argument for paying prisoners for their output even under a mandatory work program, or paying them more than the minimum required to induce them to work under a voluntary work program. The reason, of course, is that this will generate more output. Prisoners paid by the weight of their collected trash or the number of shoes they produce undoubtedly will collect more and produce more. Because these additional payments will tend to reduce deterrence,[21] their desirability will depend on a comparison of the value of the additional output they generate

[20] In this paragraph I am using the term *effort* to mean the level of exertion while working, not how many hours are worked. Elsewhere in the article I occasionally used this term to refer to how many hours are worked per period (that is, to λ).

[21] Actually, in the benchmark model in which every prisoner is assumed to be identical (aside from the benefit from committing the offense), compensation could be set to just offset the extra disutility from working harder. Then there would not be a reduction in deterrence. But more realistically, if the disutility from working harder varies among prisoners, then inframarginal prisoners will be better off as a result of the compensation, and therefore deterred less.

to the public cost of raising the probability of detection or the length of the sentence in order to restore deterrence.

(b) *What if prisoners like prison work?* I assumed that prisoners bear disutility from working, but this might not always be the case. Although prisoners may dislike picking up trash along a highway on a hot day or having to clean bathrooms, they might obtain some pleasure from repairing prison uniforms or preparing and serving food in the prison kitchen. If prisoners like some prison work, then the deterrent effect of a mandatory work program would decline to a degree. Similarly, a voluntary work program could result in lower deterrence if prisoners prefer to do some work even without compensation. Thus, to the extent that prisoners obtain positive utility from work, the desirability of prison work programs is reduced.[22] It seems implausible, however, that prisoners would like prison work so much as to make prison work programs unattractive due to their lessening of deterrence.

(c) *What if the goal of imprisonment is incapacitation?* Suppose that the objective of imprisonment is to incapacitate prisoners — to keep them from committing additional crimes — rather than to deter them from committing crimes. The degree to which this objective is satisfied depends solely on the probability of detection and the sentence—specifically, on the expected sentence length ps. Given the probability and the sentence, is a prison work program desirable, and if so, is a mandatory or voluntary program better? For simplicity, I will answer

---

[22] It might be questioned whether the social cost of imprisonment would necessarily rise due to the lower level of deterrence, since, if prisoners like work, the social cost of imprisonment per period would be lower than if they dislike work. To see that the social cost of imprisonment does rise, suppose that the initial level of deterrence is $ps(1 + \lambda d)$, where $d > 0$ is the disutility from work per period and the other variables are as defined previously. The social cost of imprisonment then is $ps(1 + c + \lambda d - \lambda n)$. Now suppose that prisoners like work and obtain utility $u > 0$ per period from it, and let p' and s' be the new probability of detection and sentence that maintain deterrence. Thus, $p's'(1 - \lambda u) = ps(1 + \lambda d)$, and obviously $p's' > ps$. Under the assumption that prisoners like work, the social cost of imprisonment is $p's'(1 + c - \lambda u - \lambda n)$. The claim is that $p's'(1 + c - \lambda u - \lambda n) > ps(1 + c + \lambda d - \lambda n)$ if $p's'(1 - \lambda u) = ps(1 + \lambda d)$. Using this equality, the inequality becomes $p's'(c - \lambda n) > ps(c - \lambda n)$, which is satisfied since $\overline{\lambda} n <$

these questions under the assumptions of the benchmark model in which prisoners are equally productive and bear the same disutility from work. First observe that if deterrence is not a concern, then there is no reason to choose one work program or the other, for each can achieve whatever level of work λ is desired. The mandatory work program can require this level of work, and the voluntary work program can offer a high enough wage to induce prisoners to work this much. Moreover, assuming that the first-best level of work is positive (that is, $n > d'(0)$), a prison work program will be desirable even when the objective of imprisonment is incapacitation.

(d) *What if the goal of imprisonment is rehabilitation?* Alternatively, suppose that the rehabilitation of prisoners is the goal of imprisonment, that is, helping them to return to their communities as more productive members of society. Closely related to this objective would be the desire to provide vocational training to prisoners. Given these objectives, prison work programs obviously would be desirable, especially ones that provide work similar to that which prisoners can perform outside of prison, such as working in factories producing apparel or providing food services. A mandatory prison work program would have an advantage over a voluntary one in that all prisoners would obtain the benefit of vocational training, rather than just those who elect to participate.

(e) *What is the character of prison work programs in practice?* The majority of prison work appears to be undertaken on a mandatory basis, sometimes without compensation, but often with very low compensation. For example, in Federal prisons "[s]entenced inmates are required

---

c. Thus, regardless of whether the probability of detection or the sentence rises (or both), the social cost of imprisonment rises.

to work if they are medically able," and they are paid 12 cents to 40 cents per hour.[23] Work by prisoners also is mandatory in many, if not most, states. In some prison work programs, in which prisoners work under supervision inside prisons for private sector employers, prisoners are paid wages closer to prevailing market wages, but then have deductions assessed for room and board, taxes, victim restitution, and family support,[24] often leaving them with a small fraction of their nominal wages. Prisoners who are awaiting trial for sentencing are not required to work, but may volunteer to do so.[25]

[23] See Federal Bureau of Prisons, Work Programs, available at <http://www.bop.gov/inmates/custody_and_ care/ work_programs.jsp>.

[24] See Seiter (2005, pp. 435-36).

[25] See Seiter (2005, p. 73).

Appendix

In this Appendix I demonstrate through an example that a voluntary prison work program can be superior to a mandatory one when the disutility of work varies among prisoners. (It also will be the case in this example that a mandatory work program is preferable to no work program.)

The example employs discrete distributions of the benefit from committing the offense and the disutility from work. For simplicity, I assume that it is costless to set the probability of detection and that it is set equal to 1.[26]

Let

b_L = low benefit to an individual from committing the offense; $b_L > 0$;

b_H = high benefit to an individual from committing the offense; $b_H > b_L$;

φ = fraction of low-benefit individuals; $0 < \varphi < 1$;

d_L = low disutility of time worked per period; $d_L > 0$;

d_H = high disutility of time worked per period; $d_H > d_L$; and

θ = fraction of low-disutility-of-time-worked individuals; $0 < \theta < 1$.

The other variables employed in the example have been defined in the main body of the article.

I assume in the example that it is possible to deter b_L-types by the threat of prison without an accompanying work program; that b_H-types cannot be deterred by prison even with the harshest mandatory work program; and that the harm exceeds b_H:

$$b_L < \overline{s} < \overline{s}(1 + \overline{\lambda} d_H) < b_H < h, \tag{A1}$$

where $\overline{s}$ is the maximal sentence length. I also assume that prison work is socially beneficial on average (averaged over the d_L-types and d_H-types), but in a limited way; specifically, let

$$n = \theta d_L + (1 - \theta)d_H + \varepsilon, \tag{A2}$$

where $0 < \varepsilon < \theta(n - d_L)$. This restriction on the magnitude of ε implies that[27]

$$d_L < n < d_H. \tag{A3}$$

No prison work program: The level of deterrence is s. If $s < b_L$, no one will be deterred and the level of social welfare will be

$$\varphi b_L + (1 - \varphi)b_H - h - s(1 + c). \tag{A4}$$

Clearly, in this case, the optimal s is 0, in which case social welfare will be

$$\varphi b_L + (1 - \varphi)b_H - h. \tag{A5}$$

If $s \geq b_L$, the b_L-types will be deterred, but not the b_H-types, so the level of social welfare will be

$$(1 - \varphi)[b_H - h - s(1 + c)]. \tag{A6}$$

In this case, the optimal s is as low as possible consistent with continuing to deter the b_L-types, that is, such that $s = b_L$. Then social welfare is (A6) with $s = b_L$.

I assume that social welfare (A6) with $s = b_L$ will exceed (A5), where no one is deterred. This requires that

$$\varphi h - b_L - (1 - \varphi)b_L c > 0, \tag{A7}$$

which clearly will hold for h sufficiently high.

Mandatory prison work program: Given s, λ, and w, there are three possible outcomes in the example under a mandatory work program: (1) no b_L-types are deterred; (2) only b_L-types with high disutility from work d_H are deterred; or (3) all b_L-types are deterred. In all three cases

---

[26] It will be clear that I could modify the example by making the conventional assumption that it is costly to raise the probability, provided that the cost of setting the probability equal to 1 is sufficiently low.

[27] Obviously, (A2) could not hold unless $d_L < n$. From (A2), $d_H = (n - \theta d_L - \varepsilon)/(1 - \theta)$, which will exceed n if and only if $\varepsilon < \theta(n - d_L)$.

no b_H-types are deterred. I consider the optimal choices of s, λ, and w within each of these cases and thereby determine the highest level of social welfare achievable within each case. I then make an assumption regarding the magnitude of h that results in social welfare being highest overall in the third case.

In the first case, $s(1 + \lambda d_H - \lambda w) < b_L$ and social welfare is

$$\varphi\{\theta[b_L - h - s(1 + c + \lambda d_L - \lambda n)] + (1 - \theta)[b_L - h - s(1 + c + \lambda d_H - \lambda n)]\} \tag{A8}$$
$$+ (1 - \varphi)\{\theta[b_H - h - s(1 + c + \lambda d_L - \lambda n)] + (1 - \theta)[b_H - h - s(1 + c + \lambda d_H - \lambda n)]\}.$$

Since $\bar{\lambda}n < c$, the terms in parentheses multiplying s are all positive, implying that the optimal s is 0. Hence, maximal social welfare in the first case is (A5).

In the second case, $s(1 + \lambda d_L - \lambda w) < b_L \leq s(1 + \lambda d_H - \lambda w)$ and social welfare is

$$\varphi\theta[b_L - h - s(1 + c + \lambda d_L - \lambda n)] \tag{A9}$$
$$+ (1 - \varphi)\{\theta[b_H - h - s(1 + c + \lambda d_L - \lambda n)] + (1 - \theta)[b_H - h - s(1 + c + \lambda d_H - \lambda n)]\}.$$

Since $\bar{\lambda}n < c$, the optimal s is the lowest s that still allows deterrence of the b_L-types with high disutility from work d_H. This s satisfies $b_L = s(1 + \lambda d_H - \lambda w)$. It is clear from this condition that the higher λ is and the lower w is, the lower s can be. Hence, maximal social welfare in the second case is (A9) with $\lambda = \bar{\lambda}$ and $s = b_L/(1 + \bar{\lambda}d_H)$.

In the third case, $b_L \leq s(1 + \lambda d_L - \lambda w)$ and social welfare is

$$(1 - \varphi)\{\theta[b_H - h - s(1 + c + \lambda d_L - \lambda n)] + (1 - \theta)[b_H - h - s(1 + c + \lambda d_H - \lambda n)]\}. \tag{A10}$$

By parallel reasoning to that in the second case, maximal social welfare in the third case is (A10) with $\lambda = \bar{\lambda}$ and $s = b_L/(1 + \bar{\lambda}d_L)$.

Now observe that maximal social welfare in the third case under the mandatory work program exceeds that when there is no work program and the b_L-types are deterred. In other

words, the claim is that (A10) with $\lambda = \bar{\lambda}$ and $s = b_L/(1 + \bar{\lambda}d_L)$ exceeds (A6) with $s = b_L$. After dividing (A10) and (A6) by $(1 - \varphi)$, this can be expressed as

$$\theta[b_H - h - s_{M3}(1 + c + \bar{\lambda}d_L - \bar{\lambda}n)] + (1 - \theta)[b_H - h - s_{M3}(1 + c + \bar{\lambda}d_H - \bar{\lambda}n)] \qquad \text{(A11)}$$
$$> b_H - h - s_N(1 + c),$$

where $s_{M3} = b_L/(1 + \bar{\lambda}d_L)$ and $s_N = b_L$, or, equivalently, as

$$s_N(1 + c) > s_{M3}(1 + c - \bar{\lambda}[n - \theta d_L - (1 - \theta)d_H]) = s_{M3}(1 + c - \bar{\lambda}\varepsilon), \qquad \text{(A12)}$$

where the equality follows from (A2). That the inequality in (A12) holds follows from observing that $s_N > s_{M3} > 0$ and $\varepsilon > 0$.

The result in the preceding paragraph also implies that, under a mandatory work program, the outcome in the third case is superior to the outcome in the first case. This is because the outcome in the first case is the same as that when there is no work program and no deterrence, which was assumed to be inferior to the outcome when there is no work program and the b_L-types are deterred.

Thus, under a mandatory work program, it remains to compare the outcomes in the second and third cases. The third-case outcome will be superior if (A10) with $\lambda = \bar{\lambda}$ and $s = b_L/(1 + \bar{\lambda}d_L)$ exceeds (A9) with $\lambda = \bar{\lambda}$ and $s = b_L/(1 + \bar{\lambda}d_H)$. By collecting terms multiplying h, this condition can be expressed as

$$-(1 - \varphi)h + K_{M3} > -[\varphi\theta + (1 - \varphi)]h + K_{M2}, \qquad \text{(A13)}$$

where K_{M3} represents the remaining terms in (A10) not multiplying h and K_{M2} represents the remaining terms in (A9) not multiplying h. This inequality can be rewritten as $\varphi\theta h > K_{M2} - K_{M3}$, which clearly will hold if h is high enough, which is what I will assume. Thus, the highest level of social welfare achievable under a mandatory work program is that in the third case, in which all b_L-types are deterred.

Voluntary prison work program: Because my object in this example is to show that a voluntary prison work program can be superior to a mandatory one when the disutility of work varies among prisoners, I will not undertake a full analysis here of the voluntary work program. Instead, I will show that under an assumption that is consistent with the assumptions already made, a voluntary work program will generate a higher level of social welfare than the best mandatory work program.

Consider a voluntary work program that, like the best mandatory work program in the example, deters all b_L-types. Specifically, let $s = b_L$, $\lambda = \bar{\lambda}$, and $w = d_L$. At this wage, prisoners with low disutility from work will participate in the work program (they are indifferent), but prisoners with high disutility from work will not. The level of deterrence for prisoners in the former group will be $s(1 + \lambda d_L - \lambda w) = b_L$ and the level of deterrence for prisoners in the latter group will be $s = b_L$. Thus, all b_L-types will be deterred and all b_H-types will commit the offense; and only the subset of b_H-types with low disutility from work will participate in the work program. Given the preceding observations, social welfare under this voluntary work program will be

$$(1 - \varphi)\{\theta[b_H - h - s(1 + c + \lambda d_L - \lambda n)] + (1 - \theta)[b_H - h - s(1 + c)]\}, \tag{A14}$$

where $\lambda = \bar{\lambda}$ and $s = b_L$.

This voluntary work program will be superior to the best mandatory work program if (A14) with $\lambda = \bar{\lambda}$ and $s = b_L$ exceeds (A10) with $\lambda = \bar{\lambda}$ and $s = b_L/(1 + \bar{\lambda} d_L)$. After dividing both expressions by $1 - \varphi$ and subtracting $b_H - h$, this condition becomes

$$\theta[-b_L(1 + c + \bar{\lambda} d_L - \bar{\lambda} n)] + (1 - \theta)[-b_L(1 + c)] > \tag{A15}$$

$$\theta[-[b_L/(1 + \bar{\lambda} d_L)](1 + c + \bar{\lambda} d_L - \bar{\lambda} n)] + (1 - \theta)[-[b_L/(1 + \bar{\lambda} d_L)](1 + c + \bar{\lambda} d_H - \bar{\lambda} n)].$$

The limit of (A15) as $d_L \to 0$ is

$$\theta[-b_L(1 + c - \bar{\lambda}n)] + (1 - \theta)[-b_L(1 + c)] > \tag{A16}$$

$$\theta[-b_L(1 + c - \bar{\lambda}n)] + (1 - \theta)[-b_L(1 + c + \bar{\lambda}d_H - \bar{\lambda}n)],$$

which, after some manipulation, can be expressed as

$$(1 - \theta)b_L\bar{\lambda}(d_H - n) > 0. \tag{A17}$$

This condition clearly holds since $d_H > n$ (see (A3)).

Thus, for h sufficiently high and d_L sufficiently low, the voluntary work program under consideration will be superior to the best mandatory work program. The following intuition explains this result in the example. If the harm h is high enough, maximal deterrence is desirable, meaning deterrence of all of the low-benefit types (by assumption, the high-benefit types cannot be deterred). This level of deterrence can be achieved with a shorter sentence under the mandatory work program than under the voluntary work program because the mandatory program can impose greater disutility on prisoners by making them work without compensation. Specifically, the optimal sentence is $b_L/(1 + \bar{\lambda}d_L)$ under the mandatory program and b_L under the voluntary one. Everything else equal, this favors the mandatory program, but because the optimal sentence is the minimal sentence required to deter b_L-types who have low disutility from work, the lower d_L is, the smaller is this benefit. On the other hand, the advantage of the voluntary work program over the mandatory one stems from the observation that, if the wage rate under the voluntary program is set equal to d_L, it will induce only those prisoners with low disutility from work to participate in the work program, whereas the mandatory program forces all prisoners to work. Because there is a net social loss from work by high-disutility-of-work prisoners ($d_H > n$), this difference favors the voluntary work program. Hence, if the deterrence advantage of the mandatory work program is small enough (that is, if d_L is small enough), then the voluntary work program will be preferable.

References

American Correctional Association, *A Study of Prison Industry: History, Components, and Goals* (1986).

Avio, Kenneth L., "The Economics of Prisons," *European Journal of Law and Economics*, Vol. 6 (1998), pp. 143-75.

Barnes, Harry E., "The Economics of American Penology as Illustrated by the Experience of the State of Pennsylvania," *Journal of Political Economy*, Vol. 29, No. 8 (1921), pp. 617-42.

Bureau of Justice Statistics, "Census of State and Federal Correctional Facilities, 2005," National Prisoner Statistics Program, October 2008, NCJ 222182.

Correctional Industries Association, "P.I.E. Programs," interview with Rod Miller of Jail Industries Association, in *1998 Correctional Industries Association Directory* (Baltimore, MD: Correctional Industries Association, 1999), pp. 100-08.

England, Ralph, *Prison Labor* (New York: United Nations Department of Economic and Social Affairs, 1955).

Garoupa, Nuno, "The Theory of Optimal Law Enforcement," *Journal of Economic Surveys,* Vol. 11 (1997), pp. 267-95.

Gibson, Mary, "Global Perspectives on the Birth of the Prison," *American Historical Review*, Vol. 116, No. 4 (2011), pp. 1040-63.

Levitt, Steven D., and Thomas J. Miles, "Empirical Study of Criminal Punishment," in A. Mitchell Polinsky and Steven Shavell (Eds.), *Handbook of Law and Economics,* Volume 1 (Elsevier, 2007), pp. 455-95.

Mookherjee, Dilip, "The Economics of Enforcement," in A. Bose, M. Rakshit, and A. Sinha, (Eds.), *Issues in Economic Theory and Policy, Essays in Honor of Tapas Majumdar* (New Delhi: Oxford University Press, 1997), pp. 202–49.

Polinsky, A. Mitchell, and Steven Shavell, "The Theory of Public Enforcement of Law," in A. Mitchell Polinsky and Steven Shavell (Eds.), *Handbook of Law and Economics*, Volume 1 (Elsevier, 2007), pp. 403-54.

Seiter, Richard P., *Corrections: An Introduction* (Pearson, 2005).