Slavery, Incentives, and Manumission: A Theoretical Model

Ronald Findlay
Columbia University

This paper presents a theoretical model of slavery and manumission in which the effective labor provided by slaves is a function of both the level of supervision costs incurred by the owner and the incentive payments received by the slaves. The optimal combination of supervision costs and incentive payments is determined together with the input of physical capital. The length of time it would take for a slave to purchase his freedom out of savings from his incentive payments is derived and is shown to vary inversely with the rate of interest.

This paper presents a theoretical model of slavery in which incentive payments are given to the slaves and manumission by self-purchase is possible. The paper is in three parts. The first part places the model in the historical context of discussions of slavery in classical antiquity and the antebellum South. The model of slavery with incentive payments is constructed in the second part, and the third analyzes the question of manumission.

I

Slavery is one of the oldest, most extensive and complex of economic institutions. It is remarkable, therefore, that so little formal economic analysis has been devoted to the subject. Until the recent flood of works on slavery in the old South, stemming from the famous article by Conrad and Meyer (1958), the only major reference that comes to mind is Cairnes (1969), which predates the marginalist revolution in economic

I would like to thank my colleagues G. Calvo, M. Edelstein, K. J. Lancaster, D. Mathieson, P. Passell, C. Rodriguez, S. Wellisz, Professor Stanley Engerman of the University of Rochester, and the editors and referees for their valuable comments and suggestions. They are not, of course, responsible for my own errors.

© 1975 by The University of Chicago. All rights reserved.
theory. The neglect can perhaps be traced to the unfortunate fact that economic theory and economic history have tended to be divided almost as sharply as the "two cultures" of C. P. Snow, at least until the emergence of "cliometrics."

Slavery in the ancient world of Greece and Rome was perhaps the most distinctive single feature of the economic system of that time, as is widely recognized (the Marxists, for example, identify the "ancient mode of production" with slavery). What was the nature of the ancient slavery, and how did it differ, if at all, from the slavery of the southern United States? An answer to this question was formulated long ago by Sir Alfred Zimmern (1928), the distinguished historian and classical scholar, in a brilliant but neglected essay. He identified two distinct "ideal types," or models, of slavery, which I shall refer to for convenience as "stick" slavery and "carrot" slavery. As these names imply, the first system conjures up the familiar picture of the overseer and his whip-driven gangs of human cattle in a galley or a mine. Plantation slavery in the southern United States was regarded by Zimmern as being essentially of this variety. The ancient world also had slavery of this type, particularly in the Athenian silver mines and the Nubian gold mines and also on the large estates of Sicily and southern Italy in the Roman Empire. The noteworthy economic feature of this pattern of slavery is the fact that relatively large numbers of slaves concentrated on simple, repetitive tasks, making for economies of scale in supervision costs.

In the ancient world, on the other hand, slaves were often highly skilled craftsmen, as the Attic vases in any museum attest. An astonishing range of occupations, from admirals in the Roman navy to policemen in Athens, was filled by slaves, as noted by Jones (1956). Brute force is obviously not a cheap and efficient method of control when activity is on a small scale and requires the intelligent and even creative powers of the enslaved workers. Consequently, various types of incentive systems were introduced to elicit the willing cooperation of the slave. Zimmern's (1928, p. 122) insight into the problem is well conveyed by the following quotation:

The object of the slave-master as an economic man is to give his apprentices the maximum amount of motive for working while leaving them the minimum amount of profit from their work. He owns the entire produce of their labor; but by surrendering a portion of it he can increase the total amount. There is therefore a delicate balance between the increase in the amount surrendered and the increase in the total amount. The slave-master's object is to discover how much of the produce it will be profitable for him to surrender in each case.

In spite of the usual layman's confusion about simultaneously maximizing return and minimizing cost, Zimmern has clearly understood the
notion of optimization by balancing at the margin. This paper will largely be concerned with a formal statement and analysis of the problem as envisaged by Zimmern. The significant implication of carrot slavery is that the slave earns income, which he may use to purchase his freedom—the owner, of course, being able to maintain production with slave labor indefinitely by turning over his slave-labor force. Manumission was a regular practice in the ancient world, and the fact that the slave and free populations were generally not ethnically distinct made this easier for that society to tolerate than it would have been in the antebellum southern United States. The interesting economic problem, therefore, arises out of what determines the proportion of a worker's life that is spent in slavery. Our analysis will be addressed to this problem as well.

The reader unfamiliar with the economic history of the ancient world might be skeptical about the sophisticated motivation that will be assumed in the model on the part of both slave and master. It should be borne in mind, however, that the ancient world achieved an extremely complex commercial network, and that market forces were quite pervasive. For documentation, the reader is referred to the massive study on the economic development of the ancient world by Heichelheim (1970) and the comprehensive monograph on slavery by Westermann (1955).

The present theoretical investigation, inspired by Zimmern's essay, was begun and essentially completed before I had read the controversial new study by Fogel and Engerman (1974). They argue that the conventional picture of stick slavery imagined by Zimmern to hold for southern plantations, largely on the authority of Cairnes, is quite false. They point out that while the threat of the stick was always present, the carrot was nevertheless used liberally to motivate slave labor. They were not aware of Zimmern's contribution and do not attempt to relate their picture of slavery in the old South to the system of antiquity. Hence, it would appear that the compatibility of slavery with incentives gives greater unity to the economic analysis of the institution, and the comparative study of ancient and modern slavery should be a rich field of enquiry for both economists and historians.

In a brief but extremely imaginative contribution, John E. Moes (1962) explicitly considers the "counter factual" question of whether the free play of economic forces would have caused the South to permit manumission by self-purchase on an extensive scale and thus, in the absence of any importation of new slaves, led to the gradual extinction of slavery such as occurred in the ancient world. He rightly considers that it is from this standpoint, rather than the mere demonstration of the comparability of the rate of return on slave ownership with other forms of investment, that the question of the viability of slavery in the old South on the eve of the Civil War should have been considered.

Before turning, in the next section, to the specification of the model,
a brief word on the nature of the exercise undertaken here is perhaps in order. It is purely deductive and formal in character and makes no assertion about what slavery in antiquity, the old South, or anywhere else, for that matter, was "really like." The extent to which the assumptions and implications of the model apply to any particular slave system are a matter for empirical investigation and beyond the author's competence. It is of course hoped, however, that the model might be of some relevance to historical studies.

II

Imagine that production of some commodity takes place in a workshop or farm with the labor of a group of slaves with identical economic characteristics. At the present stage, there is no need to specify the manner in which the slaves were acquired and their cost. They are simply a fixed asset to be exploited in whatever manner maximizes the owner's profit. All tools and equipment are the property of the slave owner, who has the full right to dispose of the product, which can be sold at fixed market prices. The quantity of output, denoted \( Y \), is a function of the input of capital \( K \) (conceived of as a homogenous "jelly" into which the services of concrete capital goods can be lumped, as in the neoclassical parable), and the input of "effective" labor of the slaves measured in terms of man-hours of standard efficiency is denoted \( L_D \). This production function is taken to have constant returns to scale and the usual neoclassical properties.

In the conventional picture of slavery, the owner provides the "subsistence" requirements of the slaves and makes them labor, by the threat and use of force, to whatever limit of their endurance is consistent with his long-run profit maximization. This is the polar opposite of the familiar model of the market for free labor, in which it is assumed that there is a freely negotiated contract in which the wages to be paid by the employer and the quantity and quality of the labor to be performed by the worker are specified. As Fogel and Engerman (1974, p. 155) point out, however, most actual labor systems are mixtures in which incentives and elements of coercion coexist. In fact, most modern enterprises only stipulate the time to be spent at the place of work, and substantial resources have to be invested in mechanisms of supervision and control to elicit the appropriate quantity and quality of work after the wage bargain has been struck. The internal organization of work in the firm has been a relatively neglected area of economic theory though, of course, there is the classic article by Coase (1937) and a recent paper by Alchian and Demsetz (1972).

We now propose a simple model of labor supply in which the optimum combination of stick and carrot for an enterprise using slave labor can be
found. The logic of the model applies just as well to the problem of finding the optimum combination of wage payments and supervision costs for an enterprise using free labor. The only difference between the two systems would be that the "supervisor" in the slave system is more likely to use a whip to correct negligence, while in the free system he is more likely to use threats of dismissal or slower promotion up the job ladder.

In figure 1, the curve $Oa$ shows the amount of labor, in terms of man-hours of standard efficiency, that can be obtained from the batch of slave workers as a function of the level of incentive payments, assuming that there is no coercion or "supervision" employed. The payment per man-hour, or the implicit wage rate, is indicated by the slope of a ray from the origin to any point on $Oa$. By incurring supervision costs, such as the hiring of overseers, the owner could shift the curve $Oa$ to the right—the extent of the shift increasing with the level of supervision costs incurred. By measuring supervision costs on the vertical axis and drawing each of the curves, such as $Oa$, from the origin corresponding to the associated level
of supervision costs, an envelope curve $OA$ can be obtained. For each amount of labor, $OA$ indicates the combination of incentive payment and supervision cost that minimizes the total cost required to elicit it.

Before proceeding to the formal analysis of the model, a brief word on the comparison between the situations of slave and free labor might be in order. As stated earlier, the model, in principle, could apply to either case. The fact that much harsher methods of supervision are usually applied to slaves means that, other things being equal, a given level of supervision cost would be more effective in pushing the labor-supply curve to the right. On the other hand, the desire for freedom on the part of an enslaved labor force might well result in a greater response to incentive payments if these can be accumulated for the eventual purchase of manumission. Thus, no presumption should exist either way as to the relative ratio of supervision costs to incentive payments in the two cases.

Both factors, however, lead to reducing the total cost per unit of labor in the slavery case. We should perhaps not be so surprised, therefore, when slavery is found, as by Fogel and Engerman, to be more "efficient" in this narrow sense. Had fewer social and political obstacles been placed on permitting manumission, the relative "efficiency" of slavery would no doubt have been found to be even greater.

Denoting the incentive payments as a fraction $\alpha$ of total output, supervision costs by $F$, the fixed subsistence costs of the slaves by $\vec{S}$, and the rate of interest at which capital is borrowed by $r$, the problem of the owner is to maximize his profit, which is equal to

$$(1 - \alpha)Y(K, L_D) - rK - F - \vec{S} \tag{1}$$

subject to

$$L_D = L_s[\alpha Y(K, L_D), F] \tag{2}$$

where the right-hand side is the labor-supply function graphically represented by the envelope $OA$ in figure 1.

The necessary conditions for profit maximization are obtained by setting the partial derivatives of the associated Lagrangean expression with respect to $\alpha$, $L_D$, $F$, and $K$ equal to zero. These conditions can be written as

$$\frac{\partial Y}{\partial L_D} = \frac{\partial (\alpha Y)}{\partial L_s} + \frac{1}{\partial F} = \lambda \cdots \tag{3}$$

$$\frac{\partial Y}{\partial K} = r \tag{4}$$

where the Lagrange multiplier $\lambda$ is the marginal cost of labor. Obviously, the stick and the carrot have to be used to where their effect is the same at the margin, and this common marginal cost must be equal to the
marginal product of labor, with the physical capital of the enterprise being such that its marginal product is equal to the fixed rate of interest. The optimal values of $a$, $Y$, $K$, $L_D$, and $F$ are all determined.

By transforming (3) into elasticities, it follows that

$$a = \theta \epsilon = \theta E - f \ldots,$$

where

$$\theta \equiv \frac{\partial Y}{\partial L_D} \cdot \frac{L_D}{Y}$$ is the elasticity of output with respect to labor,

$$\epsilon \equiv \frac{a Y}{L_S} \cdot \frac{\partial L_S}{\partial (a Y)}$$ is the elasticity of labor supply at the optimal point when supervision cost is held constant at its optimal level,

$$E \equiv \frac{a Y + F}{L_S} \cdot \frac{\partial L_S}{\partial (a Y)}$$ is the elasticity of the envelope labor-supply curve at the optimal point, and

$$f \equiv \frac{F}{Y}$$ is the share of supervision cost in the value of output when both are at the optimal level.

The optimal solution is depicted in figure 1, where $OL^*$ is the optimal level of labor input and $OQ^*$ of total output. On the vertical axis, the distances $OU$, $UT$, $TS$, $SR$, and $RQ^*$ measure the optimal levels of supervision costs, incentive payments, subsistence requirements (assumed constant), profit, and interest payments on physical capital, respectively. The $a$ is equal to $UT/OQ^*$, while $SR/OR$ is the proportion of the marginal product of labor that is captured as pure profit by the slave owner after deducting subsistence and supervision costs and incentive payments to the slaves. The $E$ is the elasticity of the envelope curve $OA$, and $\epsilon$ of the "short-run" curve $UV$ tangential to it at the optimal point.

III

The fact that the slave earns income in this system implies that he could eventually purchase his freedom if it were possible for him to save enough out of his earnings to compensate the slave owner for the loss of future net income from his human asset. The opportunity cost of saving for eventual manumission is consumption above the subsistence minimum, so that the slave has to balance these two considerations in making his decision. Instead of a complex model in which the desire for freedom is put into the slave's utility function, which is then maximized, we shall simply assume that he saves a constant fraction $h$ of his earnings which he puts into a fund for the eventual purchase of his freedom. This fund is able to earn interest, and for simplicity we shall assume that the rate of interest is the same as that facing the slave owner. Nothing essential is altered if the two rates are different, or zero for the slave. There are many
instances from the ancient world of the ownership or control of capital funds by slaves, and it was not unknown even in the United States.

Denoting the initial point of time at which a slave is acquired by \( t_o \), the end of his working life by \( T \) (assuming that this is exogenously determined), and the moment of manumission by \( t^* \), we have the relation

\[
\int_{t_o}^{t^*} h \alpha Y e^{t^*} dt = \int_{t^*}^{T} [(1 - \alpha) Y - rK - \bar{S} - F] e^{-r(t-t^*)} dt. \ldots (6)
\]

Evaluating these integrals and using the first-order approximations for the exponentials, it follows that

\[
[Y - rK - \bar{S} - F - (1 - h) \alpha Y]t^* = [(1 - \alpha) Y - rK - \bar{S} - F]T, \quad (7)
\]

indicating that \( t^* = T \) if \( h = 0 \); while if \( h = 1 \), we have the interesting result that

\[
\frac{t^*}{T} = \frac{(1 - \alpha) Y - rK - \bar{S} - F}{Y - rK - \bar{S} - F}, \quad (8)
\]

which means that the minimum proportion of his working life that a slave has to spend in bondage is equal to the ratio of the slave owner's income net of incentive payments to what it would be if he could obtain the same labor from the slave without having to make any such payments. It is perhaps illuminating to regard the slaves as turning over all their incentive payments to the owner, giving him a flow of income per unit time equal to the denominator of (8) and in return obtaining their freedom at \( t^* \). The owner trades off a shorter period of exploitation of his human asset against a higher return per unit of time over the reduced period of exploitation.

Substituting for \( \alpha \) from (5) and denoting \( \bar{S}/Y \) by \( s \), (8) can be transformed into

\[
\frac{t^*}{T} = \frac{\theta(1 - E) - s}{(\theta - s - f)} = 1 - \frac{\theta e}{(\theta - s - f)}, \quad (9)
\]

from which it is readily verified that \( t^* \) varies directly with \( \theta \) and inversely with \( s \) and \( f \).

An interesting exercise to attempt is the effect of variations in the rate of interest on the time required for the purchase of manumission. From (4) it is obvious that a lower rate of interest will increase the stock of capital. In terms of figure 1, the output curve shifts upward and the marginal product of labor will be increased if labor is held constant. For equilibrium to be restored, the labor input has to be increased until the slope of the envelope curve \( OA \) is equal to the slope of the new output curve. The level of output and the marginal cost of labor will both be higher in the new equilibrium position. The effect on \( t^* \) will depend on what happens to the values of the variables in (9). No simple prediction is possible unless some further assumptions are made.
Suppose the production function is Cobb-Douglas so that $\theta$ is constant. Since $Y$ increases and $S$ is constant, $s$ must fall. Observe that it is possible to write

$$f = \frac{F}{L_S} \cdot \frac{L_D}{Y}.$$  \hspace{1cm} (10)

It is reasonable to suppose that $F/L_S$ falls as we move along $OA$ reflecting economies of scale in supervision costs. Since $Y/L_D$ rises, it follows that $f$ must fall so that the denominator of the second expression for $t^*/T$ in (9) rises.

It is possible to show that $e$ must be lower at the new equilibrium point. Supervision cost increases, so the new "short-run" labor-supply curve would lie to the right of the first one if both are drawn from the same origin. If incentive payments are held fixed, more labor will be obtained because of greater supervision. The marginal incentive payment (equal to the slope of the curve) would be higher on the new curve, since more labor is being supplied at this point and the marginal disutility of effort is consequently increased. Hence, $e$ must be lower at the point on the new curve corresponding to the same total incentive payment as before. But, from figure 1 it is seen that $(\alpha Y + F)/L_S$ is higher as we move to the right along $OA$. Since $F/L_S$ falls by our previous assumption, $\alpha Y/L_S$ must be higher on the new curve. It follows that $e$ must be lower than before, since $e$ diminishes along any given curve and it was lower when $\alpha Y$ was held constant and $L_S$ increased. Hence, from (9) we conclude that a fall in the rate of interest must raise $t^*$ if the production function is Cobb-Douglas.

These results on the effects of the level of the rate of interest on the pattern of slavery are relevant in the context of comparative studies of slavery—for example, between the Brazilian and U.S. systems. The shorter life span of slaves in Brazil is sometimes attributed to the desire of owners in a capital-scarce economy to reduce the "pay-back period" of their asset by more intensive exploitation over a shorter period. As we have seen, a higher rate of interest would also, given the assumptions made, quicken the time of manumission—the owner finding it profitable to free slaves earlier in return for more intensive work on their part. This is consistent with the observed facts that manumission by self-purchase was much more common in Brazil and other Latin American countries than in the more capital-abundant United States. It should not be forgotten, however, that there are many broader social and demographic aspects involved in such comparisons, as is brought out very well in the volume of readings edited by Foner and Genovese (1969).

The analysis here has, of course, been based on the assumption that slaves are not able to borrow for the purchase of their freedom. Westermann (1955, p. 25) records an instance in which a female slave, engaged
in what is alleged to be the world's oldest profession, obtained the price of her freedom by a loan from a group of clients; her case, however, must be exceptional. Another limitation is that we assume a rigid dichotomy between slavery and freedom. As M. I. Finley (1964) has argued, however, the ancient world was characterized by a continuous spectrum of various degrees of bondage, such as the performing of unrequited labor on particular days while being free on others.

Another important problem is how the implicit contracts we have been discussing are enforced on the slave's side. Social custom and the self-interest of the owners if they are dealing with several slaves on a continuing basis, so that they have to make their offers credible, can be involved in this regard. Institutional arrangements such as the fascinating ancient practice of fictitiously selling slaves to Apollo and thereby having a public record of manumission inscribed in the temple at Delphi are worthy of further study by those interested in the interaction of legal systems and economic behavior.

As mentioned earlier, U.S. slavery, unlike the systems of antiquity and contemporary Latin America, did not generally permit manumission. However, it was by no means unknown, and Matison (1948) gives some very interesting information in this regard. He observes, consistently with the point of view of this paper, that manumissions tended to occur mostly in the cities and in cases where the slave had some special craft or skill, since this would make it more advantageous to use the carrot instead of the stick. The actual transaction of the purchase of freedom was always a risky business for the slave, and Matison records several instances of cheating by owners. Frequently, a third party such as a freed slave or a benevolent white man received the money from the slave and made the purchase of freedom on his behalf. Saint Philip's Church in Charleston apparently held in trust over 100 slaves who had bought their own freedom, an interesting parallel with the temple of Apollo at Delphi. Recent books by Starobin (1970) and Wade (1964) also contain much relevant information on slavery and manumission in an urban and industrial setting.

References


