The Economics of Labor Coercion[*]

Daron Acemoglu   Alexander Wolitzky

MIT and CIFAR   MIT

September 12, 2010

Abstract

The majority of labor transactions throughout much of history and a significant fraction of such transactions in many developing countries today are "coercive," in the sense that force or the threat of force plays a central role in convincing workers to accept employment or its terms. We propose a tractable principal-agent model of coercion, based on the idea that coercive activities by employers, or "guns," affect the participation constraint of workers. We show that coercion and effort are complements, so that coercion increases effort, but coercion always reduces utilitarian social welfare. Better outside options for workers reduce coercion, because of the complementarity between coercion and effort: workers with a better outside option exert lower effort in equilibrium and thus are coerced less. Greater demand for labor increases coercion because it increases equilibrium effort. We investigate the interaction between outside options, market prices, and other economic variables by embedding the (coercive) principal-agent relationship in a general equilibrium setup, and study when and how labor scarcity encourages coercion. General (market) equilibrium interactions working through the price of output lead to a positive relationship between labor scarcity and coercion along the lines of ideas suggested by Domar, while those working through the outside option lead to a negative relationship similar to ideas advanced in neo-Malthusian historical analyses of the decline of feudalism. In net, a decline in available labor increases coercion in general equilibrium if and only if its *direct* (partial equilibrium) effect is to increase the price of output by more than it increases outside options. Our model also suggests that markets in slaves make slaves worse off, conditional on enslavement, and that coercion is more viable in industries that do not require relationship-specific investment by workers.

Keywords: coercion, feudalism, labor scarcity, principal-agent, slavery, supermodularity.
JEL Classification: D23, D74, D86, J01, P16.

"In the context of universal history, free labor, wage labor, is the peculiar institution"—M. I. Finley (1976).

1   Introduction

Standard economic models of the labor market, regardless of whether they incorporate imperfections, assume that transactions in the labor market are "free". For most of human history, however, the bulk of labor transactions have been "coercive," meaning that the threat of force was essential in convincing workers to take part in the employment relationship, and thus in determining compensation. Slavery and forced labor were the most common form of labor transactions in most ancient civilizations, including Ancient Greece, Egypt, Rome, several Islamic and Asian Empires, and most known pre-Colombian civilizations (e.g., Meltzer, 1993, Patterson, 1982, Lovejoy, 2000, Davis, 2006). Slavery was also the basis of the plantation economies in the Caribbean (e.g., Curtin, 1990, Klein and Vinson, 2007), in parts of Brazil and Colombia, and in the United States South (e.g., Patterson, 1982, Fogel and Engerman, 1974, Wright 1978), while forced labor played a major role in Spanish Latin America in mining and in the encomiendas as well as in the subsequent hacienda system that developed in much of Latin America (e.g., Lockhart and Schwartz, 1983, Lockhart, 2000). Although formal slavery has been rare in Europe since the middle ages, until the 19th century feudal labor relations, which include both forced labor services from serfs and various special dues and taxes to landowners, were the most important type of employment relationship except in cities (e.g., Blum, 1998). Even today, the United Nations' International Labor Organization (ILO) estimates that there are over 12.3 million forced laborers worldwide (Andrees and Belser, 2009).[1]

The prevalence of slavery and forced labor in human history raises the question of when we should expect labor to be transacted in free markets rather than being largely or partly coerced. In a seminal paper, Domar (1970) provides one answer: slavery or serfdom should be more likely when labor is scarce so that (shadow) wages are high. This answer is both intuitive and potentially in line with the experience in the Caribbean where Europeans introduced slavery into islands that had their population decimated during the early phases of colonization. In contrast, the "neo-Malthusian" theory of feudal decline, exemplified by Habakkuk (1958), Postan (1973), Leroy Ladurie (1974), and North and Thomas (1971), claims that coercive feudal labor relations started their decline when labor became scarce following the Black Death and other demographic shocks that reduced population and raised per capita agricultural income throughout Europe in the 16th century. Similarly, Acemoglu, Johnson and Robinson (2002) show that Europeans were more likely to set up labor-coercive and extractive institutions when population density was high and labor was relatively abundant. The relationship between labor scarcity/abundance and coercion is also important in understanding the causes of continued prevalence of forced labor in many developing countries.

---

[1]The ILO estimates that of these 12.3 million, 20% are coerced by the state (largely into military service), 14% are forced sex workers, and the remaining 66% are coerced by private agents in other industries, such as agriculture, mining, ranching, and domestic service (Andrees and Belser, 2009). Our model applies most directly to the last category.

In this paper, we develop a simple model of labor coercion. In partial equilibrium, our model is a version of the principal-agent framework, with two crucial differences. First, the agent (worker) has no wealth so that there is a limited liability constraint, and the principal can punish as well as reward the agent. Second, the principal chooses the amount of "guns" (coercion), which influences the reservation utility (outside option) of the agent.[2] The first of these changes has been explored in several papers (e.g., Chwe, 1990, Dow, 1993, Sherstyuk, 2000). The second is, to our knowledge, new, and is crucial for our perspective and our results; it captures the central notion that coercion is mainly about forcing workers to accept employment, or terms of employment, that they would otherwise reject.[3]

Our basic principle-agent model leads to several new insights about coercive labor relations. First, we show that coercion always increases the effort of the agent, which is consistent with Fogel and Engerman's (1974) view that Southern slavery was productive. Second, we show that coercion is always "socially inefficient," because it involves an (endogenously) costly way of transferring resources (utility) from workers to employers.[4] Third, perhaps somewhat surprisingly, we find that workers with a lower (ex ante) outside option are coerced more and provide higher levels of effort in equilibrium. The intuition for this result illustrates a central economic mechanism: in our model—and, we believe, most often in practice—effort and coercion are "complements". When the employer wishes to induce effort, he finds it optimal to pay wages following high output, so he must pay wages frequently when he induces high effort. Greater ex ante coercion enables him to avoid making these payments, which is more valuable when he must pay frequently, hence the complementarity between effort and coercion. This observation also implies that more "productive" employers will use more coercion, and thus a worker will be worse off when matched with a more productive firm. This contrasts with standard results in models of noncoercive labor markets where ex post rent sharing typically makes workers matched with more productive employers better off. It also implies that coerced workers may receive high expected monetary compensation, despite having low welfare, which is consistent with the finding of both Fogel and Engerman (1974) and ILO (2009) that coerced laborers often receive income close to that of comparable free laborers.

---

[2] Throughout the paper, "guns" stand in for a variety of coercive tools that employers can use. These include the acquisition and use of actual guns by the employers as a threat against the workers or their families; the use of guards and enforcers to prevent workers from escaping or to force them to agree to employment terms favorable to employers; the confiscation of workers' identification documents; the setting up of a system of justice favorable to employers; investment in political ties to help them in conflictual labor relations; and the use of paramilitaries, strike-breakers and other non-state armed groups to increase their bargaining power in labor conflicts. In all instances of coercion mentioned here, for example, in the Caribbean plantation complex, African slave trade, the mita, the encomienda, the feudal system, and contemporary coercion in Latin America and South Asia, employers used several of these methods simultaneously.

[3] This view of coercion is consistent with the historical and contemporary examples given above as well as with the 1930 ILO Convention's definition of compulsory labor as "work or service which is exacted from any person under the menace of any penalty and for which the said person has not offered himself voluntarily," (quoted in Andrees and Belser, 2009, p. 179). Discussing contemporary coercion, Andrees and Belser (2009) write that "a situation can qualify as forced labor when people are subjected to psychological or physical coercion... in order to perform some work or service that they would otherwise not have freely chosen" (p. 179).

[4] It also offers a straightforward explanation for why gang labor disappeared after the Reconstruction, which was a puzzle for Fogel and Engerman. In this light, our model is much more consistent with Ransom and Sutch's evidence and interpretation of slavery and its aftermath in the South (Ransom and Sutch, 1975, 1977).

The above-mentioned partial equilibrium results do not directly address whether labor scarcity makes coercion more likely. To investigate this issue, and the robustness of our partial equilibrium results to general equilibrium interactions, we embed our basic principal-agent model of coercion in a general (market) equilibrium setting, with two distinct equilibrium interactions. The first is a *labor demand effect*: the price an employer faces for his output is determined endogenously by the production—and thus coercion and effort—decisions of all employers, and affects the marginal product of labor and the return to coercion. The second is an *outside option effect*: agents' outside options affect employers' coercion and effort choices, and because agents who walk away from a coercive relationship (by paying a cost determined by the extent of coercion) may match with another coercive employer, effective outside options are determined by the overall level of coercion in the economy. Labor scarcity encourages coercion through the labor demand effect because labor scarcity increases the price of output, raising the value of effort and thus encouraging coercion; this is reminiscent of Domar's intuition.[5] On the other hand, labor scarcity discourages coercion through the outside option effect, because labor scarcity increases the marginal product of labor in (unmodeled) competing sectors of the economy, which increases outside options and thus discourages coercion. This is similar in spirit to the neo-Malthusian theory of feudal decline, where labor scarcity increased the outside opportunities for serfs, particularly in cities, and led to the demise of feudal labor relations. We show that whether the labor demand or the outside option effect dominates in general equilibrium is determined by whether the overall *direct* (i.e., partial equilibrium) effect of labor scarcity on the *difference* between price of output and the outside option is positive or negative. This finding provides a potential answer to the famous critique of the neo-Malthusian theory due to Brenner (1976) (i.e., that falling populations in Eastern Europe were associated with an increase in the prevalence of forced labor in the so-called "second serfdom," see Ashton and Philpin, 1985), because, for reasons we discuss below, the fall in population in Eastern Europe was likely associated more with higher prices of agricultural goods than with better outside options for workers.

The tractability of our principal-agent framework also enables us to investigate several extensions. First, we introduce ex ante investments and show that there is a type of "holdup" in this framework, where workers underinvest in skills that increase their productivity in their current coercive relationship (since workers that are more productive with their current employer are coerced more) and overinvest in skills that increase their outside option (since coercion is decreasing in their outside option). This extension provides a potential explanation for why coercion is particularly prevalent in effort-intensive, low-skill labor, and relatively rare in activities that require investment in relationship-specific skills or are "care-intensive" (as argued by Fenoaltea, 1984). Second, we investigate the implications of coercion when it affects the interim outside option of the agent and also when ex post punishments are costly. Third, we show that when coercion choices are made before matching, our model generates an economies of scale effect in line with the idea suggested in Acemoglu, Johnson and Robinson (2002)

[5] A recent paper by Naidu and Yuchtman (2009) exploits the effects of labor demand shocks under the Master and Servant Acts in 19th-century Britain and finds evidence consistent with the labor demand effect.

that greater labor abundance makes extractive institutions and coercion more profitable. Finally, we investigate the implications of "trading in slaves," whereby employers can sell their coerced agents to other potential employers, and show that such trade always reduces agent welfare and may reduce social welfare (because slave trade shifts the productivity distribution of active employers in the sense of first-order stochastic dominance, and with greater productivity comes greater coercion).

Despite the historical importance of coercion, the literature on coercive labor markets is limited. Early work in this area includes Conrad and Meyer (1958), Domar (1970), Fogel and Engerman (1974), and Ransom and Sutch (1977). Bergstrom (1971) defines a "slavery equilibrium," a modification of competitive equilibrium in which some individuals control the trading decisions of others, and shows existence and efficiency property of equilibria. Findlay (1975) and Canarella and Tomaske (1975) present models that view slaves as "machines" that produce output as a function of payments and force. Barzel (1977) performs a comparative analysis of slavery and free labor, under the assumption that slaves, but not free workers, must be monitored, and that slaves (exogenously) work harder.

In more recent work, Basu (1986) and Naqvi and Wemhöner (1995) develop models in which landlords may "coerce" their tenants by inducing other agents to ostracize them if they do not agree to favorable contract terms with the landlord. Genicot (2002) develops a model in which workers are collectively better off without the option of voluntarily entering into bonded labor agreements, because this stimulates the development of alternative forms of credit. Conning (2004) formalizes and extends Domar's hypothesis in the context of a neoclassical trade model with a reduced-form model of slavery. Lagerlöf (2009) analyzes a dynamic model of agricultural development in which slavery is more likely at intermediate stages of development.

The paper most closely related to ours is Chwe's (1990) important work on slavery. Chwe analyzes a principal-agent model closely related to our partial equilibrium model. There are several differences between Chwe's approach and ours. First, his model has no general equilibrium aspects and does not investigate the relationship between labor scarcity and coercion, which is one of our central objectives. Second, and more importantly, in Chwe's model, the principal cannot affect the agent's outside option, whereas all of our main results follow from our fundamental modeling assumption that coercion is about affecting the outside option of the agent (i.e., coercing an individual to accept an employment contract that he or she would not have otherwise accepted).[6] For example, this modeling assumption is important for our results on efficiency (in Chwe's model, coercion is typically efficiency-enhancing).

The rest of the paper is organized as follows. Section 2 introduces our model. Section 3 characterizes the solution to the principal-agent problem with coercion and presents several key comparative static results. Section 4 studies our general equilibrium model and investigates the relationship between labor scarcity and coercion. Section 5 presents several extensions. Section 6 concludes. Appendix A contains the proofs of Propositions 1 and 2 and relaxes some of the simplifying assumptions used in

---

[6]Interestingly, Chwe agrees with this perspective and writes "one forces another person to be her labourer if one does so by (even indirectly) changing her reservation utility" (p.1110), but then assumes that the principal cannot change the agent's reservation utility. In particular, the agent in Chwe's model has exogenous reservation utility $\bar{U}$, and she receives at least this payoff in expectation under any contract she accepts.

the text. Appendix B, which is available online, considers a generalization of our basic principal-agent model to include multiple levels of output.

2   Model

In this section, we describe the environment. We start with the contracting problem between a "coercive" producer and an agent, and then describe how market prices and outside options are determined. We consider a simplified version of our model in the text and present the general version in Appendix A.

2.1   The Environment

There is a population of mass 1 of identical (coercive) *producers* and a population of mass $L < 1$ of identical *agents* (forced laborers, slaves, or simply workers); L is an inverse measure of labor scarcity in the economy. All agents are risk neutral. Throughout the text we focus on the case where producers are homogeneous and each producer has a project that yields $x > 0$ units of a consumption good if successful and zero units if unsuccessful.[7] Each producer is initially randomly matched with a worker with probability L.

We first describe the actions and timing of events after such a match has taken place. A producer with productivity x who is matched with an agent chooses a level of guns, $g \geq 0$, at cost $\eta \chi(g)$, and simultaneously offers (and commits to) a "contract" specifying an output-dependent wage-punishment pair (w^y, p^y) for $y \in \{h, l\}$, corresponding to high (x) and low (0) output, respectively. Wages and punishments have to be nonnegative, i.e., $w^y \geq 0$ and $p^y \geq 0$, corresponding to the agent having no wealth, and we assume that inflicting punishment is costless for the producer.[8] We also assume that $\chi(0) = 0$ and $\chi(\cdot)$ is twice differentiable, strictly increasing, and strictly convex, with derivative denoted by χ' that also satisfies $\chi'(0) = 0$ and $\lim_{g \to \infty} \chi'(g) = \infty$. The parameter $\eta > 0$ corresponds to the cost of guns, or to the cost of using coercion more generally, which is mainly determined by institutions, regulations and technology (e.g., whether slavery is legal). Following the contract offer of the producer, the agent either accepts or rejects the contract. If she rejects, she receives payoff equal to her (intrinsic) outside option, $\bar{u}$, minus the level of guns, g, i.e.,

$$\bar{u} - g,$$

and the producer receives payoff 0. The interpretation of the agent's payoff is that, if she rejects the contract, the principal inflicts punishment g on her before she "escapes" and receives her outside

---

[7]In Appendix A, we consider the case where there is a distribution of productivity among the producers.

[8]The constraint $w^y \geq 0$ is a standard "limited liability" constraint in principal-agent theory, though due to the possibility of coercion and punishment in the present context this should not be interpreted as the agent having any kind of legal protection.

option, $\bar{u}$.[9] This formulation introduces our main assumption that "coercion" involves using force or the threat of force to convince an agent to accept an employment relationship that she might have rejected otherwise. In light of this, g will be our measure of *coercion* throughout the paper. We can also think of $\bar{u} - g$ as the agent's "extrinsic" outside option, influenced by the coercion of the producer, as opposed to $\bar{u}$, which could be thought of as the "intrinsic" outside option, determined by factors outside the current coercive relationship.

If the agent accepts the contract offer, then she chooses effort level $a \in [0, 1]$ at private cost $c(a)$. Here a is the probability with which the project succeeds, leading to output x. We assume that $c(0) = 0$ and that $c(\cdot)$ is strictly increasing, strictly convex, and twice differentiable, with derivative denoted by c', and we also impose that $\lim_{a \to 1} c(a) = \infty$ to ensure interior solutions. Suppose that the market price for the output of the producer is P. Thus when the agent accepts contract (w^y, p^y) and chooses effort a, guns are equal to g, and output realization is y, the producer's payoff is

$$Py - w^y - \eta \chi(g),$$

and the agent's payoff is

$$w^y - p^y - c(a).$$

An *equilibrium contract* (for given market price, P, and outside option, $\bar{u}$) is the subgame perfect equilibrium of the above-described game between the producer and the agent.[10] Given the timing of events, this equilibrium contract is a solution to the following maximization problem:

$$\max_{\left(a, g, w^h, w^l, p^h, p^l\right) \in [0,1] \times \mathbb{R}_+^5} a\left(Px - w^h\right) + (1 - a)\left(-w^l\right) - \eta \chi(g) \tag{1}$$

subject to

$$a\left(w^h - p^h\right) + (1 - a)\left(w^l - p^l\right) - c(a) \geq \bar{u} - g, \tag{IR$_0$}$$

and

$$a \in \arg \max_{\tilde{a} \in [0,1]} \tilde{a}\left(w^h - p^h\right) + (1 - \tilde{a})\left(w^l - p^l\right) - c(\tilde{a}). \tag{IC$_0$}$$

Here (IR$_0$) can be interpreted as the "individual rationality" or "participation constraint" of the agent. If this constraint is not satisfied, then the agent would reject the contract—run away from the match with the producer. (IC$_0$) is the "incentive compatibility" constraint, ensuring that a is the agent's best response in the subgame following the contract offer and her acceptance of the contract. There is no loss of generality in letting the producer choose a from the set of maximizers, since if the producer

---

[9]This formulation implies that g affects the the agent's utility if she "escapes" rather than her utility when she accepts the employment contract. Thus it is the threat of force, not its actual exercise, that matters. Nevertheless, since g is chosen at the beginning, this threat is only credible (feasible) when the producer undertakes the investments in coercive capacity ("guns").

[10]As our definition of general equilibrium, Definition 1 (below) makes clear, not every equilibrium contract may be part of a general equilibrium, since the levels of P and $\bar{u}$, which we take as given here, may not correspond to equilibrium values. One might thus alternatively refer to an equilibrium contract as an "optimal contract". We use the term equilibrium contract throughout for consistency.

expected an effort level choice from this set that did not maximize his payoff, then he would have a profitable deviation. Thus solutions to this program coincide with subgame perfect equilibria.[11]

In the text, we make the following additional assumption on $c(\cdot)$:

Assumption 1 $c(\cdot)$ *is three times differentiable and satisfies*

$$(1-a)\,c'''(a) \geq c''(a) \ \text{ for all } a.$$

Assumption 1 guarantees that the program characterizing equilibrium contracts, (1) subject to (IR$_0$) and (IC$_0$), is strictly concave,[12] and is adopted to simplify the exposition. Appendix A shows that all of our substantive results hold without this assumption.

2.2 Market Interactions and General Equilibrium

To complete the description of the model, we next describe how the market price, P, and an agent's outside option, $\bar{u}$, are determined.

Denote the unique values of a and g that maximize (1) subject to (IR$_0$) and (IC$_0$) by a^* and g^*.[13] Then, average production among matched producers is

$$Q \equiv a^*x. \tag{2}$$

The *aggregate level of production* is thus QL, and we assume that *market price* is given by

$$P = P(QL), \tag{3}$$

where $P(\cdot)$ is a strictly positive, decreasing and continuously differentiable demand schedule. Equation (3) captures the idea that greater output will reduce price. Equation (2) makes the equilibrium price P a function of the distribution of efforts induced by producers.

An agent's outside option, $\bar{u}$, is determined according to a reduced-form matching model. When an agent escapes, she either matches with another coercive producer or escapes matching with coercive producers altogether (e.g., running away to the city or to freedom in the noncoercive sector). We assume that the probability that an agent who exercises her outside option matches with a randomly drawn, previously unmatched, coercive producer is $\gamma \in [0, 1)$, and the probability that she matches with an outside, noncoercive producer is $1 - \gamma$. In this latter case, she obtains an outside wage $\tilde{u}(L)$, which depends on quantity of labor in the coercive sector. We interpret $\tilde{u}(L)$ as the wage in the noncoercive

---

[11]Observe also that, given (w^h, w^l, p^h, p^l), the agent's maximization problem is strictly concave, which implies that the principal cannot induce the agent to randomize (and this remains true if the principal offers a lottery over (w^h, w^l, p^h, p^l)). Therefore, our implicit assumption that the principal offers a deterministic contract is without loss of generality.

[12]Assumption 1 is a slight weakening of the sufficient condition for concavity of the producer's problem imposed by Chwe. Chwe provides a justification for this apparently ad hoc assumption: define $f(\cdot)$ by $f(-\log(1-a)) \equiv c(a)$, so that $f(\rho)$ is the cost to the worker of ensuring success with probability $1 - e^{-\rho}$; then one can verify that $(1-a)\,c'''(a) \geq 2c''(a)$ if $f'''(\cdot) \geq 0$. Our condition simply weakens this to $(1-a)\,c'''(a) \geq c''(a)$.

[13]Uniqueness is again guaranteed by Assumption 1. In Appendix A, we not only relax Assumption 1, but also allow for heterogeneous producers and mixed strategies.

sector, and assume that $\tilde{u}(L)$ is continuously differentiable and strictly decreasing, consistent with $\tilde{u}(L)$ being the marginal product of labor in the noncoercive sector when the noncoercive production technology exhibits diminishing returns to scale and the quantity of labor in the noncoercive sector is proportional to the quantity of labor in the coercive sector.[14] In practice, both the parameter γ and the exogenous outside option $\tilde{u}(L)$ measure the possibilities outside the coercive sector. For example, in the context of feudalism and forced agricultural labor relations, the existence of cities to which coerced workers may escape would correspond to a low γ and a high $\tilde{u}(L)$.

This formulation implies that the outside option of an agent in a coercive relationship, $\bar{u}$, satisfies

$$\bar{u} = \gamma(\bar{u} - g^*) + (1 - \gamma)\tilde{u}(L). \tag{4}$$

Let G be the average number of guns used by (matched) coercive producers. Since producers are homogeneous, this is simply given by

$$G \equiv g^*. \tag{5}$$

We refer to G as the *aggregate level of coercion* in the economy. Equation (4) can now be written as

$$\bar{u} = \tilde{u}(L) - \frac{\gamma}{1 - \gamma}G. \tag{6}$$

Intuitively, (6) states that an agent's outside option in the coercive sector equals her payoff from exiting the coercive sector minus the aggregate level of coercion, G, as given by (5), times a constant, $\gamma/(1 - \gamma)$, which is increasing in the difficulty of exiting the coercive sector.

Given this description, we now define a (general or market) equilibrium for this economy, referred to as an *equilibrium* for short.[15] Henceforth, we use the terminology "general equilibrium" even though the demand curve $P(\cdot)$ is exogenous.

Definition 1 *An* equilibrium *is a pair* (a^*, g^*) *such that* (a^*, g^*) *is an equilibrium contract given market price* P *and outside option* $\bar{u}$, *where* P *and* $\bar{u}$ *are given by (3) and (6) evaluated at* (a^*, g^*).

Throughout, we impose the following joint restriction on $P(\cdot)$, L, x, $\tilde{u}(\cdot)$, and $c(\cdot)$:

Assumption 2

$$P(Lx)x > \tilde{u}(L) + c'(0).$$

Assumption 2 states that, even if all producers were to set $a = 1$ and $g = 0$, the marginal product of effort would be greater than the agent's outside option plus her cost of effort at $a = 0$. Our analysis below will show that Assumption 2 is a sufficient (though not necessary) condition for all matched

---

[14]For example, the total amount of labor in the economy may be $\tilde{L}$, of which some fraction $\tilde{\gamma}$ is initially matched to coercive producers. Then the quantity of labor in the coercive sector equals $L \equiv \tilde{\gamma}\tilde{L}$, and the quantity of labor in the noncoercive sector equals $(1 - \tilde{\gamma})\tilde{L} = (1 - \tilde{\gamma})L/\tilde{\gamma}$, which justifies simply writing $\tilde{u}(L)$ for the marginal product of labor in the noncoercive sector.

[15]Here, we restrict attention to pure-strategy equilibria without loss of generality, since the program characterizing equilibrium contracts is strictly concave. A more general approach is developed in Appendix A.

producers to induce their agents to exert positive effort (i.e., generate positive expected output) in equilibrium. Therefore, imposing this assumption allows us to focus on the economically interesting case and simplify the exposition considerably.

In the next section, we take the market price, P, and outside option, $\bar{u}$, as given and characterize equilibrium contracts. We then turn to the characterization of (general) equilibrium in Section 4, which will enable us to discuss issues related to the effects of labor scarcity on coercion, as well as to verify the robustness of the partial equilibrium effects in the presence of general equilibrium interactions.

3 Equilibrium Contracts and Comparative Statics

3.1 Equilibrium Contracts

Recall that an equilibrium contract is a solution to (1) subject to (IR_0) and (IC_0). Thus an equilibrium contract is simply a tuple $\left(a^*, g^*, w^h, w^l, p^h, p^l\right) \in [0,1] \times \mathbb{R}_+^5$. Our first result provides a more tractable characterization of equilibrium contracts when they involve positive effort ($a^* > 0$). Throughout the paper, we use the notation $[z]_+ \equiv \max\{z, 0\}$.

Proposition 1 *Suppose $Px > \bar{u} + c'(0)$. Then any equilibrium contract involves $a^* > 0$ and $g^* > 0$, and an equilibrium contract is given by $\left(a^*, g^*, w^h, w^l, p^h, p^l\right)$ such that*

$$\left(a^*, g^*\right) \in \arg\max_{(a,g)\in[0,1]\times\mathbb{R}_+} Pxa - a\left[(1-a)c'(a) + c(a) + \bar{u} - g\right]_+ - (1-a)\left[-ac'(a) + c(a) + \bar{u} - g\right]_+ - \eta\chi(g),$$

(7)

with $w^l = p^h = 0$, $w^h = (1-a^*)c'(a^*) + c(a^*) + \bar{u} - g^* \geq 0$, and $p^l = a^*c'(a^*) - c(a^*) - \bar{u} + g^* \geq 0$.

Proof. See Appendix A. ∎

Remark 1 The condition $Px > \bar{u} + c'(0)$ is automatically satisfied when Assumption 2 holds, since $P \geq P(Lx)$ and $\bar{u} \leq \tilde{u}(L)$. Thus Proposition 1 always applies under our maintained assumption. The qualifier $Px > \bar{u} + c'(0)$ is added for emphasis, since, as the proof illustrates, it ensures that $a^* > 0$, which is in turn important for this result. Problem (7) is not equivalent to the maximization of (1) subject to (IR_0) and (IC_0) when the solution to the latter problem involves $a^* = 0$.

Remark 2 Proposition 1 states that $(1-a)c'(a) + c(a) + \bar{u} - g \geq 0$ and $-ac'(a) + c(a) + \bar{u} - g \leq 0$ in any equilibrium contract; so in any equilibrium contract, the right-hand side of (7) equals

$$Pxa - a(1-a)c'(a) - ac(a) - a\bar{u} + ag - \eta\chi(g).$$

(8)

Straightforward differentiation shows that under Assumption 1 this expression is strictly concave, and thus an equilibrium contract is characterized by a unique pair (a^*, g^*) (see Proposition 2). Hence, in what follows we refer to (a^*, g^*) as an "equilibrium contract".

Remark 3 The maximization problem (7) is (weakly) supermodular in $(a, g, x, P, -\bar{u}, -\eta)$ (see Proposition 11 in Appendix A). We also show in Lemma 3 in Appendix A that "generically" (more precisely, for all parameter values, except for possibly one value of each parameter), the expression $-ac'(a) + c(a) + \bar{u} - g$ is *strictly* less than 0 in any equilibrium contract, and in this case, (7) is *strictly* supermodular in $(a, g, x, P, -\bar{u}, -\eta)$ in the neighborhood of any equilibrium contract. This gives us strict rather than weak comparative statics everywhere. Supermodularity will be particularly useful when we relax Assumption 1 in Appendix A.

To obtain an intuition for Proposition 1, suppose that the solution to (1) (subject to (IR$_0$) and (IC$_0$)) indeed involves $a > 0$. Then recall that c is differentiable and that the first-order approach is valid in view of the fact that there are only two possible output realizations ($y \in \{0, x\}$), which implies that, given the contract offer (w^y, p^y), the agent's maximization problem in (IC$_0$) is concave. Moreover, since $\lim_{a \to 1} c(a) = \infty$, the solution involves $a < 1$. This implies that (IC$_0$) can be replaced by the corresponding first-order condition, where we write $u^h \equiv w^h - p^h$ and $u^l \equiv w^l - p^l$ for the agent's payoff (without effort costs) following the good and bad outcomes:

$$u^h - u^l = c'(a).$$

To punish and pay the agent simultaneously would waste money, so we have $p^y = 0$ if $u^y \geq 0$, and $w^y = 0$ if $u^y \leq 0$. This implies that $w^h = [u^h]_+$ and $w^l = [u^l]_+$, so (1) can be written as:

$$\max_{(a, g, u^h, u^l) \in [0,1] \times \mathbb{R}_+ \times \mathbb{R}^2} a(Px - [u^h]_+) - (1 - a)[u^l]_+ - \eta\chi(g) \tag{9}$$

subject to

$$au^h + (1 - a)u^l - c(a) \geq \bar{u} - g \tag{IR$_1$}$$

and

$$u^h - u^l = c'(a). \tag{IC$_1$}$$

Next, using (IC$_1$) to substitute for u^l in (IR$_1$) shows that this problem is equivalent to maximizing (9) subject to

$$u^h - (1 - a)c'(a) - c(a) \geq \bar{u} - g. \tag{IR$_2$}$$

Finally, using (IR$_2$) to substitute u^h out of (9), and using (IR$_2$) and (IC$_1$) to substitute u^l out of (9), yields (7). Furthermore, it is intuitive that any solution to (7) will necessarily involve $w^l = 0$ and $w^h \geq 0$ if $a > 0$, so that the contract does not punish the agent for a good outcome and does not reward her for a bad outcome.

Note that (7) is strictly concave in g for given a. This, combined with the assumption that $\chi'(0) = 0$ and $\lim_{g \to \infty} \chi'(g) = \infty$, implies that the first-order condition with respect to g, for a given equilibrium level of a,

$$\chi'(g) = \frac{a}{\eta}, \tag{10}$$

is necessary and sufficient whenever (7) is differentiable (and (7) is differentiable in a and g whenever $u^l < 0$, which, as explained in Remark 3, holds almost everywhere; both of these claims are proved in Lemma 3 in Appendix A). This immediately implies that a producer that wishes to induce higher effort will use more guns. Put differently, as noted in Remark 3, (7) is weakly supermodular in a and g everywhere and strictly so whenever $u^l < 0$. Though mathematically simple, this result is both important for our analysis and economically somewhat subtle. One might have presumed that high effort might be associated with less or more coercion. Our model implies that it will always be associated with more coercion. The logic is as follows: (IR$_2$) implies that coercion is valuable to the producer because, regardless of effort, it allows a one-for-one reduction in wages when the agent is successful (i.e., in w^h, since $u^h = w^h$ in an equilibrium contract). An agent who exerts high effort succeeds more often and therefore must be rewarded more often. This makes coercion more valuable to the producer.

Next, recall from Remark 3 that (7) is also supermodular in $(a, g, x, P, -\bar{u}, -\eta)$ (and thus it exhibits increasing differences in (a, g) and $(x, P, -\bar{u}, -\eta)$). This implies that changes in productivity, x, market price, P, outside option, $\bar{u}$, and cost of guns, η, will have unambiguous effects on the set of equilibrium contracts. This observation enables us to derive economically intuitive comparative static results from standard monotonicity theorems for supermodular optimization problems (e.g., Topkis, 1998).[16]

Proposition 2 *There exists a unique equilibrium contract* (a^*, g^*). *Moreover,* (a^*, g^*) *is increasing in* x *and* P *and decreasing in* $\bar{u}$ *and* η.

Proof. See Appendix A. ∎

Proposition 2 is intuitive. Higher x and P both increase the value of a successful outcome for the producer and thus the value of effort. Since effort and coercion are complements, both a and g increase. The effect of P on coercion (g) captures the *labor demand effect*, which was suggested by Domar (1970) and will be further discussed in the next section. Similarly, higher cost of guns, η, reduces the value of coercion. By complementarity between effort and coercion, this reduces both a and g.

Proposition 2 also shows that (a^*, g^*) is decreasing in $\bar{u}$ $(= \tilde{u}(L) - \gamma G/(1-\gamma))$, which is the essence of the *outside option effect*. This result is at first surprising; since higher g offsets the effect of higher $\bar{u}$ (recall (IR$_0$) or (IR$_2$)), one might have expected g and $\bar{u}$ to covary positively. This presumption would also follow from a possible, perhaps mistaken, reading of Domar (1970), on the hypothesis that labor scarcity corresponds to higher $\bar{u}$. However, Proposition 2 shows that the opposite is always the case.[17] The intuition for this result is interesting: An individual with a worse outside option (lower $\bar{u}$) will be induced to work harder because her participation constraint, (IR$_2$), is easier to satisfy, and

---

[16]Throughout "increasing" stands for "strictly increasing," and "nondecreasing" for "weakly increasing".

[17]The only previous analysis of the relationship between coercion and outside options is provided by Chwe (1990) who shows that better outside options lead to higher payoffs after both output realizations for the agent. Because higher payoffs are associated with less ex post punishment, this can be interpreted as better outside options leading to less "coercion". In Chwe's model, this result depends on the agent's risk-aversion (i.e., on "income effects").

agents working harder will be successful more often and will be paid more often. This increases the value of coercion to the producer and equilibrium coercion.

3.2 Discussion of Assumptions

It is useful to briefly discuss the role of various assumptions in leading to the sharp characterization result in Proposition 1 and to (7), which will play a central role in the rest of our analysis. Seven assumptions deserve special mention. First, we assume that the coercive relationship starts with a match between the producer and the agent, and the only reason for the producer to offer an "attractive" contract to the agent is to prevent her from running away. This is important for our analysis, since it implies that producers do not compete with each other in order to attract agents. We believe that this is often a realistic assumption in the context of coercion. Serfs in Europe and forced laborers in Latin America were often tied to the land and employers did not need to attract potential workers into serfdom. Slaves throughout the ages were often captured and coerced. According to Andrees and Belser (2009), even today many forced employment relationships originate when employers are able to lure workers into such relationships, for example by promising good working conditions that do not materialize once workers arrive at a plantation or mine, at which point they are not allowed to leave.

Second, we use a principal-agent model with moral hazard and a "limited liability" constraint, so that the worker cannot be paid a negative wage. We view both of these assumptions as central for a good approximation to actual coercive relationships. Inducing agents to exert effort is a crucial concern in coercive employment relationships, and clearly these agents cannot make (unlimited) payments to their employers, since they are trapped in the coercive relationship without other sources of income. From a theoretical point of view, both of these assumptions are important for our results (and we view this as a strength of our approach in clearly delineating results that depend on distinctive features of coercive relationships). Relaxing either of these two assumptions would imply that the employer could implement the "first-best" level of effort, a_{FB}, given by $Px = c'(a_{FB})$, either by dictating it or by choosing large enough negative payments after low output (given risk neutrality). In particular, in this case the problem of a coercive producer, with productivity x, could be written as $\max_{g,w^h} a_{FB}(Px - w^h) - \eta\chi(g)$ subject to $a_{FB}w^h - c(a_{FB}) \geq \bar{u} - g$. Since the constraint will necessarily hold as equality, this problem can be written as $\max_{g \geq 0} a_{FB}Px - (\bar{u} - g + c(a_{FB})) - \eta\chi(g)$. This problem is no longer strictly supermodular and coercion will always be independent of both $\bar{u}$ and P. Therefore, all of our results depend on the principal-agent approach and the importance of effort and moral hazard (and limited liability).[18]

Third, we allow the principal to use punishment $p \geq 0$. The presence of such punishments is another realistic aspect of coercive relationships. Moreover, they play an important role in our theoretical results by ensuring that the participation constraint, (IR_0) or (IR_2), holds as equality. In the absence

---

[18] Another justification for viewing moral hazard as central to coercion is the presence of ex post inefficient punishments in many coercive relationships. In our model, as well as in all standard principal-agent and repeated game models, no punishments would be observed in equilibrium under perfect monitoring.

of such punishments, the participation constraint can be slack, in which case there would be no role for using g to reduce the (extrinsic) outside option of the agent, and one could not talk of coercion making agents accept employment terms that they would otherwise reject. One could construct different versions of the principal-agent problem, where the participation constraint holds as equality even without punishments, and we conjecture that these models would generate similar insights. Our formulation is tractable and enables us to focus on the key role of coercion in inducing agents to accept employment terms that they would have rejected in the absence of force or threats of force.

Fourth, we impose Assumption 2 throughout, which implies that productivity in the coercive sector is (sufficiently) greater than $\tilde{u}$ and thus greater than agents' (intrinsic) outside option, $\bar{u}$. This makes coercive relationships viable, and corresponds to situations in which coercive producers have access to valuable assets for production, such as land or capital. This type of unequal access to assets of production is a key feature supporting coercive relationships such as serfdom, forced labor, or slavery.

Fifth, we assume that coercion is undertaken by each producer, and thus corresponds to the producer's use of armed guards, enforcers or threat of violence against its laborers. In practice, much coercion is undertaken jointly by a group of producers (for example, via the use of local or national law enforcement, or the judiciary system, as was the case in the US South both before and after the Civil War, e.g., Key, 1949, or Ransom and Sutch, 1977). Moreover, even coercion by each individual producer presumes an institutional structure that permits the exercise of such coercion. A comprehensive study of coercion requires an analysis of the politics of coercion, which would clarify the conditions under which producers can use the state or other enforcement mechanisms to exercise coercion and pass laws reducing the outside option of their employees. Our analysis is a crucial step towards this bigger picture, since it clarifies the incentives of each producer to use coercion before incorporating details of how they will solve the collective action problem among themselves and cooperate in coercive activities. The working paper version shows how our results generalize to the case in which coercion is exercised collectively.

Sixth, we assume risk-neutrality. The effects we focus on in this paper do not disappear in the presence of risk-aversion, though adding risk-aversion complicates the analysis. Nevertheless, there is at least one important way in which making the agent risk-averse reinforces our central intuition that effort and coercion are complementary. Consider the case where $u^l < 0$, so that the sole purpose of coercion is to reduce w^h. By (IR$_2$) and convexity of $c(\cdot)$, u^h is increasing in a (for fixed g), and increasing g allows the principal to reduce u^h one-for-one. When the agent is risk-averse, the wage that the producer must pay after high output to give the agent utility u^h is convex in u^h, since u^h is a concave function of w^h. Reducing u^h is then more valuable to the principal when a is higher, which provides a second source of complementarity between a and g in the principal's problem.[19]

---

[19]The reason that this argument is not completely general is that u^l may equal 0 in an equilibrium contract. However, if the agent's utility function for money, $u(w)$, satisfies $u'''(w) < 0$, it can be shown that the producer's problem, the analogue of (7), is strictly supermodular in (a, g) regardless of u^l (proof available from the authors upon request). In fact, the producer's problem remains strictly supermodular in this case even in the absence of limited liability and/or punishments.

Finally, we assume only two levels of output. This is for simplicity, and Appendix B shows how our results can be generalized to an environment with multiple levels of output.

3.3   Further Comparative Statics

In this subsection, we use Proposition 2 to examine the consequences of coercion for productivity, welfare and wages. We first look at the implications of coercion on worker effort and productivity by changing the cost of coercion, η. Throughout the paper, when we make comparisons between a coercive equilibrium contract ("coercion") and "no coercion," the latter refers to a situation in which either we exogenously impose $g = 0$ or, equivalently, $\eta \to \infty$. Given this convention, the next corollary is an immediate implication of Proposition 2 (proof omitted):

Corollary 1 *Coercion (or cheaper coercion, i.e., lower η) increases effort.*

This result may explain Fogel and Engerman's (1974) finding that productivity was high among slaves in the US South in the antebellum period. It is also intuitive. Coercion and effort are complements, so equilibrium contracts induce less effort when coercion becomes more difficult or is banned.

The next corollary is immediate from the analysis in subsection 3.1 and shows that coercion is unambiguously bad for the welfare of the agent.[20]

Corollary 2 *Coercion (or cheaper coercion, i.e., lower η) reduces agent welfare.*

Proof. Since, as shown above, (IR_0) holds as equality, the welfare of the agent is equal to $\bar{u} - g$. The result then follows from the fact that g is decreasing in η. ∎

Even though coercion reduces agent welfare, it may still increase some measures of "economic efficiency" or net output. In fact, Fogel and Engerman not only documented that slaves in the US South had relatively high productivity, but argued that the slave system may have been "economically efficient." While some aspects of the slave system may have been "efficiently" designed, the next two corollaries show that coercion in our model *always* reduces utilitarian *social welfare*, and may even reduce *net output* (here utilitarian social welfare is the sum of the producer's and worker's utilities, i.e., $Pxa - a(1-a)c'(a) - ac(a) - a\bar{u} + ag - \eta\chi(g) + \bar{u} - g$; and net output is output net of effort costs, i.e., $Px - c(a)$).

First, we show that coercion can lead to effort *above* the first-best level that would prevail in the absence of information asymmetries and limited liability constraints (i.e., a_{FB} given by $c'(a_{FB}) = Px$); this implies that coercion can in fact reduce net output. The argument leading to this result is simple. Since $\lim_{a \to 1} c(a) = \infty$, the first-best effort, a_{FB}, is strictly less than 1. We show that as $\eta \to 0$, any equilibrium contract involves an effort level a arbitrarily close to 1 (see the proof of Corollary 3), and since $\lim_{a \to \infty} c(a) = \infty$, in this case coercion necessarily reduces net output. The

---

[20]This result also contrasts with Chwe's (1990) framework, where the agent receives her outside option (reservation utility) regardless of coercion.

intuition for this is that coercion allows the producer to "steal utility from the agent," as shown by
(IR$_0$) or (IR$_2$). Moreover, since the agent is subject to limited liability, the transfer of utility from the
agent to the producer will take place "inefficiently," by inducing excessive effort. The next corollary
formalizes this argument.

Corollary 3 *There exists $\eta^{**} > 0$ such that if $\eta < \eta^{**}$, then effort a is strictly greater than a_{FB} .*

Proof. From the proof of Proposition 2 in Appendix A, there exists $\eta^* > 0$ such that, for all $\eta \leq \eta^*$,
$u^l < 0$ and a^* solves

$$\max_{a \in [0,1]} Pxa - a\left((1-a)c'(a) + c(a) + \bar{u} - (\chi')^{-1}\left(\frac{a}{\eta}\right)\right) - \eta\chi\left((\chi')^{-1}\left(\frac{a}{\eta}\right)\right), \qquad (11)$$

where we used the fact that, when $u^l < 0$, $\chi'(g^*) = a^*/\eta$. From Proposition 1, $a^* > 0$ for all $\eta > 0$.
Since (11) is differentiable and $\lim_{a \to 1} c(a) = \infty$, the first-order condition

$$(1-a)\left(c'(a) + ac''(a)\right) + c(a) + \bar{u} = (\chi')^{-1}\left(\frac{a}{\eta}\right) + Px$$

is necessary. Now consider $\eta \to 0$. For any $a < 1$ the left-hand side is finite, while, since χ is convex
and satisfies $\lim_{g \to \infty} \chi'(g) = \infty$, and a^* does not converge to 0 as $\eta \to 0$ (because a^* is decreasing in
η, by Proposition 2) the right-hand side converges to ∞. This implies that a^* must also converge to
1 as $\eta \to 0$. Since $a_{FB} < 1$, this completes the proof of the corollary. ∎

The next corollary shows that utilitarian social welfare is *always lower* under coercion.

Corollary 4 *Social welfare in any equilibrium contract under coercion ($\eta < \infty$) is strictly lower than
social welfare in any equilibrium contract under no coercion.*

Proof. Let (a^*, g^*) be an equilibrium contract under coercion. Let SW^C be social welfare under
coercion given (a^*, g), and SW^N social welfare under no coercion.

$$\begin{aligned}
SW^C &= Pxa^* - a^*(1-a^*)c'(a^*) - a^*c(a^*) - a^*\bar{u} + a^*g^* - \eta\chi(g^*) + \bar{u} - g^* \\
&< Pxa^* - a^*(1-a^*)c'(a^*) - a^*c(a^*) - a^*\bar{u} + \bar{u} \\
&\leq \max_{a \in [0,1]} Pxa - a(1-a)c'(a) - ac(a) - a\bar{u} + \bar{u} \\
&= SW^N,
\end{aligned}$$

where the second and third lines are immediate since $g^* > 0$ by (10) and $a^* \leq 1$, and the fourth line
follows because the maximand in the third line is the same as the maximand in (7), i.e., as (8), with
g set to zero, which, by definition, characterizes the equilibrium contract under no coercion. ∎

The intuition for Corollary 4 is simple: coercion is a costly means of transferring utility from the
agent to the producer. Therefore, it is necessarily overused in equilibrium. Despite the simplicity of
this intuition, the result contained in Corollary 4 has not appeared in the literature to the best of our

knowledge, because the central role of coercion in affecting the participation constraint has not been modeled.

Another immediate implication of Proposition 2 is the following (proof omitted):

Corollary 5 *A coerced worker is better off when matched with a less productive producer (i.e., a producer with lower x).*

The intuition (and the proof) is simply that producers with higher x use more coercion and thus give lower welfare, $\bar{u} - g$, to their agents. Once again, though straightforward, this corollary has interesting economic implications. One of these, discussed further in subsection 5.4, is that trading in slaves makes agents worse off, even conditional on their being coerced.

Finally, we consider the cross-sectional relationship between coercion and expected incentive pay, assuming that cross-sectional variation is generated by variation in x. Proposition 1 implies that an equilibrium contract always involves $w^l = 0$ and $w^h = \bar{u} - (\chi')^{-1}\left(\frac{a}{\eta}\right) + (1-a)\,c'(a) + c(a)$. An increase in x leads to an increase in a, and affects w^h only through its effect on a, so

$$\frac{\partial w^h}{\partial a} = -\frac{1}{\eta \chi''(g)} + (1-a)\,c''(a).$$

The sign of this derivative is ambiguous: the direct effect of an increase in a is to increase w^h (as (IR$_2$) binds). But an increase in a also increases g (through (10)), which reduces w^h (again through (IR$_2$)). If the first effect dominates, then an increase in x leads to higher g and w^h; if the second effect dominates, then an increase in x leads to higher g and lower w^h. The former case is particularly interesting because it provides an explanation for Fogel and Engerman's observation that workers who are subjected to more coercion are not necessarily less well-paid; but, in contrast to their interpretation, our result also shows that this has nothing to do with the efficiency of slavery. We state this result in the next corollary (proof in the text):

Corollary 6 *Cross-sectional variation in x leads to a positive correlation between g and w^h if $\partial w^h/\partial a > 0$ for all a, and to a negative correlation between g and w^h if $\partial w^h/\partial a < 0$ for all a.*

4 General Equilibrium

In this section, we characterize and discuss the comparative statics of equilibria as defined in Definition 1. Our main objectives are twofold. The first is to understand the relationship between labor scarcity and coercion, which was one of our main motivating questions. The second is to investigate the robustness of the partial equilibrium insights derived in the previous section.

We first recall that Assumption 2 ensures that $P(QL)\,x > \bar{u} + c'(0) = \tilde{u}(L) - \gamma G/(1-\gamma) + c'(0)$. This implies that Proposition 1 applies and characterizes the set of equilibrium contracts given P and $\bar{u}$ (or, alternatively, G). Then a (general) equilibrium is simply a pair (a^*, g^*) satisfying (7), where P and $\bar{u}$ are given by (3) and (6) evaluated at (a^*, g^*).

Even though both endogenous (general equilibrium) objects, P and $\bar{u}$, depend on the strategy profile of producers via two aggregates, Q and G, defined in (2) and (5), the structure of equilibria in this game is potentially complex, because the game can have multiple equilibria and exhibits neither strategic complements nor strategic substitutes. When a set of producers choose higher (a, g), this increases both Q and G, but the increase in Q reduces P (since the function $P(\cdot)$ is decreasing) and discourages others from increasing their (a, g), while the increase in G reduces $\bar{u}$ and encourages further increases in (a, g). These interactions raise the possibility that the set of equilibria may not be a lattice and make the characterization of equilibrium comparative statics difficult. Nonetheless, under Assumption 1 and our additional assumption that producers are homogeneous, the set of equilibria is a lattice and we can provide tight comparative statics results. In Appendix A, we show how this analysis can be extended to the general case where the set of equilibria may not be a lattice.

If a is an optimal effort for the producer, then the strict concavity of (7) and (10) imply that $\left(a, (\chi')^{-1}\left(\frac{a}{\eta}\right)\right)$ is the unique solution to (7). Homogeneity of producers then implies that all producers choose the same equilibrium contract (a, g). This implies that $Q = a^* x$ and $G = g^*$, so that from (10) we have $G = (\chi')^{-1}\left(\frac{a^*}{\eta}\right) = (\chi')^{-1}\left(\frac{Q}{x\eta}\right)$.[21] Write $a^*(P, \bar{u}, x, \eta)$ and $g^*(P, \bar{u}, x, \eta)$ for the unique equilibrium contract levels of effort and guns given market price P, outside option $\bar{u}$, productivity x, and cost of coercion η, and

$$\phi(Q, x, \gamma, L, \eta) \equiv a^*\left(P(QL), \tilde{u}(L) - \frac{\gamma}{1-\gamma}(\chi')^{-1}\left(\frac{Q}{x\eta}\right), x, \eta\right)x,$$

so that $\phi(Q, x, \gamma, L, \eta)$ is the (unique) level of aggregate output consistent with each producer choosing an equilibrium contract given aggregate output QL and the unique level of G consistent with aggregate output QL and equilibrium contracts. The main role of our assumptions here (homogeneous producers and Assumption 1, which will be relaxed in Appendix A) is to guarantee the existence of such a unique level of G. When the parameters can be omitted without confusion, we write $\phi(Q)$ for $\phi(Q, x, \gamma, L, \eta)$.

It is clear that if Q is an equilibrium level of aggregate output, then Q is a fixed point of $\phi(Q)$. The converse is also true, because if Q is a fixed point of $\phi(Q)$ then $\left(a^* = Q/x, g^* = (\chi')^{-1}(Q/x\eta)\right)$ is an equilibrium. This implies that the set of equilibrium (Q, G) pairs is a lattice, because $(\chi')^{-1}(Q/x\eta)$ in increasing in Q. In what follows, we focus on the comparative statics of the extremal fixed points of $\phi(Q, x, \gamma, L, \eta)$. Even though $\phi(Q, x, \gamma, L, \eta)$ is not necessarily monotone in Q, we will show that it is monotone in its parameters, which is sufficient to establish global comparative static results.

Proposition 3   *1. An equilibrium exists, the set of equilibria is a lattice, and the smallest and greatest equilibrium aggregates (Q, G) are increasing in γ and decreasing in η.*

   2. If $\tilde{u}(L) = \tilde{u}_0$ for all L, then the smallest and greatest equilibrium aggregates (Q, G) are decreasing in L.

---

[21] Equation (10) holds only when $u^l < 0$. Lemma 3 in Appendix A shows that $u^l < 0$ everywhere except possibly for one vector of parameters and that comparative statics are exactly as if (10) holds everywhere.

3. *If $P(QL) = P_0$ for all QL, then the smallest and greatest equilibrium aggregates (Q, G) are increasing in L.*

Proof. $\phi(0, x, \gamma, L, \eta) \geq 0$ and $\phi(x, x, \gamma, L, \eta) \leq x$, and $\phi(Q, x, \gamma, L, \eta)$ is continuous since, given Assumption 1, the program characterizing equilibrium contracts in strictly concave. Note also that $\phi(Q, x, \gamma, L, \eta)$ is increasing in γ and decreasing in η. The first part then follows from this observation combined with Corollary 1 from Milgrom and Roberts (1994) (which in particular states that if $F(x, t) : [0, \bar{x}] \times \mathbb{R}_+^M \to [0, \bar{x}]$ is continuous in x and increasing in t for all x, then the smallest and greatest fixed points of $F(x, t)$ exist and are increasing in t). The second and third parts follow with the same argument since when $\tilde{u}(L) = \tilde{u}_0$ for all L, $\phi(Q, x, \gamma, L, \eta)$ is decreasing in L, and when $P(QL) = P_0$ for all QL, $\phi(Q, x, \gamma, L, \eta)$ is increasing in L (recalling that $\tilde{u}(L)$ is decreasing in L). ∎

Proposition 3 shows that a (general) equilibrium exists and its comparative statics are well-behaved—the smallest and greatest equilibrium aggregates are increasing in the difficulty of leaving the coercive sector, and are decreasing in the cost of coercion. Part 2 illustrates the *labor demand effect*. Because $\tilde{u}(\cdot) = \tilde{u}_0$, labor scarcity has no impact on outside options and only increases price, encouraging coercion in line with Domar's thesis. Part 3 illustrates the *outside option effect*. In this case, $P(\cdot) = P_0$ and labor scarcity simply increases $\tilde{u}$ (and thus $\bar{u}$), reducing coercion. Note, however, that comparative statics with respect to x are ambiguous, because for a fixed level of Q, the corresponding level of G is decreasing in x.[22]

Proposition 3 provides global comparative statics and addresses the labor scarcity issue when either the labor demand effect or the outside option effect is shut down. The next proposition characterizes the conditions under which the labor demand or the outside option effect dominates when both are present. Let us first observe that (8) can be alternatively be rewritten as

$$(P(QL) x - \tilde{u}(L)) a - a(1-a) c'(a) - ac(a) + a \frac{\gamma}{1-\gamma} G + ag - \eta \chi(g). \tag{12}$$

This expression shows that the return to effort is increasing in L if an increase in L decreases market price by less than it decreases the outside option, that is, if

$$QP'(QL) > \tilde{u}'(L). \tag{13}$$

An argument identical to the proof of Proposition 3 shows that an increase in L increases the smallest and greatest equilibrium aggregates (Q, G) if (13) holds for all Q. However, the converse statement that an increase in L decreases the smallest and greatest equilibrium aggregates (Q, G) if the opposite of (13) holds for all Q is not possible since (13) always holds for sufficiently small Q. Nevertheless, equivalents of these results hold locally as shown in the next proposition. Let $(Q^+(L), G^+(L))$ and $(Q^-(L), G^-(L))$ denote the smallest and greatest equilibrium aggregates given labor L, and let us use $(Q^\bullet(L), G^\bullet(L))$ to refer to either one of these two pairs.

---

[22] Another way to see this is to define ϕ as a function of a rather than Q, and to note that for fixed a, the corresponding price $P(axL)$ is decreasing in x. Some comparative statics with respect to x are provided in Appendix A.

Proposition 4 *Suppose that* $Q^{\bullet}(L_0) P'(Q^{\bullet}(L_0) L_0) > \tilde{u}'(L_0)$ *(where* $Q^{\bullet}(L_0)$ *is either* $Q^{+}(L)$ *or* $Q^{-}(L)$*). Then there exists* $\delta > 0$ *such that* $(Q^{\bullet}(L), G^{\bullet}(L)) > Q^{\bullet}(L_0), G^{\bullet}(L_0)$ *for all* $L \in (L_0, L_0 + \delta)$ *(and* $(Q^{\bullet}(L), G^{\bullet}(L)) < Q^{\bullet}(L_0), G^{\bullet}(L_0)$ *for all* $L \in (L_0 - \delta, L_0)$*).*

Conversely, suppose that $Q^{\bullet}(L_0) P'(Q^{\bullet}(L_0) L_0) < \tilde{u}'(L_0)$. *Then there exists* $\delta > 0$ *such that* $(Q^{\bullet}(L), G^{\bullet}(L)) < Q^{\bullet}(L_0), G^{\bullet}(L_0)$ *for all* $L \in (L_0, L_0 + \delta)$ *(and* $(Q^{\bullet}(L), G^{\bullet}(L)) > Q^{\bullet}(L_0), G^{\bullet}(L_0)$ *for all* $L \in (L_0 - \delta, L_0)$*).*

Proof. We will only prove the first part of the proposition for $(Q^{\bullet}(L), G^{\bullet}(L)) = (Q^{-}(L), G^{-}(L))$. The proof for $(Q^{\bullet}(L), G^{\bullet}(L)) = (Q^{+}(L), G^{+}(L))$ and the proof of the second part are analogous.

Fix $L_0 > 0$ (and x, γ, η). First, note that $\phi(Q, x, \gamma, L, \eta)$ is continuous in Q and L (because (8) is strictly concave) and thus uniformly continuous over the compact region $[0, Q^{-}(L_0)] \times [0, 2L_0]$. By hypothesis, $Q^{-}(L_0) P'(Q^{-}(L_0) L_0) > \tilde{u}'(L_0)$. Since $P(\cdot)$, $\tilde{u}(\cdot)$, and $(\chi')^{-1}(\cdot)$ are continuously differentiable, (12) then implies that for any $\varepsilon > 0$, there exists $\bar{\delta} > 0$ such that $\phi(Q, x, \gamma, L, \eta)$ is (strictly) increasing in L on $\{(L, Q) : |L - L_0| \leq \bar{\delta}$ and $|Q - Q^{-}(L_0)| \leq \varepsilon\}$. Recall that $Q^{-}(L_0)$ is the smallest fixed point of $\phi(Q, x, \gamma, L_0, \eta)$. Let $\Psi(L) \equiv \min_{Q:0 \leq Q \leq Q^{-}(L_0) - \bar{\delta}} (\phi(Q, x, \gamma, L, \eta) - Q)$, which is well-defined and continuous in L by Berge's maximum theorem. Moreover, $\Psi(L_0) \geq \bar{\varepsilon}$ for some $\bar{\varepsilon} > 0$, since if it were equal to 0, then this would imply that there exists $\tilde{Q} \in [0, Q^{-}(L_0) - \bar{\delta}]$ such that $\phi(\tilde{Q}, x, \gamma, L, \eta) - \tilde{Q} = 0$, contradicting the fact that $Q^{-}(L_0)$ is the smallest fixed point. As $\Psi(L)$ is continuous, for any $\varepsilon > 0$ there exists δ such that for any $L \in (L_0 - \delta, L_0 + \delta)$, we have $|\Psi(L) - \Psi(L_0)| < \varepsilon$. Choose $\varepsilon = \bar{\varepsilon}$ and denote the corresponding δ by $\hat{\delta}$, and let $\tilde{\delta} = \min\{\hat{\delta}, \bar{\delta}\}$. Then for any $L \in (L_0 - \tilde{\delta}, L_0 + \tilde{\delta})$, $\phi(Q, x, \gamma, L, \eta) - Q > 0$ for all $[0, Q^{-}(L_0) - \tilde{\delta}]$, and thus $Q^{-}(L) > Q^{-}(L_0) - \tilde{\delta}$.

We next show that for any $L \in (L_0, L_0 + \tilde{\delta})$, $Q^{-}(L) > Q^{-}(L_0)$. To obtain a contradiction suppose this is not the case. Since we have already established that $Q^{-}(L) > Q^{-}(L_0) - \tilde{\delta}$, there exists $\hat{L} \in (L_0, L_0 + \tilde{\delta})$ such that $Q^{-}(L_0) - \tilde{\delta} \leq Q^{-}(\hat{L}) \leq Q^{-}(L_0)$. Moreover, given our choice of $\tilde{\delta}$, $\phi(Q, x, \gamma, L, \eta)$ is (strictly) increasing in L on $\{(L, Q) : |L - L_0| \leq \tilde{\delta}$ and $|Q - Q^{-}(L_0)| \leq \tilde{\delta}\}$, and thus $Q^{-}(\hat{L}) = \phi(Q^{-}(\hat{L}), x, \gamma, \hat{L}, \eta) > \phi(Q^{-}(\hat{L}), x, \gamma, L_0, \eta)$. Since $\phi(Q, x, \gamma, L_0, \eta)$ is continuous in Q and $\phi(Q, x, \gamma, L_0, \eta) \geq 0$ for all Q, Brouwer's fixed point theorem now implies that $\phi(Q, x, \gamma, L_0, \eta)$ has a fixed point Q^{*} in $[0, Q^{-}(\hat{L})]$. Since $Q^{-}(\hat{L}) > \phi(Q^{-}(\hat{L}), x, \gamma, L_0, \eta)$, it must be the case that $Q^{*} < Q^{-}(\hat{L}) \leq Q^{-}(L_0)$, where the second inequality follows by hypothesis. But this contradicts the fact that $Q^{-}(L_0)$ is the smallest fixed point of $\phi(Q, x, \gamma, L_0, \eta)$ and completes the proof that $Q^{-}(L) > Q^{-}(L_0)$. Since $G^{-}(L) = (\chi')^{-1}(Q^{-}(L)/(x\eta))$ and χ' is strictly increasing, $G^{-}(L) > G^{-}(L_0)$ follows immediately. ∎

In general, labor scarcity both increases price, encouraging coercion through the labor demand effect, and increases the marginal product of labor in the noncoercive sector, discouraging coercion through the outside option effect. Proposition 4 shows that the impact of a small change in labor scarcity is determined by whether its direct (partial equilibrium) effect on price or outside options is greater. This proposition thus helps us interpret different historical episodes in light of our model.

For example, Brenner (1976) criticized neo-Malthusian theories that predict a negative relationship between labor scarcity and coercion because in many instances, most notably during the "second serf-dom" in Eastern Europe, population declines were associated with more—not less—coercion. Proposition 4 suggests that during the periods of population declines in Western Europe, particularly during the Black Death, which were the focus of the neo-Malthusian theories, labor scarcity may have significantly increased outside options, thus reducing $P - \tilde{u}$ and discouraging coercion. In contrast, it is plausible that the effects of population declines in Eastern Europe were quite different and would involve higher (rather than lower) $P - \tilde{u}$, for at least two reasons. First, the decline in population in this case coincided with high Western European demand for Eastern European agricultural goods. Second, the increase in the outside option of East European workers is likely to have been muted due the relative paucity and weakness of cities in this region. This discussion illustrates that the general model that allows both for labor demand and outside option effects leads to richer, and more subtle, comparative statics, and also highlights that the predictions depend in a simple way on whether labor scarcity has a larger effect on the market price of output or workers' outside options.

5 Extensions

5.1 Ex Ante Investments and Effort- vs. Skill-Intensive Labor

A natural conjecture would be that, in addition to the inefficiencies associated with coercion identified above, coercive relationships discourage ex ante investments by workers. For example, one could use the benchmark model of relationship specific investments introduced in Grossman and Hart (1986), equate "coercion" with the ex post bargaining power of the producer, and thus conclude that coercion should discourage investments by the worker, while potentially encouraging those by the producer. Our model highlights that the effect of coercion on investments is more complex. In particular, coercion will encourage agents to undertake investments that increase their outside options, while at the same time giving them incentives to *reduce* their productivity within the relationship. This is because, as shown in Section 3, greater outside options increase the agent's payoff more than one-for-one (by also reducing coercion), while greater productivity inside the coercive relationship reduces the agent's payoff (by increasing coercion). Conversely, the producer will undertake those investments that increase productivity more than the outside option of the agent. One implication of this result is that the presence of coercion may encourage agents to invest in their general human capital, though for different reasons than the standard Becker approach would suggest. Coercion also discourages them from investing in relationship-specific human capital; in fact, it gives them incentives to sabotage their producer's productive assets.

We model these issues by adding an interim stage to our game between matching and the investment in guns, g. At this stage, matched agents and producers make investment decisions, denoted by i and I, respectively. For simplicity, we will analyze the case in which such investment opportunities are

available only to one party (either the agent or the producer),[23] and will focus on partial equilibrium. Investments potentially affect both productivity x within the relationship and the worker's outside option $\bar{u}$, which we now write as either $x(i)$ and $\bar{u}(i)$, or $x(I)$ and $\bar{u}(I)$, depending on which side makes the investment. Suppose that investment i costs the agent $\zeta(i)$, while investment I costs a producer $\tilde{\zeta}(I)$, and that both cost functions are increasing and convex. We also further simplify the discussion throughout by assuming that $sign(Px'(i) - \bar{u}'(i))$ and $sign(Px'(I) - \bar{u}'(I))$ do not depend on i and I, i.e., each investment always has a larger effect on either Px or $\bar{u}$. This last assumption enables us to clearly separate two different cases, with distinct implications.

Let us first analyze the situation in which only agents have investment opportunities. As a benchmark, note that if there is no coercion (i.e., $\eta = \infty$), a matched worker anticipates receiving expected utility $\bar{u}(i) - \zeta(i)$ after choosing investment i, and therefore chooses i to solve

$$\max_{i \in \mathbb{R}_+} \bar{u}(i) - \zeta(i). \tag{14}$$

Returning to the analysis in Section 3, it is clear that when there is the possibility of coercion (recall that guns, g, is chosen after the agent's investment, i), the agent will receive utility $\bar{u}(i) - g(i) - \zeta(i)$, and she therefore chooses i to solve

$$\max_{i \in \mathbb{R}_+} \bar{u}(i) - g(i) - \zeta(i). \tag{15}$$

To characterize the solution to this program, we need to determine $g(i)$, i.e., how the choice of guns by the producer responds to changes in i. Equation (10) implies that $g(i) = (\chi')^{-1}(a(i)/\eta)$, which we differentiate to obtain

$$g'(i) = \frac{a'(i)}{\eta \chi''(g(i))}. \tag{16}$$

Next, note that the producer's expected profit in an equilibrium contract may be written as

$$(Px(i) - \bar{u}(i))a - a(1-a)c'(a) - ac(a) + ag - \eta\chi(g). \tag{17}$$

Therefore, $sign(a'(i)) = sign(Px'(i) - \bar{u}'(i))$, and thus $sign(g'(i)) = sign(Px'(i) - \bar{u}'(i))$. Combining this with (14) and (15) then immediately yields the following result (proof in the text):

Proposition 5 *Equilibrium investment by the agent under coercion, i^C (the solution to (15)) is smaller [greater] than the no-coercion investment, i^N (the solution to (14)), if $Px'(i) - \bar{u}'(i) > 0$ [if $Px'(i) - \bar{u}'(i) < 0$].*

Proposition 5 implies that under coercion agents will underinvest (compared to no-coercion) in tasks that increase their within-relationship productivity relative to their outside option. This is because when the difference between the agent's productivity and her outside option increases, the producer chooses a contract with higher effort and coercion, which reduces agent welfare.

---

[23] This implies that we are abstracting from indirect effects resulting from the interaction of investments by agents and producers.

Proposition 5 relates to the argument by Fenoaltea (1984), that slavery is often observed in "effort-intensive" tasks, but not in "care-intensive" tasks. Fenoaltea attributes this association to the psychological difficulty of using punishments to motivate care. Our result provides an alternative explanation, under the assumption that "care-intensive" tasks are those where relationship-specific investments by the worker are more important—in this interpretation a corresponds to "effort," while i is associated with "care," and we have in mind tasks where $Px'(\cdot) - \bar{u}'(\cdot) > 0$.[24]

Next, consider the situation where only producers undertake investments. Without coercion, a producer who chooses I and a receives expected payoff

$$(Px(I) - \bar{u}(I)) a - a(1-a)c'(a) - ac(a) - \tilde{\zeta}(I), \tag{18}$$

while with coercion he receives expected payoff

$$(Px(I) - \bar{u}(I)) a - a(1-a)c'(a) - ac(a) + ag - \eta\chi(g) - \tilde{\zeta}(I). \tag{19}$$

Clearly, the producer will choose $I = 0$ if $Px'(I) - \bar{u}'(I) \leq 0$, regardless of whether we allow coercion; this is a version of the standard result in human capital theory in which producers never provide general skills training. If, on the other hand, $Px'(I) - \bar{u}'(I) > 0$, then with the same arguments as in Section 3, it can be verified that (19) is supermodular in $(I, a, g, -\eta)$. Now the comparison between producer investment under coercion and no-coercion can be carried out by noting that (18) is equivalent to (19) as $\eta \to \infty$. Supermodularity of (19) then immediately gives the following result (proof in the text):

Proposition 6 *Equilibrium investment by the producer under coercion, I^C (the solution to (19)), is greater [smaller] than the no-coercion investment, I^N (the solution to (18)), if $Px'(I) - \bar{u}'(I) > 0$ [if $Px'(I) - \bar{u}'(I) < 0$].*

Proposition 6 has a similar interpretation to Proposition 5: investment incentives are determined by whether investment has a greater impact on productivity or the outside option. In contrast to the agent, the producer has greater incentives to invest when relationship-specific productivity increases more than the outside option. The general principle here is related to the result in Acemoglu and Pischke (1999) that employers will invest in general human capital when there is "wage distortion," so that these investments increase worker productivity inside the relationship by more than their outside wage.

5.2 Labor Scarcity and the Returns to Investment in Guns

In this subsection, we highlight another general equilibrium mechanism linking labor scarcity to coercion. The underlying idea is related to Acemoglu, Johnson and Robinson's (2002) argument that

---

[24]This is consistent with Fenoaltea's discussion, which emphasizes the association between "care-intensive" and skill-intensive tasks. For example, he notes that uncoerced galley crews were sometimes used because "the technically superior rowing configuration did require skilled oarsmen" (1984, p. 642), and that "at least the skilled branches of factory production" were care-intensive (p. 654).

labor coercion is more profitable when there is abundant labor to coerce (because, when labor is scarce, coercion can exploit "economies of scale"). This channel is absent in our model so far, because each employer employs at most one worker and coercion decisions are made conditional on the match. An alternative timing of events is to assume that producers invest in guns before the matching stage. This minor change in timing introduces the above-mentioned economies of scale effect and implies that investment in guns will be less profitable when producers are relatively unlikely to match with workers, that is, when labor is scarce. To bring out this particular general equilibrium effect, we abstract from the labor demand and outside option effects by assuming that $P(\cdot) \equiv P_0$ and $\tilde{u}(\cdot) \equiv \tilde{u}_0$.

The only difference from our baseline analysis is that producers choose g before they learn whether they are matched with an agent. Thus they have a two-stage decision problem: they first choose g before matching, and then after matching, they propose a contract to the agent, which, as before, can be summarized by (a, g). Even though this is a two-stage decision problem, there is no loss of generality in formulating it mathematically as choosing a and g simultaneously, that is:

$$
\max_{(a,g)\in[0,1]\times\mathbb{R}_+} L \left(aP_0x - a\left[(1-a)\,c'(a) + c(a) + \tilde{u}_0 - \frac{\gamma}{1-\gamma}G - g\right]_+ \right.
$$
$$
\left. - (1-a)\left[-ac'(a) + c(a) + \tilde{u}_0 - \frac{\gamma}{1-\gamma}G - g\right]_+ \right) - \eta\chi(g),
$$

with the interpretation that a is the level of effort that will be chosen following a match with an agent (and we have again substituted out the incentive compatibility and participation constraints under Assumption 2). Rewriting this as

$$
\max_{(a,g)\in[0,1]\times\mathbb{R}_+} aP_0x - a\left[(1-a)\,c'(a) + c(a) + \tilde{u}_0 - \frac{\gamma}{1-\gamma}G - g\right]_+
$$
$$
- (1-a)\left[-ac'(a) + c(a) + \tilde{u}_0 - \frac{\gamma}{1-\gamma}G - g\right]_+ - \frac{\eta}{L}\chi(g),
$$

we see that changing the timing of the model by requiring producers to choose g before matching is formally identical to replacing the cost of guns η with η/L. This implies that the analysis of the set of equilibria is similar to that in Proposition 3 and yields (proof omitted):

Proposition 7 *Consider the modified model presented in this subsection. Then, an equilibrium exists and the set of equilibria is a lattice. Labor scarcity reduces coercion, that is, a decline in L reduces the smallest and greatest equilibrium aggregates (Q, G). Moreover, the smallest and greatest equilibrium aggregates (Q, G) are increasing in P_0, γ, and x, and decreasing in $\tilde{u}_0$ and η.*

This proposition thus formalizes another channel via which labor abundance can encourage coercion, and thus complements Proposition 3 in the previous subsection. Naturally, if we relax the assumption that $P(\cdot) \equiv P_0$ and $\tilde{u}(\cdot) = \tilde{u}_0$, the implications of labor scarcity for coercion will be determined by competing effects as in Proposition 4.

5.3 Interim Participation Constraints

We have so far assumed that the agent cannot "run away" once she accepts a contract offer. In practice, coerced agents may attempt to run away not only ex ante but also after the realization of output. This would amount to one more "IR" or participation constraint for the agent, which we refer to as the *interim participation constraint*. The presence of such an intern participation constraint introduces a potential "useful" role of coercion. Intuitively, the agent may prefer ex ante to commit to not running away after output is realized; this is similar to the logic that operates in models such as Chwe's, where coercion does not affect the ex ante participation constraint but does enable punishment. Interestingly, however, we will see that, under fairly weak assumptions, this effect is dominated by the negative impact of coercion on welfare, and comparative statics remain unchanged.

Formally, in this subsection we again focus on partial equilibrium and assume that the agent will run away after the realization of y if her (equilibrium-path) continuation utility is below $\bar{u} - \Phi(g)$, where $\bar{u}$ is the outside option of the agent introduced above and $\Phi(g) \geq 0$ is an increasing function of guns; the interpretation of this is that the producer can inflict punishment $\Phi(g)$ if the agent runs away after the realization of y, just as she can inflict punishment g if the agent runs away before the realization of y.[25] This introduces an "interim IR constraint" in addition to the "ex ante IR constraint," (IR$_0$), in Section 2. This implies that both $w^h - p^h$ and $w^l - p^l$ must be greater than $\bar{u} - \Phi(g)$, though naturally only the constraint $w^l - p^l \geq \bar{u} - \Phi(g)$ can be binding in an equilibrium contract. Therefore, the considerations discussed in this subsection introduce the additional constraint

$$w^l - p^l \geq \bar{u} - \Phi(g). \tag{IIR}$$

Now, by an argument similar to that in Section 3, an equilibrium contract is a solution to the maximization of (1) subject to (IC$_0$), (IR$_0$), and now (IIR).

Suppose first that $\Phi(g)$ is sufficiently large for all g. It is then clear that (IIR) will always be slack and the producer's problem is identical to that in Section 3, and Proposition 2 applies. Thus, we assume in this subsection that $\Phi(g)$ is "not too large," and in particular assume that $\Phi(g) \leq g$, which implies that (IIR) is binding and "replaces" (IR$_0$):

Lemma 1 *Consider the model with the interim participation constraint and suppose that $\Phi(g) \leq g$. Then (IR$_0$) is slack and (IIR) binds.*

Proof. The first-order approach still applies, so we can replace (IC$_0$) with (IC$_1$). By (IIR), $w^l - p^l \geq \bar{u} - \Phi(g)$. Substituting this into (IC$_1$) gives $w^h - p^h \geq c'(a) + \bar{u} - \Phi(g)$. Therefore:

$$a\left(w^h - p^h\right) + (1-a)\left(w^l - p^l\right) - c(a) \geq \bar{u} - \Phi(g) + ac'(a) - c(a) \geq \bar{u} - \Phi(g) \geq \bar{u} - g,$$

---

[25] For example, g may be the pain that the producer can inflict on a worker if she runs away on the first day on the job, while $\Phi(g)$ may be the pain that the producer can inflict on the worker once she has set up a home on the producer's plantation.

where the second inequality follows by convexity of $c(a)$ and the fact that $c(0) = 0$, and the third follows by the assumption that $\Phi(g) \leq g$. This chain of inequalities implies (IR$_0$). Given that (IR$_0$) is slack, the fact that increasing p^l relaxes (IC$_1$) implies that (IIR) must bind. ∎

Provided that $\Phi(g) \leq g$, Lemma 1 then allows us to substitute (IC$_1$) into (IIR), which implies that equilibrium contracts are characterized by

$$\max_{(a,g)\in[0,1]\times\mathbb{R}_+} aPx - a\left[c'(a) + \bar{u} - \Phi(g)\right]_+ - (1-a)\left[\bar{u} - \Phi(g)\right]_+ - \eta\chi(g). \tag{20}$$

(20) is supermodular in $(a, g, x, P, -\bar{u}, -\eta)$, so comparative statics go in the same direction as in our baseline model, though they may hold with weak rather than strict inequalities:

Proposition 8 *If $\Phi(g) \leq g$ for all g, then (a^*, g^*) are nondecreasing in x and P and nonincreasing in $\bar{u}$ and η.*

The intuition for this is that, regardless of whether (IR$_0$) or (IIR) are binding, increasing g reduces the amount that the producer must pay the agent after high output is realized, leading to complementarity between effort and coercion as in Section 3.

While the possibility that (IIR) rather than (IR$_0$) may be binding does not affect our comparative static results, it does suggest that coercion may play a socially useful role. In particular, ex post punishments may be useful in providing incentives to an agent who is subject to limited liability, and (IIR) limits the use of such punishments; one may then conjecture that increasing g may increase social welfare if it relaxes (IIR). We next show that this conjecture is not correct. To see this, note that if (IIR) is binding and coercion is not allowed (i.e., $\eta = \infty$), then $\bar{u} - \Phi(g) = \bar{u} - \Phi(0) = \bar{u}$, the producer's problem becomes $\max_{a\in[0,1]} aPx - ac'(a) - a\bar{u} - (1-a)[\bar{u}]_+$, and (utilitarian) social welfare is

$$SW^N = \max_{a\in[0,1]} aPx - ac'(a) - a\bar{u} - (1-a)[\bar{u}]_+ + \bar{u};$$

while with coercion the producer's problem is given by (20), and social welfare corresponding to an equilibrium contract involving (a, g) is

$$SW^C = aPx - ac'(a) - a\bar{u} + a\Phi(g) - (1-a)[\bar{u} - \Phi(g)]_+ - \eta\chi(g) + \bar{u} - g.$$

If $\Phi(g) \leq g$ for all g, then it is clear that $SW^N \geq SW^C$, with strict inequality if $g > 0$. Thus, coercion reduces social welfare, if $\Phi(g) \leq g$. Formally, we have the following (proof in text):

Corollary 7 *Suppose that $\Phi(g) \leq g$ for all g, and let SW^C be social welfare corresponding to an equilibrium contract with coercion. Then $SW^N \geq SW^C$, with strict inequality if $g^* > 0$.*

The intuition for Corollary 7 follows by comparing (7) to (20). Both (7) and (20) are supermodular in (a, g)—because of the term ag in (7) and the term $a\Phi(g)$ in (20). Our result in Section 3 that coercion reduces social welfare exploits the fact that coercion enters into worker welfare through the

term $-g$, which is always larger in absolute value than ag. If $\Phi(g) \leq g$, then g is greater than $a\Phi(g)$, so the same argument implies that coercion reduces social welfare.

It is worth noting that the analysis here is also informative about the case where ex post punishments are costly. In particular, the model in this subsection corresponds to the case where the producer requires at least g guns in order to inflict punishment $\Phi(g) - \bar{u}$. Thus, it also shows that our comparative static and welfare results continue to hold in this case.[26]

5.4 Trade in Slaves

In this subsection, we briefly investigate the effects of "trade in slaves," whereby agents subject to coercion are sold from one producer to another. We also investigate the related issue of comparative statics in the presence of "fixed costs of coercion," for example, in the presence of a fixed cost (price) of obtaining a worker to coerce or of violating legal or ethical proscriptions against coercion.

To analyze trading in slaves, we assume, as in Appendix A, that producers differ in their productivities, which are drawn independently from a distribution $F(x)$. Let us also assume that, initially, matched producers are a random selection from the population of producers, and thus their productivity distribution is given by $F(x)$. Since $L < 1$, some producers are left unmatched. Now suppose that, before the investment in guns, agents can be bought and sold among producers. Since more productive (higher x) producers have a greater willingness to pay for coerced agents, this trading will ensure that all of the agents will be employed by the most productive L producers. Consequently, the distribution of productivity among matched producers will be $\tilde{F}(x) = 0$ for all $x < x_{1-L}$ and $\tilde{F}(x) = F(x)/L$ for all $x \geq x_{1-L}$, where x_{1-L} is the productivity of the producer at the $(1 - L)$th percentile of the productivity distribution. This implies that trade in slaves is equivalent to a first-order stochastically dominating shift in the distribution of productivity in the context of our model. We know from Proposition 2 that in partial equilibrium (i.e., with P and $\bar{u}$ fixed), this increases effort and coercion, reducing worker welfare. There is a potential offsetting effect, resulting from the fact that the reallocation of agents across producers increases productivity and aggregate output, Q, and thus decreases price $P(QL)$. In this subsection, we abstract from this effect by assuming that $P(QL) = P_0$. Under this assumption, we show that trade in slaves increases the amount of coercion and reduces agent welfare, and that it may also reduce (utilitarian) social welfare despite the fact that agents are now allocated to the most productive producers.[27]

---

[26] An alternative, perhaps more direct way of modeling "costly punishments" is to assume that inflicting punishment (utility) $u^l < 0$ costs the producer $\xi(u^l)$, where $\xi \geq 0$ and $\xi' \geq 0$. In this case, (8) becomes

$$Pxa - a\left((1-a)c'(a) + c(a) + \bar{u} - g\right) - (1-a)\xi\left(ac'(a) - c(a) - \bar{u} + g\right) - \eta\chi(g),$$

which is still supermodular in $(a, g, x, P, -\bar{u}, -\eta)$ provided that ξ is concave. Hence, our baseline comparative statics hold in this alternative model if there are decreasing marginal costs of punishment.

[27] Note that our analysis of "trade in slaves" ignores the possibility of bringing additional slaves into the market, as the Atlantic slave trade did until its abolition in 1807. We show that slave trade reduces worker welfare even if one ignores its effect on the number of coerced workers (which is outside our model).

Proposition 9 *Assume that $P(QL) = P_0$ for all QL. Introducing slave trade in the baseline model increases coercion (G) and reduces agent welfare. More formally, the smallest and greatest equilibrium levels of coercion (resp., average agent welfare) under slave trade are greater (resp., smaller) than the smallest and greatest equilibrium levels of coercion (resp., average agent welfare) under no slave trade. In addition, social welfare may decline under slave trade.*

The proof is omitted. Most of the proof follows from our analysis of comparative statics with heterogeneous producers provided in Proposition 14 in Appendix A. The result that social welfare may decline under coercion follows from a parametric example, which is contained in the working paper version (available upon request). The intuition for why trade in slaves (or, equivalently, an exogenous first-order stochastic dominance increase in the distribution of productivities, $F(\cdot)$) may reduce social welfare, note that increasing productivity x has two offsetting effects on welfare. First, fixing a, increasing x by Δ increases expected output by Δa (a first-order positive effect). Second, increasing x increases a and thus g. The increase in g has a first-order negative effect on worker welfare, but no first-order effect on producer welfare by the Envelope Theorem. Thus, it is straightforward to construct an example in which increasing productivity reduces social welfare by choosing parameters that induce a low choice of a.

We can also incorporate the price paid for a coerced agent or other fixed costs of coercion in a straightforward manner. In particular, we simply need to change the cost of coercion to $\eta\chi(g) + \kappa\mathbf{1}\{g > 0\}$, where $\kappa > 0$ is a constant and $\mathbf{1}\{g > 0\}$ is the indicator function for the event $g > 0$. This introduces a natural nonconcavity into the producer's optimization problem, which leads the producer to either set $a = a^N$ and $g = 0$ ("no-coercion"), or set $a = a^C$ and $g = (\chi')^{-1}(a^C/\eta)$ ("coercion"), paying the fixed cost of coercion and then setting the level of coercion according to equation (10). As we have seen, $a^N = \arg\max_a Pxa - a(1-a)c'(a) - ac(a) - a\bar{u}$. Comparing this with (8) and observing that $a^C > a^N$ implies the following result:[28]

Proposition 10 *Coercion is more likely when P or x are higher, or when $\bar{u}$, η, or κ are lower.*

Proposition 10 shows that comparative statics of the decision of whether to use coercion are identical to comparative statics on the level of coercion.

6 Concluding Remarks

Standard economic models assume that transactions in the labor market are "free". For most of human history, however, the bulk of labor transactions have been "coercive," in the sense that the threat of force was essential in convincing workers to take part in the employment relationship and accept the terms of employment. The small existing literature on slavery and forced labor does not model what

---

[28] In this result, "coercion is more likely when z is higher," means that there is a larger set (in the set-inclusion sense) of parameters other than z for which coercion is optimal when z is higher.

we view as the fundamental effect of coercion on labor transactions—coercion makes workers accept employment relations (terms of employment) that they would otherwise reject. This paper provides a tractable model incorporating this feature and uses it to provide a range of comparative static results useful for understanding when coercive labor relations are more likely to arise and what their welfare implications are.

At the heart of our model is the principal-agent relationship between a potentially coercive producer and an agent. Coercion, modeled as investment by the producer in "guns," affects the participation constraint of the agent. We first analyzed this principle-agent problem and derived partial equilibrium comparative statics, and then embedded it in a general equilibrium framework to study the relationship between labor scarcity/abundance and labor coercion. Both our partial and general equilibrium analyses rely on the complementarity between effort and coercion. This complementarity (supermodularity) is not only mathematically convenient, but also economically central: greater effort implies that the principal will have to reward the agent more frequently because success is more likely. But this also implies that greater coercion, which reduces these rewards, becomes more valuable to the principal. As a consequence, agents with higher marginal productivity in the coercive sector will be coerced more (the *labor demand effect*) and agents with better outside options will be coerced less (the *outside option effect*).

We show that, consistent with Fogel and Engerman (1974), coercion increases effort. However, our formulation also implies that, in contrast to Fogel and Engerman's interpretation, this does not imply that coercion is or may be "efficient". On the contrary, coercion always reduces (utilitarian) social welfare, because the structure of the principal-agent model dictates that coercion hurts the agent more than the additional effort it induces helps the principal. Our model also shows that coercion changes both workers' and producers' incentives to make ex ante investments in their relationships, and points out a new channel via which trading coerced workers makes them worse off.

A major question in the economics of coercion, both from a historical perspective and for understanding the continued prevalence of forced labor today, is the effect of labor scarcity/abundance on coercion. Domar (1970) argues that labor scarcity encourages coercion by increasing the cost of hiring workers in the market. The neo-Malthusians, on the other hand, link the decline of feudalism and serfdom to falling population in Western Europe starting in the second half of the 15th century, which led to new opportunities for free agriculture. Relatedly, Acemoglu, Johnson and Robinson (2002) suggest that Europeans were more likely to set up coercive institutions in colonies with abundant labor, because setting up such institutions was not profitable when labor was scarce. Our general equilibrium analysis shows why these diverse perspectives are not contradictory. Labor scarcity creates a labor demand effect: it increases the marginal product of workers in the coercive sector, and thus encourages employers to use greater coercion and extract higher effort from their workers. It also creates an outside option effect: it increases the outside option of the workers in the noncoercive sector, and reduces coercion because employers demand lower effort and use less coercion when workers have greater outside options. Finally, it creates an economies of scale effect: it makes it more likely

that up-front investments in coercive instruments will go to waste due to lack of labor. Whether the labor demand effect or the outside option effect dominates simply depends on whether the population change has a larger direct effect on the market price or the workers' outside options.

We view this paper as a first step towards a systematic analysis of coercion in the labor market and its implications for the organization of production and economic development. Despite the historical and current importance of forced labor and other coercive relations, many central questions in the economics of coercion have not previously been answered or even posed. Our framework provides a first set of answers to these questions and can serve as a starting point for different directions of study. Theoretical and empirical investigations of the dynamics of coercion, of why coercive relationships persist in many developing countries even today, of the effects of coercion on technology choices and organizational decisions, and of how coercive production impacts trade are important areas for future research. A particularly fruitful area of future research is a more in-depth analysis of the politics of coercion. We presumed the presence of an institutional environment that permitted coercion by producers. In many instances, coercion comes to an end, or is significantly curtailed, when political forces induce a change in the institutional environment. Combining our microeconomic model of coercion with a model of endogenous institutions would be one way of making progress in this direction.

Appendix A: General Case

Equilibrium Contracts and Partial Equilibrium Comparative Statics

In this appendix we generalize our model by allowing for heterogeneity among producers and relaxing the assumption that the program characterizing equilibrium contracts is concave. Formally, we now drop Assumption 1 and assume that each producer's productivity, x, is independently drawn from a distribution $F(x)$ with support $[\underline{x}, \bar{x}]$ with $\underline{x} > 0$. We present most proofs in this general model.

Without Assumption 1, the program characterizing equilibrium contracts may have multiple solutions. Nonetheless, Proposition 1 applies with the sole modification that the condition $Px > \bar{u} + c'(0)$ is replaced by $P\underline{x} > \bar{u} + c'(0)$. We therefore present the general proof here.

Proof of Proposition 1. First, observe that it is suboptimal to set both $w^h > 0$ and $p^h > 0$, or $w^l > 0$ and $p^l > 0$, as reducing both w^h (w^l) and p^h (p^l) by $\varepsilon > 0$ would strictly increase profits without affecting (IR$_0$) or (IC$_0$), so $w^h = [u^h]_+$ and $w^l = [u^l]_+$. With two possible outcomes, the first-order approach is valid, so (IC$_0$) can be rewritten as (IC$_1$), which also exploits the fact that the first-order condition must hold as equality, because $\lim_{a \to 1} c(a) = \infty$ and therefore $a^* < 1$. Then the producer's problem can be written as

$$\max_{\left(a, g, u^h, u^l\right) \in [0,1] \times \mathbb{R}_+ \times \mathbb{R}^2} a \left(Px - [u^h]_+ \right) - (1-a)[u^l]_+ - \eta\chi(g) \tag{A-1}$$

subject to

$$au^h + (1-a)u^l - c(a) \geq \bar{u} - g \tag{IR$_1$}$$

and

$$u^h - u^l = c'(a). \tag{IC$_1$}$$

Substituting (IC$_1$) into (IR$_1$) yields

$$u^h - (1-a)c'(a) - c(a) \geq \bar{u} - g. \tag{IR$_2$}$$

(IR$_2$) must bind, as otherwise decreasing u^h or increasing a would increase profit. Using (IC$_1$) and (IR$_2$) to substitute u^h and u^l out of (A-1) now yields (7).

It remains only to show that $g^* > 0$, $u^l \leq 0$, and $u^h \geq 0$, and finally that $a^* > 0$ in any equilibrium contract (that $w^h = (1-a)c'(a) + c(a) + \bar{u} - g > 0$ and $p^l = -ac'(a) + c(a) + \bar{u} - g \geq 0$ then following immediately from (IR$_2$) and (IC$_1$)). First, the result that $g^* > 0$ in any equilibrium contract with $a^* > 0$ follows from (7), since it must be that $\chi'(g^*) \in \left[\frac{a^*}{\eta}, \frac{1}{\eta}\right]$ in any equilibrium contract. The result that $u^l \leq 0$ and $u^h \geq 0$ is established in the next lemma.

Lemma 2 *In any equilibrium contract with $a > 0$, we have $u^l \leq 0$ and $u^h \geq 0$.*

Proof. Note first that the Lagrangian for (A-1) subject to (IR$_2$) and (IC$_1$) is

$$a\left(Px - [u^h]_+\right) - (1-a)[u^l]_+ - \eta\chi(g) + \lambda\left(au^h + (1-a)u^l - c(a) - (\bar{u}-g)\right) + \mu\left(u^h - u^l - c'(a)\right). \tag{A-2}$$

Now suppose to obtain a contradiction that $u^l > 0$. Then $[u^l]_+ = u^l$ and is differentiable. Moreover, in this case $u^h > 0$, as $u^h > u^l$ (since $a > 0$), and thus $[u^h]_+ = u^h$ is also differentiable. Clearly, $\frac{d[u^h]_+}{du^h} = \frac{d[u^l]_+}{du^l} = 1$. Then differentiating (A-2) with respect to u^h and u^l, we obtain

$$1 = \lambda + \frac{\mu}{a}, \text{ and} \tag{FOCu^h}$$

$$1 = \lambda - \frac{\mu}{1-a}. \tag{FOCu^l}$$

These first-order conditions always hold, as setting u^l or u^h to ∞ or $-\infty$ cannot be optimal if $a \in (0,1)$. Combining (FOCu^h) and (FOCu^l) then implies that $\mu = 0$.

Now differentiating (A-2) with respect to a and using (IC$_1$) yields

$$Px - \left([u^h]_+ - [u^l]_+\right) = \mu c''(a). \tag{FOCa}$$

(FOCa) holds with equality by our assumptions that $a > 0$ and that $\lim_{a\to 1} c(a) = \infty$. The fact that $\mu = 0$ then implies that $[u^h]_+ - [u^l]_+ = u^h - u^l = Px$. Since $[u^h]_+ - [u^l]_+ = u^h - u^l = Px$ and $\mu = 0$, the Lagrangian (A-2) becomes

$$-u^l - \eta\chi(g) + \lambda\left(u^l + Px - c(a) - (\bar{u}-g)\right),$$

which is maximized at $a = 0$, contradicting our assumption that $a > 0$. This completes the proof of the claim that $u^l \leq 0$.

Finally, to show $u^h \geq 0$, suppose to obtain a contradiction that $u^h < 0$. Then (A-2) is differentiable with respect to u^h at the optimum, and its derivative with respect to u^h is $a\lambda + \mu$. Since $a > 0$, this can equal 0 only if $\lambda = 0$ and $\mu = 0$. But then the maximum of (A-2) over $a \in [0,1]$ is attained at $a = 1$, which violates (IC$_1$) by our assumption that $\lim_{a \to 1} c(a) = \infty$. ∎

It remains only to check that a solution to the producer's problem with $a^* > 0$ exists if $P\underline{x} > \bar{u} + c'(0)$. Consider the producer's problem of first choosing a and then maximizing profit given a. The producer's optimal profit given fixed a is continuous in a. Therefore, it is sufficient to show that the producer's optimal profit given a is increasing in a for all sufficiently small $a > 0$. The producer's problem, given $a > 0$, is maximizing (7) over $g \in \mathbb{R}_+$:

$$\max_{g \in \mathbb{R}_+} a\left(Px - [(1-a)c'(a) + c(a) + \bar{u} - g]_+\right) - (1-a)\left[-ac'(a) + c(a) + \bar{u} - g\right]_+ - \eta\chi(g). \quad \text{(A-3)}$$

The right-derivative of (A-3) with respect to a is

$$Px - \left([u^h]_+ - [u^l]_+\right) - a(1-a)c''(a)\left(\frac{d[u^h]_+}{du^h} - \frac{d[u^l]_+}{du^l}\right), \quad \text{(A-4)}$$

where $d[u^h]_+ / du^h$ and $d[u^l]_+ / du^l$ denote right-derivatives. We have shown above that $u^h \geq 0$ and $u^h > u^l$ at an optimum with $a > 0$, and $d[u^h]_+/du^h - d[u^l]_+ / du^l \leq 1$, so (A-4) is no less than $Px - u^h - a(1-a)c''(a)$. By (IR$_2$), as a converges to 0, u^h converges to at most $\bar{u} + c'(0)$. Therefore, provided that $Px > \bar{u} + c'(0)$, (A-4), and thus the derivative of (A-3) with respect to a, is strictly positive for sufficiently small a. This establishes that $a^* > 0$ and completes the proof. ∎

We now state and prove the generalization of Proposition 2, which allows for multiple equilibrium contracts:

Proposition 11 *The set of equilibrium contracts for a producer of type x forms a lattice, with smallest and greatest equilibrium contracts $(a^-(x), g^-(x))$ and $(a^+(x), g^+(x))$. Moreover, $(a^-(x), g^-(x))$ and $(a^+(x), g^+(x))$ are increasing in x and P and decreasing in $\bar{u}$ and η.*

Proof. First note that the maximization problem (7) is weakly supermodular in $(a, g, x, P, -\bar{u}, -\eta)$. This follows because $c'(a) > 0$, and therefore the first bracketed term in (7) $((1-a)c'(a) + c(a) + \bar{u} - g)$ is greater than the second $(-ac'(a) + c(a) + \bar{u} - g)$. Hence, (7) equals either $Pxa - c(a) - \bar{u} + g - \eta\chi(g)$ (if both bracketed terms are positive), $Pxa - a((1-a)c'(a) + c(a) + \bar{u} - g) - \eta\chi(g)$ (if only the first is positive), or $Pxa - \eta\chi(g)$ (if neither are positive). In each of these cases, (7) is weakly supermodular in $(a, g, x, P, -\bar{u}, -\eta)$, so the fact that (7) is continuous in $(a, g, x, P, -\bar{u}, -\eta)$ implies that it is weakly supermodular in $(a, g, x, P, -\bar{u}, -\eta)$.

The supermodularity of (7) in $(a, g, x, P, -\bar{u}, -\eta)$ implies that the set of equilibrium contracts for a producer of type x forms a lattice and that the smallest and greatest equilibrium contracts, $(a^-(x), g^-(x))$ and $(a^+(x), g^+(x))$, are nondecreasing in x and P and nonincreasing in $\bar{u}$ and η (see, e.g., Theorems 2.7.1 and 2.8.1 in Topkis (1998)).

In the rest of the proof, we show that these comparative static results are strict in the sense that whenever we have a change from $(x, P, \bar{u}, \eta)$ to $(x', P', \bar{u}', \eta')$, where $x' \geq x$, $P' \geq P$, $\bar{u}' \leq \bar{u}$, and $\eta' \leq \eta$, with at least one strict inequality, we have "increasing" instead of "nondecreasing" and "decreasing" instead of "nonincreasing". In the process of doing this, we will also establish the "genericity" result claimed in Remark 3. These results are stated in the following lemma, the proof of which completes the proof of the proposition. ∎

Lemma 3 *Let* $a(x, P, \bar{u}, \eta)$ *and* $g(x, P, \bar{u}, \eta)$ *denote the smallest [or greatest] solution to the maximization problem (7). Then* $a(x', P', \bar{u}', \eta') > a(x, P, \bar{u}, \eta)$ *and* $g(x', P', \bar{u}', \eta') > g(x, P, \bar{u}, \eta)$ *for any* $(x', P', \bar{u}', \eta')$ *where* $x' \geq x$, $P' \geq P$, $\bar{u}' \leq \bar{u}$, *and* $\eta' \leq \eta$, *with at least one strict inequality. Moreover, let* $u^l(x, P, \bar{u}, \eta)$ *be a solution at the parameter vector* $(x, P, \bar{u}, \eta)$. *If* $u^l(x, P, \bar{u}, \eta) \geq 0$, *then* $u^l(x', P', \bar{u}', \eta') < 0$.

Proof. For brevity, we will prove this lemma for a change in x, holding P, $\bar{u}$, and η constant. The argument for the other cases is analogous.

Consider a change from x to $x' > x$. The first part of Proposition 2, which has already been established, implies that $a(x', P', \bar{u}', \eta') \geq a(x, P, \bar{u}, \eta)$ and $g(x', P', \bar{u}', \eta') \geq g(x, P, \bar{u}, \eta)$, which we shorten to $a(x') \geq a(x)$ and $g(x') \geq g(x)$. First suppose that $u^l(x) < 0$. Then (7) can be written as

$$(a^*, g^*) \in \arg \max_{(a,g) \in [0,1] \times \mathbb{R}_+} Pxa - a\left[(1-a)c'(a) + c(a) + \bar{u} - g\right]_+ - \eta \chi(g). \tag{A-5}$$

Suppose to obtain a contradiction that $u^h = (1-a)c'(a) + c(a) + \bar{u} - g \leq 0$. Then (A-5) implies $a = 1$. Since $\lim_{a \to 1} c(a) = \infty$, $u^h \leq 0$ and $a = 1$ violate (IR$_1$). Therefore, we have $(1-a)c'(a) + c(a) + \bar{u} - g > 0$, and (A-5) is strictly supermodular in (a, g, x) (or more generally in $(a, g, x, P, \bar{u}, \eta)$) in the neighborhood of x and is also differentiable in (a, g). Then since $a > 0$ and $a < 1$ (again since $\lim_{a \to 1} c(a) = \infty$), a satisfies the first-order necessary condition

$$Px - ((1-a)c'(a) + c(a) + \bar{u} - g) - a(1-a)c''(a) = 0 \tag{A-6}$$

Moreover, since (A-5) is differentiable, g satisfies (10) and thus $g > 0$. Next, note that an increase in x strictly raises the left-hand side of (A-6) and thus a and g cannot both remain constant following an increase in x. Since they cannot decline, it must be the case that an increase in x strictly increases the smallest and the greatest values of both a and g.

We have thus established that any change from x to $x' > x$ will give weak comparative static results only if $u^l(\tilde{x}) = 0$ for all $\tilde{x} \in [x, x']$ (from Lemma 2, $u^l(\tilde{x}) > 0$ is not possible). Suppose, to obtain a contradiction, that $u^l(\tilde{x}) = 0$ for all $\tilde{x} \in [x, x']$. Then (IC$_1$) and (IR$_2$) imply that

$$u^l = -a(\tilde{x})c'(a(\tilde{x})) + c(a(\tilde{x})) + \bar{u} - g(\tilde{x}) = 0. \tag{A-7}$$

Consider variations in a and g along (A-7) (i.e., holding $u^l = 0$). Since $a > 0$ and c is differentiable, this implies $dg/da = -ac''(a) < 0$. Now $a(x') \geq a(x)$, $g(x') \geq g(x)$, and (A-7) holds for $\tilde{x} \in \{x, x'\}$,

so $dg/da < 0$ implies that $a(x) = a(x')$ and $g(x) = g(x')$. Next, since $a(x)$ and $g(x)$ are optimal, any such variation along (A-7) should not increase the value of (7). At $\tilde{x} = x$, this is only possible if

$$Px - a(x)(1 - a(x)) c''(a(x)) - (1 - a(x)) c'(a(x)) - c(a(x)) - \bar{u} + g(x) + \eta \chi'(g(x)) a(x) c''(a(x)) = 0,$$
(A-8)

where we have used the fact that $u^h > 0$ wherever $u^l = 0$, by (IC$_1$) and the fact that $a > 0$. Repeating this argument for $\tilde{x} = x'$ yields

$$Px' - a(x)(1 - a(x)) c''(a(x)) - (1 - a(x)) c'(a(x)) - c(a(x)) - \bar{u} + g(x) + \eta \chi'(g(x)) a(x) c''(a(x)) = 0.$$
(A-9)

However, in view of the fact that $x' > x$, (A-9) cannot be true at the same time as (A-8), yielding a contradiction. We conclude that (i) all comparative statics are strict (even if $u^l(x) = 0$, we must have $u^l(x') < 0$, and thus a change from x to x' will strictly increase a and g), and (ii) for any pair of $x' > x$, if $u^l(x) = 0$, we must have $u^l(x') < 0$, and thus $u^l = 0$ is only possible at one level of x, that is, at the lowest level $x = \underline{x}$. ∎

It can also be shown that modified versions of Corollaries 1-6 for smallest and greatest equilibrium contracts also hold in this general case. We omit the details to save space.

Existence of General Equilibrium

To establish existence of an equilibrium, we now allow all producers to use "mixed-strategies" (so that we are looking for a "mixed strategy equilibrium" or "Cournot-Nash equilibrium distribution" in the language of Mas-Collel). We will then use Theorem 1 from Mas-Collel (1984). To do this in the simplest possible way, we rewrite producer profits slightly: Since $\lim_{g \to \infty} \chi(g) = \infty$ and $a \leq 1$, there exists a positive number $\bar{g}$ such that setting $g > \bar{g}$ is dominated for any producer. We now specify that a producer with productivity x chooses $(q, g) \in [0, \bar{x}] \times [0, \bar{g}]$ to maximize

$$\pi(x) \equiv qP - \frac{q}{x}\left[\left(1 - \frac{q}{x}\right)c'\left(\frac{q}{x}\right) + c\left(\frac{q}{x}\right) + \bar{u} - g\right]_+ - \left(1 - \frac{q}{x}\right)\left[-\frac{q}{x}c'\left(\frac{q}{x}\right) + c\left(\frac{q}{x}\right) + \bar{u} - g\right]_+ - \eta\chi(g),$$
(A-10)

where we have rewritten a as q/x. By our assumption that $\lim_{a \to 1} c(a) = \infty$, any solution to the problem of a producer with productivity x satisfies $q < x$. It is also clear that (A-10) is continuous in (p, q). This reformulation gives each producer the same action set, and our assumption that $\lim_{a \to 1} c(a) = \infty$ ensures that every equilibrium contract in the reformulated game is feasible in the original game. Formally, a (mixed) strategy profile is a measure τ over $[\underline{x}, \bar{x}] \times ([0, \bar{x}] \times [0, \bar{g}])$ such that the marginals τ_x and $\tau_{(q,g)}$ of τ on $[\underline{x}, \bar{x}]$ and $[0, \bar{x}] \times [0, \bar{g}]$ satisfy: (1) $\tau_x = F$; (2) $\tau(\{x, q, g\} : q \leq x) = 1$.

Given a strategy profile τ, the corresponding aggregates are

$$Q(\tau) \equiv \int_0^{\bar{x}} q\tau_q(q)\,dq, \text{ and}$$
(A-11)

$$G(\tau) \equiv \int_0^{\bar{g}} g\tau_g(g)\,dg,$$
(A-12)

where τ_q and τ_g are the marginal densities over $[0, \bar{x}]$ and $[0, \bar{g}]$, respectively. P and $\bar{u}$ are defined as functions of $Q(\tau)$ and $G(\tau)$ by equations (3) and (6), as in the text. Let $\pi(x, q, g, Q, G)$ denote the payoff to a producer with productivity x who chooses (q, g) when facing aggregates Q and G.

Definition 2 *An* equilibrium *is a (mixed) strategy profile τ that satisfies*

$$\tau(\{\pi(x), q, g\} : \pi(x, q, g, Q, G) \geq \pi(x, \tilde{q}, \tilde{g}, Q, G) \ \text{for all} \ (\tilde{q}, \tilde{g}) \in [0, \bar{x}] \times [0, \bar{g}]) = 1,$$

where Q and G are given by (A-11) and (A-12) evaluated at τ.

Proposition 12 *An equilibrium exists.*

Proof. The proof follows closely the proof of Mas-Collel's Theorem 1, with the addition that each distribution consistent with producers' playing best responses satisfies point 2 above (since $q < x$ in any equilibrium contract of a producer with productivity x, by our assumption that $\lim_{a \to 1} c(a) = \infty$), so any fixed point must satisfy point 2 as well. ∎

While Proposition 12 establishes the existence of an equilibrium in mixed strategies, our analysis in the next subsection will also show that in some special cases (for example, when $P(\cdot) = P_0$ or when $\gamma = 0$), we can establish that the equilibrium is in pure strategies (as was the case in the text).

General Equilibrium Comparative Statics

As discussed in Section 4, the set of equilibria may not form a lattice, which makes it impossible to consider comparative statics on extremal equilibria in (Q, G). Instead, we show that the comparative statics of Proposition 3 apply separately to the extremal equilibrium values of Q and G.

Our approach is based on the analysis of the following function:

$$\hat{\phi}(Q, \gamma, L, \eta) = \left\{ \begin{array}{c} Q' : \text{there exists a density } \tau \text{ on } \pi(x) \times ([0, \bar{x}] \times [0, \bar{g}]) \text{ such that} \\ \tau(\{\pi(x), q, g\} : \pi(x, q, g, Q, G(\tau)) \geq \pi(x, \tilde{q}, \tilde{g}, Q, G(\tau))) = 1 \\ \text{and } Q' = Q(\tau) \end{array} \right\}.$$

$\hat{\phi}$ maps Q and parameter values to those Q' that are mixed-strategy equilibrium values of output in the modified game where price is fixed at $P(QL)$. It is clear that the set of fixed points of $\hat{\phi}$ equals the set of mixed-strategy equilibrium values of Q. We first establish comparative statics on the extremal elements of $\hat{\phi}(Q, \gamma, L, \eta)$ with respect to the parameters and Q, and then use these results to establish comparative statics on the extremal fixed points of $\hat{\phi}(Q, \gamma, L, \eta)$ in a standard way.

Towards establishing comparative statics on the extremal elements of $\hat{\phi}(Q, \gamma, L, \eta)$, consider the modified game where price is fixed at $P(QL) = P_0$ for all QL. Let $(q^-, g^-)(P, \bar{u}, x, \eta)$ and $(q^+, g^+)(P, \bar{u}, x, \eta)$ denote the smallest and greatest equilibrium contract levels of (q, g) given price P, outside option $\bar{u}$, productivity x, and cost of coercion η. Let

$$\tilde{\phi}(G, P_0, x, \gamma, L, \eta) = \left[\int_{\underline{x}}^{\bar{x}} g^- \left(P_0, \tilde{u}(L) - \frac{\gamma}{1 - \gamma} G, x, \eta \right) dF(x), \int_{\underline{x}}^{\bar{x}} g^+ \left(P_0, \tilde{u}(L) - \frac{\gamma}{1 - \gamma} G, x, \eta \right) dF(x) \right].$$

$$\text{(A-13)}$$

We write $\tilde{\phi}(G)$ for $\tilde{\phi}(G, P_0, \tilde{u}, \gamma, \eta)$ and $g^-(x, G)$ $(g^+(x, G))$ for $g^-\left(P_0, \tilde{u}(L) - \frac{\gamma}{1-\gamma}G, x, \eta\right)$ $\left(g^+\left(P_0, \tilde{u}(L) - \frac{\gamma}{1-\gamma}G, x, \eta\right)\right)$ when the parameters are understood. It is clear that if G is an equilibrium aggregate level of coercion (in the modified game), then G is a fixed point of $\tilde{\phi}(G)$. The converse is also true, because if $G \in \tilde{\phi}(G)$ then, by the intermediate value theorem, there exists $x^* \in [\underline{x}, \bar{x}]$ such that

$$G = \int_{\underline{x}}^{x^*} g^-\left(P_0, \tilde{u}(L) - \frac{\gamma}{1-\gamma}G, x, \eta\right) dF(x) + \int_{x^*}^{\bar{x}} g^+\left(P_0, \tilde{u}(L) - \frac{\gamma}{1-\gamma}G, x, \eta\right) dF(x);$$

and the strategy profile in which producers of type $x \leq x^*$ choose $(q^-, g^-)\left(P_0, \tilde{u}(L) - \frac{\gamma}{1-\gamma}G, x, \eta\right)$ and producers of type $x > x^*$ choose $(q^+, g^+)\left(P_0, \tilde{u}(L) - \frac{\gamma}{1-\gamma}G, x, \eta\right)$ is an equilibrium. Thus, the fixed points of $\tilde{\phi}(G)$ are exactly the equilibrium values of G.

The following lemma shows that if G_0 is the smallest (greatest) fixed point of $\tilde{\phi}(G)$, then G_0 is the smallest (greatest) element of the set $\tilde{\phi}(G_0)$.

Lemma 4 *If G^- is the smallest fixed point of $\tilde{\phi}(G)$, then $G^- = \int_{\underline{x}}^{\bar{x}} g^-(x, G^-) dF(x)$. If G^+ is the greatest fixed point of $\tilde{\phi}(G)$, then $G^+ = \int_{\underline{x}}^{\bar{x}} g^+(x, G^+) dF(x)$.*

Proof. Suppose G^- is the smallest fixed point of $\tilde{\phi}(G)$. Then $G^- \geq \int_{\underline{x}}^{\bar{x}} g^-(x, G^-) dF(x)$, since $g^-(x, G)$ is increasing in G and any other solution $g(x, G)$ to (7) satisfies $g(x, G) \geq g^-(x, G)$. Thus to obtain a contradiction, suppose that $G^- > \int_{\underline{x}}^{\bar{x}} g^-(x, G^-) dF(x)$. Since $\int_{\underline{x}}^{\bar{x}} g^-(x, G) dF(x)$ is increasing in G and $\int_{\underline{x}}^{\bar{x}} g^-(x, 0) dF(x) \geq 0$, by Tarski's fixed point theorem (e.g., Theorem 2.5.1 in Topkis, 1998), there exists $G' \in [0, G^-)$ such that $G' = \int_{\underline{x}}^{\bar{x}} g^-(x, G') dF(x)$, yielding a contradiction.

Next, suppose G^+ is the greatest fixed point of $\tilde{\phi}(G)$. Similarly, $G^+ \leq \int_{\underline{x}}^{\bar{x}} g^+(x, G^+) dF(x)$, so to obtain a contradiction suppose that $G^+ < \int_{\underline{x}}^{\bar{x}} g^+(x, G^+) dF(x)$. Since $\lim_{g \to \infty} \chi'(g) = \infty$ and $a \leq 1$, there exists $\bar{G}$ such that $\bar{G} > \int_{\underline{x}}^{\bar{x}} g^+(x, \bar{G}) dF(x)$. So, since $\int_{\underline{x}}^{\bar{x}} g^+(x, G) dF(x)$ is increasing in G, again by Tarski's Fixed Point Theorem, there exists $G' \in (G^+, \bar{G})$ such that $G' = \int_{\underline{x}}^{\bar{x}} g^+(x, G') dF(x)$, yielding another contradiction and completing the proof of the lemma. ∎

We can now derive comparative statics on the extremal elements of $\hat{\phi}(Q, \gamma, L, \eta)$. We say that "$z$ is increasing in $F(\cdot)$," if a first-order stochastic dominance increase in $F(\cdot)$ leads to an increase in z.

Lemma 5 *The smallest and greatest elements of $\hat{\phi}(Q, \gamma, L, \eta)$ exist and are increasing in γ and $F(\cdot)$, and decreasing in Q and η. If $P(QL) = P_0$ for all QL, then the smallest and greatest elements of $\hat{\phi}(Q, \gamma, L, \eta)$ are increasing in L. If $\tilde{u}(L) = \tilde{u}_0$ for all L, then the smallest and greatest elements of $\hat{\phi}(Q, \gamma, L, \eta)$ are decreasing in L.*

Proof. Recall that the smallest and greatest elements of $\hat{\phi}(Q, \gamma, L, \eta)$ are the smallest and greatest equilibrium values of Q' when price is fixed at $P(QL)$. We claim that, in the modified game where price is fixed at an arbitrary P_0, the smallest and greatest equilibrium values of Q exist and are increasing in P_0, γ, L, and $F(\cdot)$, and decreasing in η. Given this claim, the results for γ, $F(\cdot)$, and η

follow immediately, and the result for Q follows from the claim combined with the fact that $P(QL)$ is decreasing. If $P(QL) = P_0$ for all QL, then the result for L also follows immediately from the claim. If $\tilde{u}(L) = \tilde{u}_0$ for all L, then the smallest and greatest equilibrium aggregates (Q', G') are constant in L when price is fixed at P_0, so the result for L follows from the fact that $P(QL)$ is decreasing. It therefore remains only to prove the claim.

Thus, consider the modified game where $P(\cdot) \equiv P_0$. From (7), for all G, $\tilde{\phi}(G)$ is increasing in P_0, x, γ, L, and $F(\cdot)$, and decreasing in η (in the strong set order). Since $\tilde{\phi}(G)$ is increasing, Theorem 2.5.2 in Topkis (1998) implies that the smallest and greatest fixed points of $\tilde{\phi}(G)$, and thus the smallest and greatest equilibrium values of G (in the modified game), exist and are increasing in P_0, γ, L, and $F(\cdot)$, and decreasing in η. Now Lemma 4 and the supermodularity of (7) in (q,g) imply that the smallest (greatest) equilibrium value of output, Q^- (Q^+), corresponds to the smallest (greatest) fixed point of $\tilde{\phi}(G)$, G^- (G^+); that is, the smallest and greatest equilibrium values of Q are given by $Q^- = \int_{\underline{x}}^{\bar{x}} \eta \chi'(g^-(x,G^-)) x\, dF(x)$ and $Q^+ = \int_{\underline{x}}^{\bar{x}} \eta \chi'(g^+(x,G^+)) x\, dF(x)$. Therefore, the comparative statics on G^- and G^+ together with the comparative statics on g^- and g^+ described in Proposition 11 imply that the smallest and greatest equilibrium values of Q (in the modified game) exist and are increasing in P_0, γ, L, and $F(\cdot)$, and decreasing in η, which proves the claim. $\blacksquare$

It is now straightforward to derive comparative statics on the extremal fixed points of $\hat{\phi}(Q,\gamma,L,\eta)$, which equal the extremal equilibrium values of Q:

Proposition 13 *The smallest and greatest mixed-strategy equilibrium values of Q are increasing in γ and $F(\cdot)$ and decreasing in η. If $P(QL) = P_0$ for all QL, then the smallest and greatest mixed strategy equilibrium values of Q are increasing in L. If $\tilde{u}(L) = \tilde{u}_0$ for all L, then the smallest and greatest mixed strategy equilibrium values of Q are decreasing in L.*

Proof. From Lemma 5, greater γ or $F(\cdot)$, or lower η, shifts $\hat{\phi}(Q,\gamma,L,\eta)$ up for all Q; while greater L shifts $\hat{\phi}(Q,\gamma,L,\eta)$ up for all Q if $P(QL) = P_0$ and shifts $\hat{\phi}(Q,\gamma,L,\eta)$ down for all Q if $\tilde{u}(L) = \tilde{u}_0$. Since $\hat{\phi}(Q,\gamma,L,\eta)$ is monotone in Q (by Lemma 5), Theorem 2.5.2 in Topkis (1998) implies that the smallest and greatest fixed points of $\hat{\phi}(Q,\gamma,L,\eta)$ are increasing in γ and $F(\cdot)$, and decreasing in η; and are increasing in L if $P(QL) = P_0$ and decreasing in L if $\tilde{u}(L) = \tilde{u}_0$. $\blacksquare$

Repeating the above argument interchanging the roles of Q and G gives the following:

Proposition 14 *The smallest and greatest mixed-strategy equilibrium values of G are increasing in γ and decreasing in η. If $P(QL) = P_0$ for all QL, then the smallest and greatest mixed-strategy equilibrium values of G are increasing in L and $F(\cdot)$. If $\tilde{u}(L) = \tilde{u}_0$ for all L, then the smallest and greatest mixed strategy equilibrium values of G are decreasing in L.*

Proof. The only difference between this result and Proposition 13 is that the comparative static with respect to $F(\cdot)$ now only applies when $P(QL) = P_0$ for all QL. To see why, observe that instead of assuming that $P(\cdot)$ is fixed and studying the map $\tilde{\phi}(G, P_0, x, \gamma, L, \eta)$ in (A-13), we now assume that

$\gamma = 0$ and study the map

$$\bar{\phi}\left(Q, x, L, \eta\right) = \left[\int_{\underline{x}}^{\bar{x}} q^{-}\left(P\left(QL\right), \tilde{u}\left(L\right), x, \eta\right) dF\left(x\right), \int_{\underline{x}}^{\bar{x}} q^{+}\left(P\left(QL\right), \tilde{u}\left(L\right), x, \eta\right) dF\left(x\right)\right].$$

An argument identical to the proof of Lemma 4 shows that the smallest and greatest fixed points of $\bar{\phi}$ correspond to the smallest and greatest equilibrium levels of Q when $\gamma = 0$, and that the corresponding smallest and greatest equilibrium levels of G are $G^{-} = \int_{\underline{x}}^{\bar{x}}\left(\chi'\right)^{-1}\left(\frac{q^{-}\left(P\left(Q^{-}L\right), \tilde{u}(L), x, \eta\right)}{x\eta}\right) dF\left(x\right)$ and $G^{+} = \int_{\underline{x}}^{\bar{x}}\left(\chi'\right)^{-1}\left(\frac{q^{+}\left(P\left(Q^{+}L\right), \tilde{u}(L), x, \eta\right)}{x\eta}\right) dF\left(x\right)$, respectively. $\bar{\phi}$ is decreasing in Q and η, and increasing in $F\left(\cdot\right)$; is increasing in L if $P\left(QL\right) = P_0$ for all QL; and is decreasing in L if $\tilde{u}\left(L\right) = \tilde{u}_0$ for all L. By the same argument as in the proof of the claim in Lemma 5, the smallest and greatest equilibrium levels of Q are decreasing in η and increasing in $F\left(\cdot\right)$, and are increasing in L if $P\left(QL\right) = P_0$ and decreasing in L if $\tilde{u}\left(L\right) = \tilde{u}_0$. It then follows from our characterization of G^{-} and G^{+} that they are decreasing in η; and are increasing in L and $F\left(\cdot\right)$ if $P\left(QL\right) = P_0$ and decreasing in L if $\tilde{u}\left(L\right) = \tilde{u}_0$. However, G^{-} and G^{+} need not be increasing in $F\left(\cdot\right)$ if $P\left(QL\right)$ is decreasing, because the direct (positive) effect of increasing $F\left(\cdot\right)$ may be offset by the indirect (negative) effect that Q^{-} and Q^{+} are increasing in $F\left(\cdot\right)$; this accounts for the difference relative to Proposition 13. The remainder of the proof is analogous to the proof of Lemma 5 and Proposition 13. ∎

References

Acemoglu, D., Johnson, S., and Robinson, J.A. (2002), "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 117, 1231-1294.

Acemoglu, D. and Pischke, J.S. (1999), "The Structure of Wages and Investment in General Training," *Journal of Political Economy*, 107, 539-572.

Andrees, B. and Belser, P. (eds.) (2009), *Forced Labor: Coercion and Exploitation in the Private Economy*, Boulder: Lynne Rienner Publishers.

Aston, T.H. and Philpin, C.H.E. (1987), *The Brenner Debate: Agrarian Class Structure and Economic Development in Pre-Industrial Europe*, Cambridge: Cambridge University Press.

Basu, K. (1986), "One Kind of Power," *Oxford Economic Papers*, 38, 259-282.

Barzel, Y. (1977), "An Economic Analysis of Slavery," *Journal of Law and Economics*, 20, 87-110.

Bergstrom, T. (1971), "On the Existence and Optimality of Competitive Equilibrium for a Slave Economy," *Review of Economic Studies*, 38, 23-36.

Bloom, J. (1998), *The End of the Old Order in Rural Europe*, Princeton: Princeton University Press.

Brenner, R. (1976), "Agrarian Class-Structure and Economic-Development in Pre-Industrial Europe," *Past and Present*, 70, 30-75.

Canarella, G. and Tomaske, J. (1975), "The Optimal Utilization of Slaves," *Journal of Economic History*, 36, 621-629.

Chwe, M. (1990), "Why Were Workers Whipped? Pain in a Principal-Agent Model," *Economic Journal*, 100, 1109-1121.

Conning, J. (2004), "On the Causes of Slavery or Serfdom and the Roads to Agrarian Capitalism: Domar's Hypothesis Revisited," manuscript, Hunter College.

Conrad, A.H. and Meyer, J.R. (1958), "The Economics of Slavery in the Antebellum South," *Journal of Political Economy*, 66, 95-122.

Curtin, P.D. (1990), *The Rise and Fall of the Plantation Complex*, Cambridge: Cambridge University Press.

Davis, D.B. (2006), *Inhuman Bondage: The Rise and Fall of Slavery in the New World*, Oxford: Oxford University Press

Domar, E.D. (1970), "The Causes of Slavery or Serfdom: A Hypothesis," *Journal of Economic History*, 30, 18-32.

Dow, G.K. (1993), "Why Capital Hires Labor: A Bargaining Perspective," *American Economic Review*, 83, 118-134.

Fenoaltea, S. (1984), "Slavery in Comparative Perspective: A Model," *Journal of Economic History*, 44, 635-638.

Findlay, R. (1975), "Slavery, Incentives, and Manumission: A Theoretical Model," *Journal of Economic History*, 44, 923-933.

Finley, M.I. (1976), "A Peculiar Institution?" *Times Literary Supplement*, 3877, 819-821.

Fogel, R.W. and Engerman, S.L. (1974), *Time on the Cross: The Economics of American Negro Slavery*, 2 volumes, Boston: Little Brown.

Genicot, G. (2002), "Bonded Labor and Serfdom: A Paradox of Voluntary Choice," *Journal of Development Economics*, 67, 101-127.

Grossman, S.J. and Hart, O.D. (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691-719.

Habakkuk, H.J. (1958), "The Economic History of Modern Britain," *Journal of Economic History*, 18, 486-501.

International Labor Organization (2009), "Report of the Director-General: The Cost of Coercion,".

Key, V.O. Jr. (1949) *Southern Politics: In State and Nation*, New York; Vintage Books.

Klein, H.S., and Vinson, B. (2007), *African Slavery in Latin America and the Caribbean*, New York: Oxford University Press.

Lagerlöf, N. (2009), "Slavery and Other Property Rights," *Review of Economic Studies*, 76, 319-342.

Leroy Ladurie, E. (1977), *The Peasants of Languedoc,* translated by John Day, Illinois: University of Illinois Press.

Lockhart J. and Schwartz, S.B. (1983), *Early Latin America: A History of Colonial Spanish Latin America and Brazil*, Cambridge University Press, Cambridge.

Lockhart J. (2000), *Of Things of the Indies: Essays Old and New in Early in American History*, Stanford: Stanford University Press.

Lovejoy, P.E. (2000), *Transformations of Slavery: A History of Slavery in Africa*, Second Edition, Cambridge: University of Cambridge Press.

Mas-Collel, A. (1984), "On a Theorem of Schmeidler," *Journal of Mathematical Economics*, 13, 201-206.

Meltzer, M. (1993), *Slavery: A World History*, New York: DeCapo Press Edition.

Milgrom, P. and Roberts, J. (1994), "Comparing Equilibria," *American Economic Review*, 84, 441-459.

Naidu, S. and Yuchtman, N. (2009) "How Green Was My Valley? Coercive Contract Enforcement in 19th Century Industrial Britain," Harvard University mimeo.

Naqvi, N. and Wemhöner, F. (1995), "Power, Coercion, and the Games Landlords Play," *Journal of Development Economics*, 47, 191-205.

North, D.C. and Thomas, R.T. (1971), "The Rise and Fall of the Manorial System: A Theoretical Model," *Journal of Economic History*, 31, 777-803.

Patterson, O. (1982), *Slavery and Social Death: A Comparative Study*, Cambridge: Harvard University Press.

Postan, M.M. (1973), *Cambridge Economic History of Europe: Expanding Europe in the 16th and 17th Centuries*, Cambridge University Press, Cambridge.

Ransom, R.L. and Sutch, R. (1975), "The Impact of the Civil War and of Emancipation on Southern Agriculture," *Explorations in Economic History*, 12, 1-28.

Ransom, R.L. and Sutch, R. (1977), *One Kind of Freedom: The Economic Consequences of Emancipation*, Cambridge: Cambridge University Press.

Sherstyuk, K. (2000), "Performance Standards and Incentive Pay in Agency Contracts," *Scandanavian Journal of Economics*, 102, 725-736.

Topkis, D.M. (1998), *Supermodularity and Complementarity*, Princeton University Press.

Wright, G. (1978), *The Political Economy of the Cotton South: Households, Markets, and Wealth in the Nineteenth Century*, New York: Norton.

Appendix B: Multiple Output Levels (Not for Publication)

In this Appendix, we allow for an arbitrary finite number of outputs. Let $f(y|a)$ be the probability that output equals y given effort a, and assume that $f(y|a)$ is twice-differentiable with respect to a. Let $\bar{y}$ be the highest possible output and let $\underline{y}$ be the lowest possible output. We normalize the price of output to 1 and do not consider producer heterogeneity.[29] Equilibrium contracts are given by the solution to the following problem:

$$\max_{(w(\cdot),p(\cdot),a,g)} \sum_y (y - w(y)) f(y|a) - \eta\chi(g),$$

subject to

$$\sum_y (w(y) - p(y)) f(y|a) - c(a) \geq \bar{u} - g, \tag{IR$_{B0}$}$$

and

$$a \in \arg\max_{a\in[0,1]} \sum_y (w(y) - p(y)) f(y|a) - c(a). \tag{IC$_{B0}$}$$

We continue to focus on the case where equilibrium contracts involve $a > 0$.[30] Then, from Theorem 1 in Jewitt (1988), the first-order approach to this problem is valid provided that:

1. $\sum_{z\leq\bar{z}}\sum_{y\leq z} f(y|a)$ is nonincreasing and convex in a for each $\bar{z}$;

2. $\sum_{y\leq\bar{z}} yf(y|a)$ is nondecreasing and concave in a for each $\bar{z}$;
3. $\frac{f_a(y|a)}{f(y|a)}$ is nondecreasing and concave in y for every a.

Jewitt provides an interpretation of these conditions. Note Condition 3 implies the usual (and relatively weak) monotone likelihood ratio property (MLRP). Therefore, MLRP holds throughout this appendix. Jewitt argues that the remainder of the third condition is the most restrictive; in addition to MLRP, "[it requires that] variations in output at higher levels are relatively less useful in providing 'information' on the agents effort than they are at lower levels of output," (Jewitt, 1988, p. 1181). Note that Jewitt's condition on utility functions is not needed when the agent is risk-neutral.

Given these three conditions, and simplifying the exposition by imposing $u^l < 0$,[31] we can apply the first-order approach and (writing $u(y)$ for $w(y) - p(y)$) rewrite the problem as

$$\max_{(w(\cdot),p(\cdot),a,g)} \sum_y \left(y - [u(y)]_+\right) f(y|a) - \eta\chi(g)$$

subject to

$$\sum_y u(y) f(y|a) - c(a) \geq \bar{u} - g \tag{IR$_{B1}$}$$

---

[29] If producers differ according to productivity parameter x and a producer with productivity x produces output xy with probability $f(y|a)$, all supermodularity results from Section 3 will continue to hold.

[30] We do not spell out the assumptions on primitives under which equilibrium contracts involve $a > 0$. Assumption 2 suffices for this in the two-outcome case, and similar sufficient conditions can be developed for the case with multiple output levels, but this is orthogonal to our focus here.

[31] Under suitable restrictions on $f(y|a)$, which we neglect here for ease of exposition, $u_l < 0$ in any equilibrium contract for "generic" parameter values as in the two-outcome case considered in the text.

and

$$\sum_y u\left(y\right) f_a(y|a) = c'(a). \tag{IC$_{B1}$}$$

The associated Lagrangian is

$$\sum_y \left[\left(y - [u\left(y\right)]_+\right) f\left(y|a\right) - \eta\chi(g) + \lambda\left(u\left(y\right) f(y|a) - c(a) - (\bar{u} - g)\right) + \mu\left(u\left(y\right) f_a(y|a) - c'(a)\right)\right].$$

Differentiating under the integral with respect to $u(y)$ and rearranging implies that if $u\left(y\right) > 0$, then

$$1 = \lambda + \mu\frac{f_a\left(y|a\right)}{f\left(y|a\right)}; \tag{B1}$$

if $u\left(y\right) < 0$, then

$$0 = \lambda + \mu\frac{f_a\left(y|a\right)}{f\left(y|a\right)}; \tag{B2}$$

and if $u\left(y\right) = 0$, then

$$0 \leq \lambda + \mu\frac{f_a\left(y|a\right)}{f\left(y|a\right)} \leq 1. \tag{B3}$$

By MLRP, (B1), (B2), and (B3) imply that $u\left(y\right) = 0$ for all $y \notin \{\underline{y}, \bar{y}\}$, $u\left(\underline{y}\right) \leq 0$, and $u\left(\bar{y}\right) \geq 0$. This is a standard "bang-bang" result given MLRP and risk-neutrality. To simplify notation, let $u^h \equiv u\left(\bar{y}\right)$ and let $u^l \equiv u\left(\underline{y}\right)$, paralleling the notation in subsection 3.1. The producer's maximization problem can then be written as

$$\max_{\left(a,g,u^h,u^l\right)\in[0,1]\times\mathbb{R}_+\times\mathbb{R}^2} \sum_y yf\left(y|a\right) - f\left(\bar{y}|a\right)[u^h]_+ - \eta\chi(g)$$

subject to

$$f\left(\bar{y}|a\right)u^h + f\left(\underline{y}|a\right)u^l - c(a) \geq \bar{u} - g \tag{IR$_{B2}$}$$

and

$$f_a\left(\bar{y}|a\right)u^h + f_a\left(\underline{y}|a\right)u^l = c'\left(a\right). \tag{IC$_{B2}$}$$

(IC$_{B2}$) can be rewritten as

$$u^l = \frac{c'\left(a\right) - f_a\left(\bar{y}|a\right)u^h}{f_a\left(\underline{y}|a\right)}. \tag{B4}$$

Substituting (B4) into (IR$_{B2}$) and using the fact that (IR$_{B2}$) binds at the solution gives

$$\left(f\left(\bar{y}|a\right) - \frac{f_a\left(\bar{y}|a\right)}{f_a\left(\underline{y}|a\right)}f\left(\underline{y}|a\right)\right)u^h + \frac{f\left(\underline{y}|a\right)}{f_a\left(\underline{y}|a\right)}c'\left(a\right) - c(a) = \bar{u} - g,$$

which may be rewritten as

$$u^h = \frac{c\left(a\right) - \frac{f\left(\underline{y}|a\right)}{f_a\left(\underline{y}|a\right)}c'\left(a\right) + \bar{u} - g}{f\left(\bar{y}|a\right) - \frac{f_a\left(\bar{y}|a\right)}{f_a\left(\underline{y}|a\right)}f\left(\underline{y}|a\right)}. \tag{B5}$$

Finally, substituting (B5) into the principal's objective gives

$$\max_{(a,g)\in[0,1]\times\mathbb{R}_+} \sum_y yf(y|a) - f(\bar{y}|a)\left[\frac{c(a) - \frac{f(y|a)}{f_a(y|a)}c'(a) + \bar{u} - g}{f(\bar{y}|a) - \frac{f_a(\bar{y}|a)}{f_a(y|a)}f(y|a)}\right]_+ - \eta\chi(g), \qquad (B6)$$

To establish the supermodularity of (B6) in $(a, g, -\bar{u}, -\eta)$, it suffices to show that the cross-partial of the right-hand side of (B6) with respect to a and g is always nonnegative. This is immediate if $[u^h]_+ = 0$. If $[u^h]_+ > 0$, the right-hand side of (B6) may be rewritten as

$$\sum_y yf(y|a) - \left(\frac{1}{1 - \left(\frac{f_a(\bar{y}|a)}{f(\bar{y}|a)}\right) / \left(\frac{f_a(y|a)}{f(y|a)}\right)}\right)\left(c(a) - \frac{f(y|a)}{f_a(y|a)}c'(a) + \bar{u} - g\right) - \eta\chi(g), \qquad (B7)$$

where we have divided the numerator and denominator of the middle term by $f_a(\bar{y}|a)$. The cross-partial of (B7) with respect to a and g is nonnegative if and only if the derivative of $\left(\frac{f_a(\bar{y}|a)}{f(\bar{y}|a)}\right) / \left(\frac{f_a(y|a)}{f(y|a)}\right)$ with respect to a is nonnegative. We have thus established:

Proposition 15 *Suppose that $a > 0$, Conditions 1-3 hold, and $\left(\frac{f_a(\bar{y}|a)}{f(\bar{y}|a)}\right) / \left(\frac{f_a(y|a)}{f(y|a)}\right)$ is increasing in a. Then equilibrium contracts are given by (B6), and (B6) is supermodular in $(a, g, -\bar{u}, -\eta)$.*

The condition that $\left(\frac{f_a(\bar{y}|a)}{f(\bar{y}|a)}\right) / \left(\frac{f_a(y|a)}{f(y|a)}\right)$ is increasing in a is not very restrictive. To see why it is sufficient for supermodularity of equilibrium contracts, note that, by (B5) (which is determined by (IR$_{B_2}$) and (IC$_{B_2}$)), increasing g by Δ allows the principal to reduce w^h (i.e., payment after the highest output level) by $\frac{\Delta}{f(\bar{y}|a) - \frac{f_a(\bar{y}|a)}{f_a(y|a)}f(y|a)}$. Since the principal pays w^h with probability $f(\bar{y}|a)$, increasing g by Δ benefits the principal by $\frac{f(\bar{y}|a)}{f(\bar{y}|a) - \frac{f_a(\bar{y}|a)}{f_a(y|a)}f(y|a)}$, which is increasing in a if $\left(\frac{f_a(\bar{y}|a)}{f(\bar{y}|a)}\right) / \left(\frac{f_a(y|a)}{f(y|a)}\right)$ is increasing. It is instructive to compare this with the following slight generalization of the two-outcome case discussed in Section 3.1: Suppose there are only two outcomes $y < \bar{y}$, but that $f(\bar{y}|a)$ need not equal a. Then, $f(\bar{y}|a) - \frac{f_a(\bar{y}|a)}{f_a(y|a)}f(y|a) = f(\bar{y}|a) + f(y|a) = 1$ (as $f_a(\bar{y}|a) = -f_a(y|a)$), so $\frac{\Delta}{f(\bar{y}|a) - \frac{f_a(\bar{y}|a)}{f_a(y|a)}f(y|a)} = \Delta$. Therefore, increasing g by Δ benefits the principal by $f(\bar{y}|a)$, which is increasing in a under MLRP. In particular, we see that $\left(\frac{f_a(\bar{y}|a)}{f(\bar{y}|a)}\right) / \left(\frac{f_a(y|a)}{f(y|a)}\right)$ is increasing in a if $f(\bar{y}|a) + f(y|a)$ does not depend on a, regardless of the number of outcomes.

Additional References

Jewitt, I. (1988), "Justifying the First-Order Approach to Principal-Agent Problems," *Econometrica*, 56, 1177-1190.