1997

Labor – Free or Coerced? An Historical Reassessment of Differences and Similarities

Robert J. Steinfeld
University at Buffalo School of Law, steinfel@buffalo.edu

Stanley L. Engerman
University of Rochester

Labor – Free or Coerced? A Historical Reassessment of Differences and Similarities

Robert J. Steinfeld and Stanley L. Engerman*

I

Over the last twenty years, historians have developed a more complicated picture of the forms of labor in use in the eighteenth and nineteenth centuries. Where previously they tended to divide labor simply into free wage labor on one side, and slave or serf labor on the other, this new view has added contract labor, indentured servitude, apprenticeship, and peonage to the map of forms of labor in use in earlier centuries.[1] But the new map continues to classify these different forms of labor as coming under the basic categories free or unfree, and continues to assume that each of these forms of labor is a distinctive "type" which is supposed to have a natural core set of characteristics that sets it off from the others, and marks it definitively as either "free" or "unfree". The modern conceptual map of labor looks something like the following:

Free labor:	Wage labor
Unfree labor:	Apprenticeship, indentured servitude, contract labor, peonage, serfdom, slavery

This paper will argue that the above picture and the assumptions upon which it rests still need to be radically revised to capture the reality of labor relations in the past and perhaps also the present. First, it will argue that "types" of labor, like "wage labor" and "contract labor", never did possess a set of fixed, natural characteristics, but were defined by a range of characteristics and that depending upon the precise characteristics they possessed in any particular place, such labor might be considered either "free" or "unfree". By the standards of twentieth-century American constitutional law,[2] for example, nineteenth-century English wage labor" was "unfree" labor, while nineteenth-century American "contract

* For comments in the preparation of this paper we wish to thank Seymour Drescher, David Eltis, Lou Ferleger, David Galenson, Sherwin Rosen, and the participants at the conference, particularly Pieter Emmer and Ralph Shlomowitz.

1. A number of the ideas presented in this paper have been developed more fully elsewhere. See Steinfeld, "Myth of the Rise of Free Labor", *Invention of Free Labor*; see Engerman, "Coerced and Free Labor", "Economics of Forced Labor", and "Land and Labour Problem".
2. See Clyatt v. United States, 197 U.S. 207 (1905), pp. 215-216; Bailey v. Alabama, 219 U.S. 219 (1911), p. 243; Pollock v. Williams, 322 U.S. 4 (1944), p. 18.

labor" was "free" labor. As this example makes clear, the different "types" of labor were, in certain respects, not nearly as discrete and discontinuous as the standard picture implies. At the boundaries the "types" of labor frequently blur and merge. Moreover, since labor force characteristics include a number of different dimensions (entry, performance, and exit among them), at times not all these aspects can be fit into the same side of the free-unfree dichotomy.

We shall argue that the classification of labor into "free" or "unfree" is an arbitrary not a natural classification. All "dependent" labor is elicited as the result of pressures brought to bear on it, pressures that are socially and legally constructed. Since variations exist in the economic, legal, political, and social aspects of these constructions, the spectrum of labor systems does not imply any equivalence in effects upon workers or hirers, and such differences remain a source of debate and struggle. Nevertheless, to call some labor elicited under certain pressures "voluntary" and other labor elicited under different kinds of pressures "involuntary" represents arbitrary line drawing that obscures the coercive content of some labor regimes, and ignores the important role of the legal system and laws in land and labor (including immigration) policy.

Further, we will suggest that a better way to understand the history of labor is in terms of a range of coercive practices utilized in different places and at different times, all constructed (but sometimes in different ways) by law, rather than by the present construct of discontinuous "types" of labor arranged under the headings "free" and "unfree". We will offer a number of examples of the ways in which different regimes can be made to achieve similar results. Finally, we will suggest that historians of labor should focus on a more particularistic set of inquiries about the particular practices that different societies and polities permitted and prohibited, the reasons for those decisions, and thus the nature and the outcome of the different varieties of what can be generally regarded as forms of labor coercion.

II

Part of the difficulty involved in the attempt to separate free from unfree labor arises from the fact that there are several different aspects to the package of labor inputs that need to be considered. The absence of legal or physical coercion in one dimension need not preclude such coercion in another. Various different forms of control may be used in the different dimensions of the labor relationship. Therefore, even when the absence of direct controls and coercion are as complete as they are under conditions of "freedom," some coercive elements may remain, which makes a definition of free labor problematic.

To better understand what is meant by various types of free and coerced labor,

including slavery and serfdom, it will be useful to trace the steps involved in the acquisition and use of labor, and to describe the problems faced by users of labor as well as laborers at each step. We can regard the ruling class as desiring to have a large supply of labor with low labor costs, where and when it wants it. Low labor costs need not necessarily mean low wages or other payments to laborers, since productivity must be considered, and expensive labor with large amounts of human capital may be preferred to cheap labor with only limited skills.

Among free populations, labor force participation is influenced by the desire to consume. This desire could be due to the basic need to acquire a minimal, physically necessary, amount of goods, although in general people desire to acquire, for whatever reason, amounts above what is physically necessary. Traditionally, among mercantilists and other believers in the "backward bending supply curve of labor," it had been argued that for an increased labor force participation poverty and low living standards were needed – otherwise most people would see no need to work.[3] The advantage of free labor, it was often argued, came not only from the incentives it provided by letting workers benefit from increased outputs, but also from the effectiveness of starvation as an inducement to work. It is, of course, possible to maintain aspects of this argument even today, as some do, using changing definitions of poverty and of what a necessary standard of living is. This relation among free labor, low incomes, and labor force participation is a reminder that income levels do not always track the legal status of workers or the nature of the distribution of power within society. Thus, it has been argued that, in the past, slaves had higher incomes than did free farmers or free wage laborers, and this persisted throughout most of the nineteenth century. Even under slavery, laborers may have been able to influence the determination of some part of their labor income and effort, as when slaves worked under the task system or other similar systems which allowed them to allocate some labor time, although, as with serfs, less working time for owners sometimes meant more working time overall to achieve desired consumption levels.[4]

Clearly, subsistence income levels were not determined exclusively by nature, since they incorporate much in the way of social conventions and of individual desires above the physiologically required minimum. At various times, where labor was nominally free, attempts were made, by introducing taxation, limiting the land to be settled, increasing immigration, and other similar devices influencing labor incomes and land availability, to lower incomes and change the amount

3. For discussions of this in the case of the British Industrial Revolution, see Coats, "Changing Attitudes to Labour"; and Mathias, *Transformation of England*, pp. 131-167. See also De Vries, "Industrial Revolution and the Industrious Revolution".
4. For recent discussions of slave influence on working arrangements and consumption patterns, see the essays in Hudson Jr, *Working Toward Freedom*.

and/or nature of the work free workers were "willing" to do.[5] Such considerations, for example, were central to the long-standing debate over the British Poor Laws, where the standard, and long-standing, objection to any subsidy for even the temporarily "deserving unemployed" was that workers should be maintained at poverty levels and could not be permitted to receive unnecessary subsidies from other members of society, particularly if labor could be obtained from those physically able to work.

The steps required to get labor to work when and where it is wanted include the processes of migration to place of work, getting labor into the particular work site, giving them incentives to be productive once on the job, and preventing workers from voluntarily leaving the job or region in which their labor is needed. Long-distance, intercontinental, migration to a new place of work, has been accomplished (1) by free migrations (when incomes reached an adequate level and incentives to move were there, as when industrialization over time in different parts of Europe forced people off the land), (2) by "voluntary" indentured servitude, where the initial contract (presumably but debatably entered into by voluntary choice) to pay for transport required a certain period of controlled labor and included the possibility of sale, and (3) by slavery, with the sale of labor on one continent for movement to another. In terms of the settlement of the New World prior to 1820, it was slavery that was most important, although the numbers were not as large as were to be the free migrations from Europe in the later nineteenth and early twentieth centuries.[6] In increasing numbers over time, however, there was a return flow of European migrants to their original homes, so that some aspects of the labor adjustment were more "temporary" than permanent, a situation quite different from that of slaves. These movements could be either for attachment to a specific employer in an area, or to provide a large pool of laborers in a particular area on which employers could then draw.

A type of geographic relocation historically more important for serfs, ex-serfs, and free workers than for slaves, arose with the expansion of industrial work in urban areas. Given the prior predominance of rural populations, some movement was necessary to permit modern development, a movement that led to some higher incomes but also to some dramatic changes in living and working conditions. Migration to urban areas could be either permanent or seasonal, with

5. "By creating conditions which restricted the alternatives available, increased labor input as well as changes in the composition of output were achieved. At the limit, the Malthusian subsistence level would determine the minimum necessary labor input, but it was possible to obtain similar results by artificial methods of social and political control. While individuals maintained legal property rights in themselves, controls over land and capital led to economic outcomes resembling those when property rights in persons belonged to others." Engerman, "Coerced and Free Labor", p.18.
6. See Eltis, "Free and Coerced Transatlantic Migrations".

quite different effects on family patterns, cultural continuities, and work habits.

Even when a sufficient population was available in a given area, problems could remain in getting people into the productive labor force or getting people to work at a desired work site, be it farm or factory. To accomplish this the ruling classes relied upon poverty, while individual employers could use positive incentives such as higher wages, shorter hours or better working conditions, or certain forms of legal or physical coercion involving specific controls or compulsions for particular work. Thus serfdom, which generally drew its labor force from people already resident, utilized the legal power of the lord to make serfs work for the lord's benefit, most frequently in agricultural pursuits, but, at times, also in industrial or service enterprises.[7] Similarly, there were controls introduced by the Spanish settlers over native-Americans that had different characteristics than African slavery, and did not involve long-distance involuntary movement of the Indians. After a brief period of slavery, the systems of *encomienda* and *repartimiento*, and then debt peonage, were used to compel working time on the estates of the Iberian estate owners. In addition, Indians were often required to make payments in the form of tribute to the government and/or to private landowners.[8] The coercion under serfdom and in the Spanish-American systems was intended to force work of a particular kind at a specific location, but the fact that these were already populated areas meant that the long-distance involuntary migration characteristic of modern slavery was not necessary.

Once in the workplace, high levels of worker output were not guaranteed, and incentives were needed to induce the desired amount of production. Some incentives, to the free workers or those coerced into the workplace, are regarded as positive (higher wages, including the use of piece wages; good working conditions, etc.), some as negative (fear of firing; physical punishments; removal of benefits; fines, etc.). The mix (since several different methods could be used together) could vary over time with the nature of legal institutions, and the kinds of incentives they intended to permit. Users of labor were frequently concerned with the morale of the labor force and its effect on output levels, whatever the legal status of labor. The most widely-cited argument on comparative incentives was Adam Smith's remark that slave labor would be less productive than free because of the lack of incentives for slave workers.[9] While this reflected the legal basis of the system, which did not require the payment of wages or the provision of other positive incentives, the point was also well-known to slave-owners, who frequently sought methods to provide incentives to their slaves to increase production, once they were on the plantation. At times the physical, psychologi-

7. On late serfdom, see Blum, *End of the Old Order*.
8. On the Spanish-American labor systems, see Burkholder and Johnson, *Colonial Latin America*, pp. 108-116; and Lockhart and Schwartz, *Early Latin America*, pp. 86-146.
9. Smith, *The Wealth of Nations*, p. 684.

cal, and material treatment of slaves and serfs was not as extreme as the law would permit, because of the desire to obtain high outputs from the workers.[10] Further, the need to direct and coordinate labor when working is a central concern of entrepreneurs, which requires the worker to forego certain rights during the contracted period. It has been argued, therefore, that during the work period itself the various forms of labor institutions approximate each other.[11] Clearly, the patterns of controls and punishments in the workplace would differ depending upon the legal status of labor, particularly the freedom to move from one job to another, as well as the specific nature of the job. This latter point could explain the different incentives and treatment of military in contrast to civilian labor.

An important concern of so-called free labor was the right to leave employment whenever and as soon as desired. There were some limitations on this agreed to under the terms of voluntarily accepted contracts, and there is a long history of court cases regarding specific performance, the need for reimbursement, and the measurement of loss due to premature quitting. An important example of exit restrictions has been the British Master and Servants Acts, which were carried into most of the British colonies in the nineteenth century.[12] These made the leaving of a job by "free workers" before the expiration of the agreed-upon time period a criminal offense, punishable by imprisonment. Similar examples of restrictions on quitting can be seen in what some regard as critical industries even in the twentieth century; for example, a British worker who left a munitions factory without permission in World War I would be forced to suffer six weeks of unemployment.[13] Similar legal provisions existed in other countries, particularly in the earlier stages of industrialization. In mid-nineteenth century Russia a "free" worker who left a factory without authorization was regarded as a criminal, basically similar to the status of a serf who had run away from the estate.[14] Thus even when there was freedom in hiring and in labor mobility, the ability to quit when desired, without criminal penalties, was not always available to workers.

10. See Fogel and Engerman, *Time on the Cross*, pp. 107-157; and the ensuing dicussion of slave material living standards.
11. For an early argument to this effect, see Cicero, *De Officis/On Duties*, pp. 68-69: "Similarly, the work of all hired men who sell their labor and not their talents is servile and contemptible. The reason is that in their case wages actually constitute a payment for slavery." An even earlier observation to this effect is to be found in Weber's description of Babylonian law regarding free hired labor – in *Agrarian Sociology of Ancient Civilizations*, p. 96 – "the period of work was regarded as temporary slavery".
12. In addition to the writings of Steinfeld, see also Craven and Hay, "Criminalization of 'Free' Labour".
13. See Wolfe, *Labor Supply and Regulation*, particularly pp. 217-234.
14. See Zelnik, *Labor and Society in Tsarist Russia*, pp. 21-43.

Finally, there were questions concerning the ability of even free workers to move to new locations in seeking new jobs, since such actions were sometimes restricted, used to increase the labor pool in an existing area. In the case of serfdom, labor was legally prevented from departing without the lord's permission, while the same was true under slavery. The slaveowner, however, was free to sell the slave to another person who would then control the slave's location. Neither slave nor serf was free to move of their own volition. With free labor, however, there were restrictions on mobility enforced in many different societies, and under several different arrangements. Land for sale and settlement was often restricted in countries of relatively low population density such as Australia and the United States, thus limiting the numbers of potential new settlers who might take up land and avoid wage work, providing benefits to landowners and capitalists in older, more settled areas. Similar effects restricting mobility can result from private contracts which include provisions serving to raise the costs of departure decisions. Frequently, after the ending of slavery, ex-slaves would be given small plots of land for themselves if they were willing to work for their landlords, thus presumably increasing labor availability and reducing its costs. Limitations on the incomes and nationality or ethnicity of potential immigrants and of settlers on land have existed, favoring dominant groups. Immigration laws, by influencing overall rates of settlement, have had a marked impact on returns to labor and labor supplied, even if they have not themselves had a direct, legislated effect on the incomes and locational patterns of workers. Given the frequent pattern of migrants settling disproportionately in urban areas, however, significant effects on wages and the level and structure of output resulted. We can also note that the British aid to the unemployed (poor relief) was initially intended to keep people in their original place by not letting them receive benefits if they moved elsewhere. Any such movement was also discouraged by the imposition of residence requirements to prevent aid payments in areas of in-migration.[15] By these measures the labor pool in the former area was increased, while the potential area of in-migration saved on benefit payments, albeit with a lowered labor supply available.

III

The discussion in the previous section was intended to point to several problems in maintaining the rigid distinction between free and coerced labor. First, coercion may not exist at each step of the labor acquisition and use process, but only at one or two of the stages. Thus the specific forms that coercion takes will

15. See Slack, *English Poor Law*; and Rose, *Relief of Poverty*.

differ, and what is regarded as coercion can vary over time and place. Second, the various indicators may not all go in the same direction. In particular, the distinction between the legal nature of the coercion and the relative material benefits to the coerced has long been observed. Poverty may be considered as one form of coercion of those legally free, but what some regard as poverty might have been the result of choices made voluntarily, though within a framework of constraints established by numerous factors including law. Third, it is possible to distinguish between economic coercion and legal coercion. Economic coercion results from the operation of the market, influenced by individual tastes in the face of economic constraints, but also by the legal regulations in effect. The market is itself structured by law. Legal coercion results from a more direct imposition of legal penalties and police controls. Fourth, in our focus on coercion for economic production we have omitted discussion of other reasons for coercing populations, such as military purposes, for criminal correction, for handling children and others deemed unable to care for themselves, as well as for dealing with vagrants and the unemployed who should have been capable of caring for themselves. Coercion in the economic sphere may be only a limited part of the total of society's coercive pattern.

IV

The present discussions of forced labor, concentrating on the dichotomy of free and slave is a residue of the eighteenth and nineteenth-century attacks on slavery.[16] Slavery at the time was seen as a unique evil, entailing the buying and selling of people and the absolute domination of one person by another. Whether from a religious or secular perspective, the radically unequal power and legal rights of the two parties were seen as inappropriate for modern times and modern societies. Even on grounds of economics slavery was seen as inefficient and unproductive and a system incompatible with modern economic development. In contrast, non-slave (free) labor was regarded as more productive, with its incentives generating more output per worker, as well as more morally acceptable, since it ended the legal domination over individuals by other persons. The debates over slavery posed a contrast that permitted no intermediate cases or complications in personal or legal status, a presentation useful for that particular attempt at reform, but one which has proven to be misleading as applied to the broad range of cases. For the legal, political, and economic systems have been able to provide constraints, which even if not legally as severe as slavery, serve

16. On the issues noted in this paragraph see, in particular, Davis, *Problem of Slavery in the Age of Revolution*, and *Slavery and Human Progress*; and Drescher, *Capitalism and Antislavery*.

to limit individual choice. While neither ideally "free" (however that might be defined) nor as coercive as slavery, these systems do constrain, by intent or otherwise, what individuals might do. Neither slave nor free, these hybrids contain elements of both. Some are based on initially voluntary agreements that constrain behavior for a period up to several years, some result as responses to legal or illegal actions undertaken by individuals, and some arose from a desire to constrain economic choices of landowners as well as workers on the land, but all do serve to limit worker rights and incomes, generally based on the control over the labor of individuals, not the full ownership of their bodies.[17]

V

The conventional dichotomy of free and coerced labor rests on a set of assumptions that have been with us for a long time. Labor is supposed to be arranged in a menu of "types": wage labor, contract labor, peonage, slavery, serfdom. Each "type" is supposed to have a set of fixed characteristics that define it as either "free" or "unfree". Wage labor is the principal "type" of free labor. Contract labor is one "type" of unfree labor. What separates free labor from unfree labor is that unfree labor is subject to physical or legal compulsion while free labor is not.

In this section, we will look more closely at these two "types" of labor in the United States and in England to show that nineteenth century wage work was sometimes elicited through legal compulsion, and nineteenth-century contract labor was sometimes not subject to this kind of pressure. Following the conventional distinction we would have to say that in these cases the wage labor was unfree labor while the contract labor was free labor. As noted above, for almost a century after the Industrial Revolution, most English wage workers could be legally compelled to perform their engagements.[18] If they tried to quit before they had completed their term of engagement they could be imprisoned and reimprisoned until they returned to work out their time.[19] During their period of engagement they could be sent to the house of correction for breaches of conduct, for disobedience, sloppy work or leaving before the work day had ended. The labor of English wage workers, in other words, was elicited under

17. The distinction between owning labor and owning the body of the laborer can be found in Blackstone, *Commentaries*, 2, pp. 401-402, and was later a part of the proslavery defense.
18. See particularly 6 Geo. III, c.25 (1766), and 4 Geo. IV, c.34 (1823).
19. The King v. The Inhabitants of Barton-Upon Irwell, 2 M. & S. 329 (1814); Ex Parte Baker, 7 El. & Bl. 697 (1857); Unwin v. Clarke, 1 L.R. 417 (1866); Cutler v. Turner, 9 L.R. 502 (1874).

the kinds of threats that under modern American constitutional law make the labor "involuntary servitude."[20]

On the other hand, contract laborers imported into the United States beginning in the 1860s could not be legally compelled to perform their agreements, even though their employers had paid their transportation expenses.[21] The law of Northern states simply didn't authorize criminal punishment for contract breaches.[22] Needless to say American employers did strive to hold their contract workers to their agreements. They incorporated numerous provisions into contracts to try to ensure their performance. Sometimes, they had groups of workers agree to be held jointly and severally liable for breaches by any member of the group. If one absconded before fulfilling her agreement, the remaining group members would be liable for the debt.[23] Or they would have a propertied person from the contract laborer's home town personally guarantee the passage debt in case the contract worker ran away.[24] Employers used these devices as substitutes for criminal punishment to try to force workers to perform their agreements.

The conventional picture of labor forms arranged under the categories free and unfree has reified labor "types." Neither contract labor nor wage labor were "things" with fixed, natural sets of characteristics, one of which was freedom or unfreedom. They were social/legal practices which could be constructed in a range of different ways. Freedom or unfreedom were coincidental rather than essential features of both. The state could decide to give criminal sanctions to employers for breach or could decide not to. Moreover, this decision was only one of numerous social and legal decisions that went into the construction of these practices. A hundred similar decisions determined the fine grained coercive content of these practices along a continuum, rather than in terms of a single yes/no (coerced/free) decision about criminal sanctions. Suppose, for example, that courts had refused to enforce contracts in which immigrants agreed to be held jointly and severally liable for breach. Judges might have taken this position on the ground that these kinds of coercive measures represented unwarranted impositions on immigrants. Or suppose the opposite. Suppose employers were legally entitled to double damages for premature departure, could garnishee future wages to collect these damage awards, as well as hold new employers responsible

20. See Clyatt v. United States, pp. 215-216; Bailey v. Alabama, p. 243; Pollock v. Williams, p. 18.
21. See in general Erickson, *American Industry and the European Immigrant*, pp. 1-63. See also Steinfeld, *Invention of Free Labor*, p. 172.
22. See, for example, Parsons v. Trask and Others, 73 Mass (7 Gray) 473 (1856).
23. Creamer, "Recruiting Contract Laborers", p. 46.
24. *Ibid.*, pp. 49, 53.

for the debt. A range of different possible practices of contract labor and wage labor are possible giving employers many different degrees of coercive power.

An implicit explanatory scheme underlies the contemporary map of labor "types" in which each "type" falls on one side or the other of the free/coerced dichotomy. Employers are viewed as facing an established menu of fixed labor "types". It is assumed that economic conditions explain why they choose a particular "type" in a particular place at a particular time. Unfree labor was used where the wage-labor market had not yet been created or where it had broken down, where as in the colonial periphery, labor was in short supply and land was abundant. Where there was a large population of propertyless laborers and where land was expensive, as in the mature metropolitan economies, free wage labor was the "natural" form of labor. Under those circumstances, the "dull compulsion of economic relations" could be relied upon to supply adequate labor at economical wage rates. Legal compulsion was unnecessary.

But the English and American experiences raise serious questions about this explanatory scheme. American employers did resort to contract workers largely because the domestic wage labor market had partially broken down. The military demands of the Civil War had made skilled labor very difficult to obtain at economic wage rates. But it is in such circumstances that employers are supposed to turn to unfree labor. Yet American contract labor was free labor. Why?

American contract workers were free because in the northern states into which they were imported the law did not provide criminal sanctions for contract breach. The contract rules which limited employers to money damages in northern states have to be understood as decisions taken by these polities through their legal systems to authorize certain kinds of coercion in labor relations but to prohibit others. To explain these kinds of decisions, it is necessary to undertake much more particularistic inquiries into local political (including the extent of the suffrage), legal, and cultural conditions as much as into local economic ones. The contract rules in effect in northern states were the product of a long and complex history of conflict. It is enough here to say that the outcome of this history was a collective determination reached in those states and later in the United States as a whole that slavery was illegitimate and should be prohibited and that forms of labor like slavery should be prohibited along with it.[25] The 1864 "Act to Encourage Immigration", under which some contract laborers were imported, itself makes clear why criminal sanctions for breach would play no part in this legislation. Section 2 of the act declares: "nothing herein contained shall be deemed to authorize any contract contravening the Constitution of the United States, or creating in any way the relation of slavery or servitude".[26]

25. See Parsons v. Trask.
26. *U.S. Statutes at Large*, XIII (1863-1865), pp. 385-387.

The same kind of analysis is called for in the case of wage labor, even the wage labor of the metropolitan core. Measured by comparative wage rates, American labor was in shorter supply than English wage labor throughout most of the nineteenth century. The relative abundance of English labor and the high cost of English land on one hand, and the relative scarcity of American labor and low cost of American land on the other would lead us to expect that if legal compulsion were used, it would have been used against American wage workers rather than English wage workers. But exactly the opposite turns out to have been the case. Where labor was scarcer and land less expensive in the United States, wage labor was free. Where labor was more abundant, and land more expensive in England, wage labor was unfree.

What needs explaining in the case of wage labor as in the case of contract labor is not the "type" of labor used, but the legal rules that give employers rights to certain coercive devices, permit them to use others, but prohibit them from using a range of still other possible forms of coercion. And it does not appear that these rules can be explained solely by a simple economic formula based on the characteristics of local land and labor markets. Any explanation for these rules must take into account numerous additional factors, including relative political power, the role of the judiciary, social norms, and the bargaining strength of labor.

VI

To this point, we have not directly challenged another previously mentioned basic assumption which underlies the contemporary picture of the forms of labor: the idea that labor is divided into free and unfree, and that it is natural to draw the line between the two at the point at which employers are able to compel labor by resorting to physical force or legal compulsion.

But just as we earlier criticized the soundness of the picture of discrete types of labor each with a fixed set of core characteristics, each functionally adapted to a particular stage or type of social/economic organization, at this point, we have to take a further step. We need to rethink the basic soundness of the binary opposition free/unfree labor. The line between voluntary and coerced labor has been drawn in different ways in different places at different times. For the most part, for example, the nineteenth century English did not consider their wage workers unfree. But there is a good reason that this line has been drawn in different ways. Nearly all forms of labor not performed for sheer pleasure can be characterized in either way.

When we speak about compulsion in labor relations, it is important to see that we are generally talking about situations in which the compelled party is only

offered a choice between unpleasant alternatives and chooses the alternative which represents the lesser evil to him or her.[27] This kind of compulsion is present in both slavery and free wage labor. In slavery, for example, labor is not normally elicited by directly imparting motion to a slave's limbs through overpowering physical force. It is compelled by forcing slaves to choose among very unpleasant options, between, for example, death, dismemberment, torture, endless confinement on the one hand, or back breaking physical labor on the other. The labor of free wage workers is similarly elicited by offering workers a choice, for example, between life on an inadequate welfare stipend established by the polity, or in the extreme, starvation, on the one hand, and performing more or less unpleasant work for wages for an employer on the other. In the case of both the slave and the free worker, the parties may be said to have been coerced into performing the labor, or to have freely chosen the lesser evil (made a beneficial bargain, in effect, given the background set of alternatives). Either characterization is available. And that is the reason that some choices among evils can be characterized as voluntary decisions while other choices among evils can simultaneously be characterized as coerced. Where the line is drawn, from a logical standpoint, is arbitrary. One tradition of thought, for example, portrayed slavery in completely consensual terms, as an agreement in which the slave had voluntarily granted absolute control over himself to his master in exchange for his life. In fact, in certain areas at certain times, slavery and serfdom were formally initiated through agreements. (This was never a feature of New World slavery, however.)

Needless to say, the choices presented in slavery are much harsher than the choices normally presented in free wage labor, and so we may rightly say that the degree of coercion in one form is normally much greater than in the other, but there are no logical grounds for saying that the performance of labor in one case is coerced, while in the other it is voluntary. As a matter of logic we have to say either that both are involuntary in different degrees, or that both involve the free choice of a lesser evil.

Often in drawing the distinction between free and unfree labor it is tacitly conceded that there are elements of coercion involved even in free labor. The distinction then comes to turn on the different kinds of coercion involved in free vs. unfree labor. Unfree labor is coerced through the threat or actuality of "physical violence" (state authorized in most systems of unfree labor), or "legal compulsion" (state administered). Free labor is supposed to be coerced through "economic compulsion" (market based).

27. Hale, "Force and the State", pp. 149-161, and Hale, *Freedom Through Law*, pp. 189-196. Much of the analysis in the following paragraphs derives from Hale's work.

Two implicit assumptions are made that justify the separation of labor coerced in these different ways into the opposing categories free and unfree. First, the different "types" of coercion are assumed to have fundamentally different sources and characteristics. The source of "legal" and in most systems of unfree labor "physical" coercion is law. The forces of the market are supposed to be the source of "economic" coercion, and they are supposed to operate impersonally and indirectly. They "coerce" only in the way that nature may be said to coerce, if you do not work you starve, rather than in the way "law" coerces, personally and directly. The forces of the market can be interfered with, but in so far as they are left undisturbed, no one controls them, they work naturally. Law is artificial, made and controlled by men. Second, the different types of coercion are assumed to operate with radically different degrees of harshness. "Physical" violence and "legal" compulsion are viewed as much severer than "economic" coercion.

This stark dichotomy, economic vs.legal coercion, which places labor subject to one in one category and labor subject to the other in an opposite category, operates to obscure the common sources and characteristics of the two. "Economic" coercion always has its source in a set of legal rights, legal privileges and legal powers, which place one person in a position to force another person to choose among disagreeable alternatives, just as "legal" compulsion and, in most systems of coerced labor, "physical" compulsion do. Its exercize is also quite personal and quite direct.

Consider "economic" coercion in the bargaining relationship between employers and employees. Robert Hale showed that law was the ultimate source of an employer's power to force a worker to choose among unpleasant alternatives. His argument is worth quoting at length.

> The owner [of private property] can remove the legal duty under which the non-owner labors [not to consume or otherwise use that property]. He can remove it, or keep it in force, at his discretion. To keep it in force may or may not have unpleasant consequences to the non-owner – consequences which spring from the law's creation of a legal duty. To avoid these consequences, the non-owner may be willing to obey the will of the owner, provided that the obedience is not in itself more unpleasant than the consequences to be avoided. Such obedience may take the trivial form of paying five cents for legal permission to eat a particular bag of peanuts, or it may take the more significant form of working for [a factory] owner at disagreeable toil for a slight wage. In either case the conduct is motivated [...] by a desire to escape a more disagreeable alternative [...] In the case of the labor what would be the consequence of refusal to comply with the owner's terms? It would be either absence of wages, or obedience to the terms of some other employer [...] Suppose, now, the worker were to refuse to yield to the coercion of any employer, but were to choose instead to remain under the legal duty to abstain from the use of any of the money which anyone owns. He must eat. While there is no law against

> eating in the abstract, there is a law which forbids him to eat any of the food which actually exists in the community – and that law is the law of property. It can be lifted as to any specific food at the discretion of its owner, but if the owners unanimously refuse to lift the prohibition, the non-owner will starve unless he can himself produce food. And there is every likelihood that the owners will be unanimous in refusing, if he has no money [...] Unless, then, the non-owner can produce his own food, the law compels him to starve if he has not wages, and compels him to go without wages unless he obeys the behests of some employer. It is the law that coerces him into wage-work under penalty of starvation – unless he can produce food. Can he? Here again there is no law to prevent the production of food in the abstract; but in every settled country there is a law which forbids him to cultivate any particular piece of ground unless he happens to be an owner. This again is the law of property. And this again will not be likely to be lifted unless he already has money. That way of escape from the law-made dilemma of starvation or obedience is closed to him [...] In short, if he be not a property owner, the law which forbids him to produce with any of the existing equipment, and the law which forbids him to eat any of the existing food, will be lifted *only* in case he works for an employer. It is the law of property which coerces people into working for factory owners [...][28]

So-called "economic" coercion is an artifact of law, not of nature, and operates personally and directly on individuals. There is, however, a crucial respect in which it does differ from what we normally call "legal" compulsion. "Legal" or "physical" compulsion normally confronts an individual with a highly restricted set of unpleasant alternatives from which to choose. The choice, for example, might be between going to prison, becoming a fugitive from justice, or remaining at work. In most cases of "economic" coercion, the universe of alternatives is at least potentially much wider. A worker thinking about leaving his employer can try to find another job, apply for food stamps and Section 8 housing, open a small business repairing cars in the back yard homestead in Alaska, move to Georgia where there are more jobs, emigrate to Australia, go back to school to retrain, join the army, move in with relatives, play guitar in the subway, and so forth. What is different about "economic" coercion is that it is often difficult to evaluate the universe of options which an individual faces. If most of the above options are unfeasible, as they often are, and the real choices are much narrower, remain at work, look for another job in the same line, go on welfare, become homeless, then the degree of coercion operating in this situation may even be greater than the degree of coercion operating when the threat is a mere fourteen day prison sentence. (Many nineteenth century English wage workers served such short term sentences.) On the other hand, if Australia is desperate for workers, is willing to subsidize passage over, provide housing, and guarantee a good paying job, and the unemployment rate in this country

28. Hale, "Coercion and Distribution in a Supposedly Non-Coercive State", pp. 472-473.

is 2%, then the degree of coercion operating in this situation may be practically nonexistent. And so it is difficult to characterize this worker's situation in the abstract, because the feasibility and unpleasantness of all the potential options involves highly context specific information. It is hard even to compile a definitive list of what all the "real" options may or may not be. It is much easier to evaluate degrees of coercion when the alternatives confronting an individual are clearly defined and strictly circumscribed. This is why "economic" coercion is so slippery and contestable.

But this difference does not make "economic" coercion "natural". The background conditions which constitute the options available to individuals and determine the degree of "economic" coercion operating in any situation are pervasively shaped by law.[29] Whether welfare is a good, bad or impossible option depends upon welfare law, on eligibility rules and payment standards. Emigrating to Australia may be hard or easy depending upon Australian immigration rules. But it is only an option at all because American workers enjoy the background legal privilege of leaving the country whenever they wish to take up work elsewhere. Alaska may be a better or worse option depending upon whether the state government has a homestead program, but it is an option because Americans enjoy the legal privilege of settling in any state in the union they wish to. Playing guitar in the subway may be more or less feasible depending upon vagrancy rules and vagrancy rule enforcement. Another job may be a better or worse option depending upon how great the demand for labor is. But it is only because a worker possesses the background legal privilege of selling or withholding his labor in the first place, i.e. is not under a legal duty to work where an employer has need of labor (as in England under the Statute of Labourers, or as in slavery), that increased demand for labor can make this a better or worse option. If a preparatory course and a license are required in order to repair automobiles, or if nuisance law or zoning regulations prohibited it, opening a small shop in your yard might be less feasible or completely out of the question. How desperate workers as a group might be for work depends in good part on how widely property is distributed. But even this situation is conditioned to a great extent by a larger set of taken for granted background legal rules. Property is so unevenly distributed in the United States in good measure because inheritance laws make it possible for accumulations of property to be preserved over generations, and so on.[30]

29. Kennedy, "The Stakes of Law, or Hale and Foucault!", pp. 330-334, and "The Role of Law in Economic Thought", pp. 965-967.
30. On the "artificial" or "positive" nature of the law of inheritance see, for example, Blackstone, *Commentaries*, II, pp. 10, 12: "The most universal and effectual way, of abandoning property, is by the death of the occupant; when, both the actual possession and intention of keeping possession ceasing, the property, which is founded upon such possession and intention, ought

In extreme cases, such as South Africa, where ruling elites adopted vagrancy legislation and legal restrictions on land ownership in an effort to drive Africans into wage labor, law conditioned (created) "economic" coercion may produce outcomes very similar to those achievable under regimes of direct "legal" compulsion like slavery.[31] But in all market societies, an extensive set of background legal rules establish to a significant degree the real alternatives available to working people, and hence the extent of the so-called "economic" coercion they confront. In most market societies, these numerous background rules are not put in place for a single purpose or by a single group. Nevertheless, whatever the formal reasons for their adoption, they affect the universe of possibilities open to working people. And in many cases, at least certain of these rules (welfare and vagrancy rules for example) are consciously adopted with an eye toward inducing laboring people to enter into or remain in wage work. Depending upon their precise details, the magnitude of "economic" coercion facing ordinary people can be quite different in different free market societies.

Legal rules governing bargaining behavior in the market also play a crucial role in the creation of so-called "economic" power.[32] In free markets, the state must always make a detailed set of decisions about which bargaining tactics will be permitted and which kinds prohibited. It simply has no choice. The legal system will always be confronted with claims that one action or another of a bargaining party has injured the other party. If workers are given the privilege or right of freely combining to withhold their labor, employers may enjoy less coercive power in the relationship because workers have been given legal permission to damage them through certain forms of activity. If workers are permitted to mount secondary boycotts, employers may have less power still, because workers have been given legal permission to damage them in another way. If employers are prohibited from replacing workers who go out on strike, they may enjoy even less coercive power. If employers possess the legal privilege of combining to negotiate with workers, they may enjoy more power vis a vis their employees, and so forth ad infinitum. Change one or two rules and you may change the degree of coercion slightly. Change all the bargaining rules in favor of one party and you may change the degree of coercion radically. There are numerous possible ways free market bargaining relationships can be legally

also to cease of course [...] Wills therefore and testaments, rights of inheritance and successions, are all of them creatures of the civil or municipal laws, and accordingly are in all respects regulated by them [...]" See also Fried, *Robert Hale and Progressive Legal Economics*, p. 41: "The Supreme Court [...] in its 1898 decision in Magoun v. Illinois Trust, [upheld] a progressive inheritance tax against an equal protection claim, on the ground that as inheritance itself was a special privilege conferred by the state, the state was free to condition it as it saw fit."

31. See Engerman, "Coerced and Free Labor".
32. The following paragraph draws heavily on Kennedy, "The Stakes of Law, or Hale and Foucault!", and "The Role of Law in Economic Thought".

constructed, all of them placing the parties in different positions to coerce one another.

In an earlier section (V), we looked at another set of legal rules that define another dimension of power in labor relationships, rules governing what contract remedies are available to the parties, and what remedies they may establish for themselves as part of their agreement. A nineteenth century American common law rule of contract provided that if a worker left his job before he had fulfilled his contractual term of service, he forfeited his legal claim to wages for the time he had already worked.[33] Where an employer had held back wages, such a rule forced a worker to choose between staying and forfeiting his accumulated wages. If he stayed because he simply could not afford to lose his back wages, the legal rule would have operated to coerce him into remaining in the same way a rule that forced him to choose between staying and serving fourteen days in prison would have. Is the compulsion of the American rule any less "legal" than the compulsion of the English rule of criminal sanctions?

So-called "economic" coercion, the coercion of the market, is not like the coercion of nature, impersonal and indirect, in numerous different ways it is a product of law, just as so-called "legal" coercion is. Thus, the effort to separate labor into opposing types on the basis of a simple opposition between "economic" and "legal" coercion does not hold up. When we examine the economic/legal distinction carefully, it dissolves into a complex account of the numerous different ways in which coercion is constituted by law.

The second assumption made about "economic" coercion that justifies placing labor coerced by the market into the category free, while labor coerced by "law" is designated unfree is that "economic" coercion is less harsh than other types of coercion. In the twentieth century, a kind of implicit scale of coercion has been developed arranged abstractly by "types", with "economic" coercion placed at one end as mild, and "legal" and "physical" coercion at the other as severe. American law certainly seems to be based on this kind of abstract scale. But doesn't the coerciveness of all "types" of coercion depend upon particular circumstances? Can't economic coercion sometimes be more coercive than legal coercion? Depending upon the precise circumstances, for example, imprisonment may or may not be more coercive than so-called "economic" threats. The threat of starvation may certainly operate more powerfully than a short prison term. 19th century English wage workers sometimes expressed a preference for a short prison term rather than forego the opportunity to obtain higher wages by changing employers. And who is to say that a fourteen day prison term is invariably a more disagreeable alternative than losing one's life savings, or home.

33. Nearly all states adhered to this rule during the first half of the nineteenth century. See Holt, "Labour Conspiracy Cases in the United States", and "Recovery by the Worker Who Quits"; see also Karsten, "'Bottomed on Justice'".

It is clear that there are numerous degrees of "legal" coercion just as there are numerous degrees of "economic" coercion. It is more accurate to think of the two in terms of two parallel scales, or a combined scale, in which both run from very severe to rather mild, rather than in terms of a single scale arranged by abstract "types". In characterizing labor relations in terms of a continuum, however, we don't mean to imply that there have not been real differences in these forms of coercion. Dramatic differences have existed and these have been the basis for profound political and social disagreements and occasional violent conflict.

VII

Where does all this leave us. First, where the opposition free labor/coerced labor is used, we have to see the line drawn between the two as arbitrary. One of our tasks should be to try to explain how such arbitrary lines came to be constructed, how certain forms of coercion were subtly legitimized by being sanctioned by law. The nineteenth century English, as we have said, considered their wage workers free for the most part. Twentieth century American lawyers and historians view the decision of a worker to remain in or return to a job because his employer threatened him with imprisonment as a coerced decision in which continued work is involuntary servitude. On the other hand, twentieth-century American lawyers and historians view the decision of a nineteenth-century worker to remain at a job because his employer threatened him with the loss of his back wages as a voluntary decision in which the continued work was consensual not coerced.

Second, if the categories voluntary/coerced, free/unfree labor are, in certain important respects, empty conceptual shells, our task becomes to offer historical accounts of labor relations in terms of specific coercive practices, all, in different ways, the product of law, rather than in terms of discontinous "types" of labor arranged under the categories free and coerced. These accounts should attempt to explain why a particular polity adopted the rules it did, why it granted to employers or workers particular legal rights, particular legal privileges, or imposed particular legal duties, because these rules determined which coercive devices the parties would have a right to, which would be legally tolerated and which would be prohibited – from among the rich variety of possibilities. Political and legal struggle between different groups over these rules must certainly be part of these explanations. These accounts should not only describe struggles *over* rules, but also struggles *under* rules. How individuals and groups defined and pursued their interests was also a function of the set of rules within

which they operated at any particular time. Change rules, and interests might well come to be redefined and pursued differently.

In writing this kind of history, we have to reject the view that history unfolds in stages with each stage a functionally coherent whole composed of a fixed block of practices and institutions all required by the first principles of that form of social/economic organization. Rather, the lesson to be learned from nineteenth century English wage labor is that the rules governing social/economic life are established in ad hoc ways out of a variety of existing and new materials, and may be changed piecemeal, the outcomes depending on the relative political power of the participants, extant social views about permissible and impermissible forms of coercion, the normative persuasiveness of different groups, and a host of other factors.