The Causes of Serfdom:

Domar's puzzle revisited.

Jean-Jacques Rosa

Sciences Po Paris

March 10, 2009. Revised, May 2, 2011

13,561 words

Abstract

Domar (1970) has singled out the land/labor ratio as the main cause of both slavery and serfdom. But he then recognizes that scarcity of labor is necessary but not sufficient for serfdom to exist, and that an exogenous "political factor" is required to determine the status of labor. I show that in an agrarian subsistence economy the complementary conditions of serfdom are (a) oligopsony power in labor demand, sustained by (b) an oligopolistic supply of violence by large landowners. These conditions explain historical episodes that Domar accounted for, those he left unexplained, and imply that serfdom and slavery are mutually exclusive.

JEL numbers: P26, P48, P23, J42.

Keywords: Oligopsony, rents, price elasticity, cartels, slavery, serfdom.

1. Introduction

Serfdom is usually described as a situation where labor is tied to the land, examples of which are to be found in the late Roman Empire, Western Europe during the early Middle Ages, and Russia in the 16th and 17th centuries. In a well known analysis, Domar (1970) attempted to revive an older tradition linking this status of labor to the land/labor ratio: facing an increasing abundance of land, landowners would restrict labor mobility in order to suppress competition for labor and by this means prevent the falling trend of land rents. Domar however shows convincingly that plausible as the land/labor ratio explanation is it cannot explain other historical cases, and especially the vanishing of serfdom in Europe following the Black Death, or the absence of serfdom in the land abundant and labor scarce Northern American Colonies. He suggests the need for, but does not provide, a political theory of serfdom to replace the factor ratio approach.

After briefly summarizing Domar's reasoning (Section 2), I shall suggest a more complete economic theorizing: while a simple factor ratio approach cannot explain by itself the occurrence of serfdom, the theory of monopsony defines the necessary conditions for such an organization of production to exist. Moreover, when developed in its dominant firm variant, it can be used to make explicit the contribution of the political authority to the enforcement of such an organization, and the reason why that authority could choose to intervene (Section 3).

To understand further the reason for the political authority intervention, one has to show however how the structure of the economy determines both the optimal tax choice and the optimal size of the political unit. Serfdom is the result of a certain structure of the economy and of the polity.

In such a broadened political economy framework the historical cases that Domar found puzzling are easily explained (Section 4). And it is then possible to show precisely in what respect slavery fundamentally differs from serfdom (Section 5).

As a conclusion, objective conditions favoring freedoms are shown to be the same than the conditions for serfdom, but in reverse. They are the opposite side of the same coin: the intensive use of market exchanges, a low concentration of resource ownership, and an expanding size of political units, provide the conditions for a vanishing of serfdom and the spread of free labor.

2. The land/labor deadlock.

Forty years ago, Evsey D. Domar presented in the Journal of Economic History (March 1970, 18-32) an elegant and potentially powerful theory about the causes of slavery or serfdom.[1] Or more exactly, in his own terms, he revived an earlier hypothesis suggested by the Russian historian Kliuchevskii in his description of the Russian experience of the sixteenth and seventeenth centuries.[2]

Domar summarized the economic theory that he build on the Kliuchevskii's historical account as follows: Russian land owners lived off rents collected from their estates. This proved difficult however because land in Russia was abundant relative to labor, and was made even less scarce by Russian conquests and territorial expansion. Since the scarce factor of production was not land but labor, competition among employers raised wages to the (high) level of labor productivity, laminating rents[3]. It was thus the ownership of peasants and not of land that could yield an income to the propertied class.

Such an ownership of men, however, was not likely to be obtained by economic forces alone as far as peasants could easily migrate and use the large amount of new land available on the eastern and southeastern marches of the empire to set up new independent farms, given the low level of capital and competence necessary for such ventures[4]. With peasants free to move, the competition among owners will raise wages to the level of marginal labor productivity, erasing landowners' rents.

Under the pressure of the landowners' class that it wanted to satisfy, the tsarist government gradually restricted the freedom of peasants to move. They became enserfed by the middle of the seventeenth century and more completely so during the following decades (Domar, op. cit. p. 19). With labor tied to land, competition between land owners then ceased and the employers could appropriate all or most of the peasants' income above some subsistence level[5].

---

[1] « The causes of slavery or serfdom: a hypothesis », op. cit. The terms slavery and serfdom are used by him interchangeably. More on this later.

[2] V. Kliuchevskii, Kurs russkoï istorii [A course of Russian history](Moscow: Gosudarstvenne sotsial'no-economischeskoe izadatel'stvo, 1937, 1906) translated as A History of Russia (New York: Russell and Russell, 1960) by C.J. Hogarth. Domar also quotes several predecessors in the concluding section of his paper.

[3] Marginal productivity of labor is close to average productivity under conditions of land abundance and free labor, and thus little surplus will remain for the owners.

[4] In other terms there were no barriers to entry in the agricultural sector, absent political intervention.

[5] As far as custom allows and serfs do not run away, which they did, in effect a common occurrence in Russia.

It seems clear then that serfdom was the result of a high land to labor ratio. But having developed this parsimonious theory, Domar then proceeded, with great scientific integrity, to undermine (or complement?) it. Serfdom, he notes, depends critically on the political will of governments to bestow rents on the owners' class. He thus has to acknowledge that:

"The presence of this exogenous political variable seriously weakens the effectiveness of my model: it makes the presence of free land by itself neither a necessary nor a sufficient condition for the existence of serfdom. It is not a necessary condition because so long as marginal productivity of labor is high, serfdom will continue to exist even if free land is no longer present. […] Free land is not a sufficient condition because […] without governmental action free land will give rise to free farmers rather than to serfs." (p. 19).

Hence Domar's example: suppose that, land being kept constant, the population increases. Then the free wage will fall, the difference with the serf's subsistence wage will diminish, and the landowner has no more incentive to enserf the peasant, because no rent can be extracted from him. Peasants will be free. But let population decline instead, the free wage will increase as a consequence. Absent governmental constraint limiting their mobility, the peasants will be better off because free to use the larger per capita amount of land.

Thus, diametrically opposite changes in the land/labor ratio lead to exactly the same condition for labor, as far as there is no deliberate governmental intervention, leaving all the task of explanation of the labor institutions to the sole political factor that Domar considers exogenous to the problem. This indeterminacy is explained below (see page 8).

It follows that the land/labor model cannot predict the net effect of a change in the land/labor ratio on the status of the peasants, that is, on serfdom or freedom. The really determining factor is thus the governmental intervention suppressing the mobility of labor in order to introduce a wedge between the productivity of labor and a wage at or near the subsistence level, allowing in other words an exploitation of labor by the landowners.

As historical examples contradicting his land/labor hypothesis, Domar mentions the case of post Black Death Western Europe, and also that of North American States versus the Southern Confederates (which did confirm the theory). In both cases the land/labor ratio was quite high and increasing, but no serfdom resulted.

Commenting on Domar, Paul Krugman reminds us what the Black Death puzzle is: In Europe circa 1100, with population scarce, serfdom was useful to the ruling class. By 1300 it wasn't, because population had much increased, the wages tended to decrease towards subsistence level, and as a consequence, serfdom had been allowed to drift away. But after 1348 – the beginning of the Black Death - it should have been worthwhile again. Yet it wasn't effectively reimposed despite a fall of 1/3 of the population. "There were attempts to restrain wages and limit labor mobility, as well as attempts to tax the peasants…. but all-out feudalism didn't return. Why?" ("Serfs up!!", Entry of 5/8/2003, on *The Unofficial Paul Krugman blog*).

Was serfdom then only a question of political will, and was it, as such, discretionary, putting the analyst in a need to examine things on a case by case basis and search for the particular details of political equilibrium?

Domar did not answer that question. What he emphasized was that a necessary (and sufficient?) condition for extracting rents from labor in an agrarian economy was a governmental intervention meant to enforce collusion among employers, an agreement not to compete for labor, and the best way to do that is to restrict the mobility of peasants (serfdom assigning every peasant to a specific land, and to its owner). This intervention is all the more necessary that suppliers of labor have more employment (and self-employment) opportunities: this is the case when the land is abundant. But it can be also the case when employment opportunities in other sectors of the economy are numerous and within "easy" reach for labor.

In both cases, the supply of labor to the agricultural sector becoming very elastic, because of occupational mobility, the institution of serfdom has for aim and for consequence to increase the market power of employers by forbidding peasants to take advantage of the competition for labor amongst employers. Indeed, serfdom is a specific legal status for labor. It is defined and enforced by a political authority.

He thus concluded that a political explanation was central to the explanation of serfdom. The question we have to ask then is: why did not all states or central authorities chose to enforce a legal status of serfdom everywhere?

To find a way out of Domar's impasse we thus have to develop – as he suggested but in an economic rather than a sociological or political way - the analysis of political conditions leading to the adoption of such a status for labor. This is easier done today than in 1970, due to the development of the discipline of public choice. But first, we have to define precisely serfdom in order to understand what its basic economic characteristic is, a prerequisite for selecting the adequate theorizing capable of isolating its determinants.

In what follows we show that to break away from the Domar deadlock we have not only to provide a theory of political choice, but also to give a precise definition of serfdom in order to understand the economic conditions – the economic model - underlying the political possibility - and interest - of serfdom. Both the political theory and the economic adequate model were not available in the 60s, the time at which Domar wrote.

3. A Theory of Serfdom

Let us first expose the general outline of our economic and political market power theory, and then examine the broad social conditions that give rise to monopsony serfdom (a polar case), or to collusive oligopsonistic serfdom (the general case). In both, the optimal size of political units combined to the size distribution of land ownership and agricultural exploitations – or "firms" - rather than the simple land/labor ratio, determine the likelihood of serfdom as a mode of production in self-sufficient agrarian economies.

3.1. Analytical outline

1. Serfdom is usually defined by the unfree status of labor: peasants are tied to the land, assumedly by landowners' coercion, and such a characteristic reduction of individuals' rights makes it quite similar to slavery. Indeed, many authors (Domar, Engerman, etc.) put serfdom and slavery in a same category, only with a difference of degree in the suppression of individual rights[6].
   We argue that the definition of serfdom must be much broader, and take into account the structure of production and the mode of cooperation between the various production factors. We suggest the following provisional definition:

Serfdom is mostly an agrarian production system in which, de facto or de jure, voluntarily or involuntarily, labor is tied to the land of an owner and is allowed to cultivate it for himself in exchange for a (mostly in-kind) input supply payment. Laborers (serfs) also receive some protection and justice service from the landlord who is simultaneously the main political authority.

Regarding the coerced nature of serfdom Stanley L. Engerman emphasizes the wide range of possibilities:

[6] "Slavery and serfdom have generally been regarded as the extreme versions of coerced labour, one end of a spectrum, at the other end of which is what would be called free labour." (Engerman, "Slavery, serfdom and other forms of coerced labour: similarities and differences", p.19).
According to North and Thomas: "The reason why Domar's explanation fails is that serfdom in Western Europe was essentially not an exploitative arrangement .... (it) was essentially a contractual arrangement where labor services were exchanged for the public good of protection and justice." (p. 778). They add that a contract is defined by the impossibility for either party to change it unilaterally. We'll return later on to the conditions that could lead to coercion in serfdom.

"In general, most forms of slavery and serfdom have been regarded as involuntary institutions. This view is based primarily on the cases of the New World slavery of the European powers and of serfdom in Russia, where this depiction is clearly most appropriate. Yet, given the many cases of slavery and serfdom, it is of interest to see when those statuses were the outcome of voluntary and mutual decisions. Some such cases included arrangements to provide the labourer with subsistence or defence, in exchange of the giving up of some control over work and living arrangements[7]. These varieties of social security could be important where levels of income were low and highly variable or where societies were warlike and chaotic." ("Slavery, serfdom and other forms of coerced labour: similarities and differences", chapter two in M.L. Bush (ed.) *Serfdom & Slavery: Studies in Legal Bondage*, London and New York: Longman, 1996, p. 20).

"In some sense, the interesting historical question may be less that of whether coercion existed, than of why specific forms of coercion existed at particular times, and why they changed over time." (Ibid., p.18).

But in all cases there is some tying of the laborers to the land and payment in labor services to the landowners.

Serfdom thus broadly defined characterized predominantly agricultural, self-sufficient economies, with few market exchanges and very limited monetary circulation: classical Greece Sparta, or the Athenian economy before its commercial expansion, the late Roman Empire when the economy, repeatedly invaded by the barbarians on its north-eastern borders, was on a returning path to subsistence in large villas and pre-medieval manorial organizations, the Middle Ages from the dissolution of the Carolingian Empire until the revival of trade, urbanization, and monarchy in the 11-13th centuries, and Russia in the 16th and 17th centuries, well before the industrial revolution.

2. What we want to explain is the rise and decline of serfdom, the changing fortunes of this institution in time and space.

3. Let us start with the first characteristic, the agrarian production function. Agricultural production requires three inputs: land, labor and security (a fourth is physical capital but we can put it aside for the moment).

4. In serfdom there is a <u>composite exchange</u> between these factors of production:

[7] It should be noted that the modern labor contract for free workers also includes some giving up of control over individual decisions during working hours, and over some aspects of living arrangements, for instance the residential area of choice, place of work, and time schedules. There is really a continuum of situations between forced labor and completely free labor of the individual craftsman or artist.

An in-kind service of labor is exchanged for land use and protection (and justice). There is a tie-in in the supply of land and protection by the landowner.

This composite exchange can be written as:

VMPPL - Wage = Land rent + Tax > 0 (1)

Where: VMPPL is the value of marginal physical product of labor, and
Wage is the income of the labor factor. The right side is the total landowner's income.

In this complex exchange under serfdom, the serf receives, in addition to the right to cultivate the landowner's property for his own consumption, a protection against external dangers (invasions, wars, raids) as well as civil security in local disputes, a law and order service.

The rent extracted from labor by the landowner includes a classic land rent on the one hand, and a tax for financing the supply of a public good (justice and defense) on the other. It is not clear under these circumstances, that the serf is economically "exploited" for an amount equal to the rent plus tax that he pays, as far as these payments are equal to the costs of the services he obtains.[8]

Part of the rent extracted from labor is used for the production of security, either directly by the landowner himself who is also, in a pure manorial system, a knight and castle owner, and as a consequence the local political authority (detaining the monopoly of legitimate violence), or indirectly by the military service that the landowner owes to his overlord ("suzerain") in exchange for the use of the "fief" that the suzerain granted to him. To account for this complex implicit contract we suggest a new, more precise definition of serfdom:

Serfdom is a fiscal system adapted to self-sufficient economies in which the tax is mostly on labor and is collected by the landowner or his "franchisee". It is indeed a sort of franchise contract between the overlord and his vassal, attributing to the latter the use of the overlord's land in exchange for the supply of military services, and cum in-kind tax farming.

The tying of labor to the land is a means for collecting taxes from labor as we will see shortly.

[8] Cotton M. Lindsay notes that this neoclassical concept of "exploitation" is debatable since it implies that higher capital investment for instance, given a competitive supply of labor, would raise employment, leave the wage unchanged and raise the exploitation of labor. See Applied Price Theory, pp. . We will stick nevertheless to this concept here as a measure of the divergence of the value productivity of labor and the wage (or income of labor) in monopsonistic labor conditions (rather than "markets" since precisely in the serfdom case it is dubious that a labor market exists.

5. The previous condition of equation (1) implies both that:

   VMPPL > subsistence wage, or the system would not be viable, and:

   VMPPL > "Wage", the income of the serf, which means that there is an expropriation, or transfer, of some of the labor revenue to landowners.

6. For this transfer to take place, then, landowners have to wield some market power, i.e. some monopsony power in the labor "market".

   Serfdom does not depend – contrary to Domar and Kliuchevskii – directly on the land/labor ratio (or other factors to land ratios) <u>because the factors ratio does not determine in itself the degree of competition on factors markets</u>. The factors' ratio determines which factor could potentially benefit from a large producer's surplus, but not the conditions for the appropriation of this surplus, and especially not the degree of competition on factors' markets.

   A change in the land/labor ratio changes the competitive wage level. Specifically a rise in the ratio increases the equilibrium wage. But at the same time (see Appendix 2) an outward shift of the demand for labor resulting from an increase of the land/labor ratio, as well as a leftward shift of the supply of labor leading similarly to a wage increase, would both increase the price elasticity of labor supply at equilibrium.

   Thus the increase of the land/labor ratio would both increase the equilibrium wage above subsistence level, and thus the potential for surplus extraction from labor, but at the same time it would reduce the ability of landowners to extract such a rent because their monopsony power would be decreased by the higher price elasticity of labor supply. The net effect is indeterminate, which is the reason why Domar could not find any unambiguous impact of the land/labor ratio on serfdom.

7. Given an equilibrium wage level, monopsony power depends on turn on other factors: land ownership concentration in the "relevant economic space", and limited or absent markets. Monopsony power is maximal where there is only one owner of productive units (that demands labor) in a given economic space defined by the political entity borders.

   The theory of serfdom is based on the following determinants:

   (a) The productivity of labor is higher than the subsistence wage: there is an agricultural surplus that can be taxed.

(b) There are few or no productions for markets and trade.
(c) There is a high concentration of the ownership of production factors complementary to labor (land mostly).
(d) There is a convergence of the size of economic and political units.
Condition (a) is obvious, conditions (b) and (c) determine monopsony power for the landowner, and condition (d) insures that the implicit contract between landowner (or franchisee) and peasants includes the supply of protection and in-kind tax farming.

8. Depending on the relative sizes of the productive units and the political unit, two variants of the monopsony power model are defined:

a) The monopsony proper is the simple (polar) monopsony case.
b) The oligopsony, "dominant firm and fringe" model (or cartel model), is the more realistic and generally relevant version.

Simple monopsony is a special case where the economic unit and the political unit have the same dimension. There is just one monopsonist in the relevant market, so that there is no "domestic" (or internal) political problem in that case. The financing of the political authority comes from the labor rent extracted by the monopsonist (See Appendix 1 for a brief presentation of the simple monopsony model).

In the oligopsonistic or dominant firm model, there are several economic units (the Roman "villa" or medieval "demesne") in the political unit space. The exercice of monopsony power relies on politics (cartel organization) in order to reduce or suppress the competition for labor. The financing of the political authority is more diversified than in the pure monopsony model, the monopsonistic power is reduced, and the size of the political unit is itself endogenous to the structure of economic activity because a subsistence agrarian economy cannot support a large state relying on specialized bureaucratic organizations.

As explained below, in the "fringe model" the Buying Power Index (BPI) which measures the capacity to extract rents from labor depends on the share "Sh" of the "industry" (here agriculture) detained by the producers that participate in a cartel. The higher that share, the higher the cartel's power on the labor market.

Sh in turn depends on:

a) the concentration of producers and

b) the cohesion between the interests of producers and the interests of the political authority which alone can enforce the legal constraints imposed on labor (see Engerman, , and North & Thomas). That cohesion is maximal if there is a monopsony, i.e. if there is only one producer per political entity. Such a one-to-one correspondence between the productive unit and the political unit is found in two extreme systems: the pure manorial economy in which the landlord is also the political sovereign, and the soviet type planified economy in which the state, the political sovereign, has also acquired the complete set of enterprises in all sectors of the economy, and is thus a monopsonist of labor in the country.

In all other organizations of the economy, there are several independent producers within a political entity. The cohesion between the political ruler and these landowners aimed at the enforcement of the oligopsony rule is maximal when most or all of the ruler's revenues come from the core cartel of producers, i.e. an agrarian economy in which markets are undeveloped and the fringe of other demanders of labor is consequently thin. On the basis of such an economy, the size of the political unit is small, because the in-kind taxing of agricultural production cannot travel at long distances and must be consumed locally (hence the "itinerant" way of life of the king and his court under the Merovingian and Carolingian dynasties).

When markets and specialization develop, the ruler can extract trade tax, this gives rise to larger political entities (states), a larger fringe of other productions, and this decreases the share of the core cartel, Sh, in the overall economy, thus determining a decrease of BPI. The ruler then is less inclined to defend the rent extraction capacity of landowners by the legal immobilization of labor on land.

There are exceptions however in large states that enforce serfdom: the late Roman empire, the Carolingian empire, and the Russian 16th and 17th empire in formation. All are examples of transitional agrarian economies with very limited -- whether contracting (Late Roman empire) or expanding (Roman empire) -- fringes and few monetized markets.

In the late Roman empire, the disruption of the economy and security by the barbarian invaders contributed to contract markets and to focus the economy on large agricultural exploitations, evolving more and more towards self-sufficiency. These domains became the more important source of revenues for the state as the imperial bureaucracy went into decay. The enserfment of peasants thus was advantageous both to landowners and to the state (high "cohesion" of purpose).

In the Carolingian empire, the system of personal bonds (or franchise) between the overlord and the vassals for the supply of military services also relied on extraction of rents by the feudal landowners which provided the basis of in-kind financing of the empire. A similar organization prevailed in the Russia of the 16th and 17th centuries, an almost exclusively agrarian economy with few markets and other activities (artisanal and urban) in the "fringe".

The Tsar's power relied on "servitors" (feudal lords and landowners) who supplied military service in exchange for the use of some of the Tsar's property in land. The tsar was obviously willing to guaranty conditions that would allow his "servitors" to extract rents from peasants if he was to continue his conquest strategy of new lands in the east.

In all these cases a strong cohesion of interests between the ruling authority and the landowners (the land cum tax franchise) allowed the legislator to impose and enforce serfdom. As a consequence serfdom rise and fall paralleled the rise and fall of self-sufficient large agrarian producers.

9. The "relevant economic space" is defined by the political unit area enclosing a number of agricultural exploitation units. Their number depends on their average optimal size. Given a size distribution of agricultural exploitations, a smaller political unit is enclosing fewer agricultural units and is thus characterized by a high land ownership concentration. A large political unit, inversely, is usually characterized by a lower concentration of economic units, because there are more of them. Landowners' market power is thus presumably high in small political entities, and low in larger ones, on purely economic structure grounds. However, where the size distribution of agricultural ownership and productive units is more concentrated, even a large political unit could wield high market power.

   In the agrarian environment of pure manorial economies (before the rise of monarchical states), small political units prevail. It follows that the concentration of land is high in such economies (for a given "exogenous" or economically based size distribution of property) relative to that of large economies, where trade and specialization are more developed, and thus, for reasons explained below, political entities are larger. Small, self-sufficient agrarian economies are thus high monopsony power areas, and rely on serfdom, while larger exchange economies are characterized by lower monopsony power, and thus use free labor.

10. Oligopsony and the size of political entities.

    The political unit size depends in turn on the type of taxes levied: a monetary economy in which taxes are mostly levied on transactions implies a large (centralized) political unit, both because successive taxation of the same goods along a trade route would discourage trade and contract government revenues, thus favoring political integration of adjacent territories (David Friedman, "A Theory of the Size and Shape of Nations"), and because taxes on trade can be paid in money and thus easily centralized by a large authority. Specialization and trade thus give a

> competitive advantage to larger political entities while in-kind taxes – the only ones obtainable in a self-sufficient agrarian economy - cannot be used far away from the production site, due to the traditional technology high transport costs of bulky agricultural goods. Moreover, they do not raise the problem of successive taxation by neighboring political authorities that plague small political units levying taxes on trade. They clearly are not a handicap for small political units.

Monopsony, however, is seldom observed in the real world, outside a pure domanial and autarchic economy. An oligopsony is much more likely, with a small or larger (and variable) number of buyers among which the peasants have a potential employment choice.

A more realistic analysis like this allows us to understand better which conditions can lead to labor exploitation. The case of oligopsony also allows a role for the political authority. In an oligopsony, several buyers of labor have to collude in order to extract monopsony rents from labor by forming a cartel. But collusion, through the creation of a cartel, is difficult (Mancur Olson, *The Logic of Collective Action*, 1965). Usually a government intervention is warranted to organize and sustain a cartel.

3.2. The dominant firm with a fringe.

In an oligopsonistic industry, member firms' interest is to collude in order to create a cartel, suppress overbidding for labor among them, and achieve a price-quantity solution as close as possible to the monopsony case. And, as explained by Stanley L. Engerman, this requires the cooperation of some political authority:

"Both slavery and serfdom rested upon individual power in the control of labour and persons, but it required direct or supportive control by the state or other governmental apparatus or, at the least, a successful cartel among the owners to limit runaways and to prevent the raiding of other owners to acquire labourers. There was a need to have some legal authority limiting intra-elite (actual and potential) disputes that could benefit subject workers (such as that which occurred in England after the Black Death), with sufficient judicial and/or military controls to have these laws enforced. Perhaps at an earlier time continued control could have been achieved by combinations of large or small lords in a military standoff, but it was essential that there was no direct free bidding by lords to attract the serfs away from other lords. Under slavery, it was required that there be no negotiations directly between owners and slaves, all transfers among owners were to be made only by transactions in the market, with payments to previous owners and not to the slaves."

(in “Slavery, serfdom and other forms of coerced labour: similarities and differences”, chapter two in M.L. Bush (ed.) *Serfdom & Slavery: Studies in Legal Bondage*, London and New York: Longman, 1996, p. 21).

And also:

“Serfdom in Europe has been regarded either as coercive in its origin or else as initially the result of an implicit contract for military support from a lord, support that continued even after changes in relative bargaining strength, and would later become unnecessary and/or unfeasible with the rise of larger central states. The state and the nobles may have come into conflict in their attempts to control labour and the overall population, and the shifting power balance might have meant, depending on the particular circumstances, that serfdom was either terminated or reintroduced.” (Stanley L. Engerman, op. cit. p. 20).

The wedge between the productivity of labor and the wage paid is then chosen to maximize aggregate rents of cartel members. Not all firms however will participate in the cartel: it is likely that a fringe of competitive buyers not included in the cartel will remain, because the transaction cost of an agreement is rising very quickly with the number of participants. These fringe firms will accept the price that a dominant firm (or a cartel of major firms) pays as the market determined price.

Accordingly a cartel can be analyzed in the same way as a dominant firm with a fringe of competitive buyers (Carlton-Perloff, *Modern Industrial Organization*, IVth edition, p. 147). In such a “cartelopsony” the fringe firms which accept the dominant firm price will buy the labor input up to the point where their demand equals the cartel price. The dominant buyer’s problem is to adjust its purchases to maximize profit subject to the behavior of the other buyers.

Note however that the potential rent varies with the land/labor ratio if land is the only other factor of production. But an increase in the land/labor ratio has different effects on the potential rent whether it is due to an increase of land or a decrease of labor available (See Appendix B).

An increase in land availability, with constant returns to scale, determines an increase in the number of exploitations, and thus a decrease in Sh, an increase in the fringe producers’ numbers, and thus an increase in the price elasticity of labor supply to the cartel (see below the analysis of the dominant firm model with a fringe), and thus also a decrease in rents. Indeed the shift to the right of the demand for labor, confronted to a non perfectly elastic supply, should lead to a wage increase. The wage increase, however, should in turn provide a political incentive for an increased coercion of labor on the part of a central political authority, under certain conditions. The outcome then is ambiguous.

While a decrease in labor induces a leftward shift of the supply of labor, and thus an increase in the price elasticity of labor supply (Appendix B), and thus a decrease in land rents, the increase of wages, again, should lead political powerful landowners to coerce labor by forbidding mobility.

There is thus an ambiguity: increasing land/labor should increase competition for labor, increase wages and decrease rents but at the same time this evolution should stimulate a landowners' reaction and tentative immobilization of labor under certain conditions. Accordingly, relative land abundance and labor scarcity are compatible historically both with labor freedom in post Black Death Europe, and in the northern states of the U.S., and to the opposite condition of slavery in the southern Us states experience, and coercive serfdom during the Russian territorial expansion of the 16th and 17th centuries.

Other determining factors are clearly needed which exert some influence on Sh and not only on the equilibrium competitive wage and the elasticity of the supply of labor. These are (a) the concentration of other production factors ownership, (b) the availability of many employment opportunities outside subsistence agriculture or competition between potential employers, and (c) the presence of a large state which could rely on several efficient taxes and would have little use for an in kind food tax, and thus would have no incentive to organize and enforce a landowners' cartel.

Indeed a cartel seldom includes the whole list of existing producers if only because of the transaction costs of an agreement among them, which rises exponentially with the number of participants and their eventually diverging interests. It follows that the more realistic model is that of a dominant buyer (or « incomplete cartel ») that has some market power, but is surrounded and followed by other "firms" that behave as price-takers.

The market power index « BPI » (or « Buying Power Index »), and the monopsonistic rent, give a measure of the extortion intensity. In this case BPI is (Blair & Harrison):

$$BPI = S / [\varepsilon + \eta (1 - S)] \qquad (9)$$

Where S is the share of the cartel (or dominant firm) in the production of the sector or economy,

ε is the price-elasticity of the overall labor supply,

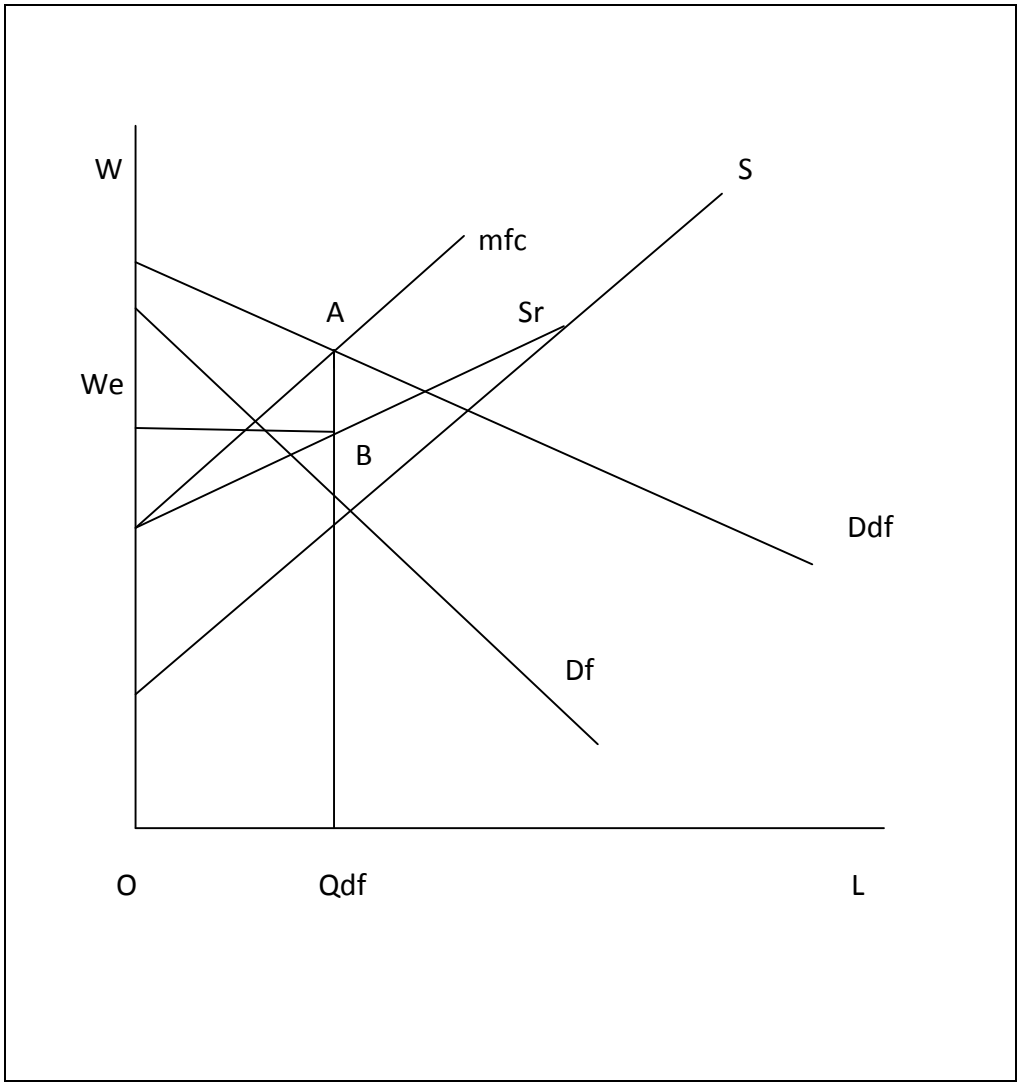
and η is the price elasticity of the « fringe » firms' demand for labor.

The dominant firm (or the cartel) tries to maximize BPI, but of course a higher BPI induces entry in the medium or long term and reduces the rent to a quasi rent.

The resulting equilibrium is shown in Figure 2 (adapted from Blair and Harrison, ch. 3).

Figure 1

The dominant buyer equilibrium



Df represents the demand for labor of the competitive fringe firms.

Ddf is the demand for labor of the dominant firm.

S is the supply curve of labor to the whole market.

The dominant firm recognizes that the fringe firms will demand some labor at each wage that it will fix. Thus it incorporates their behavior by subtracting Df from S, to obtain the residual supply (to the dominant firm) Sr.

Then, as is standard in monopsony analysis, a curve marginal to Sr, labeled here mfc, represents the marginal cost of labor to the dominant buyer. It is the marginal factor cost, mfc.

The dominant firm (or cartel) determines the equilibrium by demanding the quantity Qdf where the mfc just equals its demand for labor Ddf (at point A). The corresponding wage is determined on the residual supply curve of labor for the dominant firm, for that quantity: We.

For that wage the demand for labor of the fringe firms is determined on Df.

The total of the labor demands of the cartel and of the fringe is determined on the total labor supply curve, S.

The rent of the dominant firm is Qdf times the difference between We and the productivity of labor, AQdf:

$$\text{Rent} = (AB) \times (OQdf) \qquad (10)$$

We can then derive the "buying power index" (BPI) which measures the magnitude of the rent as a percentage of the wage paid by a dominant firm (or cartel) facing a fringe of competitive buyers.

$$BPI = Sh/ [\varepsilon + \eta(1 - Sh)] \qquad (11)$$

The BPI of landowners depends on the overall elasticity of labor supply S (that is ε) and the elasticity of demand for labor of the fringe firms Df (that is η) as well as the dominant buyer's market share (Sh). (See the Appendix 1 for the derivation).

Note that Sh is a concentration of landownership measure, and is <u>not</u> unambiguously related to the land/labor ratio. A high land/labor ratio could give rise to a high as well as a low concentration of land ownership. It could thus make collusion and cartel building either difficult or easy.

That's why it is doubtful if the land/labor ratio can, by itself, explain the existence of large potential labor rents, an economic condition for cartel building and serfdom systems.

Indeed, BPI depends on Sh, ε, η, i.e. on the structure of the economy, that is, on landownership concentration, which impacts the share of the cartelized firms among the

total population of firms, and on the elasticity of supply of labor , ε, as well as the elasticity of demand for labor by the fringe firms, η, firms that do not belong in the cartel.

3.3. Monopsony power and the structure of the economy

Now if we interpret the dominant firm as a cartel, the breadth of the cartel, its share of the sector's production, Sh, which depends on the collusion capacity of the cartel members, directly and positively affects the rents extracted from labor. This dimension is the result of the work of the "political" factor mentioned by Domar: developing membership in and enforcing cohesion of a cartel is a typical political activity requiring some coercive capacity.

We must thus consider Sh as a variable, because the logic of collective action (Olson, 1965, 1982) implies that reaching a cartel agreement will be difficult, absent some political constraint and central authority intervention on individual producers. Indeed, a centralized political authority helps police the behavior of cartel members and it thus makes the cartel larger and more effective. The cartel agreement will also be easier to reach when the number of producers in the sector is smaller.

The overall labor supply elasticity (and thus the residual labor supply elasticity) will be lower (and thus BPI higher) when the mobility and rights of labor are restricted as discussed above. The residual labor supply elasticity is also larger when the fringe firms' demand for labor is higher. An increase in the fringe firms' number increases their demand for labor and will increase the residual supply of labor elasticity, as well as it will decrease Sh. A decrease in central political authority will weaken the cartel and Sh. An increase of the environment instability (the risks of moving, repeated invasions, piracy and banditry and so on) and a decrease of the fringe activities, notably urban ones, will thus decrease the residual labor supply elasticity, making for a higher BPI. It will also be the case when employment opportunities outside the agricultural sector (in industrial and urban productions for instance) decrease, thus lowering the fringe firms' (here industrial and urban) demand for labor.

Indeed, the fringe firms can be understood in a narrow sense as other agricultural producers (landowners) not included in the cartel, or, in a broader sense, as all other producers in various activities susceptible to use labor. A more complex and diversified economy, and especially an urban economy, will determine a larger fringe demand for labor and a more elastic residual supply to the dominant firm.

These factors, taken together, determine the rent level that can be extracted from labor and the market power of landowners. And since a higher potential rent enhances the profitability of cartelization, it will determine actions to make it effective, for instance through legal rules tying labor to the land, as was the case in the Russian “second serfdom”. As far as a joint political action is required to enforce the cartel, a public choice analysis is necessary to analyze its occurrence.

In other terms there is a potential rent determined by basic economic conditions, and there is a supply of cartel management which is a function of the potential rent, on the one hand, and, in part, of the political equilibrium (the centralization of political authority for instance), on the other hand. The supply of cartel management is not arbitrary because many interest groups demand interventions by the State, and the latter’s capacity for redistribution is limited. A choice must be made in the efforts of the State to distribute favors to various groups, and an equilibrium in the political influence and political transfers between these groups results (see Becker, 1983). Where state revenues come mainly from landowners’ exploitation of labor, the government is more willing to immobilize labor in order to maximize rent extraction.

Given the political structure defined in these terms, the parameters of BPI (equation 11) define the probability and intensity of a serfdom regime:

a. A small numbers of landowners, or more precisely a high concentration of land and agricultural exploitations, making collusion easy. A high degree of collusion will make governmental intervention more likely and more successful.
b. An increasing and inelastic supply of labor, which depends in part on a low occupational mobility of labor, either due to the general environment or to specific and restrictive legislation.
c. But also shifts of the supply of labor (due for instance to population dynamics) displacing the labor market equilibrium towards the low price-elasticity section of the demand for labor.
d. And a high demand for labor (VMPPL) which results, among other factors, from a high land/labor ratio.

To sum up: an inelastic labor supply, an inelastic and limited demand for labor from the fringe activities, and a small number of landowners owning a large fraction of the agricultural land, make for a large potential rent and determine as well as political legislation restricting the mobility of labor, thus a typical serfdom status.

Conclusion of section 3: The three factors of serfdom theory

In the general case developed above, that of a dominant cartel, there are three basic conditions for serfdom to exist:

A. The conditions of production allow a labor surplus to exist,
B. There are few markets and employment opportunities outside the local domain (the fringe is limited or absent). The landowner has monopsony power.
C. The landowners also supply protection and military services to the overlord, and benefit in return from the overlord enforcement of their cartel vis-à-vis labor: mobility is forbidden.

Serfdom prevails, ceteris paribus, where the economic producers are self-sufficient, land abundant relative to labor, trade limited, and political entities small.

4. Historical puzzles explained.

The contraction of monetary markets, where they existed beforehand, is a general outcome of an unstable and chaotic environment (invasions, war, piracy). It leads to de-specialization, a lowering of the productivity, and population scarcity. Population scarcity in turn does not allow specialization nor urbanization (as in the cases of the late Roman Empire where invasions and pestilence reduced significantly the populations, and of the early Middle Ages. See Derville, 2002, and Bloch, 1947).

Market power of the buyer-landowner is the condition for extracting rents from labor. When the landowner is also the provider of military and civil protection – because a local castle and a fragmented polity is the best protection strategy against decentralized menaces and anarchy in an environment such as that of Europe in the feudal age (7th or 9th to 11th centuries, but Moses Finley sees it as prevalent as soon as the 6th century) – labor is not really free to migrate because of the general insecurity of outer empty spaces after the breakdown of the Carolingian empire and the rise of Viking, Saracen, and Hungarian invasions.

In such a fragmented economic and political space the conditions for a “natural” monopsony are met: serfdom has not to be “forced” on labor because there is only one demander of labor in the local “domain economy”, the local lord, and the peasants have no real alternative employment opportunity. To make serfdom legally compulsory is then not necessary, and many authors have emphasized the “voluntary” submission of peasants to the serf status, more akin under these circumstances to a voluntary (implicit) contract.

Given the subsistence agrarian economy in which peasants produce almost exclusively food products, the rent levied by the lord for financing its production of protection is simply a share of the crop, plus some “corvée” labor for artisanal production and services to the castle. Both constitute an in-kind tax. Serfdom is then “spontaneous”.

When population increases and security sets in, trade increases, the optimal tax shifts to a trade tax paid in money because there are many more monetary transactions, the tax revenues can be centralized, the states grow, the agricultural industry concentration within a single political entity thus decreases, and with it the monopsony power of landowners. Serfdom as a result of monopsonistic exploitation and as a tax raising device disappears.

We can now return to Domar’s puzzles. Why would a same increase in the land/labor ratio lead to opposite results: freedom for peasants in the post Black Death Europe and in the Northern States of the USA, or on the contrary to the “second serfdom” in XVIth century Russia, and in the late Roman empire?

I. Freedom : The North American case.

The North American colonies were characterized by a few features that precluded serfdom:

A) Many small agricultural exploitations instead of a few large landowners,

B) Many commercial interests and productions linked to procurement for the Caribbean economies in food, industrial products, wood, fish, boats, all small scale artisanal activities.

C) A competitive opening of western lands after the French selling of Louisiane, in a democratic and relatively egalitarian process for the acquisition of land. There were also many cattle raisers instead of a few labor intensive monoculture domains.

D) And contrary to Russia, no class of warriors-landowners existed who had to be compensated with land in the absence of other tax revenues.

There were accordingly many competitive demanders of labor, and a cartel was too difficult to organize.

One can find in Findlay and O’Rourke [pp. 234-238, and especially in their table 5.3 presenting data on the exports of the US regions in 1768-72 (in pounds sterling)] many data about the structure of the North American colonies.

While the West Indies produced almost exclusively sugar (3.186.450 pounds out of total exports of 3.910.600) and almost all shipped to England, New England exported Fish/Whale for 214.258 out of a total of 439.101, 89.953 of Livestock/meat, and 65.271 of Wood.

".. rapid population growth and a diversified economy led to substantial urbanization and the emergence of a trading and financial sector based on Boston, with its own merchant elite, which was the entry point for British imports from where they were distributed to the other mainland colonies and the West Indies. Earnings from shipping services were almost as much as total commodity exports, at least £ 427.000 (McCusker and Menard, 1991, p. 110) out of a total of 600.000 for the thirteen colonies (Findlay, p. 237), and insurance services and commissions should be added for about 220.000 for the whole and probably mostly from New Egland, the most urbanized and financiarized of the colonies. This fact leads McCusker and Menard (p. 92) to the observation that "the New England became the Dutch of England's empire.""(Findlay and O'Rourke, p. 235).

"The other early mainland colonies to be developed were the Upper South of Virginia and Maryland, which soon emerged together as a classic monoculture export economy growing the "noxious weed" tobacco as a highly lucrative cash crop around Cheeasapeake Bay." (Findlay … ibid.). Indeed it was the dominant export of the region : £ 756.128 out of a total export of 1.046.883.

The Lower South exported chiefly rice (305. 533) and Indigo (111.864) to England out of a total export of 551.949 , while the Middle Colonies, more diversified, nevertheless exported mostly Grains (379.380 out of a total of 526.545).

To summarize: early in the development of the US, the northern colonies followed a more diversified, urbanized, small scale, industrial and service economy path than the southern ones, which went the classic way of monoculture economies, implying large agricultural plantations using a large number of slave peasants.

"The pattern of development made the economy of New England complementary with that of the West Indies, exporting corn, codfish, and other supplies to sustain both the free and slave populations of the islands." (Findlay, op cit p. 234).

Thus the North had a diversified, mostly small scale fishing, industrial and service economy where the South had large scale monoculture exploitations. And where agriculture was concerned, the dominant structure of the North was that of the small scale farm, of the high labor/land variety.

As Domar wrote:

".. so long as agricultural skills can be easily acquired, the amount of capital for starting a farm is small, and the per capita income is relatively high (because of the ample supply of land), a good worker should be able to save or borrow and start on his own in time. Most of

the farms will be more or less family-size, with an estate using hired labor (or tenants) here and there in areas of unusually good (in fertility and/or in location) land, or specializing in activities requiring higher-than-average capital intensity, or skillful management. But until land becomes rather scarce, and/or the amount of capital required to start a farm relatively large, it is unlikely that a large class of landowners, such as required by the Muscovite government, could be supported by economic forces alone. The American North in the Colonial period and in the nineteenth century would be a good example of an agricultural structure of this type." (in Capitalism, Socialism and Serfdom, p. 227). (my emphasis).

The opposite case: the American Southern States

The Southern economic strategy points to the importance of the fringe for the viability of a forced labor system. In their book "A deplorable scarcity", Bateman and Weiss (1981) show that the southern planters were indeed consciously reluctant to industrialize and invest in industrial ventures because they feared that it would weaken the slave economy. In fact it would have increased the competition for labor and the employment opportunities for slaves by increasing the industrial "fringe", de-concentrating the production, thus lowering the capacity for collusion of planters and landowners.

A highly cohesive political authority will also increase the barriers to mobility of labor in order to enhance rents where supply elasticity is relatively high, thus decreasing it. It can also increase barriers to entry of fringe activities, especially industrial and urban. It is because of the density of market exchanges and the diversity of productions and employment opportunities that "the air of cities makes free".

II. Freedom: Europe after the Black Death

First, the inward shift of labor during the Black Death led to an increased elasticity and a decrease of the land rents, and this can explain the non reestablishment of serfdom, which was already a disappearing institution during the demographic expansion (and decreasing wage levels) of the 11th- 12th centuries.

Other factors however are at play: the increase in the size of the political authority with the new rise of state kingdoms, led to a decrease of landownership concentration and cartel power of landowners in the relevant eco-political space. Simultaneously the revival of security, and thus trade, led to the re-specialization of production, urbanization, and commercial exchanges and facilitated the reestablishment of taxes in money, which in turn

made the contribution of feudal lords less necessary to the royal authority and even contrary to it. Conflicts between the local lords and the kings were detrimental to the enforcement of serfdom and labor immobility, the Russian way. Moreover, labor mobility benefitted the efficiency of the economy, increasing trade and thus taxes on trade, which incited kings to erode the serfdom legal status.

"Sh" consequently declined as new activities developed, while ε increased too and the rents on labor decreased, all factors leading the political authority to reduce support for serfdom and even to weaken the feudal lords.

III. Serfdom: The case of the Late Roman Empire

In the late Roman empire, the economic and political structure evolved in the opposite direction: a population decrease (thus an increase in the land/labor ratio), reduced specialization and trade, shrinking the tax base and ability of the central government to finance the army, and led to a return to taxation in kind which led in turn to a fractionalization of the military, the tying of labor to land serving both to extract rents on the available economic base, and to finance the local military defense.

The vanishing of markets also made slavery unprofitable and the new trend was of a decrease of slavery, slaves being converted into serfs. This goes hand in hand with a decline of the central political authority, thus of the capability to collude and enforce slavery. First barbaric kingdoms instead of Roman empire, then the even more fractionalized system of feudalism (based on large agricultural and rather autarchic domains) relied almost exclusively on an in-kind labor tax serving to finance local defense, which was provided by the local land owners. The local power, defensive, of the castle and knight was the only one capable of protecting a small number of local peasants and to oppress them when necessary (see Fenoaltea's comment on North and Thomas, see also the Olson analysis of the "local bandits" who are both the saviors and the curse of the peasants. See Hopkins, Moses Finley, on the tying of peasants to land in the late Roman Empire, and its reasons).

The commercial economy of the late empire (see Temin's argument that it was a market economy and that the labor market was essentially free, despite slavery) was progressively strangled by repeated barbarian invasions, which increased insecurity and led to a decline of trade and a return to self-sufficient agricultural domains. Taxes in money could not be raised anymore as the tax base was shrinking by the barbarian annexation of whole regions, and the central authority decided to tie labor to land in order to have some in kind taxes collected by the large farmers-owners who were transformed into tax farmers.

The political authority was thus decentralized and serfdom became the rule as large landowners also provided some protection and justice services. The polar case is reached after Charlemagne, when practically no central authority survived.

IV. Serfdom : 16th century Russia

In the period of Russian eastward imperial expansion during the 16th and 17th centuries, the economy was almost entirely agrarian and produced exclusively for subsistence. The Tsar could not consequently collect any tax in money and had to rely on « servitors » for constituting his army. He paid them with grants of some of his lands and of conquered new territories, on a provisional basis. Given the increasing availability of land with new conquest, the incentive for peasants to flee the land they cultivated for the servitors increased, due to competition between servitors for labor which was increasingly scarce relative to land. The Tsar had a direct interest in the success of his « franchisees » in extracting rents from the peasants if he was to keep an active military for his conquest.

He acted in his own interest and in the interest of his servitors by enserfing the peasants, to make rent extraction from labor easier (the serf status determined a reduction of the price elasticity of labor supply to a locally monopsonistic landowner or servitor, and thus an increase in BPI). This was all the more necessary that with the conquest, the number of servitors had to be increased, and a cartel of landowners became more and more difficult to enforce. The intervention of the higher level central political authority was thus warranted to suppress landowners competition for labor.

In both the Roman case and the Russian case one can consider that it is the weakness or absence of the “competitive fringe” that made the in-kind tax on labor optimal. In the late Roman Empire it was a progressive weakening of commercial activity by insecurity due to the invasions that made enserfment and extraction of the rural labor surplus necessary. In the Russian expansion it was the overwhelmingly rural character of the economy and the absence of large cities and diversified production that made it compulsory if the territorial expansion was to be pursued.

On the contrary, it is the important development of commercial and urban activities, and the dispersion of land ownership that made taxes on trade advantageous in the North American colonies and precluded the use of serfdom. Labor mobility was too essential to a rapidly changing commercial economy to immobilize labor.

Conclusion: Four cases, three factors.

According to our theory of serfdom,

Sh = f (high concentration of property, small fringe size and markets, Political cohesion among landowners and central political author, if any).

Note that the small fringe size and markets mean that the economy is basically of the agrarian, self-sufficient type (with the possible exception of the Soviet Union, even though it was mostly agrarian in the 1930s, and without many markets). Note also that the political cohesion among landowners is due in part to their small number (high concentration of property) and to their supply of military services to the higher political authority when there is one. In the case of the feudal pure manorial economy the landowner is also the political local authority and the cohesion is of course maximal.

To the three puzzling cases raised by Domar, -- Russia in the 16th and 17th centuries, the American North, and Western Europe after the Black Plague --, we add the case of the late Roman empire which abandoned free farming and slavery for serfdom, we find that three explanatory variables are at play:

A) the number of landowners, or the concentration of land ownership measured by the share of the production controlled by the land cartel, Sh, and the associated price elasticity of residual supply of labor (in the dominant firm oligopsony model), ε. It could be high or low.

B) The employment opportunities for labor in the "fringe" sector, which includes urban activities, non-cartel farm production, industrial and commercial activities. These opportunities could be high or low.

C) Complementarity or competition between large landowners of the cartel and the superior state authority, where existing. Complementarity (or cohesiveness) would be signaled by a > 0 sign while competition would be given a < 0 sign.

The four cases are presented below:

Table 1.

Three determinants of serfdom

The four puzzles explained

	Concentration of ownership A	**Fringe size** B	**Political interest cohesion** C
1. Russia	High	Low	> 0
2. Late Roman Empire	High	Decreasing	>0
2. American North	Low	High	< 0
3. Western Europe (post Black Death)	Low/High	High	< 0

It is easy to see that only case 1 (Russia) and 2 (late Roman empire) meet all three conditions for serfdom, while cases 3 (American North) had all signs going against serfdom, and case 4 (Western Europe in the 14th and 15th centuries) had at least two out of three signs going against serfdom.

The land/labor ratio of the Kliuchevskii/Domar theory impacting only indirectly the first factor (A) could not determine by itself the labor status. Indeed a high land/labor ratio could be compatible with either a high concentration ratio of land (Russia) or a low concentration of land (American North).

Thus the "puzzles" resulted from both an inadequate and incomplete theory.

5. Slavery and serfdom are mutually exclusive.

In the introduction we noted that usually slavery and serfdom were considered as differing in degree but essentially responding to similar determinants and constituting a same phenomenon. The above analysis, on the contrary, leads to a vision of slavery and serfdom as mutually exclusive phenomena. In the case of the late Roman Empire it is evident that slavery disappeared and was replaced, as well as free farming, by serfdom. But no serfdom existed in the American Southern states, and in the Europe of the early Middle Ages, what was left of slavery of the Roman heritage disappeared too when serfdom spread.

This is due to the specific conditions of slavery as defined by Moses Finley: a high productivity of labor due to important market outlet for the agricultural production, a local scarcity of labor, and a possibility to obtain foreign labor.

Indeed, Athens, then Rome had large commercial exchanges of agricultural and mining products while Sparta used the labor of helots only for its own consumption.

The first condition assumes that there exists a large market economy for the products of slave labor, whether agricultural of mineral. But precisely our theory of serfdom assumes that there are no markets in the economy and that the agricultural production is mainly for consumption internal to the domain. Where slavery thrives serfdom should thus disappear.

And the third condition also differentiates a slave economy from an economy relying on serfdom. In a slave economy, labor is constantly imported from foreign wars, or capture, in an outer anarchical space. Stolen labor (even when it is paid for by the ultimate user) is the basis of the slave economy. For instance slavery developed in Rome with external military conquests, and modern slavery relied on superior European armament used in Africa against low tech military equipments.

Slavery relies on the mobility of the workforce, allocated by the price mechanism. On the contrary the serf economy relies mainly on local labor and there is no market for labor, at least no such market independently of the exchange of land ownership. While serfdom is a device used to forbid competition for labor, slavery is based on a competitive bidding for labor by slave owners. Serfdom is a tax device in a self-sufficient economy, while slavery is a private ownership right (however asymmetric) in a market economy. Slavery has no component of a tax to pay for public goods.

Serfdom relies on local labor that is refrained from leaving while slavery relies on imported foreign labor, and usually no local labor (for a variety of reasons).

As Robin Blackburn puts it ("Slave exploitation and the elementary structures of enslavement", Chapter 9 in M.L. Bush (Ed.) Serfdom & Slavery. Studies in legal bondage, Longman. London and New York 1996.):

"The slave plantations of the New World like the slave-worked latifundia of Antiquity, were mostly situated in climatic zones where work on secondary crops could proceed in winter; (…) When access to markets was disrupted, then the planter could, as Gorender (1982) puts it, 'withdraw into the shell of natural economy' by devoting his slave crew to self-supply. The slave plantation was a hybrid entity oriented to the market exchange but capable of surviving for a time without it. But if markets for slave produce disappeared for good then the slave plantation would not be reproduced, since new slaves and supplies could not be purchased. The value of the plantation would become a dead weight. Slaves might either be sold to those better placed or converted in serf-like peasants. The element of "natural economy" on the slave plantation boosted its capacity to survive economic difficulties but, in

the long-run, its existence was premised on fairly extensive market exchanges. The decline of slavery in the later Roman Empire was certainly linked to the contraction of trade, just as the rise of colonial slavery on the Americas was associated with the Atlantic commercial boom “, (p. 177).

The conclusion is that slavery and serfdom are not a same phenomenon differing only in degrees of the constraints on labor, but are quite different phenomena pertaining to quite different economies and societies: serfdom is peculiarly adapted to conditions of a self sufficient agriculture in a non market economy, while slavery is developing in a market economy.

While both legal statuses include harsh constraints imposed on labor, the more severe conditions of slavery, compared to those of serfdom, are a result of very different economic conditions:

In slavery, the initial theft of human beings and initial capture by force make for much harsher conditions.

Given an abundant external supply of slave forced labor, the price elasticity of labor supply will be lower, allowing a more intense exploitation, also facilitated by the low cost of replacing slaves, while, on the contrary, in serfdom labor is scarce, costly, difficult to replace and should thus be economized and not exploited to the limit.

In slavery slaves can be worked to death (see the demographic trends in the Caribbean plantations). Then ε is small and the degree of exploitation (BPI) is high. It means that the “remuneration” of labor is low, near the subsistence level or even lower (Nazi extermination camps, Gulag, etc.) and the use of force and severe punishment is high.

This condition for slavery obtains when the war industry is powerful, the military competitive advantage strong (Europeans in Africa from the 16th century on), and political power centralized.

A market economy is needed for the exchange of slaves between warriors and ultimate slave owners (see the case of the German industry during WWII and the use of prisoners in the extermination camps). It is also possible since there is a strong centralized political authority.

On the contrary in serfdom the political authority is local and not differentiated from the landowners. No serf market is possible or necessary.

Having contrasted freedom and serfdom we conclude with a parallel, but symmetric comparison between serfdom and slavery. The similarity of the conditions for serfdom with the conditions of democracy and individual freedoms is striking, only with the signs reversed.

According to Olson (Power and Prosperity) the conditions for democracy (which imply individual rights, and thus "freedoms") are: (a) dispersion of power and thus of resource ownership, and (b) absence of extortion (or theft).

In our analysis we have noted that the concentration of land ownership and the monopsonistic power allowing an extraction of the labor rents are conditions essentials to serfdom.

CONCLUSION

Domar was right to show that a rising land/labor ratio jeopardized the landowners' rents, but that government intervention to restrict labor mobility could reestablish them.

He did not explain, though, the conditions under which the "political factor" would intervene. We suggest that it is not discretionary but depends on the structure of the economy and the correlative size of the political unit.

The structure of the polity is linked to the structure of the economy and both determine the desirability of serfdom for the owners of land and the producers of protection.

In a modern economy the employment opportunities are much too numerous and too varied (a) for substantial labor market monopsony to prevail, and (b) for it to be advantageous to raise taxes through rent extraction from labor, and/or delegate defense tasks to large landowners. Taxes on transactions, or later on income, are much more efficient in an exchange, monetized economy, and the military power in large nations and diversified economies is both centralized and professionalized (even where military conscription is used).

Serfdom then is a labor-fiscal status characteristic of the transition (in both directions) from autarchic agrarian economies to partly urbanized and diversified exchange economies. As such it is most unlikely to occur again.

References

Batchelder, R.W., Freudenberger, H., 1983. On the Rational Origins of the Modern Centralized State, Explorations in Economic History, 20, 1-13.

Bateman, F., Weiss, T., 1981. A Deplorable Scarcity: The Failure of Industrialization in the Slave Economy, Chapel Hill, NC: The University of North Carolina Press.

Becker, G.S., 1983. A Theory of Competition Among Pressure Groups for Political Influence, The Quarterly Journal of Economics, 98, 3, 371-400.

Blackburn, R., 1996. Slave exploitation and the elementary structure of enslavement, Chapter 9 in M.L. Bush (Ed.) Serfdom & Slavery: Studies in legal bondage, London and New York: Longman.

Blair, R.D., Harrison, J.L., 1993. Monopsony: Antitrust Law and Economics, Princeton: Princeton University Press.

Bloch, M., 1947. Comment et pourquoi finit l'esclavage antique, Annales: Histoire, économie, société, 1, 33-44.

Bush, M.L., (Ed.), 1996. Serfdom & Slavery, Studies in Legal Bondage, London and New York: Longman.

Carlton, D.W., Perloff, J.M., 2005. Modern Industrial Organization, IVth edition, Boston: Pearson, Addison-Wesley.

Derville, A., 2002. L'économie française au moyen-âge, Paris : Ophrys.

Domar, E.D., 1970. The causes of slavery or serfdom: a hypothesis, The Journal of Economic History, March: 18-32. Also reprinted in Domar, E., Capitalism, Socialism, and Serfdom, Cambridge: Cambridge University Press.

Engerman, S.L., 1996. Slavery, serfdom and other forms of coerced labour: similarities and differences, chapter 2 in Bush, M.L., (Ed.), Serfdom & Slavery: Studies in Legal Bondage, London and New York: Longman.

Fenoaltea, S., 1975. The Rise and Fall of a Theoretical Model: The Manorial System. The Journal of Economic History, 35, No. 2, 386-409.

Filson, D., Fruits, E., Borcherding, T., 2001. Market Power and Cartel Formation: Theory and Empirical Test, The Journal of Law and Economics, Vol. 44, No. 2, 465-480.

Findlay, R., O'Rourke, K.H., 2007. Power and Plenty: Trade, War, and the World Economy in the Second Millennium, Princeton: Princeton University Press.

Finley, M.I., 1980. Ancient Slavery and Modern Ideology, London: Chatto and Windus.

Friedman, D., 1977. A Theory of the Size and Shape of Nations, The Journal of Political Economy, 85, No. 1, 59-77.

Geary, P. J., 1988. Before France and Germany: The Creation and Transformation of the Merovingian World, Oxford: Oxford University Press.

Gorender, J., 1982. O Escravismo Colonial, Sao Paulo.

Hopkins, K., 1980. Taxes and Trade in the Roman Empire (200 B.C.- A.D. 400), The Journal of Roman Studies, Vol. 70, 101-125.

Kliuchevskii, V., 1937, 1906. Kurs russkoï istorii [A course of Russian history](Moscow: Gosudarstvenne sotsial'no-economischeskoe izadatel'stvo), translated as A History of Russia (New York: Russell and Russell, 1960) by C.J. Hogarth.

Krugman, P., 2003. Serfs up! The Unofficial Paul Krugman Home Page, May 8.

Lagerlöf, N.-P., 2009. Slavery and Other Property Rights, The Review of Economic Studies, 76, 319-342.

Lagerlöf, N.-P., 2002. The Roads To and From Serfdom, Working Paper, Concordia University.

Lindsay, C.M., 1984. Applied Price Theory, New York: Dryden Press.

McCusker, J.J., Menard, R.R., 1991. The Economy of British America 1607-1789, Chapel Hill, NC: University of North Carolina Press.

North, D.C., Thomas, R.P., 1971. The Rise and Fall of the Manorial System: A Theoretical Model, The Journal of Economic History, Vol. 31, No. 4, 777-803.

Olson, M. Jr., 1965. The Logic of Collective Action, Cambridge: Harvard University Press.

Olson, M. Jr., 2000. Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships, New York: Basic Books.

Temin, P., 2001. The Labor Supply of the Early Roman Empire, Working Paper 01-45, Department of Economics, MIT.

Waterson, M., 1984. Economic Theory of the Industry, Cambridge: Cambridge University Press.

APPENDIX A. The simple monopsony model.

Let's assume for simplicity that free labor is as productive as serf labor (if not it would simply introduce a difference in the calculus of the rent without changing the reasoning).

Let us define the production of some agricultural product Q as:

$$Q = Q(L, M) \qquad (2)$$

where L is labor and M machinery.

The supply of labor is:

$$W = W(L) \qquad (3)$$

The profit of the firm is:

$$\Pi = P.Q - W(L).L - p.M \qquad (4)$$

where P is the output price, and p the machinery price. The amount M is fixed in the short run.

To maximize profit let:

$$d\Pi / dL = P.\, dQ/dL - [\, W(L) + L.\, dW(L)/dL \,] = 0 \qquad (5)$$

The first term is the value of the marginal productivity of labor, VMP_L, and the bracketed term is the marginal factor cost of labor, MFC_L .

The value of L that equates the VMP_L and the MFC_L defines the maximization of profits and thus the equilibrium in the monopsonistic industry.

The measure of the buying power, the unit rent or Buying Power Index (BPI) is:

$$BPI = (VMPP_L - W) / W \qquad (6)$$

Since at the defined equilibrium the $VMPP_L$ is equal to the MFC_L , we can substitute the later for the former in the BPI formula:

$$BPI = [W + L.dW(L)/dL - W] / W = (L / W). dW(L)/dL \qquad (7)$$

The term on the right is just the reciprocal of the price elasticity of the supply of labor.

Thus:

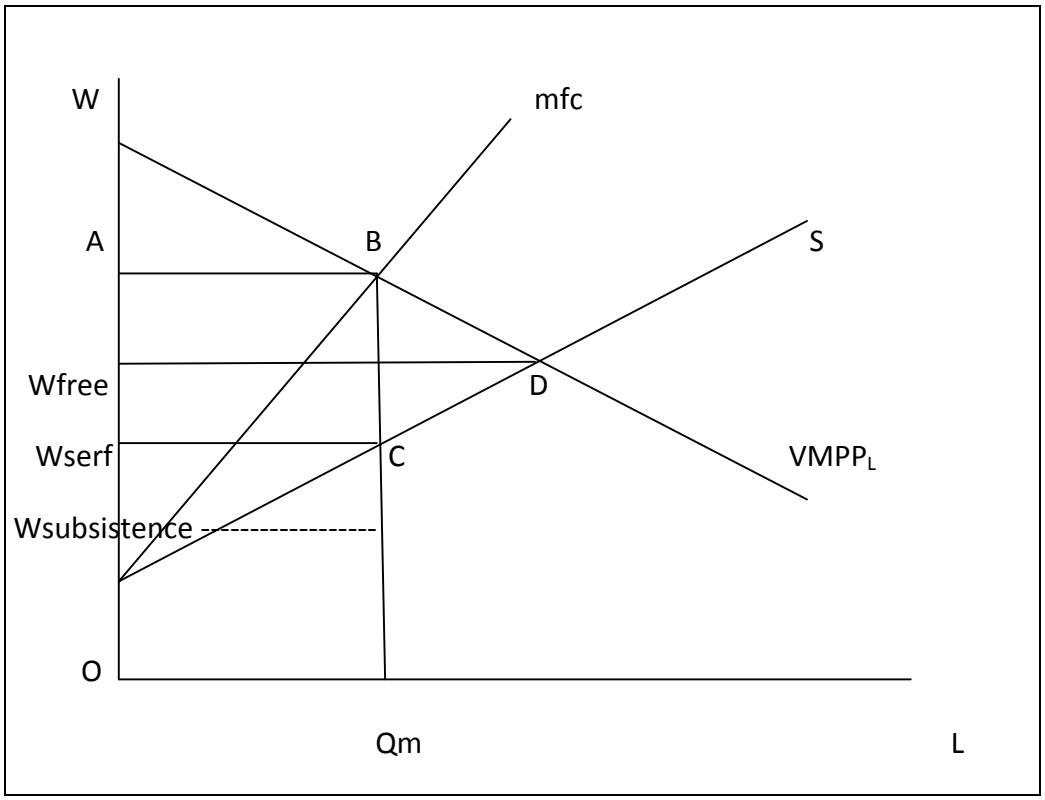
$$BPI = 1 / \varepsilon \qquad (8)$$

The market power of the employer is the reciprocal of the price elasticity of labor supply. In sectors where there is only one employer, that elasticity defines the rent. It is also the well known definition of an optimal Ramsey tax.

As noted above, that elasticity is larger when peasants have more diverse employment opportunities, which depend on the geographical environment, the density of economic activity and the general mobility of labor both functionally and geographically. The monopsonistic buyer of labor can try to reduce this elasticity or ask the state for help. The efficiency of the state's intervention is measured by the decrease of ε that it is able to produce.

Figure 2

The simple monopsony case



The rent from serfdom is ABCWserf, and the serf remuneration Wserf is less than the competitive market wage, Wfree, and even less than the marginal value productivity of labor, OA.

In such a simple model the determinant of the rent extracted from labor is just the price elasticity of the supply of labor given that the economy is structured around a single landowner. No political intervention is necessary because the sole landowner has market power over labor.

Could a change in the land/labor ratio change the BPI as Kliuchevskii and Domar thought? Assuming a fixed quantity of land, a change of the ratio would result from an increase or a decrease in the labor supply, a shift in the schedule S. A lowered land/labor would result from an outward shift of S, and a rise of the ratio from an inward shift of S. An outward shift of S would lower ε while an inward shift would raise ε on the supply curve. Thus a decrease of the land/labor ratio would increase the BPI of the landowners, more numerous laborers competing for employment, while an increase of the ratio would decrease the BPI of the

landowners, in conformity with the implicit Domar's model, landowners competing for labor. But at the same time, as already noted, a decrease of the land/labor ratio (an outward shift of S) would lower the equilibrium wage, thus the rent that could be extracted from labor, and thus the incentive to enserf labor. The result is indeterminate.

The same line of reasoning applies to a shift of the demand for labor, $VMPP_L$ and with the same conclusion.

APPENDIX B. The dominant firm (or cartel) and fringe equilibrium

Let's take the simple case of a dominant firm, which has some market power, and thus determines its equilibrium in the same way than the simple monopsony of the previous section. There are also, however, an unspecified number of smaller other firms, "fringe" firms that initially adopt a competitive behavior (Blair and Harrison, ch. 3).

The total supply of labor to the industry is equal to the sum of uses of the dominant and of the fringe firms:

$$L = L_D + L_F \qquad (1)$$

The amount of labor absorbed by the fringe firms, at each value of W, has to be deducted from the total amount of labor supplied at each wage level to obtain the supply curve of labor to the dominant firm. The dominant firm is thus confronted to a "residual supply schedule" of labor, and its derivative the marginal residual supply curve of labor. It will determine its equilibrium at the intersection of the marginal residual curve of labor supply and the value marginal product of labor, and from this point determine the wage level on the residual marginal supply curve of labor.

We want to define the price elasticity of the residual supply curve of labor and its relation with the labor market structure, specifically with the price elasticity of the total supply of labor, the price elasticity of the demand for labor of the fringe firms, and the share of the dominant firm in the sector's production.

From (8) we get the changes when the wage rate varies:

$$dL_D / dW = dL / dW - dL_F / dW \qquad (2)$$

Multiplying both sides by W/L_D we get:

$$(L_D/dW).(W/L_D) = (dL/dW).(W/L_D) - (dL_F/dW).(W/L_D) \quad (3)$$

The left member is the wage elasticity of the dominant firm's supply of labor (the "residual" supply), ε_D . So:

$$\varepsilon_D = (dL/dW).(W/L).(L/L_D) - (dL_F/dW).(W/L_F).(L_F/L_D) \quad (4)$$

Or:

$$\varepsilon_D = \varepsilon.(L/L_D) + \eta_F.(L_F/L_D) \quad (5)$$

Where ε is the wage elasticity of labor supply to the industry, and η_F is the wage elasticity of the <u>demand</u> for labor of fringe firms (since it is a negative term the minus sign in (11) becomes a positive sign in (12)).

L/L_D and L_F/L_D are respectively $1/S_D$, the share of the dominant's firm employment in the total employment of the industry, and $(1 - S_D)/S_D$.

Then,

$$\varepsilon_D = \varepsilon/S_D + \eta_F.(1 - S_D)/S_D \quad (6)$$

And since, as we saw earlier, $BPI = 1/\varepsilon_D$, then:

$$BPI = S_D/[\varepsilon + \eta_F(1 - S_D)] \quad (7)$$

It follows that the ability of the dominant firm to extract rents from labor increases when its share of the market is high. That means that other "fringe" firms are few (if of a standard dimension). The fewer the firms in the industry, the more concentrated it is, the higher th

e market power.

The rents are smaller when the wage elasticity of labor supply of the industry is higher, as well as the wage elasticity of the demand of labor of the fringe firms.

In an industry where the number of firms is high, S_D is small and $(1 – S_D)$ is large, monopsony power is small, and it is important for the firms to extract rents, that the authorities reduce η_F , the elasticity of demand for labor of other firms with which it is difficult to collude because of their number. In such a case an overall state legislation assigning to each producer, however small a fixed number of peasants (the amount that they happen to use already, for instance), is going to minimize the denominator of (7,) and thus bring the oligopsonistic case towards the simple monopsonistic one, maximizing BPI.

The market power of employers then depends both on the market supply elasticity of labor (which depends on opportunities of employment in other sectors) and on the state intervention to immobilize the labor used by each producer in the industry.

Industries with a small number of employers and few outside opportunities of employment in other sectors will have more market power. And if there are many small employers in the sector, a state intervention to assign labor to each producer, thus “freezing” the labor market in the industry, is going to be necessary to increase market power to its purely economic maximum.

The conclusion is that the state intervention does not create all the market power. It helps bring it in large numbers industries to its economic maximum value. The state constraint is a substitute for costly or impracticable collusion. It is a low cost solution for policing a potential cartel.

Indeed the non free status of labor is a kind of club good. It is difficult or impossible to impose it partially. It has to be general to succeed.

APPENDIX C. The comparative statics of labor monopsony

(Demand and supply shifts and the price elasticity of labor supply).

Figure 3

Price elasticities:

The comparative statics