Defining preindustrial. We need to circumscribe the scope of preindustrial. At some level, it is easy: preindustrial economies are characterized by low urbanization rates, high share of agriculture in GDP, low literacy rates, and of course low overall GDP per capita. However, many of today’s poor countries share precisely these features. They are however “non-industrial” or “non-industrialized” rather than “preindustrial” economies: this is because they are part of the modern world, systematically included in trade and voluntary movements of factors of production (“globalization”) and have social structures which are very different from those of preindustrial societies. The life expectancy of their populations as well as the immunization and school enrollment rates exceed manifold those of “true” preindustrial societies. Not the least important is the fact that political compulsion of slave or serf labor, so ubiquitous in all preindustrial societies, is—with the exception of a few pockets—largely absent. So, our definition of preindustrial includes all societies prior to the Industrial revolution; and those that have not engaged into the industrial revolution then, only up to a point when they began to be integrated systematically, rather than episodically, into the world economy. For many of them, this coincides with the period of colonization. Thus, broadly speaking—since we are painting with a very broad brush here—we can set limits around the end of the Napoleonic wars for Western Europe and the United States, and end of the 19th century for everybody else. Twentieth century societies, even when poor and hardly industrialized, belong to a different category.

---

1 World Bank, Research Department and University of Maryland, College Park. I am grateful to Mihail Arandarenko, Leandro Prados de la Escosura and Jeffrey Williamson for excellent comments.

2 1815-1820 is a convenient date for at least three reasons. Politically, it coincides with a “rearrangement” of Europe and, as later emerged, the world. It marks the beginning of the “long 19th century”. Economically, it marks, according to the new English wage data series produced by Clark (2005), the beginning of a long-run rise in real wages which is continuing to this day. In terms of history of economic thought, Ricardo’s Principles were published in 1817.
An obvious, but nevertheless important, clarification is that we are concerned here with *income* inequality, *viz.* inequality that includes all sources of income and reflects differences in households’ and individuals’ living standards. This, for example, rules out wage or rural-urban inequalities as such.  

**Implicit theory.** We do have an implicit theory about income inequality in preindustrial economies. The Kuznets hypothesis (formulated in 1955), the bread-and-butter of inequality economics, posits that inequality charts an inverted U shape as economy transforms from predominantly agricultural to predominantly industrialized or modern. The same hypothesis, albeit without the mechanism which generates the inverted U-shaped curve, was formulated 120 years before Kuznets by Tocqueville. From both we should retain that inequality is supposed to emerge only when societies are richer, and thus inequality in preindustrial societies may be expected to be low. But differently, we also have an image of preindustrial societies as combining abject poverty in the bottom with extravagant wealth on the top.  

Could both these images be right? As we shall argue below, yes—and this is one of the key features which distinguishes inequality in pre-modern from inequality in modern times.

But in order to speak about inequality in preindustrial societies, we must also assume that preindustrial societies were “modern” in the sense that they were (predominantly) market-oriented economies with non-negligible monetized sectors—and when non-monetized, goods and services given or received for political or power reasons could be valued at some meaningful “market” prices. This is a position not universally accepted.

---

3 Wage inequality has meaning only if calculated across all earners; income inequality includes the entire population.

4 In Kuznets own words: “One might thus assume a long swing in the inequality characterizing the secular income structure: widening in the early phases of economic growth when the transition from the preindustrial civilization was most rapid, becoming stabilized for a while; and then narrowing in the later phases” (p. 276).

5 "If one looks closely at what has happened to the world since the beginning of society, it is easy to see that equality is prevalent only at the historical poles of civilization. Savages are equal because they are equally weak and ignorant. Very civilized men can all become equal because they all have at their disposal similar means of attaining comfort and happiness. Between these two extremes is found inequality of condition, wealth, knowledge—the power of the few, the poverty, ignorance, and weakness of all the rest." (1997, pp.42-3).

6 For Rome, Goldsmith (1984, p. 287) notes extraordinarily high income of the monarch relative to Great Britain in the early 19th century.
In a famous debate about the later Roman Empire (and, by extension about all ancient economies) and “modernity”, there were two camps: that of “primitivists” led by Polanyi (1944), Finley (1985) and Schiavone (2000) and that of “modernists” (Rostovtzeff 1957 [1926], Walbank 1946). The first believed that Rome lacked most of the modern concepts that we associate with market economy. Market relations, even when present, were of peripheral importance, and market economy, itself a recent phenomenon, is perhaps, in a historical sense, only a brief episode (Polanyi 1944). For the “modernists”, the links between a preindustrial society like Rome and modern capitalism were obvious. Both Rostovtzeff and Walbank write of Roman “bourgeoisie”.

Whatever our opinion about the respective merits of “primitivists” and “modernists” it is important to realize that once we attempt to make some tentative estimates of economic inequality in preindustrial societies, we ipso facto accept that, while preindustrial societies might have been poorer and with different social structure compared to the modern societies, the differences are of magnitude not of a kind. For if such key concepts of market economy as prices, wage-labor and private property are vague, insufficiently understood by population, not sanctioned by custom or law, then applying modern economic categories may be meaningless. Every attempt to empirically study preindustrial societies using today economist’s tools, must assume that “ancient” and “modern” are fundamentally the same—so that the “ancient” can be described and understood using economic concepts developed from Adam Smith onwards.

Where do data for preindustrial inequality come from? Since World War II, empirical studies of income distribution are based on household surveys (nationally representative samples of households who are anonymously interviewed about their household characteristics, spending patterns, and income). The earliest household surveys are from late 18th century England. There were a few and sporadic surveys in 19th century (continental Europe, rural Russia) but they spread broadly after the end of the Second

---

7 Private property must enter this list with a caveat. No one would deny that socialist societies, where private property was limited, were not modern. Moreover, they regarded themselves as the epitome of modernity. Similarly, societies with largely communal ownership of land (as in Africa) are modern too. Thus, private property of the means of production seems to be less of a requirement for a modern society than e.g., monetization. Rawls (1971), who can hardly be seen as non-modernist, allows in his Theory of Justice both private and non-private ownership of the means of production (see pp. 54, 240-1)
World War, and, as far as Africa and China are concerned, surveys became available only more recently, from the early 1980s. Obviously, such surveys were not conducted in any preindustrial society—even if censuses (driven by governments' tax needs) were.

However, to gauge income distribution in preindustrial societies there are relatively abundant sources that economists can use although the sources are often buried in hard-to-access archives and books, written in not widely known languages and alphabets,\(^8\) requiring large amounts of both money and effort to be brought to “light” in a usable form. And then lots of heroic assumptions to be “translated” into modern economic categories. This has severely limited the use of ancient sources, and this is probably why only a fraction of such sources has been used so far.

The most comprehensive contemporary sources are tax data and government censuses undertaken in order to supply governments with information about taxation and war-waging capacity of its populace (number of men, houses, horses, grain). Early documentary evidence includes government edicts (e.g., Diocletian’s edict on maximum prices and wages from 301)\(^9\), as well as numerous Roman papyri preserved in the dry climate of Egypt. English Domesday survey of 1086 is perhaps the best known of such sources. From the Byzantine Empire, we have a few preserved praktika that provide descriptions of household characteristics, inventories of possessions and taxes paid although they cover only limited areas (towns, ecclesiastical communities).\(^10\) Ottoman censuses (defterlar) from approximately 14\(^{th}\) century onward, conducted to assess wealth and military capacity of the newly-conquered territories, provide detailed information on settlements (hamlets, villages, small and larger towns) but then present it in average amounts for each settlement (not per individual household). If inequality within settlements is not huge, and the number of settlements included large, censuses can be used to assess overall income distribution within a country or a region. A much used

---

\(^8\) Ottoman censuses are written in Turkish but using Arabic (rather than as today Latin) script. To process them, requires knowledge of an often archaic Turkish and an alphabet into which this language is no longer written. See Cosgel (2002, 2004).

\(^9\) Recently studied by Allen (2007),

source is the Florentine *Catasto* from 1427.\textsuperscript{11} Spanish *Ensenada Cadastre*, similar to modern-day household surveys, was carried out in the 1750s for the purposes of a never-implemented fiscal reform. It has recently been used by researchers, and will be no doubt analyzed more once it is digitalized. Inequalities for the cities of Paris, Amsterdam and London were studied from tax data for respectively 1292-1313, 1732-42 and 1797-1801.\textsuperscript{12} However, they refer to wealth inequality (there is no attempted “conversion” to income), cover very truncated data sets, focused either on the rich—those subject to taxation—or the poor (McCants, 2007), and of course include single cities only.\textsuperscript{13} The data on Latin America, produced by various Spanish *Visitas* which collected detailed information on population, age, land-ownership, and agricultural output have been published in numerous volumes but not used for estimates of income distribution.\textsuperscript{14}

What is common to these sources is that they are in principle surveys of stocks (people and wealth) and require a huge effort of price imputation; first, to “transform” a stock into a meaningful annual yield (income), then to convert produced quantities, expressed in local “natural” units (e.g., Egyptian *modii* of wheat), into kilograms, and finally to convert all of these into monetary units. Then, the researcher needs to resort to even more heroic assumptions to calculate other sources of income, from husbandry, vineyards, honeybee cultivation, fruits and plants, services provided by farmers, and not least importantly, from manufacturing activities like pottery, glass or cloth-making, or provision of urban services from the shoe-maker to the teacher (for which, at least some wage data are generally available). Particularly vexing is the issue of measurement units, volumes or weights with often confusingly similar or identical names, which nevertheless imply different physical amounts from one region to another; or when money-units are provided, the issue of silver or gold conversion between them. But such sources, however

\textsuperscript{11} The data were originally collated by Herlihy and Klapisch-Zuber (1985). Currently, they are available on the Internet.
\textsuperscript{12} Sussman (2005) for Paris; Schwarz (1979) for London, McCants (2007) and Soltow (1989) for Amsterdam.
\textsuperscript{13} Incidentally, all examples but one used by Pareto in the formulation of his famous “iron law” of income distribution come from various European tax data from the end of the 19th century (see Pareto (1997)).
\textsuperscript{14} For Peru, books with detailed notes from *Visitas* for the years 1562, 1567 and 1604-05 have been published.
frustrating, can and do provide very useful evidence about ancient living standards and
distribution of income.

The second contemporary evidence is provided by social tables. This is what William
Petty termed “political arithmetick”. They aim to describe the structure of a society by
listing all salient social classes (or professions) and estimating their average (per
household or, less often, per person) average incomes. For modern economists, these
sources are much easier to use because the classification into presumably socially
important groupings and estimates of their money-equivalent incomes provide us with
most of what we need to know for the derivation of income distribution. England was the
pioneer in the production of social tables, beginning with the famous one of Gregory
King for 1688 (which contains 33 social groups with their population sizes and average
incomes), and continuing with Massie (1759) and Colquhoun (1801-3). Much more
recent authors produced similar social tables for a number of countries. These new social
tables are of course not contemporary sources but they were produced, using bits of
dispersed primary or most often secondary sources, by economic historians who
specialize in various eras and countries, and they represent our best guess of social
structure at remote points in time. The work of Milanovic, Lindert and Williamson, 2009
(hereafter MLW), who do the first systematic attempt to measure and analyze
preindustrial inequality, is largely based on such (contemporary and recent) social tables.

**Empirical evidence.** To translate preindustrial inequality into modern economics, we
must not only hold that preindustrial societies were largely monetized (and whatever was
not monetized could be ultimately expressed in money), but also hold that their
inequality can be meaningfully handled by Gini, Theil or any other currently used
inequality measure. Otherwise we lack the same yardstick with which to compare past
and present.

15 None of the social tables, nor the results obtained from them, is without its critics: for a critique of King’s
social table, see Arkell (2006); that of Colquhoun, see Schwarz (1979); for a critique of Lindert-Williamson
use of English social tables, see Feinstein (1988).
Using mostly social tables from 30 preindustrial societies, MLW calculate Gini coefficients. They find that the preindustrial Ginis range from mid-20 to around 65 with the mean of 45 and the standard deviation of 11.\(^{16}\) This is almost the same as the range of Ginis in modern societies. In fact, the modern equivalents of the preindustrial societies included in MLW sample (e.g., Turkey for Byzantium, Syria for the Levant, today’s United Kingdom for the 1688 England and Wales etc.) have an average inequality of 40 Gini points with a standard deviation of 10. However to make such simple comparison and to leave it at that would be erroneous. Preindustrial and modern societies were very different, even when compared in the language of modern economics.

First, it is very likely that the income gradient (how income increases as we move from poorer to richer income classes) was much flatter in preindustrial that in modern economies.\(^{17}\) Using Jan Pen’s (1971) metaphor of dwarfs and giants, where people are visualized as marching in a sixty-minute parade, from the poorest to the richest, with everyone’s height reflecting his income, preindustrial societies can be seen as societies of dwarfs who would take some 40 to 45 minutes to file past. They contained large groups of people (most of the time, vast majority of the population) living at, or just above, the subsistence minimum. Percentage differences in income among this vast mass of people were small. Income gradient was flat up to a very high point in income distribution. But then and quickly, as we approach the very end of the parade, the gradient would suddenly increase, much more so than in modern societies. Thus unlike a modern parade that would be characterized by a steady increase of the gradient, in preindustrial societies, the middle was not much different from the bottom. There was dearth of people whom we would (using modern terminology) identify with the middle class. (It is worth pointing out that this “middle class” is not defined in terms of absolute income, or what we would consider today to be middle class requirements, but entirely in terms of the period average income.) We can thus see why both of our preconceived notions—of

\(^{16}\) Gini is the most commonly used measure of inequality that ranges from a theoretical zero (everybody has the same income) to a theoretical maximum of 100 (everybody but one person has a zero income, and the richest person takes the entire income of the community).

\(^{17}\) See MLW (2009).
generalized equality and drastic income disparity among the ancient—are true: they just refer to different parts of income distribution.

This difference in structure implies that the same calculated measures of inequality have different meanings. Ginis, as we have already indicated, were broadly in the same range then and now. But Gini of 40, estimated independently for the Roman Empire by MLW (2009) and Scheidel (2009), had an altogether different meaning than the same Gini in the contemporary United States. Roman Empire’s mean income was about twice the physiological subsistence ($s$). If we require that all members of a society have at least the subsistence minimum—for otherwise the society will tend to shrink and disappear—then a very low level of mean income, regardless of how tiny the upper class, limits the extent of measured inequality. Simply put, the extent of inequality is limited by the size of average income. That ceiling is more binding when a society is poor. To realize this, assume that society’s mean income is just a fraction above $s$. If all but a tiny elite live merely on $s$, the elite cannot be extravagantly rich because total income is low, and Gini or Theil indexes, which take into account incomes differences between all individuals, cannot be very high either. This is the idea underlying Inequality Possibility Frontier (see Figure 1) defined by MLW (2009) and Milanovic (2006).

The frontier gives maximum Gini (or Theil) which is compatible with a given level of mean income and maintenance of society as a going concern. Maximum Gini is equal to $\frac{(\alpha-1)}{\alpha}$ where $\alpha$=mean income divided by $s$, or the number of subsistence minima contained in the mean. As can be seen from the formula, maximum feasible Gini rises in mean income ($\alpha$) but at a decreasing rate. If average income is twice the subsistence ($\alpha=2$), the maximum Gini will be 50. Thus, we see that the Roman inequality of 40 exhausted some 80 percent of maximum feasible inequality. But for the modern-day United States, where the mean income stands at more than one hundred $s$, the maximum Gini is 99. The actual inequality will have exhausted only 40 percent of its maximum.

---

18 MLW estimate refers to the year 14 (at the death of Octavian), Scheidel’s estimate to mid-second century.
value. Hence, the social meaning of the same Gini is entirely different. To sustain high inequality, societies must be relatively rich.

**FIGURE HERE**

We have left the issue of defining what subsistence minimum is deliberately vague. Depending on whether we pitch this physiological (note: not social, not relative) minimum higher or lower, the IPF will move down or up, but the same logic will hold.

The difference in the income structure (income gradient) also shows why some other measures, like top-to-bottom ratio or top 1% share, may not be very useful in the preindustrial context. They show the extent of the gap between the richest and the poorest, but they disregard the entire distribution in-between, which in the past has been much more equal than in today’s societies.

**IPF** imposes a consistency check on our inequality calculations, a fact which is particularly useful for preindustrial societies where the evidence is scant. As illustrated in the figure, once we know mean income of a society, and estimate its Gini, we know that this estimate must be within, and at most at, the frontier. If it is not, there is something wrong with either income or inequality estimate, or such a society is doomed to a dwindling population and ultimately extinction. It is not surprising that MLW find that all six cases of ancient societies with inequalities close to the frontier were colonies: India in 1750 and 1947, Kenya in 1914 and 1927, Nueva España (Mexico) in 1790, and Maghreb in 1880. Colonizers were clearly much less concerned about the welfare of the populations they ruled, or did not have to fear them as much as native rulers.

**Preindustrial inequality and modern debates.** Empirical evidence on preindustrial inequality has direct bearing on several contemporary debates. Evidence from the two most advanced economies at the time (England and Holland) paints a picture of increasing inequality from 16th century to the beginning of the Napoleonic wars. 19

---

19 The exception is Soltow (1968) who finds English inequality to have been flat throughout 18th century.
Premodern growth seem to have exacerbated inequality even in the areas that were characterized by an already high inequality of wealth and income (South Midlands in England; Allen 1992). Using social tables, Lindert (2000) and Lindert and Williamson (1982, 1983) document increase of inequality in England between 1750 and 1801. All four observations available for England and Wales in the MLW database (1290, 1688, 1759 and 1801-3) show both mean income and inequality rising with time. Similarly, van Zanden (1995), and Soltow and van Zanden (1998) find that income inequality increased in Holland during its “Golden age”: between 1561 and 1732, the urban area Gini rose from 53 to 59, rural area Gini from 35 to 38. According to a pioneering study by Hoffman, Jacks, Levin and Lindert (2002), “real” European inequality between 1500 and early 19th century increased even more because prices of wage-goods, consumed by the poor, rose relatively to prices of “luxuries”.

The upswing of the Kuznets curve seems to be strongly in evidence in all these cases. But what drove it? Was it a “classical explanation” (as van Zanden 1995 terms it), namely a shift in the functional distribution of income toward property owners (and their rising concentration) and away from labor—a mechanism that Marx would have recognized easily? Or was it, as argued by Lindert and Williamson (1985) and Williamson (1982, 1985), due to the “wage-stretching” which continued well into the 19th century and involved labor-saving technological progress and increased pay-ratios for the skilled labor in the presence of strong, and mostly unskilled, population pressure. Education responded only very slowly, and the process continued for a couple of centuries until massive European emigration reversed it. The latter is a very neoclassical mechanism familiar to every economist, working on poor or rich countries today. The focus is on the functioning of factor markets not on the division of society into capitalists and workers.

If countries where the Industrial revolution originated went through a period of sustained increase in inequality prior to the Industrial revolution, does it shed some light

---

20 From Campbell (2007).
21 For Spain, Prados de la Escosura (2008) uses functional distribution of income, and also finds a clear Kuznets upswing from 1850 to around 1914.
on the relationship between higher inequality and the take-off? A number of recent writings (most famously, Pomerantz 2000, Frank 1998, and more recently Wen 2009, Shuie and Keller, 2008) have contrasted China and Western Europe in 17th and 18th century, trying to understand why these two large areas that seemed in many respects similar (e.g., market integration, level of income, technological innovations) charted in the following three centuries such different paths. Does income distribution have to do something with it? Unfortunately, we do not yet have even the intimation of an answer because the historical data for China are not available. However a recent upsurge in archival research on Chinese sources might help throw some new light on this issue.

The work of Engelman and Sokoloff (1997) has profoundly affected our conception of the role of inequality in explaining economic success of North America and relative decline of Latin America. But while there is little doubt that Latin America was more unequal (particularly in land ownership) that the North, recent historical evidence, contrasting Western Europe and Latin America, finds no perceptible difference in inequality between the two. Williamson (2009) thus wonders why Western Europe and Latin America have followed different growth trajectories? If the inequality explanation works for one set of regions (the two “New Worlds”), why does it seem not to work for another (Europe and Latin America)? Moreover, it is not evident that Latin America was “always” unequal. Prados de la Escosura (2007) and Bertola et al. (2009), argue that strong expansion of inequality occurred during the previous round of globalization (1870-1920). Prados de la Escosura (2007, p. 298) sees the explanation as consistent with factor-price equalization theorem: opening up Latin America to trade raised land rents, and since land was unequally distributed, increased concentration of incomes. The data prior to around 1870 are not available, but one may wonder whether our “acquired idea” of an always high inequality in Latin America may not be wrong, or perhaps that it was not inequality, but inequality extraction ratio that was high. Thus recasting the issue, the Latin American problem was low level of income rather than high Gini.

Although some estimates for 1870 show inequality in the Southern Cone countries to be at the same level as in Spain (Prados de la Escosura, 2008, Figure 8, p. 307).
The concluding point is to realize that studying inequality in its historical context, an area which will doubtlessly loom larger in economics, as the search for economic past progresses, is not only important because it helps us learn about history but because it helps us understand today’s economic problems. Actually, as every historian and politician knows, studying the past is about the future.
REFERENCES


Herlihy, David, and Christiane Klapisch-Zuber (1985), Tuscans and their Families. New Haven: Yale University Press. See also
The Online Catasto is a World Wide Web searchable database of tax information for the city of Florence in 1427-29. It is based on David Herlihy and Christiane Klapisch-Zuber, Principal Investigators, “Census and Property Survey of Florentine Dominions in the Province of Tuscany, 1427-1480.”


Economica, August-September, pp. 645-660. Originally published as "La courbe de la repartition de la richesse", Université de Lausanne, 1896. The same issue of Rivista di Politica Economica contains English translations of five other articles by Pareto on the same topic.


