Roman economic performance and inequality: in defense of the big picture

© **Walter Scheidel (Stanford University)**
March 2022

Abstract Macro-models make a significant contribution to our understanding of economic performance and economic inequality in the Roman empire. Comparative framing is essential for guiding our thinking on these matters by imposing some discipline on our conjectures. The application of Bayesian modeling improves on earlier estimates of Roman imperial GDP. The scope for estimating Roman income inequality is strongly constrained by GDP estimates and historical comparanda.

Challenges

How can we find out about economic inequality in ancient societies? The data are usually poor. Just how poor they are becomes painfully clear when we consider the scope of recent discoveries regarding late medieval and early modern Europe. A generation ago we knew very little about the distribution of wealth in those societies. Since then forays into countless archives have unearthed reams of relevant information, above all from fiscal records of private assets. Guido Alfani has played a major role in getting transnational research projects off the ground. This work has produced nothing short of a revolution in our understanding of pre-modern inequality, a revolution that is ongoing (Alfani 2021 is the most recent survey). Roman historians must ask themselves, what can *we* do?

Nothing of the sort. The statistics that once existed are all gone. The Roman citizen census, for example, was designed to produce them in abundance: Republican *tributum*, as a proportional levy on declared personal assets, would have been unthinkable without them. By and large, Egypt under Roman rule is the only partial exception: several tax registers that disaggregate certain kinds of landholdings have survived on papyrus (Bowman 1985; Bagnall 1992). But even those texts have their limits, by omitting the landless and also for other more specific reasons. They are certainly better than nothing, and the impression they convey of higher inequality in urban than in rural settings is consistent with what is known from much later history. But that's about it – they don't tell us much that we couldn't have guessed anyway. Then there are some land registers preserved on inscriptions from Italy, such as the tabula of Veleia, and from the Aegean and North Africa. They provide isolated snapshots but little context (Duncan-Jones 1990: 129-142).

Other sources offer us glimpses of income inequality. James Macksoud has found that the dispersion of wages in the Roman military of the Principate was somewhat greater than in ancient Chinese and the current US militaries, but not much so (Macksoud 2019). Seth Bernard has analyzed Roman-era skill premiums, which were substantial (roughly 100 percent) and thus broadly in line with those observed in other pre-modern societies (Bernard forthcoming). Once again, this is better than nothing but doesn't tell us about the urban self-employed or the vast masses of family farmers.

Can archaeology save us? The material record isn't particularly promising if we are interested in the actual distribution of income or wealth, but is arguably more useful for gauging, however impressionistically, differences between locales and change over time. Recent examples include Tim Kohler and Michael Smith's recent edited volume on 10,000 years of inequality derived from house sizes (Kohler and Smith 2018), and their multi-millennial comparison of house size dispersion in the Old and New Worlds (Kohler et al. 2017). Rob Stephan did something similar for the Roman world, measuring house size dispersion in several areas and persuasively tracking change over time in Britain (Stephan 2013). Miko Flohr calculated Gini coefficients for housing in Pompeii (Flohr 2016). Going a step further, established separate metrics for Pompeii before and after the earthquake of 62 CE: it turns out that Pompeii's housing stock was somewhat more unequally distributed in the latter phase (Gaggioli 2019). But while that is nice to know, this is not an exercise that can readily be repeated elsewhere.

Study of the distribution of variables such as body height and the presence and severity of dental and bone lesions might allow us to derive inequality metrics from the enormous body of osteological evidence from the Roman world. In practice, this tends to be quite a challenge. But at least there is a wealth of tangible material to be analyzed, which even if easier said than done is feasible in principle.

But what if we are interested in estimating the overall economic output of the Roman empire and how it was distributed across society? In her recent article in the new journal *Capitalism*,[1] Kim Bowes rejects pretty much all strategies currently on offer. Thus, macro models are bad, summarily dismissed as producing "thin history, stripped of nuance and complexity" (Bowes 2021b: 27), a charge that might of course be leveled against every broadly synthetic account ever written. Yet single-indicator studies are no good either (Bowes 2021b: 27), and "case studies render the lives of working people just as flat and one-sided as does the current bad data" (Bowes 2021b: 28).

So what are we supposed to do? According to Bowes, "the task, therefore, is to find better, thicker data, capable of telling the rich, complex stories of microhistories, but at a bigger scale that allows us to see welfare and inequality across time and space" (2021b: 28). Unfortunately, by her own admission, "the field is a long way from being able to do this" (2021b: 32). Yet if such a superior approach has not actually been employed and currently amounts to little more than a twinkle in the archaeologist's eye, what exactly enables Bowes to conclude of existing approaches that "they've probably totally misrepresented welfare and inequality as well as overall economic performance" (2021b: 32)? If the superior approach hasn't yet enlightened us, how would we know? Conversely, if that pessimistic assessment is a foregone conclusion, why do we need to worry about developing superior methods at all? But I do not wish to belabor this odd paradox. What matters most is whether such a superior approach will ever be feasible. For all I can tell that may very well not be the case.

If so, the most economical solution would be just to give up. The second-best approach might be to develop better indicator studies, because at least there is substantial material to work with. Richard Saller's critique of Andrew Wilson's biased reading of the Greenland ice core data (McConnell et al. 2018; Pavlyshyn et al. 2020) is a recent example of what such an approach might involve.

And, like it or not, the third-best course of action could be to revisit the merits and demerits of GDP-focused approaches (such as Scheidel and Friesen 2009; Scheidel 2020) – for multiple reasons. One is the usefulness of maintaining communications with the outside world – retreat into the splendid isolation of analyses that only work for ancient societies may be a noble gesture but makes students of antiquity even less relevant than we already are. Granted, the obvious counterargument would be to say that if that is just not possible we cannot and must not pretend to be relevant or able to communicate to academics in other fields in a shared language. Fair enough but too soon – for it is by no means obvious that such communications cannot legitimately be maintained.

To give just one example: it is all too easy to dismiss a concept such as GDP by noting that it was developed for and in the context of modern industrialized economies. But that is a non sequitur. True, GDP as we know it was largely the brainchild of the economist Simon Kuznets during the Great Depression. But the underlying idea that it is possible and desirable to aggregate and estimate a given society's total income and output is much older and predates modern industrial capitalism by a considerable margin. When pre-modern historians try their hand at tentative GDP estimates, they stand in a long tradition going back to William Petty in the 1660s and Gregory King in the 1690s, two Englishmen who pioneered estimates of

[1] I take this opportunity to address some inaccuracies in that paper. Bowes 2021b: 12 claims that "the Roman population headcount is hotly debated, with estimates differing by an order of magnitude". The first half is very true, the second one not at all – there are no estimates for any Roman site or group that differ "by an order of magnitude," i.e. 1,000 percent, or in fact by anywhere near that much. The time-honored use of "wheat-based subsistence" values for GDP estimates (Bowes 2021b: 29) has nothing to do with actual wheat consumption (only the share of GDP represented by wheat consumption does). I not only never claim in Scheidel 2019 that "the Roman economy experienced no real per capita growth," as Bowes 2021b: 24 avers, I explicitly state that it did (Scheidel 2019: 504, 507). I likewise never envision in Scheidel 2017 any "inevitability of a collapse that was produced by massive inequality" (Bowes 2021b: 25, and cf. again 26); instead, I spend several pages rejecting the very notion that inequality is bound to precipitate societal collapse (Scheidel 2017: 392-394).

cumulative income by class, creating what has become known as social tables. These tables, which allow an approximation of what we would now call GNI, are a much cruder instrument that contemporary GDP measures but also much better suited to pre-modern contexts, whether we are looking at seventeenth-century England or second-century Rome.

Another point in favor of such GDP estimates is that they are not merely helpful but well-nigh indispensable in reining in free-floating speculation. When I say 'speculation' I mean qualitative, verbal, non-quantifiable statements such as 'better,' 'more,' 'prosperous,' 'significant,' etc. Bowes seems to think that the one of the biggest problem with GDP approaches is that they make ancient or any pre-modern societies seem poor in comparison with modern ones (or at least that is how I read Bowes 2021b: 22-23). However, even though the risk of overdoing our pessimism is real, it is worth noting a different risk that has traditionally been much more common and serious – that of *over*estimating the accomplishments of early civilizations in terms of material wellbeing and human welfare.

For instance, I have read enough scholarship on ancient population history going back several centuries to know that the biggest problem has not been conservatism or pessimism but the readiness to assume that any number of things that may not have been possible in other early societies were possible for these elect cultures: that their members could do more, become more numerous, live longer, you name it – and for no other reason than that they were intrinsically special. This propensity, deeply rooted as it is in the veneration of the wonders of the so-called "classical world" of the ancient Greeks and Romans, has never fully gone away.

In that context, calls for restraint in assessing material wellbeing or economic performance are not an extravagance but a necessity. That is not the same as trying to make all pre-modern societies look alike – except to the extent that they were.

Roman GDP revisited

I shall spend the remainder of my paper explaining what I mean by "except." The dangers of circular reasoning certainly need to be taken seriously. Bowes notes that the metrics used to construct Roman GDP estimates are generally of poor quality and are hard to generalize from (Bowes 2021b: 12-13, 16-17). It is hard to disagree with that. However, the notion that in choosing a particular Pareto distribution for Roman wealth we are "tinkering with Pareto's math" (Bowes 2021b: 21) rests on a misunderstanding of this particular kind of power-law probability distribution. The shape of the distribution curve is not fixed but is a function of the shape parameter α, which may vary across cases. Contrary to Bowes' insinuation of special pleading, there is nothing sinister in taking account of this variability, as we did in our 2009 article (Scheidel and Friesen 2009: 79) – even if, with the benefit of hindsight, we should admittedly have made that clear(er).

She also holds that separate estimates of Roman imperial GDP on the output and consumption side converge because they are not actually independent (e.g., Bowes 2021b: 20), and that this in turn undermines estimates of inequality. In an update to the original model, published in 2020 (Scheidel 2020) and thus too late to be considered by Bowes, I tried to address the issue of circularity. I came up with an alternative distribution model for the Roman empire which (except for GDP size) is entirely determined by analogies to income/wealth ratios and inequality levels in several early modern societies. This alternative model yields an overall inequality estimate that is perfectly compatible with the results drawn from the original model even though they were derived from different premises.

In theory, that could have happened by chance: maybe our reconstruction for ancient Rome is wrong in a way that wrongly makes it seem comparable to, say, France around 1700. Or maybe all reconstructions are wrong in the same way, perhaps even the result of a vast conspiracy ranging from ancient to modern historians to obscure the true variety of historical outcomes.

Or these results could be similar simply because the societies in question were, in terms of fundamental inputs and outputs and social conditions, ultimately not that different from one another – not all the same, mind you, but not wildly different either. At the end of the day, if we cannot agree that agrarian

societies labored under meaningful constraints on output and wellbeing, then there isn't really much point in having a debate at all.

Then again, to insist that pre-modern economies were unlikely to be "wildly different" does not mean that they could not be significantly or noticeably different. In the absence of precise data, how could we hope to pin down a plausible range of variation?

This is a crucial question, and one that brings me to a problem that has barely been mentioned at all: which is that in existing models, individual variables, even if they are presented as ranges rather than single values, in practice tend to be approximated to the latter, to single values, if only for ease of calculation. That quickly becomes a serious issue whenever we combine several variables, none of which is well known. In the case of Roman GDP, these variables are population size and per capita output, and in the case of inequality, those two plus the distributional pattern. Joining roughly guesstimated values inevitably compounds margins of error.

This needs to be expressly acknowledged, and in as much as I can see any major weakness in my own earlier work on this topic it is that I neglected to do so. Fortunately there are ways of addressing this problem: kudos to Myles Lavan at St Andrews for his tireless efforts to alert his fellow ancient historians to this problem and to a possible solution – the application of Bayesian modeling, which has long been common in other fields but not, or not yet, so much in ancient studies (e.g., Lavan 2019, and now esp. Lavan et al. in press).

In the present case, how can Bayesian modeling help? Any GDP estimate rests on two variables, economic output per capita multiplied by population number. Even though in some contexts only the former might be of interest, total GDP also matters, for instance if we seek to relate economic performance to fiscal capacity or, most saliently here, for estimating the distribution of income across society: average per capita output or consumption does not tell us anything about inequality. For any estimate of overall inequality, we are compelled to draw on two exceedingly poorly known variables, which is of course exactly the kind of challenge Bayesian modeling is meant to address.

Modern estimates of the population of the Roman empire have been creeping up over time, from 46 million (Frier 2000: 812) or 54 million (Beloch 1886: 507) in 14 CE to 61 million (Frier 2000: 814) and closer to 70 million in the 160s CE (Scheidel 2007: 48). If the method for extrapolating urban population numbers from archaeological data developed by Hanson and Ortman 2017 holds water, we need to allow for the possibility of even larger totals. The latest estimate of 13.4 million residents (give or take 2 million or so) for cities of 5,000+ in the second century CE (Hanson in press) would translate to very high urbanization rates if applied to the lower of the standard population estimates for the empire as a whole: the implied (central) rates are 22 percent in an empire of 60 million and 19 percent in one of 70 million. If we wanted to lower the urbanization rate to a more conservative 15 percent, we would need to raise the overall population to almost 90 million. Given that it would seem marginally less hopeless to estimate urban numbers than rural numbers, following the lead of estimates of urban numbers might well be the most prudent course of action (at least unless or until the Hanson method can be shown to be flawed after all).

At the same time, any further increases of Roman population totals would create growing tensions with much better documented later periods. The territories once held by Rome did not re-attain a population of close to 90 million until 1700, at a time when France alone housed more than 20 million people, many more than any reasonable estimate for Roman Gaul. Even allowing for the possibility of nineteenth-century levels of population density in parts of the Roman Near East and North Africa, it would be difficult to pull the population of the Roman empire to much higher levels: for instance, when the former Roman territory hit 100 million in 1750, England and France had already reached 30 million.

Bayesian modeling requires three inputs: a lower and an upper bound, and a value the observer deems most likely. Based on this brief survey, I set these bounds at 60 million and 90 million, respectively, and retain my own previous estimate of 70 million as perhaps most likely. But I am also prepared to adjust the latter to a higher value to see what difference it makes.

Per capita GDP is the second component. The Scheidel-Friesen estimate from 2009 implies a mean of $1,250 in 2011 standardized dollars (Scheidel 2020: 346). How much lower could we plausibly go? The Roman empire was not a poorhouse: so let us be conservative and set the bottom at 90 percent, or $1,125.

How much higher we might go is a much trickier question. Upward adjustments are necessarily more elastic: while there may not be much room for less there is a lot of room for more, at least in theory.

This is where analogy is essential. I propose an upper limit of $1,700 (in 2011 standardized dollars). That would place Roman economic performance on a par with several early modern European economies, which seems a fairly generous but perhaps not impossible level of attainment (Maddison Project Database 2020, with Bolt and van Zanden 2020). In the seventeenth and in much of the eighteenth century, French per capita GDP moved in the $1,600-$1,700 range. Estimates for Britain in the sixteenth and the first half of the seventeenth century hover around $1,700, just as they do for Portugal around 1700 and Norway and Spain as recently as 1850. Pre-modern China is thought to have peaked at $1,550 around 1700.

For comparison, $1,700 equals mean per capita GDP in Japan and Poland in the 1870s, in Mexico in the 1890s, in Brazil in the 1940s, in Albania in the early 1950s, in Egypt and Romania in the early 1960s, in India in the mid-1980s, and in Vietnam in the early 1990s. By all accounts, as an average for the Roman world as a whole this value clearly sets a high bar. I deliberately choose it to dispel any notion that the game is somehow rigged in favor of a making Rome seem poorer than it might have been.

What of the most likely value? Examples for a mean per capita GDP of $1,250 include Indonesia around 1910, Brazil around 1920, Chile, India, Kenya and Pakistan in the mid-1960s, Laos in the early 1980s, Nepal in the late 1980s, Bangladesh in the mid-1990s, or Afghanistan in 2007. All of these were low-income countries, to be sure, but to varying degrees no longer untouched by modern development. Extreme poverty, by contrast, is associated with much lower output, such as current levels in Burundi and the Central African Republic at half the envisaged Roman level, in Liberia at two-thirds and in Niger at three-quarters.

My point is very simple: $1,250 is not a fantasy number designed to relegate Rome to the bottom of the heap, and $1,700 is a generous upper limit that not only places the Roman empire in the company of major western European states of the early modern period but even approximates Roman maxima with levels in various developing countries, several of which had already entered the early stages of industrialization. Anyone who wishes to maintain that these parameters fail to do justice to an ancient society that was profoundly agrarian and had no access to fossil fuels had better offer a compelling reason for that dissent: I return to this below.

For a scenario of 60 million people (minimum), 70 million (most likely) and 90 million (maximum), a mean per capita GDP of $1,125 (minimum), $1,250 (most likely) or $1,700 (maximum), and an urbanization rate for sites of 5,000+ residents of 15 percent (minimum), 17 percent (most likely) and 22 percent (maximum), a Monte Carlo simulation with 5,000 iterations yields a total GDP of $100 billion and a 95 percent confidence interval (which captures 19 of every 20 outcomes) ranging from $81 billion to $122 billion (Fig. 1).[2] The mean is about one-seventh higher than our 2009 estimate of $87 billion.

[2] I choose a triangle distribution (which centers the "most likely" version) for the first two categories and a uniform distribution for the third (on which I am more agnostic). My thanks to Myles Lavan for his invaluable help.

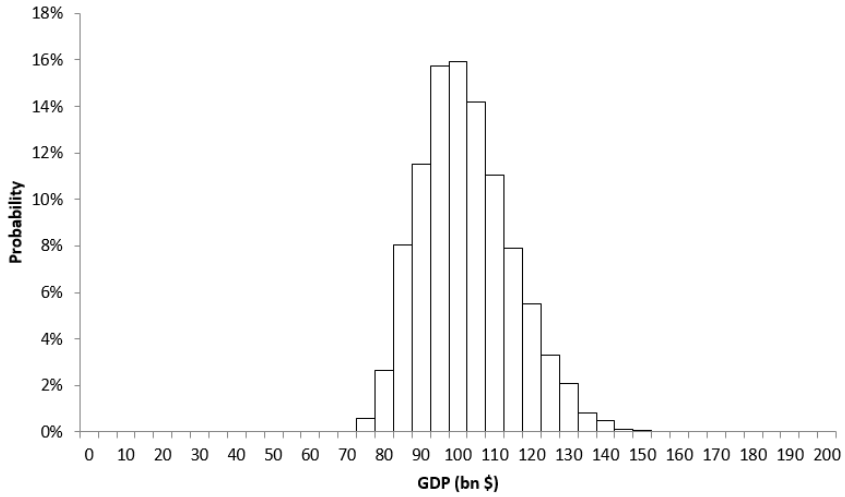

Fig.1 Simulation output: GDP of the Roman empire

In a next step, I explore upward adjustments for "most likely" population size and per capita GDP. I do this in a deliberately schematic fashion. For instance, we might be tempted to tie urbanization rates to per capita GDP by assuming the two to be positively correlated – the more urbanized a society, the higher its average per capita GDP. Even so, data from early modern Europe indicate that we cannot rely on this ostensibly plausible relationship to hold true across the board.

I proceed with six scenarios: the one already presented (1), a split-the-difference version (2), two deliberately extreme bounding cases (3-4), and two others that posit a strict correlation between urbanization rates and per capita GDP (5-6).[3] The first four are Monte Carlo simulations, the last two straightforward calculations (an empire of 60 million with a high urbanization rate of 22 percent that implies high per capita GDP, and the inverse).

(1)
Central estimate: 60m/70m/90m x $1,1250/1,250/1,700 = $100bn ($81bn-$122bn)

(2)
Split-the-difference: 60m/75m/90m x $1,1250/1,1413/1,700 = $106bn ($86bn-$128bn)

(3)
Small population, low per capita GDP: 60m/63m/90m x $1,125/1,183/1,700 = $95bn ($76bn-$119bn)

(4)
Large population, high per capita GDP: 60m/87m/90m x $1,125/1,642/1,700 = $117bn ($93bn-$140bn)

(5)
High urbanization rate, high per capita GDP: 60m x $1,700 = $102bn

(6)
Low urbanization rate, moderate per capita GDP: 90m x $1,250 = $113bn

[3] The "most likely" values in (3) and (4) are 10 percent (3) and 90 percent (4) of the difference between minimum and maximum values higher than the minimum value.

All the moderate scenarios (1-2, 5-6) produce very similar central GDP values of $100 billion to $113 billion, up to 30 percent higher than our 2009 estimate. That, and the fact that even the low scenario result (3) is somewhat higher than the latter, reinforces the impression that the 2009 estimate was somewhat too low. At the same time, even the maximalist scenario (4) does not produce a central value that is more than one-third higher than the original.

In other words, Bayesian modeling, operating within fairly generous bounds – while 60 million people may not be that many, 90 million is a lot, as is $1,700 – does not produce results that are markedly different from the original estimate but are nevertheless consistently somewhat higher. I conclude that the application of single-value estimates, as well as an underestimate of probable population size, kept our earlier estimate somewhat lower than it should have been.[4]

At the same time, and that is the principal take-away, it seems hard to move beyond a range of somewhere around $100 billion to $120 billion that implies an imperial GDP of 21 to 25 billion sesterces or so (compared to the earlier estimate of 18 billion). Are there more radical (that is, more upwardly biased) alternatives worth considering?

One option would be what we might call the "Rome ~ France scenario," in which 60m/75m/90m x $1,600/$1,700/$1,800 yields a mean of $127 billion, something like $26 billion sesterces (and an interval from $110 billion to $145 billion). Yet even that is not more than 45 percent higher than our original estimate, or a little than a quarter higher than my new central estimate (1).

The purpose of this exercise is not to make all pre-modern societies look the same. Instead, its purpose is to ask, what would it take to make the Roman world exceptional? If we set population number and per capita GDP in ranges that make sense in terms of what we know about early modern economies and more recent developing countries, we end up with a fairly wide range within which Roman GDP most likely fell. Otherwise we would need to assume that the Roman empire, in its entirety, was highly unusual, i.e. either much poorer or much richer. Archaeology, we may all agree, leaves us no room for the former. So what about the latter – richer than seventeenth-century France?

For the Roman empire to have enjoyed an even larger GDP, we need to explain how that could have happened. We know enough about productive technology to know that Roman capabilities were not ahead of those of our comparison cases; if anything they lagged behind. But GDP does not just fall from the sky. Did the Romans benefit from especially efficient institutions that generated intensive growth? We have no reason to believe so, unless we fall back on slavery, and in that case slavery would have a lot of heavy lifting to do – not just to raise per capita GDP in slave-rich areas (which perforce included lots of low-income enslaved workers) but on average across the board.

At the end of the day, any extra output had to come from somewhere. If anyone wants to argue that it existed, and that the comparative framing presented here is invalid, they need to come up with an

[4] A somewhat higher average per capita GDP estimate may be easier to reconcile with the consumption patterns and living expenses derived from Pompeiian graffiti discussed by Bowes 2021a. That said, it is not at all clear if we can use those data to derive representative metrics that can be meaningfully related to average GDP values for the entire population of the Roman empire. Bowes herself notes that the individuals in question appear to have been of somewhat elevated socio-economic status (Bowes 2021a: 575-9), whereas a slave referenced in one of these datasets had a much lower living standard (which matters not least because the Bay of Naples appears to have been an very slave-rich environment [De Ligt and Garnsey 2012], a feature that would have depressed average consumption). Other factors likewise merit attention: (1) whereas the individuals mentioned in the graffiti appear to have been adults, mean values for the empire apply to the total population, which included a sizeable share of minors with lower consumption requirements; (2) given that the non-reciprocal flow of tribute and rent into Rome ought to have inflated nominal prices there (Freyberg 1989) and given that the Campanian coast was closely linked to the city of Rome and the imperial elite, nominal prices (and wages) may well have been higher in Pompeii than in many other places (and in so far as that did not apply to grain [Bowes 2021a: 578], maritime imports may have influenced price-setting for that commodity); and (3) nominal wage rates in less central parts of the empire, most notably for the Roman military and in Egypt, strongly suggest that the nominal price levels inferred from the Pompeiian material cannot have been representative of overall averages for the empire as a whole.

alternative framework that helps us understand how such achievements became possible. I am not aware of any attempt to do so, and I doubt that it would be feasible.

From GDP to inequality

So what does all that mean for our understanding of Roman inequality? We have a rough idea of Roman imperial GDP – no longer a single value as before but a plausible (or, I would say more strongly, almost inescapable) range (see above); we have an even rougher idea of aggregate elite wealth (again in the form of an even wider range) (Scheidel and Friesen 2009); and we have a sense of how later pre-modern societies compared (Scheidel 2020). In other words, we are dealing with three variables, the first two of them crucial – and that would seem to call for another round of Bayesian modeling.

I leave that exercise for another day mostly because its payoff is bound to be limited – it is unlikely to tell us anything genuinely new. How so? Bowes complains that Gini coefficients for Roman income in the low .40s somehow result from twisting Roman metrics to fit preconceived notions about what pre-modern societies looked like (Bowes 2021b: 21). But that misses the point of such estimates for low-income economies. The range of plausible inequality outcomes for any low-income economy is much narrower than for higher-income ones: the higher per capita GDP is, the greater the potential for variation in income inequality becomes. That is empirically true – this is why Finland can have an income Gini coefficient of .27 and South Africa one of .63. But it is also a mathematical necessity, captured with exemplary clarity by the concept of the Income Possibility Frontier (IPF) developed by Branko Milanovic and associates in the 2000s (Milanovic et al. 2011; cf. Scheidel 2017: 445-456 for elaboration).[5]

My per capita GDP range for the Roman empire sets the lower bound at 2.4 times minimum subsistence, at 2.7 for the most likely value, and at 3.6 for the upper bound. The maximum possible Ginis – the IPF – supported by these multiples are .58, .63 and .72. If we lower them by one-fifth to arrive at defensible maximum real-life Ginis we end up with a range from .46 to .58, with .5 as the most likely upper bound. By comparison, the subsistence multiple in France in 1788, on the eve of the French Revolution, was something like 3.8, for a Gini of .56, or 76 percent of the theoretical maximum (Milanovic et al. 2011: 263) – and that was a deeply unequal society. All this suggests that is that a Roman empire-wide Gini was unlikely to have been higher than the high .40s or the low .50s. Higher values are effectively ruled out by constraints on inequality (we know that not everyone was destitute) and per capita GDP, as discussed above.

Those are upper bounds. Our 2009 central values of .42-.44 are somewhat lower, as one would expect of a more realistic estimate.[6] To be sure, in principle Roman Ginis could also have been lower than that, even much lower. By how much? China's extraction rate in 1880 may have been as low as 55 percent (Milanovic et al. 2011: 263; the extraction rate is the proportion of theoretically possible inequality that actually applied). Applying this value to the Roman case would yield a Gini range from .32 to 0.4. China was not then particularly unequal – a persistent phenomenon that later made Mao invent class enemies such as "landlords" who were only marginally better off than other villagers. Could we make a case that the same was true of the Roman empire? I strongly doubt it: signs of a high concentration of landownership in the Roman case (Harper 2015) speak against this, as does the high degree of de-urbanization in the late Qing period (Xu et al. 2018), which is the opposite of what we observe in the Roman empire. This should discourage us from lowering our Roman estimate below .4 or so. Even so, more comprehensive comparison would be required to move us onto more solid ground.

[5] The IPF equals the Gini coefficient of the income distribution of a society in which all persons but one survive on the bare subsistence minimum and one person captures all remaining income: thus, the IPF is a function of average per capita GDP (see Fig. 2). Any real-life Ginis for a given average per capita GDP must necessarily be lower than the IPF.

[6] Scheidel and Friesen 2009: 86. The discussion of extraction rates (87 n.88) can be simplified by adopting standard minimum subsistence levels ($300 in 1990 dollars or $470 in 2011 dollars), which I have done here and which accounts for slight discrepancies between then and now.

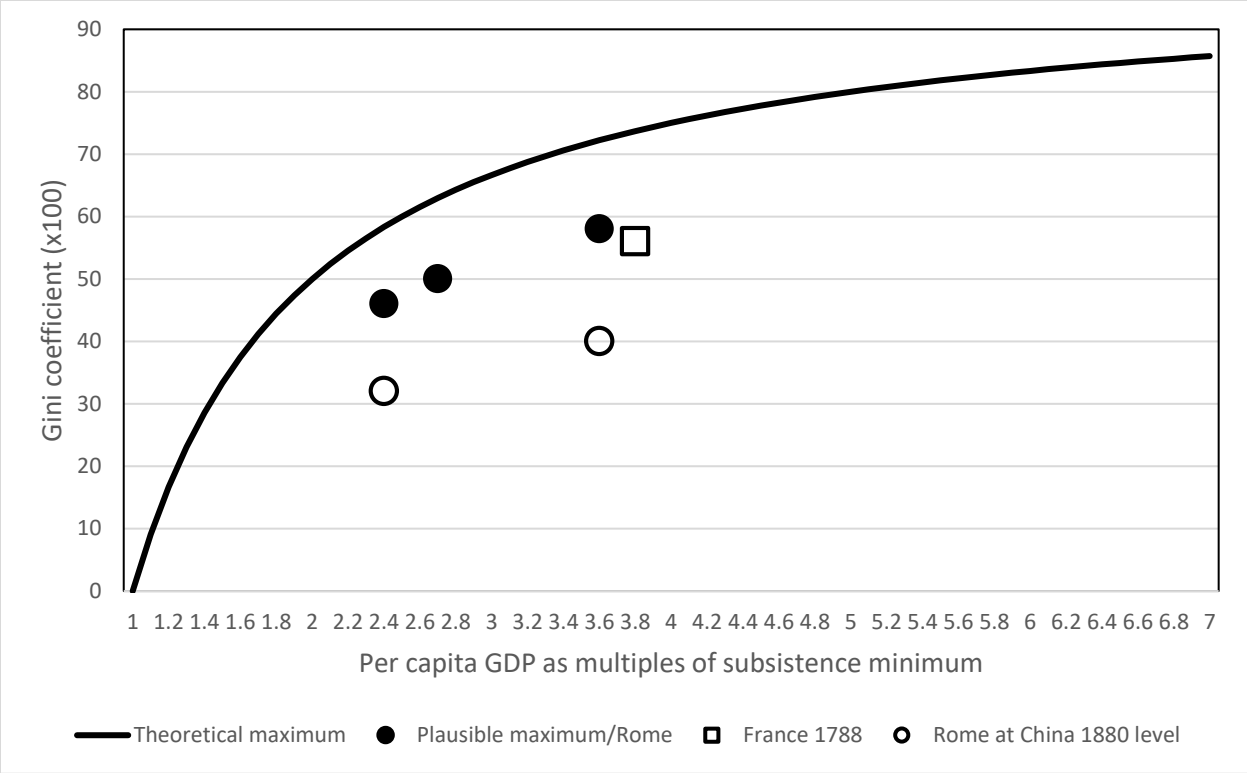

Fig. 2 Inequality possibility frontier (IPF) relative to per capita GDP and inequality estimates

And that is exactly the point. According to Bowes, the purpose of such estimates is to "re-inscribe the Roman world into a pre-modern economic monolith, not to reveal anything about the exigencies of life experienced by the Roman majority" (Bowes 2021b: 22). That is not true: there is no monolith. But there is most certainly a finite range of what agrarian societies could plausibly have experienced. By the same token, those "exigencies" were shaped by the same constraints that circumscribed that range. To dismiss, however implicitly, the notion that such constraints applied cannot be more than an empty rhetorical gambit.

Beyond charting the boundaries of the plausible, what else are such macro models good for? Properly qualified by Bayesian techniques, they might be of use as capacious frames for smaller-scale inquiries. We might ask, is what we find compatible with that framing, and why do we think we can tell? Models are not supposed to float in a rarefied sphere of abstraction, forever detached from actual ancient remains, but to contribute to the analysis of such remains, not least by encouraging explicit comparison. Our response cannot be to turn our back on the rest of history and claim that we can only understand our material on its own terms, ideally by employing methods not yet available. What is needed is exactly the opposite: testing models and mobilizing comparison to make better sense of what we see in the Roman record.

That, of course, is easier said than done. In the end, it may well turn out to be too difficult to relate macro-models to conditions on the ground, whether at archaeological sites or in samples of textual data. But that doesn't mean we shouldn't try.

I prepared this paper for the Langford conference "Socio-economic inequalities of the Roman world" held at Florida State University in Tallahassee on February 25-26, 2022. My thanks go to the organizers (Elizabeth Murphy and Robert Stephan) and the other participants for making it such a splendid event. I expect to revisit and revise my argument as the discussion continues.

References

Alfani, Guido. 2021. "Economic inequality in preindustrial times: Europe and beyond." *Journal of Economic Literature* 59: 3-44.

Bagnall, Roger S. 1992. "Landholding in late Roman Egypt: the distribution of wealth." *Journal of Roman Studies* 82: 128-149.

Beloch, Karl J. 1886. *Die Bevölkerung der griechisch-römischen Welt*. Leipzig: Duncker & Humblot.

Bernard, Seth. Forthcoming. "The premium for skilled labor in the Roman world."

Bolt, Jutta and van Zanden, Jan Luiten. 2020. "Maddison style estimates of the evolution of the world economy: a new 2020 update." Maddison Project Working Paper 15, https://www.rug.nl/ggdc/historicaldevelopment/maddison/publications/wp15.pdf.

Bowes, Kim. 2021a. "Tracking consumption at Pompeii: the graffiti lists." *Journal of Roman Archaeology* 34: 552-584.

Bowes, Kim. 2021b. "When Kuznets went to Rome: Roman economic well-being and the reframing of Roman history." *Capitalism* 2(1): 7-40.

Bowman, Alan K. 1985. "Landholding in the Hermopolite nome in the fourth century AD." *Journal of Roman Studies* 75: 137-163.

De Ligt, L. and Garnsey, P. 2012. "The album of Herculaneum and a model of the town's demography." *Journal of Roman Archaeology* 25: 69-94.

Duncan-Jones, Richard. 1990. *Structure and scale in the Roman economy*. Cambridge: Cambridge University Press.

Flohr, Miko. 2016. "Quantifying Pompeii: population, inequality, and the urban economy." In Flohr, Miko and Wilson, Andrew, eds. *The economy of Pompeii*, 53-84. Oxford: Oxford University Press.

Freyberg, Hans-Ulrich von. 1989. *Kapitalverkehr und Handel im römischen Kaiserreich (27 v. Chr. – 235 n. Chr.)*. Freiburg: Haufe.

Frier, Bruce W. 2000. "Demography." In Bowman, Alan K., Garnsey, Peter, and Rathbone, Dominic, eds., *The Cambridge ancient history, 2nd edition. Volume XI: the high empire, A.D. 70-192*, 787-816. Cambridge: Cambridge University Press.

Gaggioli, Amanda. 2019. "Quantifying the impacts of disaster on wealth inequality: a diachronic study of household wealth inequality in the Insula of Menander, Pompeii." Seminar paper, Stanford University.

Hanson, John W. In press. "New approaches to the urban population and urbanization rate of the Roman Empire, AD 1 to 200." In Lavan et al., eds. in press.

Hanson, John W. and Ortman, Scott G. 2017. "A systematic method for estimating the populations of Greek and Roman settlements." *Journal of Roman Archaeology* 30: 301-324.

Harper, Kyle. 2015. "Landed wealth in the long term: patterns, possibilities, evidence." In Erdkamp, Paul, Verboven, Koenraad, and Zuiderhoek, Arjan, eds., *Ownership and exploitation of land and natural resources in the Roman world*. Oxford: Oxford University Press: 43-61.

Kohler, Timothy A. et al. 2017. "Greater post-Neolithic wealth disparities in Eurasia than in North America and Mesoamerica." *Nature* 551: 619-622.

Kohler, Timothy A. and Smith, Michael E., eds. 2018. *Ten thousand years of inequality: the archaeology of wealth differences*. Tucson: University of Arizona Press.

Lavan, Myles. 2019. "Epistemic uncertainty, subjective probability, and ancient history." *Journal of Interdisciplinary History* 50: 91-111.

Lavan, Myles, Jew, Daniel, and Danon, Bart, eds. In press. *The uncertain past*.

Macksoud, James. 2019. "Income inequality in the Roman army." Seminar paper, Stanford University.

Maddison Project Database. 2020. https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2020?lang=en.

McConnell, Joseph R. et al. 2018. "Lead pollution recorded in Greenland ice indicates European emissions tracked plagues, wars, and imperial expansion during antiquity." *Proceedings of the National Academy of Sciences* 115: 5726-5731.

Milanovic, Branko, Lindert, Peter H., and Williamson, Jeffrey G. 2011. “Pre-industrial inequality.” *Economic Journal* 121: 255-272.
Pavlyshyn, Damian, Johnstone, Iain, and Saller, Richard. 2020. “Lead pollution and the Roman economy.” *Journal of Roman Archaeology* 33: 354-364.
Scheidel, Walter. 2007. “Demography.” In Scheidel, Walter, Morris, Ian, and Saller, Richard P., eds., *The Cambridge economic history of the Greco-Roman world*, 38-86. Cambridge: Cambridge University Press.
Scheidel, Walter. 2017. *The great leveler: violence and the history of inequality from the Stone Age to the twenty-first century*. Princeton: Princeton University Press.
Scheidel, Walter. 2019. *Escape from Rome: the failure of empire and the road to prosperity*. Princeton: Princeton University Press.
Scheidel, Walter. 2020. “Roman wealth and wealth inequality in comparative perspective.” *Journal of Roman Archaeology* 33: 341-353.
Scheidel, Walter and Friesen, Steven J. 2009. “The size of the economy and the distribution of income in the Roman empire.” *Journal of Roman Studies* 99: 61-91.
Stephan, Robert P. 2013. “House size and economic growth: Regional trajectories in the Roman world.” PhD dissertation Stanford University.
Xu, Yi, van Leeuwen, Bas, and van Zanden, Jan Luiten. 2018. “Urbanization in China, ca. 1100-1900.” *Frontiers of Economics in China* 13: 322-368.