Abstract

We survey archaeological evidence suggesting that among hunter-gatherers and farmers in Neolithic western Eurasia (11,700 to 5,300 years ago) elevated levels of wealth inequality occurred but were ephemeral and rare compared to the substantial enduring inequalities of the past five millennia. In response, we seek to understand not the de novo “creation of inequality” but instead the processes by which substantial wealth differences could persist over long periods and why this occurred only at the end of the Neolithic, at least four millennia after the agricultural revolution. Archaeological and anthropological evidence suggests that a culture of aggressive egalitarianism may have thwarted the emergence of enduring wealth inequality until the Late Neolithic when new farming technologies raised the value of material wealth relative to labor and a concentration of elite power in early proto-states (and eventually the exploitation of enslaved labor) provided the political and economic conditions for heightened wealth inequalities to endure.

JEL codes: D3, D63, E24, N53

Keywords: stationary distribution of wealth, ancient inequality, intergenerational transmission of wealth, wealth shocks, technology, institutions, states, slavery, agriculture, hunter-gatherer, horticulture, culture, power, Gini coefficient.

* Affiliations: Bowles, Santa Fe Institute and CORE Econ; Fochesato, Dondena Centre and BIDSA, Bocconi University. We would like to thank the Dynamics of Wealth Inequality Project of the Behavioral Sciences Program at the Santa Fe Institute and Caroline Seigel of the SFI library for their support of this research. Thanks also to Max Greenberg and David Romer for their careful reading of the entire manuscript and numerous suggested improvements, to Steven Durlauf for suggesting the topic and, for their contributions to this research, to Guido Alfani, Robert Allen, Jeanne Arnold, Amy Bogaard, Wendy Carlin, Jung-Kyoo Choi, Paul Halstead, Timothy Kohler, Ian Kuijt, Chiaki Moriguchi, Scott Ortman, Marko Porcic, Rick Schulting, Eckart Voland, Todd Whitelaw, Kai Wilfur, Elisabeth Wood, Henry Wright, Katherine Wright and four anonymous referees.
1. **Introduction**

Lavish burials at Sungir on the Russia Plain including two children—dating from 24,000 years ago—may indicate the antiquity of inherited wealth inequality (Pettitt and Bader 2000). Anthropologists have long recognized that some hunting-and-gathering populations were exceptionally unequal in both economic and political terms, in some cases including the exploitation of enslaved labor (Hayden 2001) and elite command of a standing army (Macmahon and Marquardt 2004). However, while occasional opulent burials like the ones at Sungir and other evidence from the Late Paleolithic (50,000 to 11,700 years ago) attest to cases of substantial economic inequality early in human prehistory, it appears to have been rare and episodic until much more recently.

In Figure 1, we present estimates of inequality in material wealth for western Eurasia from the end of the Paleolithic almost twelve millennia ago to the end of the Roman Empire a millennium and a half ago.\(^1\) The estimates are based on the size of dwellings, the size of storage areas (where these can be identified), land ownership, and the value of goods buried with the dead.\(^2\) Gini coefficients (denoted \(G\), with 0 indicating perfect equality and 1, maximal inequality) based on these measures have been calculated (insofar as is possible) to be comparable across sites and dates. This has entailed adjustments with respect to the above four asset types (grave goods, for example, are typically more unequally distributed than dwelling size) and wealth holding unit (we infer household wealth from individual grave goods by creating hypothetical male-female pairings). We have also adjusted the estimates to correct for biases due to both small sample size and the fact that the populations from which our samples are drawn differ in size and geographical scale. We also adjust for the common absence of those without wealth (for example, the landless and enslaved people) from data sets on wealth holdings.\(^3\)

---

\(^1\) A note on geographical terms and dating: By western Eurasia we refer to Europe and the Levant in the broader sense comprising the Mediterranean lands, including those in Africa, as well as Mesopotamia (the area between the Tigris and Euphrates Rivers). When drawing on archaeological data, we express dates based on the author’s own usage, either BCE (for before the Common Era, or equivalently BC) or BP (meaning before the present, which by convention means before 1950). CE refers to the Common Era (formerly termed AD).

\(^2\) Because most of the key developments in our narrative occurred prior to the development of writing, we necessarily rely on archaeological evidence that is subject to substantial uncertainty, the magnitude of which we also document in Fochesato and Bowles (2024a).

\(^3\) These adjustments and our sources and methods for calculating the Gini coefficients are described in Fochesato and Bowles (2024b).
We consider an individual’s wealth to be an aggregate of individual capabilities and material assets, expressed as a stock that yields a flow of valued services—consumption, leisure, status, power, or many offspring—to an individual or household. Disparities in each of these valued dimensions are subjects of concern from the standpoint of both distributive justice and human flourishing. We focus on disparities in material wealth (which we term just “wealth” in what follows) in part for practical reasons: archaeological evidence on stocks—of storage, burial goods, and house size—is available, while evidence on flows such as income, consumption, or reproductive success typically is not. Moreover, due to our interest in the persistence of economic disparities over time, intergenerational transmission of economic status will be critical to our interpretation of the data, and hence wealth is the appropriate measure of interest.

Figure 1. Inequality in material wealth in western Eurasia from the end of the Paleolithic to the Roman Empire. Dates on the horizontal axis are years before the Common Era (BCE). PPNA and PPNB refer to the western Asian Pre-Pottery Neolithic periods A and B periodization used by archaeologists. Our data base is \( n = 185 \) estimates of the Gini coefficient (including 3 from the late Paleolithic, not shown in this figure). The line labeled \( G = 0.695 \) indicates the mean of 182 comparable estimates of wealth inequality in state-governed societies in which slavery did not play a major role in production for the period from 300 BCE to 2000 CE. The sources used are listed in Table 1 and in the Appendix, Sections 2, and 6.
In our data set (seen in Figure 1), prior to the Bronze Age (about 3300–1200 BCE)) just a few societies exhibit modern levels of inequality. In these exceptional cases, the earliest of which include the Mesopotamian sites at Tepe Gawra, Ur, and Tell Brak 5,000 or more years ago, the degree of wealth inequality approximates the levels sustained ever since then in state-governed societies in which slavery did not play a major role in production (what we term “modern” levels of wealth inequality). The populations of the Roman Empire, a society based on enslaved labor, are exceptionally unequal.

We term these developments the origin of enduring wealth inequality, occurring in the region for which archaeological evidence is sufficient to establish such a regularity. While the intergenerational transmission of wealth will play an important causal role in our explanation, the enduring character of inequality that we seek to explain is not the elevated wealth of particular dynastic families but the persistence of substantial wealth differences among households, even as the memberships of the wealthy and less wealthy classes change over time.

Figure 1 is possible due to the recent proliferation of data allowing estimates of prehistoric wealth inequality to be made. The major sources are shown in Table 1.

Given the unavoidable uncertainties associated with archaeological data, many of the inferences that we draw from our estimates as well as our overall account are best considered plausible conjectures rather than tested hypotheses. (Economists may be struck by our having adopted, where appropriate, the practice of archaeologists in reflecting this uncertainty in our modes of expression, as in “may have been,” “appears likely,” and so on.) Moreover, the differences in the level of material wealth inequality over the 10,000-year span in Figure 1 do not measure experienced inequality. This is particularly the case for economies in which material wealth is not very important for one’s livelihood or standing in the community, which, as we will see in Figure 3 and Table 2, would include many of our observations, particularly during the Early Neolithic. In both hunter-gatherer and early farming populations, people may have cared more about the size of their family than the size of their dwelling or the other indicators of material wealth that we measure with archaeological data.  

4 We begin, in Section 2, by providing a unified framework to study the emergence of enduring (rather than episodic) wealth disparities, drawing on models of wealth dynamics. Then, in Section 3,  

---

4 By this measure (Gini coefficients of reproductive success), ethnographic hunter-gatherer populations (even foragers, that is, mobile hunter gathere) exhibit substantial levels of inequality, certainly on a par with material wealth inequality among the Early Neolithic western Eurasian farmers in our data set (Hill and Hurtado (1996), Howell (2000), Borgerhoff Mulder, et al. (2009), Ross, et al. (2023)).
we review the literature on three sets of factors that are thought to influence the level of inequality in a population. These, broadly, concern technology/ecology, stressing the storability, heritability, and taxability of the elements of people’s livelihoods and the circumscription of their environments; institutions/politics, focusing on political elites’ capacities to sustain their rule while extracting high levels of rents, tribute, or taxation; and culture/ideology, describing the beliefs, social norms and practices that fostered either compliance to wealth inequalities or resistance among the less well off.

<table>
<thead>
<tr>
<th>Source</th>
<th>Inequality</th>
<th>Geographic area (dates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windler, Thiele and Müller (2013)</td>
<td>Grave goods</td>
<td>Southeast Europe (5000–4000 BCE)</td>
</tr>
<tr>
<td>Kohler, Smith, Bogaard et al. (2017)</td>
<td>House area</td>
<td>Eurasia (10000 BCE–100 CE) North America and Mesoamerica (1100 BCE–1700 CE)</td>
</tr>
<tr>
<td>Our estimation from (Porčić 2019, 2012)</td>
<td>Grave goods</td>
<td>Balkans (6500–4200 BCE)</td>
</tr>
<tr>
<td>Basri and Lawrence (2020)</td>
<td>House area</td>
<td>Near East (9000–700 BCE)</td>
</tr>
<tr>
<td>Fochesato, Higham, Bogaard et al. (2021)</td>
<td>Grave goods</td>
<td>Thailand (2000 BCE–600 CE)</td>
</tr>
<tr>
<td>Squitieri and Altaweel (2022)</td>
<td>House area</td>
<td>Near East (3000 BCE–200 CE)</td>
</tr>
</tbody>
</table>

Table 1. Sources of evidence on wealth inequality from prehistoric data. The ongoing Global Dynamics of Inequality Project led by Tim Kohler, Amy Bogaard, and Scott Ortman will substantially add to available sources (Kohler and Thompson 2022).

There is a gap in these literatures that we call the puzzle of the origins of enduring economic inequality. This is not why persistent heightened inequality emerged, but instead: why did it not occur earlier? Before the Bronze Age, how did a combination of the three themes in the literature—technology, institutions, and culture—contribute to sustaining egalitarian societies long after the agricultural revolution, and how were these equalizing processes eventually reversed, allowing the emergence of substantial and enduring wealth inequalities?
Then, in Section 4, we review archaeological evidence from North America, Southeast Asia, and western Eurasia suggesting that prior to the Bronze Age, aside from exceptional and typically ephemeral ecological conditions, modern levels of inequality were difficult to sustain, in part due to what the archaeologist Ian Hodder has termed the “aggressively egalitarian” nature of some Neolithic communities (Hodder 2014) p.1. To imagine how this inequality-mitigating dynamic might have worked, we then model the coevolution of wealth inequalities along with egalitarian and collectivist social norms. Agent-based simulations of our model, calibrated to Neolithic conditions, show that occasional instances of substantial inequality ($G > 0.5$) could occur even in a population of “aggressive egalitarians” that sustained very modest levels of inequality over the long run. We close the section by noting the lack of convincing evidence that the introduction of farming *per se* was associated with an increase in wealth inequality.

Recent research in paleobotany by Amy Bogaard and her colleagues (introduced in Section 5) provides a key piece of our proposed resolution of the above puzzles: developments in farming technology providing novel opportunities for accumulating wealth that differentially favored those with more initial wealth. These innovations—especially ox-drawn plows—raised the value of land, draft animals, and other forms of material wealth relative to labor, which in turn were associated with important demographic, cultural, and institutional changes.

The ox-drawn plow transformed what had previously been a land-abundant and labor-limited economy to one in which material wealth was scarce relative to labor. The result was to generalize to any locale suitable for plow-based farming the previously rare and often ephemeral ecological conditions for the emergence and of substantial wealth inequality. We then draw on recent ethnographic evidence to suggest that the proposed explanation may extend beyond the plow- versus hoe-based farming distinction to any innovation—such as irrigation and the domestication of long-lived animals—that enhances the value of land and importance of material of wealth in a household’s livelihood. These ethnographic data also provide a robustness check on our estimates of the effect of the shift in technology on the distribution of wealth.

Comparing the new plow-based farming systems with the earlier hoe-based economies, we find a substantial increase in wealth inequality, a change that can be attributed to increases in both the magnitude of contemporaneous shocks to a household’s wealth and the intergenerational transmission of wealth. We close the section by exploiting an unusual new archaeological data set from Early
Neolithic to late Iron Age Thailand to provide a second robustness check for our account of the shift from a labor-limited to land-limited economic environment.

In Section 6 we develop an economic model of the process by which the ox-and-plow technological revolution could have elevated wealth inequality well before the concentration of elite political power. But the substantial wealth inequality that resulted from this process at a few Late Neolithic and Bronze Age sites might have been short-lived (like their Early Neolithic antecedents) but for a subsequent process of political centralization that complemented the growing wealth inequalities, resulting in the emergence of what we will below define as the first archaic proto-states in Mesopotamia.

Two subsequent institutional developments (which we introduce in Section 7) brought enduring wealth inequalities up to and in some cases above modern levels. The first was the continuing process of state formation, by which a unified elite more effectively monopolized the use of coercion, the first examples in our data set being observations from the Roman Empire in the first centuries of the Common Era. The second was the imposition of slavery, converting free labor to a form of material wealth that could be accumulated and transmitted over generations.

We have recast the question of the origin of inequality to be about not the creation of inequality but instead about its persistence and replication (hence the “enduring” in the title of our paper). This Darwin-inspired restatement of the question in terms of replication rather than innovation is based on the idea that attempts at personal aggrandizement are a ubiquitous feature in human societies, and when these attempts are successful at a systemic level, economic inequalities will periodically occur. And, with equal certainty, there will be opposition to the claims of aggrandizers. Our question thus concerns the emergence of conditions that favored the wealthy in these ongoing conflicts of interest between aggrandizing would-be elites and others.

2. A framework for the analysis of the origin of enduring ancient wealth inequality

The question, thus rephrased, and our answer to it thus do not strictly concern origins at all; they are about the conditions under which substantial levels of inequality could endure, requiring a dynamic analysis of change and stasis in the distribution of wealth. Our representation of the determinants of inequality in the stationary distribution of wealth in an ancient economy is an adaptation of the models
initiated by Alan Blinder, Glenn Loury and Gary Becker and Nigel Tomes.\textsuperscript{5} Here we provide a common framework to explore the effects of developments stressed in the archaeological, anthropological, and historical literature, including the relative scarcity of land and other capital goods, marital and other sorting, the demographics of polygyny, redistribution among members of a lineage (ancestral descent group), and increasing returns in both investment and acquiring power.

\textit{2.1 The dynamics of wealth inequality}

We consider a setting in which one’s own wealth depends not only on parental wealth but also on other inputs into the process of wealth generation that are available independently of parental wealth, such as the public goods aspects of institution quality, open access knowledge and natural capital, the health and skills of the labor force available for employment, and the wealth of potential marital or other partners. We proxy the contributors to wealth acquisition that are unrelated to parental wealth by the mean wealth level, which we consider to be complementary to parental wealth in contributing to an individual’s end of life wealth.

Letting a prime (\('\)) indicate the previous generation, an individual’s wealth $w_i$ varies with parental wealth $w_i'$ and mean wealth $\bar{w}$ according to

$$w_i = (1 - \beta)w + \beta w_i' + \lambda_i$$

Wealth is measured in natural logarithms and normalized so that mean wealth is invariant across generations; $\lambda_i$ is a wealth shock uncorrelated with parental wealth, with mean zero and variance $\sigma_{\lambda_i}^2$ (possibly mitigated by risk pooling practices). We term the previous generation the “parent” and the current generation the “offspring”; while in describing our model, we refer to them as “individuals,” they represent any wealth pooling household or other entity (such as an extended family or lineage).\textsuperscript{6}

In this system, the long-run (stationary) distribution of wealth is the result of two offsetting forces. The first is (uninsured) positive or negative shocks to wealth, introducing new inequalities in every

\textsuperscript{5} Becker and Tomes (1979), Blinder (1973), Blinder (1976), Loury (1981). A similar model has been used to study the dynamics of wealth inequality as recorded by ethnographers in hunter-gatherer and small scale farming populations (Borgerhoff Mulder, et al. (2009), Bowles (2010), Smith, et al. (2010)).

\textsuperscript{6} The same framework is applicable to both sexes depending on the inheritance practices of the population under study; they typically differ by gender (Borgerhoff Mulder, et al. (2019)).
generation. The second, dissipating these inequalities over time, is the process of intergenerational regression to the mean resulting from various forms of wealth redistribution (including sharing norms and possible confiscation) and the less-than-perfect processes of parent-to-child bequest and marital sorting.

Our measure of inequality, $\sigma^2_w$, is the variance of the natural logarithm of wealth, which in the stationary (that is, equilibrium) distribution of wealth is given by:

$$\sigma^2_w = \frac{\sigma^2_w}{1 - \beta^2}$$

(2)

where $\beta$ is the intergenerational wealth transmission elasticity (the derivative of adult offspring wealth with respect to parental wealth, both in natural logarithms). This means that (for $\beta < 1$) the degree of inequality in the stationary distribution is given by the magnitude of the wealth shocks, expanded by the intergenerational transmission multiplier, $(1 - \beta^2)^{-1}$. Equation (2) thus captures the fact that where intergenerational transmission is substantial, the inequalities introduced by wealth shocks in previous generations persist, augmenting the inequalities induced by contemporaneous shocks. Stationary inequality is a value of $\sigma^2_w$ such that the reduction in inequality occasioned by regression to the mean exactly offsets the dis-equalizing effect of current wealth shocks.

Values of $\beta > 1$ give rise to bifurcating processes—runaway inequality—resulting in poverty traps of the type recently studied and documented by Clare Balboni and her coauthors (Balboni, Bandiera, Burgess et al. 2022). Though we cannot estimate the relevant transmission elasticities ($\beta$'s) from archaeological data, Monique Borgerhoff-Mulder and her coauthors (including one of the current authors) collected wealth data from two generations of farmers in small-scale societies that may be similar to the ancient economies under study here (Borgerhoff Mulder, Bowles, Hertz, et al. 2009). Using matched parents and offspring with appropriate age controls, they estimated the intergenerational elasticity ($\beta$) for what they termed relational (meaning network), embodied (human), and material wealth. The mean estimated $\beta$ for material wealth in their data set is 0.476 (CI: 0.304; 0.648), suggesting significant regression to the mean on average. This estimate implies that the offspring of a wealthy individual is likely to be a bit less than half as wealthy relative to the mean as were their parents.

---

7 This commonly used unit-free measure of inequality allows a remarkably simple model of the dynamics of wealth inequality, but it is not readily applicable to empirical data sets in which many households have zero wealth. The derivation of equation (2) from equation (1) is explained in nontechnical terms in Borgerhoff Mulder, et al. (2009).
2.2 The determinants of stationary inequality: Wealth shocks and intergenerational transmission

We think of the transmission process as a set of mechanisms that jointly translate one generation’s wealth at the normal age of death to the next generation’s wealth at the same age, meaning that it encompasses more than bequest (wealth being transferred from parents to offspring). So, we decompose transmission into three processes (assumed for simplicity to be sequential): bequest, the passing on of the deceased’s wealth to the offspring in the next generation; sorting, the merging of the bequeathed wealth with another’s wealth through marriage or membership in some other wealth-sharing entity; and accumulation, the change in the individual’s level of wealth between sorting and death that results from investment, wealth shocks, and other influences over their adult life. We now consider factors affecting each of these determinants of stationary wealth inequality.

Shocks and insurance. The magnitude of (uninsured) wealth shocks (the numerator in equation (1)) due to bountiful harvests, theft, ecological calamities, health afflictions, and the like will depend on both the nature of the wealth and the extent of formal or informal shock-buffering insurance mechanisms. Ethnographic evidence suggests that food-sharing practices may have had substantial risk-mitigating effects in Neolithic economies. Among five small-scale labor-limited, land-abundant (horticultural) populations for which ethnographic data are available, over a third of the food produced or acquired by members of a family was consumed by people beyond the immediate family on average (Fochesato and Bowles 2015). Herders often tend some animals owned by others and entrust some of their own animals to the care of others so that theft or contagious illness in a herd will not eliminate the wealth of any household (Bollig 2006), sometimes adopting what anthropologists call fictive kin to broaden their risk-sharing pool (Thomas 1979). These risk-pooling practices reduce the effective variance of wealth shocks, leading to more egalitarian outcomes.

Bequest transmission. The bequest transmission process contributes to regression to the mean for three reasons. The first is that it is likely that in many of the societies under consideration (as we will see), inheritance goes to the parents’ lineage or some other wealth-sharing group rather than to parents’ own offspring. Offspring may thus benefit from having a wealthy parent both directly (via bequest) or indirectly via the greater likelihood that a wealthy parent will belong to a wealthy lineage or other group, and thus have access to the wealth of others’ parents, as a result of
group membership rather than parental bequest. But unless lineages are homogeneous in wealth, this form of bequest to non-offspring will reduce the intergenerational transmission coefficient.

Second, other forms of redistribution may be substantial, as we will see, including lavish funeral feasting and deliberate destruction of assets. Finally, the inheritance of assets is diluted due to the greater number of offspring wealthier parents have, especially if polygynous (Bowles, Ross, Borgerhoff Mulder et al. 2023). Primogeniture does not eliminate this dilution, because under primogeniture the fortunate single offspring of the wealthy has a large number of wealthless siblings, resulting on average in significant regression to the mean.

**Sorting into wealth-pooling entities.** Following bequest, pairing for marriage (or joining any other wealth-pooling entity) may vary from perfect sorting, in which post-sorting (per person) wealth is identical to pre-sorting wealth, to no sorting, in which post-sorting wealth is an equally weighted sum of own and mean wealth. Any less-than-perfect degree of sorting will contribute to regression to the mean.

**Accumulation.** Following marital or other assortment, the individual may consume a portion of their wealth while investing the rest, the cumulative returns to which result in a level of wealth at death. Unlike bequest and sorting, accumulation may counteract regression to the mean: strong increasing marginal returns to invested wealth may result in the intergenerational wealth transmission elasticity exceeding 1 over some portion of wealth levels. A similar process also arises naturally from models of savings and investment under realistic assumptions concerning the credit market—minimum project size or quantity constraints facing less wealthy borrowers—and has been widely studied since Glenn Loury’s foundational paper.

An increasing-returns-type phenomenon also occurs when wealth is invested in acquiring and sustaining political power that is then converted into greater wealth. Contest success functions—relating the probability of victory in some political or military contest to the resources devoted to the conflict—may exhibit strong increasing returns to additional levels of troop deployment (or other contest

---

8 This idea is key to what Steven Durlauf (1999) calls the “memberships theory of inequality.”
9 Blinder (1973) pointed out that inheritance to males only is equivalent in wealth dynamics to sons marrying their sisters. We have not explored how unigeniture would affect these dynamics (though simulations of the model in Blinder (1976) suggest that the effect would be small) including taking account of the possibility of strong positive feedback mechanisms (for example, economies of scale and increasing returns to investments in political power, mentioned below) not present in Blinder’s model.
10 See Loury (1981) and the subsequent literature, for example, Galor and Zeira (1993) and Mookherjee and Ray (2002). For empirical evidence, see Bardhan, et al. (2000), Banerjee and Duflo (2010), and the works cited therein.
resources) when the parties are evenly matched.\textsuperscript{11} Thus, in addition to conventional increasing returns, another empirically plausible mechanism generating economic polarization are positive feedbacks associated with the exercise of power, especially where some degree of concentration of political power raises the stakes of the political game, as was initially stressed by Eric Wolf and developed in more recent works (Acemoglu and Robinson 2008, Dow and Reed 2023, Graeber and Wengrow 2021, Wolf 2010).\textsuperscript{12}

These three positive-feedback processes enjoyed by those with greater wealth—arising from conventional economies of scale, wealth-constrained borrowing, and increasing returns to investments in acquiring political power—can function in a complementary manner, heightening stationary wealth inequality. An example in the ethnographic literature is the ownership of large herds of water buffaloes by the wealthy on the Indonesian island of Sumba, which allows them to borrow additional animals on favorable terms for the purpose of slaughter at feasting events that enhance their political power and with it their access to additional wealth (Adams 2016).

3. \textit{Inequality’s origins: Technology/ecology, politics/institutions, and culture/ideology}

Turning now to the archaeological and historical literatures on the determinants of economic inequality, we exploit a feature of prehistoric and ancient studies. This is the fact that the range of institutions, technologies, and cultures that humans experienced over this long period is vastly wider than in the more recent societies that make up the preponderance of empirical studies leveraging data made possible by introduction of income and wealth taxation over the last couple of centuries. The greater variation of settings affecting the level of inequality in the prehistoric record allows us to study the effects of radical changes along three dimensions: technology (e.g., from collecting food to producing it, to the substitution of the work of draft animals for human labor); governance (e.g., from decentralized polities without formalized political leadership to state governance and enslaved labor); and cultures (e.g., from collective and reciprocal norms to greater individualism). We are able to draw on rich literatures regarding each of these three sets of determinants of inequality.

\textsuperscript{11} Sung-Ha Hwang (2012) provides empirical estimates of contest success functions based on the outcomes and troop deployments in 315 battles in seventeenth-century Europe.

\textsuperscript{12} The centrality of politics as a determinant of wealth inequality is consistent with evidence that societies governed by states exhibited markedly higher levels of wealth inequality than non-state societies as shown in Fochesato and Bowles (2024b).
3.1. Technology/ ecology

Stressing technical aspects of how we secure our livelihoods, economists and other social scientists sometimes derive hypotheses about inequality from the characteristics of a production function, the kinds of goods produced, and the natural environment in which this takes place. Perhaps the most influential of these hypotheses dates back to the related works of Friedrich Engels and Lewis Henry Morgan in the late nineteenth century. They held that the emergence of a surplus over subsistence that was made possible by producing food rather than hunting and gathering is the key to explaining the origin of high levels of inequality (Engels 2010 [1884], Morgan 1963 [1877]). Robert Allen provides a recent statement along these lines: he interprets “the archaeological record … in terms of changes in labour productivity … and the size of the associated agricultural surplus” (Allen 2024) p.1 (Mayshar, Moav and Pascali 2022).

Recent contributions stressing technology (though often not through the Engels-Morgan lens of an agricultural surplus) include Alberto Alesina and co-authors’ exploration of a set of hypotheses due to Ester Boserup and Jack Goody. They find that, in comparison to hoe-based farming, a tradition of plow-based cultivation is associated with values and practices supporting greater gender inequalities (Alesina, Giuliano and Nunn 2013, Boserup 1965, Goody 1976). Similarly, Alain Testart, James Scott, and others have suggested that societies based on crops that are readily stored or appropriated by others (cereals rather than root crops, for example) are likely to exhibit greater political and economic inequality (Scott 2009, 2017, Testart 1982). Scott remarks: “It is surely striking that virtually all classical states were based on grain. . . . History records no cassava states, no sago, yam, taro, plantain, breadfruit, or sweet potato states” (Scott 2017) p. 21. Joram Mayshar and his coauthors recently provided convincing empirical support for this so called appropriability hypothesis concerning the origin of autocratic states (Mayshar, Moav and Pascali 2022).

Behavioral ecology provides analogous hypotheses on the sources of inequality in fitness, based on the nature of the goods constituting an animal’s livelihood (for example, clumped -- i.e. geographically concentrated and thus defensible -- resources) and the benefits of group membership (and associated costs of dispersion). Analogous to this “clumped resources” hypothesis in social science include the fact that among Native American populations on the northern Pacific coast, resources such as prime fishing

---


---
or mollusk-collecting sites were associated with elevated levels of institutionalized social inequality (Smith and Codding 2021). Fernand Braudel (1972) and others have observed that irrigated farming (making land a clumped resource) is associated with more class-stratified social structure. Comprehensive comparative evidence consistent with Braudel has been provided by Jeanet Bentzen and her coauthors, who hypothesize that “irrigation allowed landed elites in arid areas to monopolize water and arable land” (Bentzen, Kaarsen and Wingender 2017 p.1).

We borrow from this literature the hypothesis that increased inequality would be associated with major changes in methods of production such as the greater capital intensity of production brought about by the substitution the work of non-human animals for human labor (as with the ox-drawn plow), a shift from rain-fed to irrigated farming, or changes in the source of one’s livelihood, such as the prehistoric shift from wild to cultivated and tended plant and animal species.

### 3.2 Institutions/politics

Daron Acemoglu and James Robinson suggest a quite different research agenda: in the evolution of wealth inequalities, “institutions are a powerful and autonomous force, not simply an adaptation to preexisting (or autonomously evolved) technological conditions. They shape—as much as they are shaped by—the technological trajectory of societies” (Acemoglu and Robinson 2009) p. 679. Many historians (Brenner 1976, Wright 2013), sociologists of the “conflict” school (Dahrendorf 1959, Wright 1979), and others (Acemoglu and Robinson 2008, Boix 2015, Scheidel 2017) focus on what Eric Wolf called “the exercise of power and domination—through a political process.” In Wolf’s tributary mode of production the “cultivator or herdsman is allowed access to the means of production while tribute is extracted … by political or military means” (Wolf 2010:80).

We will see that the gradual concentration of political power in western Eurasia during the Late Neolithic and Bronze Age can be represented as a transition away from Wolf’s “kin-ordered mode of production” to his “tributary mode” and a movement along his continuum of the tributary mode ranging from one “in which power is held largely by local overlords and the rule at the apex is fragile and weak” to one “in which power is concentrated strongly in the hands of a ruling elite” (Wolf 2010) p.286.

David Stasavage provides an account (with empirical support) of Wolf’s first, more decentralized, form of political rule (albeit using a different terminology) in Bronze Age Mesopotamia (Stasavage 2020). Guido Alfani interprets the evidence for high and sustained (even rising) levels of wealth inequality in early modern Europe as the result of regressive fiscal institutions and the “growing
coercive power of state institutions,” that is, a move toward the more centralized node of Wolf’s tributary mode continuum (Alfani 2021).

Charles Tilly describes a cultural group selection process by which these institutions arose in Europe. In the states that prevailed in warfare, as “properly socialized citizens came to attach moral value to the payment of taxes, they monitor themselves and each other, blaming tax evaders as free riders” (Tilly 1990) p. 89). In a similar vein, Timothy Besley, following Margaret Levi ideas, stresses the complementarity of states’ tax capacity and citizens’ compliance based on state legitimacy in the form of reciprocity toward elites when they provide public goods (Besley 2020, Levi 1988).

Evidence from Highland Mesoamerica during the thirteenth to sixteenth centuries suggests that an internal source of state revenues—locally raised taxes rather than tribute extracted from other polities—constituted an impediment to concentrated elite power (Fargher, Heredia Espinoza and Blanton 2011, Feinman 2018). Richard Blanton and Lane Fargher apply similar reasoning to the formation of the first states (Blanton and Fargher 2008).

Politics is central also to Walter Scheidel’s interpretation of trends in inequality since prehistory. He writes that “four different kinds of violent ruptures have flattened inequality: mass mobilization warfare, transformative revolution, state failure, and lethal pandemics. I call these the Four Horsemen of Leveling. … There is no repertoire of benign means of compression that have ever achieved results that are even remotely comparable to those produced by the Four Horsemen” (Scheidel 2017) p.8-9.

Consistent with Scheidel’s hypothesis is the account by Peter Lindert and Jeffry Williamson of the abolition of slavery in the United States and the economic measures adopted by the Japanese emperor in the years immediately prior to and during World War II, including dividend and executive pay caps and a doubling of the marginal tax rate on high income groups. “In both the Japanese and Confederate crises” they write, “a polity that had been slow to liberalize had much of this top wealth suddenly confiscated and redistributed to those in the bottom 99 percent” (Lindert and Williamson 2016) p.198. Thomas Piketty also identifies political “shocks,” including two world wars and the Great Depression, as sources of what Lindert and Williamson term the “Great Leveling” of top incomes in the United States during the years 1910 to 1970 (Piketty 2013).

From these works we adopt the hypothesis that the evolution of economic inequality will reflect changes in the distribution of political power. Examples include the concentration of political power

---

14 Rothman and Peasnall (1999) offer an application of their approach to early Mesopotamian state formation.
during the Bronze Age Mesopotamia; the demise of slavery and feudalism and the rise of liberal
democracy as a form of rule; changes in the technology of combat such as the development of the stirrup
(whereby cavalry-based militaries became more effective) or the development of projectile weapons
(arrows or, later, handheld firearms termed the great equalizer); and changes in the domain of human
activity governed by private property rights.

An open question in this literature (which our data addresses) concerns the temporal sequence of
increased economic inequality and the concentration of political power (Scheidel 2013). Did an initial
emergence of heightened wealth disparities create an elite interest in the formation of states as a means
of protecting their wealth (Engels 2010 [1884])? Or did substantial wealth disparities appear only later
and as a consequence of state formation (Claessens and Skalnik 1978, Service 1978)?

3.3 Culture/ideology

Others scholars have suggested that inequality may be elevated and sustained by cultural change,
including the rise of “formalized ideologies in which the accumulation of wealth and high status were
seen as rightfully belonging to leaders” (Yoffee 2005)p.32.

The archaeologists Kent Flannery and Joyce Marcus hold that, following the last Ice Age, a process
of “surrendering bits of equality … grew out of perceived differences in life force, virtue, intellectual
property, generosity, debt, and prowess in combat” (Flannery and Marcus 2012) p.547. And they point
to “preexisting inequities that could serve as raw material. Included were the differences in prestige
between … people who had climbed the ritual ladder and those who had not, between the clan that
arrived first and everyone else; and between the man chosen for success by a demon, and lesser men”
((Flannery and Marcus 2012) p.547,554. Clifford Geertz represented power as something that is
symbolically performed by rulers in what he termed the “theatre state” in Bali (Geertz 1981). The
architecture of the sixteenth-century Mughal capital at Fatehpur Sikri appears to have enhanced the
ruler’s power by clearly demarcating spaces and elevations of ruled and rulers (Asher 1992).

In The Dawn of Everything, David Graeber and David Wengrow develop a sustained critique of
what they consider to be technological and environmental determinism, rightly noting (their italics) that
“while agriculture allowed for the possibility of more unequal concentrations of wealth, in most cases
this only began to happen millennia after its inception” (Graeber and Wengrow 2021) p.248. If we want
to understand the origin of economic inequality, they continue, we should be asking “what made
possible the emergence of kings, priests, overseers and judges” (p.75–76). This would seem to point to
political and institutional determinants of inequality, but they trace these proximate factors to underlying cultural causes: people lost the ability to “imagine and enact other forms of social existence” (p.484).15


We take from this literature the hypothesis that cultural practices may put a brake on the evolution of inequality, as with Hodder’s aggressive egalitarianism in Neolithic communities mentioned above, and that the long-term evolution of inequality may be affected by hierarchical orderings such as divine and mortal, by the mutual recognition of property rights even before the emergence of states capable of enforcing them; by writing and systems of measurement that reduce costs of communication across groups and distances; and by architectural features corresponding to cultural stratification and wealth inequality.

3.4 Complementary sources of inequality

The above influences on economic disparity—technology/ecology, institutions/politics, and culture/ideology—are sometimes taken as alternative sufficient explanations of the origin of enduring inequality. But the authors cited above do not advocate uni-causal explanations of something as complex as sustained inequality. They are better read as contributing factors that led to high levels of inequality and that were often complementary rather than simply additive.

This complementarity may have contributed to the resilience of egalitarian societies long after the emergence of farming. Egalitarian social norms—for example, the convention among some foragers that, upon being acquired, food should be widely shared beyond the immediate family—may influence the choice of technology, discouraging farming and storage even in environments where under different institutions both could contribute substantially to individuals’ livelihoods

---

15 See Bowles (2023) for reflections on the contributions of Graeber and Wengrow in their Dawn of Everything.
(Bowles and Choi 2013, Kaplan and Hill 1985, Woodburn 1982a). The result is that social norms could alter the evolution of technology in ways that limit economic inequality.

But far more common in the literature is the contrary case: technology/ecology, institutions/politics, and cultures/ideology acting as complements sustaining inequality. A highly productive and circumscribed natural environment—the lower Nile Valley 5,000 years ago, for example—makes the reservation option of residents very unfavorable. This contributed to the emergence and success of coercive state power, which in turn supported high levels of wealth inequality (Allen 1997, Carneiro 1970, Wengrow 2023). In like vein, Kirch advances the idea that the circumscription experienced by island populations may also have contributed to the formation of the first Hawai’ian state (Kirch 2010).

Similarly, Karl Wittfogel proposed that ecologies suitable for irrigated farming are more likely to support despotic states (Wittfogel 1957). While for the most part viewed skeptically by archaeologists, the Wittfogel hypothesis has recently found empirical support (Bentzen, Kaarsen and Wingender 2017). A related study by Robert Allen and his coauthors exploited shifts in the course of the Tigris and Euphrates Rivers as an identification strategy to provide convincing evidence that the provision of large scale irrigation facilities contributed to the concentration of political power in the region (Allen, Bertazzini and Heldring 2022).

Others have proposed explanations of the emergence of states based on Malthusian demographic and technological environments. Patrick Kirch models the ecology of state formation in Hawai`i with a critical role for diminishing marginal returns to labor in producing crops (which, incidentally, included sweet potatoes (Kirch 2010) p.186.) Gregory Dow and Clyde Reed also use a Malthusian model of production and population growth in their analysis of state formation and the emergence of substantial inequality between 15,000 and 5,000 years ago (Dow and Reed 2023).

Concerning early modern Europe, Daron Acemoglu and Simon Johnson combine technological and institutional factors in their investigation of the conditions under which, over the last millennium, technological progress contributed to inequality. They describe how a “series of technological improvements in medieval and early modern agriculture including better plows, smarter crop rotation, more use of horses and improved mills … tended to enrich and empower small elites while generating

---

16 The contributions of Dow and Reed are reviewed in Bowles and Bogaard (2023).
few benefits for agricultural workers: peasants lacked political and social power, and the path of technology followed the visions of a narrow elite” (Acemoglu and Johnson 2023) p.4, 6.

3.5 Why did farming not “cause” enduring inequality, and why did states arise so late?

The Engels-Morgan hypothesis—that the emergence of farming resulted in a surplus over the normal consumption levels of the producers, and this was captured by an elite—is appealing for its parsimony. But it does not work. The first reason is that farming, at least initially, probably did not raise labor productivity, that is, caloric output per hour of work (Bowles 2011). The first farmers (not yet living under elite domination) tended to be smaller and less healthy than the hunter-gatherers they replaced (Cohen and Crane-Kramer 2007). The new technology—farming—did not by itself create a surplus.

The second reason to doubt the Engels-Morgan hypothesis was noted in passing (but not addressed) by Gordon Childe, an important archaeologist of the emergence of stratification in western Eurasia: If farming did indeed raise labor productivity, why did farmers not reduce their hours of work? Childe wrote that the origin of the surplus

was initiated in the alluvial valleys of the Nile, the Tigris-Euphrates, and the Indus about five thousand years ago … [when] society persuaded or compelled the farmers to produce a surplus of foodstuffs over and above their domestic requirements, and by concentrating this surplus used it to support a new urban population of specialized craftsmen, merchants, priests, officials, and clerks (Childe 1946) p.18.

Melville Herskovits, who like Childe saw a surplus over customary consumption as a key development in explaining social stratification, remarked: “Just why the surplus is produced remains obscure” (Herskovits 1952) p.413.

---

17 Addressing the puzzle posed by erstwhile hunter-gatherers taking up a technology that reduced their caloric productivity would take us too far afield. A possible explanation is modeled, explored empirically, and simulated in Bowles and Choi (2019) and Bowles and Choi (2013).
18 The Anatolian site of Çatalhöyük to which we refer below appears to be an exception (Larsen, et al. (2015).)
19 A variant of the surplus interpretation of inequality proposed by Branko Milanovic observes that the maximum feasible degree of inequality is limited by the extent of potential output above the population’s necessary subsistence level, without inferring that that level of inequality or surplus production will be realized. Milanovic (2013). The capacity to produce a surplus in this view is thus a necessary but not sufficient condition for inequality.
Harry Pearson, in a paper titled “The Economy Has No Surplus,” stressed the need for a political explanation, observing that “there are definite institutional requirements if surpluses are to be made available for specific purposes” (Pearson 1957) p.335. Robert McCormick Adams pointed to both cultural and political origins of a surplus: “It is precisely the growth of the collective symbols and institutions of the primitive states that can explain the conversion of peasant leisure into foodstuffs in urban storehouses” (Adams 2005) p.45.

Important archaeological studies of the few cases of pristine states (meaning those that arose independently of other states) that might shed light on this puzzle have been written about the Zapotec (Mexico) civilization between 500 BCE and 200 CE (Marcus and Flannery 1996) and Hawai’i in the eighteenth century (Kirch 2010). Consistent with Tilly’s cultural group selection explanation of the formation of the modern national state in Europe (mentioned above), both of these cases suggest a key role for conflict between previously autonomous polities. Marcus and Flannery, who chronicled the formation of the Zapotec state, recognized the similarities: “The chieftdom of Hawai’i did not ‘evolve into a state’ … states arise when one member of a group of chieftdoms begins to take over its neighbors, eventually turning them into subject provinces of a much larger polity” (Marcus and Flannery 1996) p.157.20

Consistent with the crop appropriability and ecological circumscription reasoning above, Stasavage has proposed that “early autocracy”—the centralized pole of Wolf’s continuum of rule in the tributary mode of production—arose where environmental conditions “made agricultural production more predictable, … when people had fewer means of exiting” and where a preexisting bureaucracy could be repurposed. Extending the themes first developed in Scott’s Seeing Like a State (Scott 1998), he points out that the information asymmetries that weakened early rulers relative to their subjects were mitigated by “new technologies for writing, mapping, and measurement” (Stasavage 2020) p.79.

Childe’s pathway to stratification whereby -- “society… compelled the farmers” -- was recently echoed by Acemoglu and Robinson. They conclude that:

in order for sedentary life to emerge it … seems plausible that hunter gatherers would have had to be forced to settle down, and this would have to have been preceded by an institutional innovation concentrating power in the hand of a group that would become the political elite,

---

20 Early state formation in Mesopotamia is similar in this respect to the Hawai’ian and Zapotec cases according to Flannery and Marcus (2012), who describe it as the “product of the dynamic competition and alliance-building among several urban societies.”
enforce property rights, maintain order, and also benefit from their status by extracting resources from the rest of society. (Acemoglu and Robinson 2012)p.139

Dow and Reed similarly suggest that the source of heightened within-population inequality may have been the emergence of “technologies of confiscation where an organized elite can seize the goods, land or labor from unorganized commoners” (Dow and Reed 2023) p. 217.

But these accounts are not entirely satisfactory, for two reasons. First, no mechanism is suggested to explain this surprising redistribution of political power (why would the technologies of confiscation prevail against an “aggressively egalitarian” populace?). Neolithic and Bronze Age farmers still hunted, and using projectile weapons they could safely kill would-be usurpers of power at a distance. Disarming or subduing a population so as to allow their exploitation by state or private elites would have been difficult.

More importantly, however, the timing is not consistent with the Acemoglu-Robinson hypothesis. Based on data from both Mesopotamia and Mesoamerica, archaeologist Henry Wright concludes that, in the cases where state-like political organizations eventually did develop, “class-organized sociopolitical organization flourished for centuries before actual state formation” (Wright 1994)p.81. Substantial increases in wealth inequality—as Figure 1 makes clear—occurred beginning in the Bronze Age, that is, prior to the formation of states and the effective political mechanisms for elite coercion that they provide. The process leading to the production of an economic surplus therefore must have involved the active acceptance of novel forms of social organization by at least some of those who would not (in the long run, at least) be the primary beneficiaries of the changes.

The much later development of states entailed an analogous concentration, in this case of coercive powers. This, too, poses a puzzle, but one that is perhaps more easily resolved in light of the (by then) substantial preexisting wealth inequality and the collective benefits of the state as a coordinating device including, as we have seen, the provision of public goods and irrigation on a substantial scale (Allen, Bertazzini and Heldring 2022, Fargher and Blanton 2017, Fargher, Heredia Espinoza and Blanton 2011). Norman Yoffe and Gojko Barjamovic suggest that a degree of concentrated political power was crucial to providing the institutional and physical infrastructure of private trade at Assur. But they write that “traders moved goods for centuries before centralized states appeared in Mesopotamia” (Yoffee and Barjamovic 2018) p.22.
What is not easily explained, then, is why the concentration of political power into state-like forms occurs so late, long after the emergence of large villages like Çatalhöyük in Anatolia and Abu Hureyra on the Euphrates River. Governing such settlements without political centralization would have been an enormous challenge, even (as appears to have been the case at Abu Hureyra) where substantial wealth inequality was absent. The fact that readily appropriated crops such as cereals and circumscribed environments such as irrigated farmlands are favorable for the formation of states does not resolve the puzzle of the late emergence of state-like political organizations in western Eurasia, where cereal production and irrigation predates the formation of the first proto-states by many millennia.

What does the archaeological record provide by way of clues to the surprisingly limited concentration of political power and modest levels of wealth inequality prior to the Bronze Age?

4. Failed experiments in wealth inequality among hunter-gatherers and Neolithic farmers

Henry Wright referred to early Bronze Age and Iron Age states as “political experiments” (Wright 2006). Most of them were short-lived. Cases of substantial wealth inequality prior to what we term the first proto-states appear to have been even more episodic. Many populations during this period may have “flirted with greater inequality,” to quote Flannery and Marcus; “most of those flirtations, however, ended with a return to egalitarian behavior” (Flannery and Marcus 2012) p. 551. Understanding the instability of wealth inequality during the Neolithic will provide some clues as to how extraordinary and enduring disparities eventually emerged.

4.1. Episodic inequality prior to the Bronze Age

Even more striking than the opulence evident in the Sungir burials mentioned at the outset is how rare it is among the known cemeteries dating from the Late Paleolithic. Burials with fox or wolf teeth or shells and ochre coloring were common (Aldhouse-Green and Pettitt 1998, Vanhaeren and D'errico 2005), but the twenty-nine Late Paleolithic burial sites (including Sungir) mentioned in the standard survey report no other evidence of remotely comparable wealth (Pettitt 2011). Within the borders of modern-day France alone, forty-six cemeteries dating from a single 5,000-year period at the end of the Paleolithic have been excavated (Jeunesse 2021). In none have similarly opulent burials been found.

But not far from Sungir on the Russian Plain the size of dwellings and location of storage pits during this period does provide evidence of what archaeologist Olga Soffer terms “the origins of nascent forms of economic inequality” (Soffer 1989) p.727. She studied twenty-nine settlements of a partially
sedentary population of mammoth hunters who constructed large subterranean storage facilities. A comparison of the locations of storage pits in early sites, such as Radomyshl, to those in later sites, such as Dobranichevka, suggests the emergence of wealth inequality, but not its persistence. Soffer writes: “The central location of Radomyshl’s pit argues for open, visible, and equal access to stored resources for all inhabitants of the site. Resources at Dobranichevka, [a later site] on the other hand, were distributed around each dwelling and this distribution implies a restriction of access to the pits to the residents of the dwelling only. … [A]ccess to resources was controlled by each household [and] the amount per household was fairly equal” (Soffer 1985) p. 462.

At another site, Mezin, this apparent privatization of storage was associated with unequal holdings: six out of eight pits were clumped near [a single dwelling that] contained an abundance of portable art and of exotic jewelry, including fossil marine shells, originating some 800 kilometers south of the site. At another site, Mezhirich, Soffer found the standard archaeological traces of emerging political hierarchy: “site-wide planning” and “unified conception and construction” along with “monumental architecture” p.463. But surprisingly, Soffer found that “sites dating since 18,000 BP on the Russian Plain [including the Sungir area] … have not yielded any examples of burials accompanied by burial goods” p. 57. In subsequent periods, groups became smaller and more mobile. “The observed changes should be seen as fluctuational,” she concludes p. 475.

“Fluctuational” inequality is also evident in a newly available time series of wealth inequality spanning Early Neolithic to Iron Age Thailand, from 2000 BCE to 600 CE (Fochesato, Higham, Bogaard, et al. 2021). Gini coefficients estimated from burial goods rose by more than 50 percent with the advent of the Bronze Age in the adjacent villages of Ban Non Wat and Ban Lum Khao, which were situated in a position astride a route for long-distance exchange of copper, marble, and marine shells. The authors explain that “it was by securing control over the ownership of and access to these exotic goods, including the novel and remarkable first metal, that an aggrandizer segment of Ban Non Wat society secured elite status…. At sites … without control over exchange routes all of the Bronze Age

---

21 Soffer (1989) p. 727. If the excavated storage pits were all that existed for the six households that occupied the excavated dwellings, the apparently wealthy household controlled 96 percent of the village storage capacity. We calculated the Gini coefficient for dwelling size at Mezin: $G = 0.509$. For the other villages (taken as a whole), the dwelling size Gini coefficient is 0.224 Soffer (1985)121). These Gini coefficients are computed from the dwelling size adjusted to be comparable to the other estimates in our data set. The unadjusted estimates calculated from raw data on dwelling size are 0.31 and 0.14, respectively.

22 The Bronze Age in the Thai data spans 3000–2370 BP.
dead were comparatively poor” (Fochesato, Higham, Bogaard, et al. 2021) p. 5. But the affluence of the Ban Non Wat elite did not last: “It endured for about eight generations before there was a sharp decline in mortuary wealth, as doubtless, bronze became more readily available”(Fochesato, Higham, Bogaard, et al. 2021) p. 6.

Archaeologists have not determined the sources of the wealth of the Sungir elite adult, or of the children buried with him, or of the household that apparently came to control virtually all of the mammoth hunters’ storage at nearby Mezin, a few millennia later. But the origins of nascent wealth inequality in Bronze Age Thailand are clear: a group was able to profit from new commercial opportunities that were not available to others. The long-distance trade from which the Ban Non Wat elite profited required a substantial amount of working capital to purchase and transport the goods initiating an exchange. The better-off Neolithic households probably first engaged in this trade, and then used the resulting profits to finance further exchanges. The combination of a minimum project size with what were probably significant economies of scale (in transporting and defending the goods to be exchanged) meant that the limiting factor in acquiring wealth was location plus material wealth itself. Similarly, the financing of long-distance trade in Old Assyria would have been prohibitive to all but the relatively wealthy (Yoffee and Barjamovic 2018).

A very similar process appears to account for the beginnings of substantial wealth inequalities in another of the few archaeologically documented transitions away from a relatively egalitarian wealth distribution. This is the 1,500-year archaeological record on the Chumash, a maritime hunting-gathering people on the Northern Channel Islands off the coast of Southern California (Arnold 1993, Arnold and Graesch 2004, Graesch 2004, Kennett 2005). Prior to 1150 CE “nothing in the archaeological record … provides any reliable sign of chiefdom-like organization…or hereditary wealth and leadership” (Arnold and Graesch, 2004) p.6. But over the next century and a half the Chumash appear to have developed both hierarchical political structures and substantial wealth inequality. The new social structure -- what the archaeologists Jeanne Arnold and Anthony Graesch (p.3) describe as “ascribed status and elite leaders’ control of non-kin labor” -- is evident in the emergence of status differentiation in burials.

A preceding period of adverse climatic conditions had reduced maritime productivity (likely depleting the stocks sea mammals, mollusks, and fish on which the Chumash depended). The islanders eventually re-established their livelihoods by constructing large, sophisticated redwood-plank ocean-going canoes (some with a capacity of two tons) and developing a highly specialized industry producing beads fashioned with microliths. The beads were exported across forty kilometers of open
ocean to the mainland in return for food and exotic prestige goods. Some of the beads came to be used as currency on the mainland. “The Island Chumash, the sole manufacturers of the region’s chert microdrills [that were used to make the beads], became the operators of a ‘mint’” (Arnold and Graesch 2004) p.7.

The microliths and beads were made by independent households specializing in these pursuits, and “there is no archaeological or ethnographic evidence to suggest that elites had the power to physically coerce specialist labor” (Graesch 2004) p.135. But “aspiring elites controlled exchange of these valued products through their assumption of the responsibility of managing construction and ownership of plank canoes … [which] was never within the means of commoner households. … [A] small group of canoe owners seemingly formed the nucleus of a rising elite” (Arnold 1993) pp.105, 107. A political structure based on “hereditary chiefly families” emerged. There is also evidence that a religious association may have been developing at about this time or possibly somewhat earlier (Arnold and Graesch 2004) pp. 3, 5.

Most households engaged in some bead production, but one dwelling thought to be the residence of a canoe owner engaged less than the other dwellings in the village (Graesch 2004) p.164. The key to wealth apparently was not producing the beads but controlling the terms on which they were exchanged for mainland goods, and this required ownership of a canoe. What apparently allowed Chumash elite to set themselves apart from other islanders was an improbable confluence of two environmental conditions. First was the collapse of the productivity of their conventional maritime livelihood (which was the fallback reservation option for less well-off islanders). Second was the presence of the chert and shells necessary to make the beads on an island sufficiently distant from the mainland so that trade was possible only by means of highly capital-intensive oceangoing canoe transport subject to increasing returns to scale.

A similar set of unlikely (and in this case unsustainable) circumstances may account for our final example of episodic inequality prior to the emergence of proto-states: the salt producers and traders at Provadia near the Black Sea during the period 5400 to 4200 BCE. There, a village that had earlier subsisted entirely on farming einkorn, emmer, and hull barley included a “factory” for transforming naturally occurring salt brine into ingots that were traded throughout the Balkans and, like the Chumash beads, were used as a currency. The single oven that has been excavated to date is estimated to have produced ten tons of ingots annually. During the period that the site was occupied, the technology of ingot production was dramatically improved, vastly increasing the output per hour of labor. “The riches,
amassed by the salt producers, had to be defended; this is why the … settlement was fortified,”
including an encircling three-meter-high wall (Nikolov 2009) p. 125.

While estimates of the degree of inequality in the value of grave goods in local burials are not available for Provadia, the presence of unusual levels of exotic prestige goods in some burials is unprecedented for the time and region. Moreover, the sheer scale of the salt ingot factory and the challenge of supplying it with specialized workers from the surrounding area and also with firewood from the ever-receding forests suggest a degree of organization atypical of the farming communities in the region. Vassil Nikolov, the leading archeologist studying Provadia, concludes that “without a hierarchically structured authority, such a system could have not existed” (Nikolov 2010) p.60.

The brine springs dried up sometime toward the end of the fifth millennium BCE, forcing what Nikolov calls the “nomadization” of the population, repeating the pattern noted by Soffer on the Russian Plain much earlier. The site remained uninhabited for the next two millennia. But quite apart from the dry spell that ultimately doomed the salt producers, Nikolov believes that the wealth inequalities associated with the production of ingots would have been unsustainable: “The accumulation of substantial wealth in a small area would sooner or later have caused pressure for its redistribution” (Nikolov 2010) p.62.

Pressures for redistribution, in turn, may have provided the impetus for institutional innovation. Nikolov refers to the Provadian “hierarchical structure with organization management functions” as a “variant of the politogenesis” that was characteristic of communities in the Balkans at the time. Had the Provadia brine springs not dried up and had the salt works continued, Nikolov speculates, “authority there would have been institutionalized … and a pre-state would have been created” (Nikolov 2010) p.62. But this did not occur. Like the emergence of wealth inequality ten millennia earlier on the nearby Russian Plain, the political and economic inequalities evident in the Provadia area remained exceptional for this period and region.

4.2. Limited inequality during the Late Paleolithic and Early Neolithic

We have few comparable quantitative estimates of wealth inequality in western Eurasia prior to 6700 BCE (the end of what archaeologists of western Asia term the Pre-Pottery Neolithic B, or PPNB), but other archaeological evidence is consistent with our conjecture that Late Paleolithic and Early Neolithic wealth inequalities were limited and ephemeral. We begin with the Natufian period (12,300–
10,300 BP), during which this Levantine population harvested but did not cultivate wild species of cereals. (Some have called the Natufians proto-farmers (Bar-Yosef 1998, Unger-Hamilton 1989).)

Because (as we have seen) mortuary practices leave archaeological traces and often reveal striking differences in wealth and social status, most of the evidence that might reveal economic inequality among Natufians is based on burials, some of which early in the Natufian period were decorated in ways differing across sites (Belfer-Cohen 1995, Byrd and Monahan 1995, Kuijt 1996). However, contrary to earlier work based on incomplete evidence, Anna Belfer-Cohen writes, “evidence for social stratification in the Natufian inferred from the decorated burials is … non-existent” (Belfer-Cohen 1995) p.16. The grave goods in the earlier period were almost entirely personal ornaments, which increased in use during the later Natufian period but, as Belfer-Cohen shows, virtually disappeared from burials, consistent with a culture placing little value on differences in material wealth.

Brian Byrd and Christopher Monahan focus on early Natufian evidence, “since later Natufian burials are characterized by a virtual absence of mortuary elaboration (particularly with respect to grave goods and construction techniques).” They conclude that “there is no strong mortuary evidence for hereditary social inequality in the Natufian” (Byrd and Monahan 1995) p. 280, 251.

A particularly well-studied site is Abu Hureyra on the Euphrates where, excepting a hiatus very early in the Neolithic, evidence spans the entire period from initial settlement as a Natufian community of sedentary hunter-gatherers to the adoption of the full Neolithic farming package (Moore, Hillman and Legge 2000). Like Byrd and Monahan, A.M.T. Moore and his coauthors provide evidence of burials under the floors of dwellings. They write that “the houses were family dwellings … a family could lay claim to the space its house occupied … and its descendants could build a new house on the same spot in which to live…. Family rights to private property were firmly established in Abu Hureyra 2” (Moore, Hillman, and Legge 2000) p. 505.

This evidence comes from the PPNB (8800–6500 BCE), well into the Neolithic, when Abu Hureyra was already a farming community, but as the following passage from Moore et. al. indicates, it is consistent with the view that the advent of both cultivation and private property did not initially entail significant economic and political differences among the resident families. “The similarity in the houses across the site and the lack of differentiation between the burials of each sex suggest that in material terms the villagers were of similar status…. Abu Hureyra 2 seems to have been an egalitarian community” (Moore, Hillman, and Legge 2000) p.505. Furthermore, they note:
There was no indication from the burials that they had developed a social system based on a hierarchy of classes that was maintained from one generation to the next. Abu Hureyra, then, was an unusually large, early Neolithic village. It had not developed all of the characteristics, for example, substantial public buildings, a social hierarchy, and large-scale trade, that we associate with the towns of early historic times in Southwest Asia. (495)

The evidence thus suggests a marked absence of sustained substantial wealth inequality during the final millennia of the Paleolithic and Early Neolithic. Moreover, it appears that this was not simply an egalitarian default state inherited from a hunter-gatherer past in which self-aggrandizement and the concentration of wealth had not yet become part of the repertoire of human behavior. Instead, we think it likely that substantial disparities were rare and transient because people actively resisted the concentration of wealth and power, deploying a set of practices that appear designed to thwart self-aggrandizement. Many of these equalizing practices (for example, those pertaining to burials, which we discuss below) were novel at the time, but there are analogous practices in the ethnographic record (Boehm 1993, Knauft 1991). Referring to early farmers, Christopher Boehm writes: “their egalitarianism, like that of hunters, is the product of constant vigilance and sometimes very harsh sanctioning” (Boehm 2000) p.13. Bruce Trigger and Graeber and Wengrow provide detailed accounts of this “aggressive egalitarianism” among the Iroquois at the time of contact with Europeans (Graeber and Wengrow 2021, Trigger 1990).

4.3 “Aggressively egalitarian” communities during the transition to farming

Reflecting on the disappearance of decorated burials and grave goods in the Late Natufian, Belfer-Cohen asked rhetorically: “if decorated burials are a symbol of social ranking, what happened [to explain why] they disappeared in the Late Natufian?” (Belfer-Cohen 1995) p.15. The archaeologist Ian Kuijt proposes that Late Natufian and Early Holocene mortuary practices were part of “a system of social codes for limiting the development and centralization of power and authority within early agricultural communities” (Kuijt (1996) p.332. He writes that this “egalitarian … belief system was materially expressed through … the control and restriction of the display of material differences (lack of grave goods, homogeneous grave construction) and … the development of mortuary rituals that emphasize a community of identity and a shared ancestor” (p.331).

Among these rituals is the widely observed Late Natufian and Pre-Pottery Neolithic practice of secondary burials. Following an initial individual interment, skulls were detached and, in many cases,
cached in a second, common burial, Kuijt explains, “as a means of recognizing a collective shared past and identity” (p.322). Kuijt also reports a consistent pattern in “residential architecture with little variation in shape, size or organization of structures” (p.328).

Kuijt concludes that “in the face of considerable economic and social change from 11,000 to 9300 BP, … social codes were expanded and increasingly standardized within the Levantine region to reinforce a shared community ethos and limit the development of social inequality” (p.332). The changes that “threatened the status quo” included, first, the harvesting and then cultivation of wild cereals; the later domestication of cereals as well as of sheep, goats, and eventually cattle; and the formation of large villages. These shifts in livelihoods facilitated household-level storage (including the ownership of long-lived animals), which would have allowed the more successful families to withdraw from systems of smoothing consumption by means of multilateral food sharing and to convert other previously egalitarian risk-pooling practices such as shared ownership of animals to hierarchcal patron-client relationships.23

Another example of egalitarian practices under pressure comes from the Early Neolithic population at Jerf el Ahmar on the Middle Euphrates, who originally stored their harvest of wild cereals in subterranean communal buildings. Much of their food processing and consumption likely were initially strongly communal practices. The evidence in Figure 1 suggests that the distribution of wealth was initially highly egalitarian (a Gini coefficient of 0.158 in 9000 BCE). However, it appears that Jerf el Ahmar and similar communities in the Early Neolithic made a transition from communal food production, consumption and storage to private, individual household based systems. Reflecting on the evidence from Jerf el Ahmar and elsewhere in the region, Alexander Weide and his coauthors write:

The general absence of grave goods and a lack of indicators for the individual accumulation of material wealth during much of the Levantine Pre Pottery Neolithic situate this communal organization of cereal exploitation in a pan-regional socio-economic system that strongly emphasized group cohesion and identity at the expense of individuality. … The ‘kitchens’ at Jerf el Ahmar, rooms with accumulations of food processing and preparation implements that are only found in some buildings, … indicate that food storage, processing and consumption

23 This is observed in the ethnographic literature, for example, among the Pimbwe hoe-farmers in Tanzania, where a study of sharing networks found that “the wealthy can (unlike the poor) afford to eschew risk pooling” (Kasper and Borgerhoff Mulder (2015) p.714. Among the Sumba Island traditional agropastoralists in eastern Indonesia, those wealthy enough to own water buffaloes, horses, or pigs do their own individual consumption smoothing by bartering these for rice in times of poor harvest (Adams (2016).
retained a strong communal element.... This changed with the gradual emergence of household-centered farming communities since the early and middle PPNB between ca. 10,700–9,500 cal. BP, which developed increasingly ‘private’ systems of food storage, processing, and consumption (Weide, Green, Hodgson et al. 2022) p.630-631.

Tellingly, the last communal building constructed at Jerf el Ahmar did not contain any storage facilities. Stores may have migrated to individual household dwellings, as appears to have been the case among the mammoth hunters on the Russian Plain.

Across the region, Karen Wright documents a similar transition from “the open easy-access [outdoor] patterns of the PPNA; to visible kitchen porches bridging house and neighborhood in the mid-PPNB; to a retreat of cooking and storage facilities to more sheltered, private spaces in larger, more enclosed households in the Late PPNB” (Wright 2014) p.7. While the evidence is necessarily indirect, it appears that, over this period, the domain of private household property—the legitimate and effective right to exclude others from using the objects and spaces one owns—tended to expand (Halstead 2006).

A well-studied illustration of the threat to egalitarianism posed by the expanding domain of private ownership is the Early Neolithic hoe-farming, sheep-and goat-herding, and hunting community of Çatalhöyük.24 The village was formed at the end of the eighth millennium BCE by people who had previously inhabited small scattered settlements on the sparsely populated Konya Plain in Anatolia. There, according to Ian Hodder, the leading contemporary archaeologist on this settlement, “the focus on collective sharing and non-accumulation by individual units was already under threat at the start of the occupation” of the village (Hodder 2014) p.17. Hodder asks: “What held an egalitarian community of such a size together for over 1,000 years?” (p.6).

Similar to the “egalitarian” Natufians, burial practices at Çatalhöyük included secondary burials and removal and common caching of skulls, as was evident from what Hodder terms “the circulation of human body parts and joint burial in multiple burials” below the floors of a family’s dwellings. Under the floors of some houses these exceeded 10-15 deceased individuals during a single phase (Hodder 2014) p.8. Dental evidence indicates “that inclusion for [burial] within a house is only minimially related to biological affinity” (Pilloud and Larsen 2011) p.519, suggesting to Hodder “the separation of children from biological parents at some point soon after birth” (Hodder 2014) p.9.

24 Our data and references are for the site called Çatalhöyük East in the archaeological literature.
Beyond evidence for these mortuary practices, two studies of food preparation and consumption at Çatalhöyük illustrate Hodder’s “aggressive egalitarianism.” Amy Bogaard and coauthors suggested that the ceremonial communal sharing of the meat of hunted animals may have worked to at least partially offset the privatization of storage and food processing. In addition to storage bins for cultivated cereals deep in the interior of dwellings at Çatalhöyük, in many cases there was also a large room with displays of (initially wild) cattle heads and horns, in which Bogaard et al. conjecture that extrafamilial feasting took place. They propose that “concealment of stored plant food within the house was balanced against visible installations of animal parts that plausibly refer to the sharing of meat beyond the household…. The divisive effects of private storage were subverted through the sharing of cattle” (Bogaard, Charles, Twiss et al. 2009) p.664–65.

This pattern of sharing food acquired in substantial amounts at a given time (such as meat from a large animal or honey) while reserving plant foods and other small package items for immediate family consumption is widely observed among hunter-gatherers in the ethnographic record (Kaplan and Hill 1985). The increasing importance of cereals as a staple may have shifted the balance from public communal sharing of food to private household consumption.

A second study by Karen Wright documented the distribution of food-processing tools among Çatalhöyük households, ranging from small grinding stones to large millstones (called querns), some fashioned from boulders transported from thirty-five kilometers away (Wright 2014). The smaller tools were found in most households. The querns were expensive to produce and were exceptionally concentrated—only 30 percent of households had one, and a single dwelling, Building 77, described as “elaborate,” held two-thirds of them. Querns had potential use lives of a century or more. Much later in the Bronze Age, they were included in dowry and inheritance documents. Wright asks, “Were they transmitted through time? At Çatalhöyük, … clearly not.”

[Large] millstones were deliberately destroyed at house abandonments, before they showed heavy wear…. They were special targets for destruction which would have required substantial effort. The driving force appears to have been social, not practical. If there were forms of inheritance and material wealth transmission, millstones were not part of this. Breakage patterns seem an emphatic “statement” against transmission of these artefacts across generations (Wright 2014) p. 27.

The threats to egalitarianism that Hodder mentioned mounted after 6500 BCE, when the more elaborate dwellings “invested less in ritual ties … and focused more on independent production and the
build-up of their own surpluses. The increased focus on domestic production led to heavier investment in sheep herding and in the adoption [after 6400 BCE] of domestic cattle.” It appears likely that after 6500 BCE, sheep that had been raised communally in common herds came to be tended by individual households. “The presence of a large amount of meat (and milk) on the hoof may have allowed a slight shift from the collective pooling of resources” (Hodder 2014) pp. 11, 16.

At about this time, the “elaborate” Building 77 at Çatalhöyük burned at very high temperature, without having been emptied of the household’s tools and stores; it was subsequently abruptly abandoned. One of the few other dwellings with a quern, the also “elaborate” Building 52, was also destroyed by fire. These and other burnings could have been accidental, but intentional burning as an egalitarian leveling practice is also considered to be a possibility. Concerning Tepe Gawra, another Mesopotamian site in our data set, M. S. Rothman and B. Peasnal write, “Perhaps it was competition for … new exchange opportunities or internal resentment that caused the fire that ended Gawra’s fourth millennium BC occupation” (Rothman and Peasnall 1999) p.114.

Similar practices that reduce wealth inequality by interrupting the intergenerational transmission of wealth within households are observed among ethnographic hoe-based farmers. An important example is the proliferation of lineage- or other group-held assets, including rights to farm communally held land. Reflecting on the importance of the introduction of plow agriculture in ancient western Eurasia, Andrew Sherratt (following Jack Goody) writes, “Typical of African hoe cultivators, property reverts to the lineage of a deceased person and is not passed to his or her offspring. [In contrast,] … plough cultivation marriage produces successive recombinations of holdings in each generation… [and] land becomes a major heritable commodity” (Sherratt 1981) p. 298,. Goody reports that in The Ethnographic Atlas the kin terminology of a majority of the populations practicing hoe farming does not distinguish siblings from cousins (Goody 1976) p.19. The convention of inheritance from parents to their own children is absent in more than two-thirds of the eighty-three horticultural (hoe farming) societies in the Standard Cross Cultural Sample (Gurven, Borgerhoff Mulder, Hooper et al. 2010) p.51–53. In many ethnographically recorded societies, funeral feasting and bride wealth transfers—sometimes on a massive scale by those with substantial wealth—effectively reduced the intergenerational transfer of wealth within families and hence limited the sustainable degree of wealth inequality (Adams 2016, Hayden 2009, Parker Pearson 2000). The previously mentioned econometric estimates of the

intergenerational transmission elasticity (β) are consistent with these ethnographic observations. Among three horticultural populations for which there are data on parental and offspring material wealth, the mean integenerational wealth transmission elasticity was 0.09 [-0.08; 0.26]. The corresponding figure for farmers is 0.55 [0.41; 0.68] (Borgerhoff Mulder, Bowles, Hertz, et al. 2009).

Estimates of the Gini coefficient of wealth inequality in Çatalhöyük (based on house size) show a marked increase from 0.206 in 6650 BCE to an average of 0.499 between 6475 and 6150 BCE. But even this latter figure is somewhat limited compared with what was to follow. Modern levels of inequality were not attained in Çatalhöyük, possibly because, though under pressure from the changes brought about by farming, egalitarian practices persisted over a millennium before weakening. Wright summarizes the trajectory:

Transmission of material wealth through time would have also been a key threshold in the entrenchment of inequality. Çatalhöyük ... seems not to have reached that threshold. In sum, the data from Çatalhöyük ... hint at a society trying to maintain egalitarian principles even as it was slowly changing into something else. Mechanisms for maintaining egalitarianism may have encountered problems as the site grew large, leading to abandonment. (Wright 2014) p.29

The abundance of nearby land that could be farmed or grazed may have also limited the extent to which substantial wealth disparities could be sustained at Çatalhöyük. The apparent dispersal of shepherds to seasonal satellite camps, the marked population decline after 6500 BCE (Cessford 2005, Hodder 2014, Kuijt and Marciniak 2022), and the eventual abandonment of the site are not surprising given that, as Wright put it, “at Çatalhöyük land scarcity was not a problem; the site was almost alone in the Konya Plain” (Wright 2014) p.12. In this respect, Çatalhöyük was apparently not atypical of contemporary large villages in the region, according to Hodder: “The seventh millennium BC more widely in the Middle East is a time at which the Neolithic village pattern reaches its apogee in a series of mega-sites, similar in size to Çatalhöyük, many of which decline” (Hodder 2014) p.17.

But for the most part, the egalitarian practices we have described survive at least until the late Neolithic. Across the Near Eastern early and mid-Neolithic, the small scale of storage facilities is consistent with contemporaneous multilateral food sharing continuing to be the primary consumption-smoothing practice at that time. Ian Kuijt writes: “The drastic increase in the scale of storage in the later Neolithic was a materialization of a new economic emphasis on storage as a solution to unpredictability” (Kuijt 2017) p.85. Given the scale of storage required for effective consumption smoothing without
drawing upon communal risk pooling, it is likely the increase in storage occurred primarily among the wealthy.

Referring to early Mesopotamia as a whole, Flannery and Marcus cite examples of “the burning of elite residences in raids” (Flannery and Marcus 2012) p. 261, 272, 279 including at Tepe Gawra during a period of substantial inequality (Level XII, see Figure 1), just prior to the likely formation of the first archaic proto-state there (see the discussion below). Paul Halstead finds that in Neolithic Greece “deliberate” burning of houses and “breakage of grinding and edge tools … may have served as a leveling mechanism, limiting transmission of any accrued inequalities between generations” (Halstead 2019) p. 63. Also in Greece, he observed a privatization of food storage and eating spaces similar to that which appears to have taken place at Çatalhöyük: “during the … Early and Middle Neolithic … hearths were often located in open spaces between ‘houses’ implying that food was frequently cooked in public, inviting neighborly pressure to share,” but he notes that “Final Neolithic and Early Bronze Age cooking facilities were more consistently secluded indoors” (Halstead 2019) p. 58.

4.4. The dynamics of episodic inequality in a simulated aggressively egalitarian population

Our reading of the evidence on Late Paleolithic and Early Neolithic egalitarian practices in southwestern Asia, as well as the cases of substantial inequality at other sites—the mammoth hunters on the Russian Plain, the Chumash islanders, the Bronze Age Thailand traders, and the Provadia salt producers—is that the conditions supporting substantial wealth inequality may have been rare and ephemeral among both Neolithic farmers and ancient sedentary hunter-gatherers. We do not have data on prehistoric foragers (mobile hunter-gatherers), but it is unlikely that they would reveal more substantial or enduring inequalities than among the late Paleolithic and other sedentary hunter-gatherers for which we do have evidence.

To better understand the economic and social structures that might have supported the observed pattern—persistent low levels of inequality with occasional but apparently short-lived excursions to greater disparities—we draw on an evolutionary model of a hypothetical population similar in size and composition to a Late Paleolithic ethnolinguistic unit (600 individuals per generation), made up of many subpopulations (of twenty individuals per generation) about the size of forager bands or small villages among which there is occasional migration. The model and its parameterization were developed to understand the Early Neolithic transition from hunting and gathering to farming; when simulated along
with climate data over the period 40,000 BP to the present, it closely replicates the observed archaeological record.  

Here, we use the model for a different purpose: to explore the dynamics of wealth inequality in a Late Paleolithic or Neolithic population in which both aggressive egalitarianism and self-aggrandizement are behavioral options, reflecting patterns of self-interest, sharing, and collective discipline found in ethnographic studies of hunter-gatherers and horticulturalists (Boehm 1993, 2000, Hayden 2001, Kaplan and Gurven 2005, Kaplan and Hill 1985, Wiessner 2005). We will find that in simulations of the model, societies that are highly egalitarian over long periods sometimes experience excursions of substantial but short-lived inequalities.

The simulated agents may either farm or hunt and gather, and when they interact with others in their subpopulation, the strategy that each follows determines how the goods they have acquired are divided up. There are four stages in the model: production, distribution, cultural updating, and intergroup competition. In the production stage, individuals adopt one of two technologies: farming or hunting-gathering, acquiring a product as a result. In the distribution stage, each member of the population is paired randomly with another member of their subpopulation. In these paired interactions, individuals may adopt one of three behavioral strategies.

The first behavioral type, called the sharer, is similar to the dove in the Hawk-Dove game about sharing and self-aggrandizement (a variant of the Chicken game): a sharer concedes half of the product to the other, or the whole product if the other claims it. When two sharers are paired, they exchange equal shares of their product and as a result have equal payoffs. A second behavioral type, the grabber (similar to the hawk), claims the entire product, successfully if paired with a sharer, but, if paired with another grabber, a costly conflict results, resulting in negative expected payoffs. The third (novel) behavioral type, the civic, acts exactly like a sharer when paired with a fellow civic or a sharer (they share), but when paired with an individual who refuses to share (a grabber), they join with any other civics in the subpopulation to contest the claims of the grabber, succeeding with a probability that increases with their numbers. If they succeed, they distribute the product equally among all the civics.

---

26 The model and its simulations are described in the text and online supplementary materials of Bowles and Choi (2013).
whereas the loser bears a cost. If, instead, the civics fail, they all share the losers’ cost and the grabber gets the product. The civics are the aggressive egalitarians.

Because the payoffs of each behavioral type differ according to the type with which they are paired, expected payoffs depend on the distribution of types in each subpopulation. Agents occasionally update their strategies, switching from lower to higher payoff behaviors and occasionally just experimenting (choosing a strategy randomly). As a result, the distribution of payoffs in each subpopulation evolves over time as the distribution of behavioral types changes. The final stage of the model is between-group competition, in which the subpopulations are randomly paired and the one with the higher average payoffs wins a contest, gaining some of the loser’s wealth. Members of the losing subpopulation then differentially copy the strategies of the higher-payoff-winning subpopulation. These four stages are then repeated over many periods.

An equilibrium in this model is characterized by a stationary distributions of behavioral types. One of these is similar to the Hawk-Dove equilibrium, there are no civics. Grabbers gain high incomes by exploit sharers and they persist because the occasional civic that appears due to migration or behavioral experimentation will engage in frequent contests with aggrandizing grabbers without the assistance of other civics, losing many conflicts. The resulting lower payoffs of the civics will prevent them from proliferating in the subpopulation. Because grabbers also interact with each other, bearing the negative expected payoff as a result of the conflict, average payoffs in the grabber-sharer equilibrium are low, and payoff differences between grabbers and sharers are substantial.

The all-civic population is also an equilibrium, from which an occasional greatly outnumbered grabber (introduced by migration or experimentation) will readily be eliminated. Because within-group conflicts are as a result rare, average payoffs are close to the maximum possible. And, because civics and sharers are behaviorally identical and have the same payoffs if grabbers are absent, payoff inequality is limited.

A population without grabbers —like the all-civic subpopulation just described—will tend to prevail in contests with a subpopulation in which civics are absent due to their superior average payoffs.

---

27 The civics’ cooperative punishment strategy is based on ethnographic studies of the coordinated maintenance of social order in stateless societies (for example, Mahdi (1986) and Wiessner (2005)). Coordinated punishment contrasts with the noncooperative (individual) punishment strategies in the conventional public goods with punishment game and has been shown to be evolutionary viable under plausible parameters in Boyd, et al. (2010).
But a weakness of the civic equilibrium is that it is only neutrally stable: small displacements from the all-civic population distribution (due to immigration of sharers or civics randomly switching to sharer) will not be corrected by the within-subpopulation selection process. So, the all-civic equilibrium is subject to drift, which occasionally will result in most members of the subpopulation becoming sharers. This mostly sharer subpopulation is then readily invaded by grabbers, moving the subpopulation to the unequal sharer-grabber equilibrium.

Figure 2 shows the Gini coefficient for payoffs among the twenty members (in a single generation) of one of the subpopulations over the period from 15,000 to 6,000 years ago (they were predominantly hunter-gatherers, only 18 percent engaging in farming over this simulation). Panel A shows that while the mean and most common level of inequality is quite limited, the subpopulation does occasionally (three times in the run presented) experience Gini coefficients above 0.5. Ancient real counterparts of these cases might have left opulent burials or other evidence of wealth inequality.

In Panel B we see that the Gini coefficients in the archaeological record, while higher on average, are distributed similarly with a substantial right tail. (There is nothing unusual about the particular subpopulation shown in the figure; other subpopulations exhibit very similar results.)

Our understanding of the Early Neolithic is that wealth inequality may have been limited by the mutually policed sharing in an arrangement similar to an all-civic equilibrium, perhaps sustained by the egalitarian social norms suggested by the archaeological evidence among the Natufians and the residents of Jerf el Ahmar and Çatalhöyük. But the resulting highly unequal outcome would have been ephemeral, due perhaps, as in the model, to the likelihood that a low average payoff sharer-gabber subpopulation would lose a pairwise contest with an all-civic subpopulation.

Other possible factors contributing to the instability of subpopulations with elevated levels of wealth inequality (beyond what is represented in this model) include their lesser ability to survive an adverse environmental shock (because a subpopulation of grabbers and sharers would be less cooperative than an all civic subpopulation) or internal dissention (arising from the inequality of wealth). Additionally, egalitarian practices concerning inheritance (funeral feasting, bride wealth payments, or inheritance of wealth by one’s lineage rather than by offspring) could have also contributed to making wealth inequalities ephemeral.
A. Inequality (Gini coefficient) in a simulated Late Paleolithic and Neolithic group.

B. Frequency distribution of Gini coefficients among simulated and observed ancient populations.

Figure 2. Simulated and observed ancient economic inequality. (A) Gini coefficients for payoffs for a single subpopulation, from a simulation of the model in the text implemented by Jung-Kyoo Choi. The horizontal axis is years BCE, the vertical axis is the Gini coefficient. (B) Frequency distribution of Gini coefficients in the simulated population and in our archaeological data set (n = 94), including ninety early farmers (from Levant, Mesopotamia and Western Balkans in a period ranging from 12 thousand years ago to 3 thousand years ago) and four sedentary hunter-gatherer populations (Palestine from 21 thousand and 12 thousand years ago and 18 thousand years ago Russian Plain).

4.5. Neolithic wealth inequality: Did early farming in western Eurasia increase inequality?

In short, we do not really know if the shift from hunting and gathering to farming in western Eurasia increased inequality. Indirect evidence suggests that it may not have. If the prehistoric hunter-
gatherers and the Neolithic hoe farmers resembled their recent ethnographic counterparts, it appears unlikely that the intergenerational transmission of material wealth -- a key determinant of the degree of stationary inequality, as we have seen -- would have been greater in the hoe farming societies (Borgerhoff Mulder, Bowles, Hertz, et al. 2009, Gurven, Borgerhoff Mulder, Kaplan et al. 2009) Direct evidence is sparse. We have just four observations on hunting-gathering populations in western Eurasia (three exploiting wild cereals) with a mean Gini coefficient of 0.307, somewhat less than the mean for Neolithic hoe-based farmers, 0.355.28 (We consider the substantial changes brought about by the Late Neolithic and Bronze Age introduction of animal traction and plows in the next section.)

In this and subsequent comparisons of means, the small samples and (in many cases) non-normal distributions of the data (as in Figure 2) motivate our replacing the conventional $p$-statistics with a permutation-based alternative. This method relaxes the assumptions of normal distribution of the measures of interest and of random sampling of the observations from the reference populations (Kennedy 1995). It consists of simulating what the data would look like if wealth inequality were unrelated to the hoe-farming and hunter-gatherer distinction.

We do this by randomly assigning the two labels (hunter-gatherers and Neolithic farmers) to our ninety-four data points and checking if the mean of the now randomly labeled “Neolithic farmers” equals or exceeds the mean of the “hunter-gatherers” by more than the observed quantity, namely 0.048. We do this random assignment of labels exercise 10,000 times and record the fraction of times that a mean difference equal to or greater than the observed occurs, defining this fraction as the permutation based $p$ statistic. Conditional on each of the possible configurations of the randomized data being equally likely under the null hypothesis (the exchangeability assumption), the randomization of labels allows us to estimate how unlikely the difference in means that we observed (or greater) would be if the null hypothesis, that there is actually no relationship between the two groups and the degree of inequality, were true.

The fraction of mean differences under randomization that is equal to or greater than the empirically observed difference is the permutation-based $p$ statistic that we report in what follows (rather than the conventional $p$). For the observed difference in hunter-gatherer and Neolithic hoe-farmer mean wealth

---

28 Estimates of wealth inequality among sedentary hunter-gatherers in the Western Hemisphere, for the most part fishers exploiting a “clumped resource” ($G = 0.335 [0.151; 0.475]$), are slightly smaller than estimates for Neolithic hoe farmers in western Eurasia ($G = 0.355$, above).
inequality, we have \( p = 0.242 \), suggesting that a difference of the observed size and sign would be likely to be observed if the null hypothesis that hunter-gatherers and farmers do not differ with respect to inequality were true.

The small difference in mean wealth inequality between hunter-gatherers and early Neolithic farmers and the substantial frequency with which a difference of this magnitude or greater arises in our randomization exercise casts doubt on the common view that the origin of enduring wealth inequality is to be found in the Neolithic agricultural revolution. The Early Neolithic witnessed a fundamental innovation in technology -- producing food rather than acquiring it from uncultivated wild species -- and associated institutions, including a greater role for private property in dwellings, stores, and domesticated animals and other correlates of sedentism. But if our reading of the evidence is right, the advent of hoe farming and these concomitant changes in social organization, while possibly associated with somewhat greater wealth disparities, were not sufficient to sustain modern levels of inequality.

5. Land-intensive, labor-saving technology and the emergence of substantial wealth inequality

What, then, might account for the emergence of enduring wealth inequality? The first step in this process apparently was not the emergence of faring \textit{per se}, but instead a much later change in farming technology. In this section, we present evidence that a transition occurred in Bronze Age western Eurasia from what we term labor-limited to land-limited livelihoods and that this was associated with a substantial increase in wealth inequality. We provide examples of the two types of economy from both the ethnographic record and econometric estimates using historical data and use these estimates both as a robustness check on our inferences from the archaeological evidence and to assess the relative importance of increased wealth shocks and increased intergenerational wealth transmission as sources of the late Neolithic emergence of substantial wealth inequalities.

5.1. Labor-limited and land-limited economies\textsuperscript{29}

Clues about how this process may have unfolded are to be found in the emergence of substantial wealth inequalities among the four groups of hunter-gatherers and early farmers described in the previous section. In these cases—the Bronze Age Thai traders, the Chumash islanders, the Provadia salt

\textsuperscript{29} The distinction between labor-limited and land-limited economies was introduced in Bogaard, et al. (2019).
ingot manufacturers, and possibly the residents of Çatalhöyük—the common elements are a shift to a capital-intensive way of accumulating wealth, one characterized by economies of scale.

The Ban Non Wat traders’ need for a substantial sum of working capital to engage in long-distance trade at a viable scale could explain why most residents of that site were excluded from these profit-making opportunities. And similarly, the less wealthy Chumash islanders could not finance canoe building and ownership, while the ordinary Provadians were able to participate in salt ingot making as subordinate workers, but not as owners of the kilns.

We termed these cases *episodic* because they were environmentally contingent and perhaps not politically sustainable. The Chumash elite of canoe owners profited from a warming episode that reduced the productivity of their previous labor-intensive maritime foraging livelihoods. The effect was to diminish the reservation option of the less well-off islanders favoring the canoe owners in bargaining with the rest of the Islanders. Similarly, the wealth of the Provadian elite stemmed from an unusual geological occurrence that ended when the brine springs dried up. The Ban Non Wat elite’s wealth virtually disappeared when bronze became sufficiently available locally. The abundance of arable land in the sparsely populated Konya Plain may have placed limits on the degree of sustainable inequality at Çatalhöyük and facilitated its eventual abandonment when heightened wealth inequalities emerged.

We hypothesize that substantial wealth inequality became a common characteristic of an economy when a new means of acquiring wealth—one that was both capital intensive and subject to increasing marginal returns—became generally available rather than being contingent on unusual or ephemeral environmental conditions. In western Eurasia, this was the introduction of the cattle-drawn (and later ox-drawn) plow, and the resulting shift from a labor-intensive form of garden farming to a land-intensive labor saving technology.30

Estimates of the amount of land that can be prepared for cultivation by a household using an ox team and plow rather than hoes range from a twofold to a fifteenfold increase, depending on conditions (Halstead 1995). This increase in the average productivity of labor meant that introducing the ox-drawn plow permitted the substitution of land for labor, resulting in a substantial increase in the area farmed by a single household able to deploy an ox team. Jack Goody, writing about the plow, pointed to “the scarcity of land which it creates or aggravates, and the differential holdings of land and capital which then became important” (Goody, 1976) p.97.

---

Also contributing to a more capital-intensive livelihood was the increasing reliance on slow-maturing and long-lived tree crops like dates, olives, and grapes (Gilman 1981) and domesticated animals such as goats, sheep, and, later, cattle, as occurred at Çatalhöyük. Where offshore fishing became part of people’s livelihood, economies became more capital intensive: the Bronze Age “boats and nets required for effective exploitation of offshore fisheries involve a very considerable investment of labor in advance of production” (Gilman 1981). Some boats, at fifteen meters in length and with a working lifetime of fifty years, resembled the oceangoing canoes of the Chumash islanders. The result was to make land, animals, and other material forms of wealth more valuable compared to labor. Importantly, these forms of wealth would contribute to increases in the stationary inequality of wealth in our model because they could be accumulated and transmitted across generations and were subject to substantial shocks.

We refer to the Neolithic economies as labor-limited and the Bronze Age and later economies as land-limited.31 (For simplicity, we use the latter term to refer not just to land but to any form of material wealth, such as the salt at Provadia or the microliths and oceangoing canoes among the Chumash.) In what follows we will measure the limiting nature of labor or material wealth not by the quantities available but by the relative shadow price of labor and material inputs—that is, the relative value to a household of having access to an additional unit of each.

5.2 The transition from labor-limited to land-limited farming in western Eurasia

Labor-limited garden farming during the Neolithic involved the intensive weeding, manuring, and other aspects of the cultivation of relatively small plots using hand tools, characterized by extraordinary output per hectare (Bogaard 2004). The land-limited system of farming that eventually replaced it involved much less labor-intensive cultivation of larger tracts of land using specialized ox-drawn plows, characterized by reduced levels of land productivity and elevated labor productivity. Bogaard and her coauthors (including the current authors) have distinguished between these two systems of farming using evidence for both animal traction and crop growing conditions.32

Archaeological evidence of crop remains and the extent and nature of weeds associated with a cultivated crop can be used to identify cultivation practices. These data, along with stable carbon and

---

31 In the anthropological literature, the labor-limited and land-limited economies, respectively, would be termed horticultural and agropastoral.
nitrogen isotope analysis of crop remains, allow additional measures of cultivation intensity, for example, the extent of manuring. Further evidence of a land-limited system are costly investments in terracing or irrigation that would be cost effective only if the value of augmenting the effective supply of land was substantial relative to the value of the significant amount of labor necessary to undertake and sustain these improvements. These methods allow us to classify ninety archaeological sites as labor-limited farming and sixteen as land-limited. To identify the effect of the shift in farming technology per se, we exclude the Bronze Age land-limited economies in which there had emerged a degree of political centralization that we later will define as the archaic proto-state.

The mean Gini coefficients for the (non-state) labor-limited and land-limited populations are 0.355 and 0.577, respectively. The p-statistic (0.0001) tells us that a mean difference of this magnitude (that is, 0.222) and sign would be very unlikely to be observed were the null hypothesis -- that labor-limited and land-limited populations did not differ with respect to inequality -- true. The labor-limited economies are for the most part earlier in the Neolithic while the land-limited economies are later, but there is a span of at least a millennium in which the two coexist. The difference in means is not explained by a time trend.

5.3 Labor-limited and land-limited economies: A robustness check based on more recent evidence.

Given the necessarily sparse information about the transformation of prehistoric farming by the introduction of the ox-drawn plow, we now review more recent evidence in order to gauge the magnitude of the impact that this change in technology may have had. These data will show that farming economies differ substantially in the importance of land and other forms of material wealth as a factor of production. We measure this by \( \mu \), the coefficient of land, cattle, or other material wealth inputs in a

---

33 The criteria used to designate labor-limited and land-limited economies are explained in Appendix, Section 4. The earliest land limited are the sites in modern-day Bulgaria in which the material wealth that limited livelihoods was access to either copper or salt. The earliest (literally) land-limited farming sites are the southern Mesopotamian settlements at Tepe Gawra and Eridu (over the period 4400 BCE to 3000 BCE).

34 In the period between 5100 BCE and 3900 BCE, the mean of the 8 land-limited economies 0.473 [0.303; 0.753] is larger than the mean of the 64 labor-limited economies, 0.341[0.149;0.584]. A difference of 0.132 is unlikely were a null hypothesis that the two groups are not different with respect to inequality be true, p=0.005.

35 In an ordinary least squares (OLS) regression in which we condition the prediction of measured inequality on a time trend, the predicted difference between the land-limited and labor-limited economies is 67 percent of the difference when not conditioned on time. The details of the regression are reported in the Appendix, Section 5.
Cobb-Douglas production function.\textsuperscript{36} We will find that economies in the ethnographic and historical record that were characterized by greater importance of material wealth as a factor of production also tend to exhibit greater inequalities in material wealth. These data suggest that the Late Neolithic and Bronze Age shift in farming technology could have been associated with a substantial increase in wealth inequality.

We first use evidence in the ethnographic record on hunting-gathering as well as three kinds of farming: hoe farming (commonly termed horticulture by anthropologists), plow agriculture, and pastoralism. These data come from a study by a group of anthropologists and economists (including one of the current authors) who sought to estimate the importance of material wealth to a population’s livelihood compared to two other types of wealth, which we termed \textit{embodied} wealth (referring to an individual’s health, strength, and other individual capacities) and \textit{relational} or \textit{network} wealth (referring to the individual’s social network ties).\textsuperscript{37}

For each of the three wealth types above, we solicited ethnographers’ judgments on the percentage difference in household well-being associated with a 1 percent difference in amount of a given wealth type in the population they studied, holding the levels of other forms of wealth constant at the population mean and requiring these percentage effects to sum to one. In effect, the ethnographers were asked to estimate the exponents of a three-input constant returns to scale Cobb-Douglas function predicting well-being defined as a flow of services that the individuals in that culture value, whether it be affluent consumption, large families, or something else.

Consistent with descriptive ethnographies of these and other populations, they reported that embodied and relational wealth are relatively important for hunter-gatherers and horticulturalists, while material wealth is key in pastoral and agricultural populations. Figure 3 shows the ethnographers’ estimates. The small triangle on the left edge means, for example, that according to the ethnographer of that hunting and gathering economy (the Hadza of Tanzania), material wealth was unimportant to a family’s livelihood while for the two other components, embodied wealth was a bit more than twice as important as relational wealth (exponents of 0.7 and 0.3, respectively).

\textsuperscript{36} The exponent \( \mu \) is the percentage increase in output associated with a 1 percent increase in material wealth applied to production, the other inputs remaining unchanged.

The mean values of $\mu$ and the exponents of the other forms of wealth, shown by the coordinates of the larger shapes in the figure were virtually identical for agriculture and pastoralism. And consistent with our representation of hoe and plow farming in the Neolithic the ethnographers’ estimate of $\mu$ was much smaller in the horticultural economies and similar to hunting and gathering. As a check on these subjective (if well-informed) estimates, econometrically estimated values of $\mu$ are also available for two horticultural, one pastoral, and four small-scale agricultural economies. Table 2 summarizes these results. The econometric estimates in column 5 are remarkably close to the ethnographers’ assessments (shown in parentheses).

**Figure 3. Labor and material wealth as limits on livelihoods.** The points’ coordinates (which sum to one) indicate the relative importance of the forms of wealth at the three vertices. A point closer to a vertex indicates that the form of wealth in the vertex label is more important. Large shapes are averages. Additional information about the figure are in the Appendix, Section 2.

We now use these estimates of $\mu$ in Figure 3 and Table 2 to illustrate the positive relationship between the importance of material wealth in the production of a people’s livelihoods and the degree of inequality in material wealth, shown in Figure 4. The simple correlation of the two 0.491 CI [-0.053;0.810], so a standard deviation difference in the importance of material wealth is associated with a half a standard deviation difference in wealth inequality.

In sum, recent ethnographic data suggest that small-scale farming societies do indeed differ along the labor- and land-limited continuum and that wealth inequality in the latter is substantially greater than in the former. If the transition from a labor-limited (low $\mu$) economy in the Early
Neolithic to a land-limited (high $\mu$) non-rice agricultural or pastoral economy late in the Neolithic resembled the differences observed in Figure 4 and Table 2, then the expected change associated with the shift to a land-limited economy would have been an increase in the Gini coefficient of something like 0.242, which is close to the observed difference in means (0.222) reported above.\textsuperscript{38}

Table 2. The importance of material wealth in farm economies. Source: See Appendix, Section 2.

<table>
<thead>
<tr>
<th>Population</th>
<th>Date</th>
<th>Mode of production</th>
<th>Material wealth</th>
<th>Econometric est. $\mu$ (ethnographic est.)</th>
<th>Capital intensity</th>
<th>Source of econometric estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yamaguchi Prefecture, Japan</td>
<td>1840s</td>
<td>Land-limited (rice)</td>
<td>Land and fertilizer</td>
<td>0.45 (0.42)</td>
<td>0.82</td>
<td>Nishikawa (1978)</td>
</tr>
<tr>
<td>Nyaturu TZ</td>
<td>1950s</td>
<td>Land-limited (agropastoralist)</td>
<td>Cattle and land</td>
<td>0.76 (0.61)</td>
<td>3.17</td>
<td>Massell (1963)</td>
</tr>
<tr>
<td>India</td>
<td>1950s</td>
<td>Land-limited (all cereal)</td>
<td>Land</td>
<td>0.68 (0.66)</td>
<td>2.13</td>
<td>Bardhan (1973)</td>
</tr>
<tr>
<td>India</td>
<td>1950s</td>
<td>Land-limited (rice)</td>
<td>Land</td>
<td>0.33 (0.42)</td>
<td>0.49</td>
<td>Bardhan (1973)</td>
</tr>
<tr>
<td>Borara Ethiopia</td>
<td>2000s</td>
<td>Land-limited (pastoral)</td>
<td>Livestock</td>
<td>0.84 (0.61)</td>
<td>5.25</td>
<td>Berhanu, Colman and Fayissa (2007)</td>
</tr>
<tr>
<td>Borara Ethiopia</td>
<td>2000s</td>
<td>Labor-limited</td>
<td>Land livestock</td>
<td>0.23 (0.21)</td>
<td>0.30</td>
<td>Berhanu, Colman and Fayissa (2007)</td>
</tr>
<tr>
<td>Gambia</td>
<td>1940s</td>
<td>Labor-limited</td>
<td>Land</td>
<td>0.11 (0.21)</td>
<td>0.12</td>
<td>Bowles (2011), Haswell (1953)</td>
</tr>
</tbody>
</table>

5.4. Ancient Northeast Thailand: Another robustness check.

We have focused on western Eurasia from the Early Neolithic through the Iron Age because of its rich archaeological evidence and because sustained high levels of wealth inequality arguably first occurred there. A second robustness check for our conjecture that heightened inequalities resulted from the transition from labor-limited to land-limited production methods is provided by newly available set of estimates of wealth inequality in the upper Mun Valley of modern-day Thailand.\textsuperscript{39} The data set is

\textsuperscript{38} Using the data in Figure 4 to estimate an OLS regression of the $G$ on $\mu$, we obtain a best-fit slope equal to 0.411 [-0.047, 0.870]. Multiplying that by the difference given in Table 2 between the econometric estimates of $\mu$ for the non-rice-farming (and pastoral) and the horticultural economies (0.59), we get the estimate shown in the text (0.242). The observed difference in means is $0.577 – 0.355 = 0.222$.

\textsuperscript{39} Fochesato, et al. (2021). This evidence was not available when with Amy Bogaard we formulated and initially tested our hypotheses concerning labor-limited and land-limited economies in Bogaard, et al. (2019), Fochesato, et al. (2019)
unparalleled in the archaeological literature for the number of comparable estimates on the same wealth indicator (grave goods) from the same locality, spanning over two and a half millennia—from the first farmers to the first state. The series is also unusual in that there is a shift from a labor-limited to land-limited economy that is subsequently reversed: The increased inequality associated with the land-limited Bronze Age trading opportunities described in Section 3 were ephemeral. When bronze trading declined, the inhabitants returned to labor-limited rain-fed rice cultivation as their primary source of wealth, and evidence of heightened wealth inequalities disappears from the archaeological record.

Figure 4. The importance of material wealth ($\mu$) and material wealth inequality (the Gini coefficient) in small-scale societies. Source: See Appendix, Section 2.

Toward the end of the series, a new change in the economic system took place due to a major reduction in monsoon rains that reduced the effective supply of land. This forced the construction of irrigation systems and effected a transition similar to the one that occurred in western Eurasia from hoe farming to draft-animal-based plowing. Evidence on farming techniques and trading allow us to identify nine labor-limited sites and five land-limited sites (including those associated with the bronze trade), with mean Gini coefficients of 0.326 and 0.458, respectively.\(^4\) The sign and magnitude of the mean difference (0.132) would be very unlikely to be observed were the null hypothesis, that the two groups are not different with respect to inequality, to hold ($p = 0.0001$).

\(^4\) The Thai data estimates are comparable to the estimates presented earlier, having been adjusted by the same algorithms suggested in Fochesato and Bowles (2024b).
The difference in mean wealth inequality between the Thai labor-limited and land-limited economies is half of that associated with shift to land-limited farming in Western Eurasia prior to the emergence of archaic proto-states (0.222). The lesser increase in the wealth inequality associated with the shift to land-limited wet rice farming in Thailand may be due to the more labor-intensive nature of wet rice plow farming (in comparison to the cereals raised western Eurasia). A result was that shift to a land-limited farming economy in Iron Age Thailand may thus have been associated with a more modest increase in $\mu$, the exponent of material wealth in our hypothetical Cobb-Douglas production function, possibly accounting for the lesser increase in wealth inequality.

To explore this hypothesis, we find that the difference in mean inequality between the labor limited and material wealth limited Thai sites is quite close to what we would expect from the difference in our estimate of $\mu$ for land-limited (irrigated) rice farming and labor-limited hoe agriculture, from the estimates of $\mu$ in Table 2. The anticipated increase in the Gini coefficient based on estimates from the dry rice (and other cereal) farming in the Gambia and wet rice farming in nineteenth-century Japan and twentieth-century India is 0.115, which is close to the increase observed in the Thai data above (0.132).\(^{41}\)

5.5 The late Neolithic emergence of inequality: Shocks and intergenerational transmission

To better understand the emergence of substantial wealth inequalities, a natural question concerns the relative importance of the two effects: greater contemporaneous shocks and increased intergenerational transmission.\(^{42}\) We can rewrite equation (2) as

$$\sigma_w^2 = \sigma_x^2 (1 - \beta^2)^{-1} = \sigma_x^2 (1 + \frac{\beta^2}{1 - \beta^2}) = \sigma_x^2 (1 + \rho) \quad (5)$$

We might use this expression to compute a decomposition of a change in $\sigma_w^2$ into components due to changes in the level of shocks and the extent of intergenerational transmission of wealth. But the relevant decomposition is the change in the Gini coefficient, which is not proportional to $\sigma_w^2$, making such a decomposition impossible. However, under plausible assumptions (given in the Appendix)

\(^{41}\) This estimate does not take account of the land-limited nature of the trading economy at Ban Non Wat. The estimate is calculated as was the analogous estimate for the main data set described in section 5.3 but using rice farming estimates to calculate the difference in $\mu$ between hoe and plow farming (0.28), multiplied by the estimated coefficient of $\mu$ in the equation predicting the Gini coefficient (0.411) based on the data in Figure 4.

\(^{42}\) We are grateful to David Romer for suggesting the decomposition, and to Max Greenberg for devising this method to implement it.
another unit-free measure of inequality, the standard deviation of the logarithm of wealth ($\sigma_w$), is very close to a scalar multiple of $G$, allowing the following decomposition.

Using equation (5) we express the relative wealth inequality in two cases $A$ and $B$ as

$$\frac{\sigma_{wB}^2}{\sigma_{wA}^2} = \frac{\sigma_{AB}^2}{\sigma_{AA}^2} \frac{1 + \rho_B}{1 + \rho_A}$$

or, to convert the ratio of variances to the ratio of standard deviations, taking the square root of both sides

$$\frac{\sigma_{wB}}{\sigma_{wA}} = \frac{\sigma_{AB}}{\sigma_{AA}} \sqrt{\frac{1 + \rho_B}{1 + \rho_A}}$$

which we rewrite as $1 + \delta_w = (1 + \delta_\lambda)(1 + \delta_\rho)$ or equivalently:

$$\delta_w = \delta_\rho + \delta_\lambda + \delta_\rho \delta_\lambda$$

(6)

We identify $\delta_\rho/\delta_w$ as the fraction of the difference in inequality between sets of observations $A$ and $B$ that, absent differences in the magnitude of shocks, would have been caused by observed difference in the intergenerational transmission of wealth; $\delta_\lambda/\delta_w$ as the fraction of the difference in inequality that, absent differences in intergenerational transmissibility, would be caused by a difference in the magnitude of ( uninsured) wealth shocks, and $\delta_\rho \delta_\lambda/\delta_w$ as the residual caused by the interaction of these two differences, were both to have occurred.

Though we cannot estimate the relevant intergenerational transmission elasticities ($\beta$’s) from archaeological data, we expect they will differ substantially between the hoe-farming, labor-limited economies, and the land (or other material wealth)-limited economies. This is consistent with econometric estimates based on ethnographic data from two generations of farmers in small scale societies that may be similar to the ancient economies under study here (Borgerhoff Mulder, Bowles, Hertz, et al. 2009). The mean estimated $\beta$ for material wealth—including land, tools, and livestock—in horticultural (that is, hoe based labor-limited) farming economies is $0.089 \pm 0.056; 0.234$ while for agricultural and pastoral (land-limited) economies the corresponding value is $0.605 (0.483; 0.727)$.

---

43 Thus $\delta_w = \frac{\sigma_{wB}}{\sigma_{wA}} - 1$, $\delta_\lambda = \frac{\sigma_{AB}}{\sigma_{AA}} - 1$, and $\delta_\rho = \frac{1 + \rho_B}{1 + \rho_A} - 1$. 

48
We now apply these estimates to our archaeological data. To do this, we suppose that the difference in the intergenerational transmission elasticity between our (mostly Neolithic) labor-limited economies and the later (prior to the archaic proto state, mostly Bronze Age) land-limited economies were of the magnitude suggested by the ethnographic data just reported. A second piece of evidence is that as reported in Section 5.2, the estimated mean Gini coefficient in the land-limited economies without states of any kind is 63 percent greater than in the labor-limited (also stateless) economies.

These two pieces of information, along with the assumption that the distribution of wealth is approximately log-normal, are sufficient to estimate equation (6), decomposing the difference into the three components. We find that components of the change in wealth inequality are given by the following fractions of the total change: increased intergenerational transmission, \( \delta_r / \delta_w = 0.352 \); increased wealth shocks, \( \delta_A / \delta_w = 0.518 \); the interaction of the two changes \( \delta_r \delta_A / \delta_w = 0.130 \). We conclude that the origin of enduring wealth inequality likely was associated with a substantial increase in both the magnitude of (uninsured) shocks and the intergenerational transmission of wealth.

The implied Neolithic (labor-limited) to Bronze Age (land-limited) increase in the (uninsured) shock term is also consistent with the reasoning based on our model in Section 2 and with ethnographic data. Prior to the use of ox-drawn plows and other labor-saving means of production, labor-limited farming economies in the region engaged in a significant amount of hunting and gathering, and as a result they had a broad portfolio of food sources. With the transition to a land limited economy reliance on a more limited number of crops and animals, increased population densities (both humans and non-human) with concomitant health hazards, and the reduced scope of collective consumption smoothing practices (such as sharing animals probably increased the magnitude of wealth shocks).

6. Modeling the transition from a labor-limited to a land-limited economy

We have to explore, then, the mechanisms by which, as the archaeological data appear to show, the introduction of plow farming and other changes associated with the land-limited economy might have substantially increased wealth disparities. Equally important, we need to explain how this could have occurred prior to the wealthy being able to exert the kind of coercive powers that the later concentration of their political power in states would allow.

6.1 Production and shadow prices of inputs in labor- and land-limited economies.
The change in limiting factors—from labor to land and other forms of material wealth—is key to our argument. Farming system A is more land-limited than system B if land (or other forms of material wealth) is more valuable relative to labor in A than in B. The value of a factor of production is its marginal product, and the ratio of the value of land to the value of labor is termed the shadow price of land (relative to labor). The extent to which a production system is labor-limited depends on the goods and services constituting a population’s livelihood (wild or domesticated species, for example, tubers or tree crops), the nature of the production processes by which these are acquired (with or without the use of draft animals, for example), and the relative abundance of land, labor, and other production factors (for example, the abundance of land in the environs of Çatalhöyük compared to the scarcity of brine springs other than at Provadia). \(^{44}\)

Property rights are also important. In this framework, the heritable, proprietary nature of enslaved people, for example, makes enslaved labor in this respect similar to land or other forms of material wealth. As a result, where crops requiring substantial labor per unit of land, such as wet rice or cotton, are produced using enslaved labor, the economy may be land-limited rather than (free) labor-limited. \(^{45}\)

A puzzle about the transition in farming methods in Late Neolithic and Bronze Age western Eurasia, if we are correct, is that the increased value of land and other forms of material wealth coincided with a shift to a more land-intensive mix of inputs. This increase in the ratio of land to labor inputs, taken by itself, should have reduced the relative marginal productivity (and thus value) of land, but this is not what occurred. Looking ahead, the resolution of the puzzle, in short, is that the ox-drawn plow was a significantly labor saving innovation that made labor redundant and land scarce.

To clarify how this may have occurred, drawing on our joint work with Bogaard we introduce a production function that we think captures the salient knowledge about farming methods during this period (Bogaard, Fochesato and Bowles 2019). In our model, there are four factors of production: labor, manure, land, and animal traction.

We use the following notation: \(Q = \) quantity of output produced; \(m = \) amount of manure applied to the land; \(T = \) amount of land cultivated (in hectares); \(x = \) a measure of ox team services; and \(L = \) hours.

\(^{44}\) The crop does not uniquely determine the labor- or land-limited nature of the economy; in prehistoric western Eurasia, staple crops such as wheats and barleys were grown under both labor- and land-limited regimes (Bogaard, et al. (2016), Bogaard, et al. (2019).)

\(^{45}\) Thus, given that cotton is a labor intensive crop, the abolition of slavery may have converted the southern United States economy from a land (and other material wealth including enslaved people)-limited economy to a (free) labor-limited economy.
of labor services applied to cultivation. We assume that $T$ and $L$ are positive, and that in one technology oxen are used along with labor and in the other manure is used to augment the land input. The parameter $B$ is a positive constant indicating how productive the technology is, while the land exponent, $\varphi < 1$, is a measure (analogous to $\mu$ above) of the importance of land (possibly augmented by manure) in the production process:

$$Q = B(m + T)^\varphi (x + L)^{1-\varphi}$$

(3)

where $x$ and $m$ are zero, respectively, in the labor-limited and land-limited cases. In the equation above, $(m + T)$ represents “effective land cultivated,” recognizing the land-augmenting nature of manure, while $(x + L)$ is “effective labor devoted to cultivation” augmented by animal traction.\(^ {46} \) Table 3 and Figure 5 clarify our use of the terms labor-limited and land-limited.

Figure 5. The value of land relative to labor in hoe farming and plow agriculture. The absolute values of the slopes of the isoquants at points $i$ and $x$ are the marginal rates of substitution in production of land for labor in, respectively, a labor-limited (hoe) farming economy and a land-limited agricultural economy with animal traction, including plows. This is our measure of the value of land relative to labor.

\(^ {46} \) We have defined the units in which manure is measured so that one unit of manure is equivalent to one hectare of land, and analogously for the use of an ox team and labor. Our assumption that manure and land and ox teams and labor are perfect substitutes simplifies the presentation of the model but is not essential to its working.
In Figure 5, the household deploys the same amount of labor hours under the two types of farming but farms more land in the land-limited case. The ratio of the two marginal products, our measure of the relative scarcity of land, is

\[
\text{shadow price of land} = \frac{Q_l}{Q_L} = \frac{\varphi}{1-\varphi} \left( \frac{x+L}{m+T} \right)
\]

which is the absolute value of the slope of the isoquants in Figure 5.

<table>
<thead>
<tr>
<th>Labor per</th>
<th>Manure per</th>
<th>Land</th>
<th>Oxen</th>
<th>AP (land)/ MP (land)</th>
<th>AP (labor)/ MP (labor)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor-limited</td>
<td>ach</td>
<td>ach</td>
<td>tle</td>
<td>me</td>
<td>gh</td>
</tr>
<tr>
<td>Land-limited</td>
<td>tle</td>
<td>me</td>
<td>ach</td>
<td>any</td>
<td>w</td>
</tr>
</tbody>
</table>

Table 3. Labor-limited and land-limited farming in our model. AP and MP, respectively, are average and marginal productivity of the entity in parentheses, and \( \varphi \) is the exponent of land in equation 3, a measure of the importance of material wealth in the production process (as shown in equation 4).

Land is less valuable in the labor-limited farming system because, in the absence of plows \((x = 0)\) and with the use of manure \((m > 0)\), land is effectively relatively abundant and labor effectively scarce. It is also clear from equation (4) that, in its effects on the shadow price of land, introducing ox-drawn plows and reducing the per-hectare land-augmenting investments (setting \(m = 0\)) is equivalent in its effect on the shadow price of land to increasing the importance of land relative to labor in the production function, that is, increasing \( \varphi \).

Despite the diminishing marginal productivity of each factor, because of the positive effect of manure on the marginal product of land and the negative effect of animal traction on the marginal product of labor, as well as the lower value of \( \varphi \), labor is be more valuable in labor-limited farming systems (despite its greater relative use) and land more valuable in more land-using land-limited systems.

6.2. The dynamics of wealth inequality in Late Neolithic and Bronze Age western Eurasia

To see how the transition from labor-limited to land-limited farming may have occurred, consider a population engaged in labor-limited farming—similar to point \(i\) (for labor intensive farming) in Figure 5. How could a transition to a land-limited system illustrated by point \(x\) occur?
We base our answer on four facts. First, by definition, in labor-limited economies, labor is relatively valuable compared to land. Second, training, using, and maintaining an ox team was a substantial expense. When cultivating barley, for example, we know from early Mesopotamian texts that an ox team consumed an amount of the crop equal to the (at that time very substantial) portion set aside for the subsequent crop’s seed. Moreover, a team would have required a substantial period of training (Sherratt 1981). Third, there is a minimum project size for the primary innovation: two oxen are required and the pole and yoke apparatus for the team—would not have been cost effective for small plots. Based on historical data for Mediterranean ox-team cultivation, it appears that at least five hectares are required to offset the cost of acquiring and maintaining the team (Halstead 1995). And fourth, while modest in most labor-limited economies, wealth disparities were far from absent. By our estimates, even at “aggressively egalitarian” Çatalhöyük for the period associated with Level 6A (6500–6400 BCE), the wealthiest household had three times the average wealth.

Given this setting, there must have been many populations in an economy represented by point $i$ in which a few households had both the resources for maintaining plow oxen and access to additional land for expansion (Halstead 1995). If one of the wealthier individuals in the labor-limited economy were to consider acquiring an ox team and farming a larger area of land with less intensive labor input, they would seek to acquire rights to use more land, which would then be more valuable to the land-limited farmer-in-the-making than to the other farmers in his community. Moreover, because land is worth more to the farmer with oxen and a plow, acquiring additional land would be feasible even in the absence of coercion: there would be mutually beneficial voluntary transfers of access to land from the garden (hoe) farmer to the would-be land-intensive (plow) farmer. This is a process of wealth accumulation (i.e., land and oxen) with positive feedback, leading (as our model suggests) to elevated levels of wealth concentration.

The person “selling” or giving up the land might become the tenant or employee of the ox team owner, or part of a growing pool of labor available seasonally at harvest, during the peak labor demand of the land-intensive farming model. If, in the new economy, other wealthy farmers emulated the initial oxen-and-plow farmer, the value of land would eventually rise. As a result, available land would now be scarce, and labor would be relatively abundant, as shown by the steeper slope of the isoquant at point $x$. The economy would now be land-limited.

---

47 We refer to access to land rather than ownership of land as it is likely that access was governed by use rights sanctioned by the community rather than individually held property rights in the modern sense.
This situation would create fertile ground for the emergence of sustained inequalities. Andrew Sherratt suggests that “the use of animal traction was initially an activity reserved for the elite requiring a concentration of resources and involving an assertion of power and control over livestock, and by implication, over people” (Sherratt 2006) p. 353 (our translation). Peter Bogucki conjectures that “with time, the number of cattle-poor households increased in number, while the number of cattle-rich households forms a progressively smaller proportion of the population” 48 The population could develop a class structure: some having access to a considerable amount of land and one or more teams of oxen, and others working for them under some kind of subordinate relationship as renters, sharecroppers, clients, employees, or unfree labor.

This narrative, of course, is conjectural, but it is consistent with ethnographies of small-scale farmers in the recent past who worked with oxen (Halstead 2014). Plow farming, which allowed a single household to cultivate much more land than a family could harvest (Halstead 1995), thus simultaneously generated land scarcity and the labor abundance that provided those with an ox team access to the labor willing to be engaged at harvest time.

In well-watered areas, it is likely that the plow brought all arable land into cultivation. But even in the drier areas, where uncultivated arable land must have remained, land would have become more scarce given that, to be cultivated, a plot of land had to be within walking distance of a settlement, and so the more distant plots would be of limited value. The reservation options of the less wealthy would be limited, as simply withdrawing (possibly en masse) to some other yet uncultivated location (as we conjecture the less wealthy residents of Çatalhöyük in its later years might have done) would have been less rewarding than what the ox-and-plow farmer could offer. Daniel Potts’s assessment is: “The idea that large tracts of agricultural land could be brought under cultivation without a substantial investment in draught animals (principally oxen) is unrealistic” (Potts 2004) p.23.

The ox team and plow thus made land a limiting factor, even if some arable land (distant from settlements) remained uncultivated. Given the modest bargaining power of the newly dependent classes of workers in a labor-abundant economy (their reservation option was hoe farming), it seems unlikely that their consumption levels would have increased proportionately with the increase in average labor productivity—that is, output per worker—made possible by the introduction of animal traction. The

likely result would have been a divergence between the wealthier class and the rest of the population in their ability to save, invest, and accumulate wealth.\textsuperscript{49}

The transition to a land-limited economy may have contributed not only to substantial wealth inequalities but also to institutional changes that would eventually perpetuate and exacerbate these disparities. Large extended households with complex multifaceted economies engaged in animal-traction farming are thought to be the early precursors of temple- and palace-based landholding institutions observed in the earliest written documents from Mesopotamia (Pollock 1999, Ur 2014).

Cultural changes also accompanied the transition to a land-limited economy, illustrated by the association of cattle traction with social prestige that emerged around the mid-fourth millennium BCE in western Eurasia. Included was a wave of new prestigious rituals associated with cattle draft teams across Europe (Hadjikoumis, Robinson and Viner-Daniels 2011). “But did this translate into power?” asks Bogucki. “It appears that it did not, for there is no evidence of a formal structure of public life which would suggest political power beyond factions and hamlet organization” (Bogucki 1999) p.230. The emergence of permanent political disparities formalized in new structures of governance would come substantially later.

7. Archaic proto-states, states, slavery, and the origin of enduring wealth inequality

Ivan Ivanov conjectured that the fifth-millennium BCE gold and copper working center at Varna (in modern-day Bulgaria), which exhibits evidence of substantial wealth inequalities, was governed by a state (Renfrew 1978) p.202. We think this to be unlikely in light of the evidence (and the consensus among archaeologists) that the emergence of states, as conventionally understood, occurred much later.

To see what Ivanov's conjecture of a “state” at Varna would have entailed, we follow Max Weber by defining a state as an entity that can “monopolize the use of force,” that is, a “system of order [that] claims binding authority … over all action taking place in the area of its jurisdiction” in which “the use of force is regarded as legitimate only so far as it is either permitted by the state or prescribed by it” (Weber 1978 [1921]) p.56.\textsuperscript{50}

\textsuperscript{49} This interpretation (based on the model and evidence in Bogaard, et al. (2019)) is nicely illustrated for the case of irrigated lands in southern Mesopotamia by Allen in his fourth Ellen MacArthur Lecture at Cambridge (2022) and in Allen (2022) p.27, except that Allen’s account, like that advanced earlier by Halstead (1995) p.18, is in terms of the capacity of plow farming to produce a “surplus” rather than its effect on the relative value of factors of production.

\textsuperscript{50} Weber pointed out that defined in this way “the state has only in modern times reached its full development” (56). We do not depend on the commonly used inference that spatial hierarchies (e.g., “cities, towns, large villages, small villages”) reflect
7.1 Coercion and contested power: Weberian states and the archaic proto-state

We adopt Weber’s focus on a monopoly of the legitimate use of violence because it seems unlikely that the extraordinary levels of wealth inequality that we observe in Bronze and Iron Age populations (and the associated material and social deprivations, including enslavement) could sustained in the absence of a substantial concentration of the coercive powers available to an elite. Many archaeologists have adopted a variant of Weber’s definition of the state (Kirch 2010, Service 1975). Trigger, for example, defines a state as “a politically organized society that is regarded by those who live in it as sovereign … and has leaders who control its social political, legal, economics, and cultural activities” and “central governments possessed ultimate control over justice and the use of force” (Trigger 2003) p.47, 92.

States in this Weberian sense are thus distinct from other forms of political leadership such as chiefdoms and “big man” systems, in which leaders coordinate the activities of a group, constituting “a permanent agency of coordination” in the words of Elman Service, but lacking coercive powers (Service 1962). Referring to the literature in archaeology, Norman Yoffee makes this distinction: “In political terms chiefdoms contain hereditary … leaders (sometimes called a nobility) and centralized direction … but they have no formal machinery for forceful repression” (Yoffee 2005) p.23. Gary Feinman and Joyce Marcus distinguish archaic states “as having more power than the rank societies … particularly in the areas of waging war, exacting tribute, controlling information, drafting soldiers and regulating manpower and labor” (Feinman and Marcus 1998) p.4-5.

Many of the polities that are central to our interpretation appear to have been neither chiefdoms nor states. What is striking about what some have termed the “first states” or “archaic states” in Mesopotamia is their apparent lack of a monopoly on the legitimate use of force, hence our use of the prefix “proto.” Kin-based lineages and the exercise of what Weber called patrimonial authority “not enforced by physical coercion” were essential to these systems of rule, constituting the “patrimonial state,” in which “the political realm as a whole is approximately identical with a huge princely manor” (Weber 1978 [1921]) p.1013. To capture this aspect of early political centralization, we use the term administrative strata and thus indicate the presence of a state due to Wright (1978), because most of our archaeological data on wealth are from single sites, not from multiple contemporaneous sites in a given area.
archaic proto-state to describe a system of elite governance with substantially kin-based or other personalized lines of authority rather than a centralized anonymous bureaucratic system of governance.

Our evidence about the nature of political organization prior to written history is necessarily based on what Flannery calls “archaeological clues,” namely, “the settlement patterns and buildings whose ground plans reflect the social political and religious institutions of the archaic state” (Flannery 1998) p.15. We identify a population as having been governed by an archaic state by the presence of monumental structures such as what archaeologists term “palaces” with significant space allocated to administrative, craft production, or other nonresidential (non-storage) functions, as well as iconic or mortuary evidence of rulership. Both our definition and the nature of the evidence that we draw upon are used by archaeologists studying ancient states (Kirch 2010).51

We turn now to the nature of governance of communities in the Neolithic and the subsequent emergence of archaic proto-states and eventually systems approaching proper Weberian states. We will then see why this process is central to our narrative of the origin of sustained inequality.

7.2. Political differentiation during the Neolithic

Evidence during the Neolithic for collective decision-making and possibly political differentiation is found in large-scale communal storage facilities (Willcox and Storedeur 2012), massive gazelle traps (Moore, Hillman and Legge 2000), and nonresidential public buildings (Roux, Der Aprahamian, Brenet et al. 2000). Large-scale collective action is also reflected in communal processing and storage prior to full-blown farming (Willcox and Stordeur 2012) and in building projects ranging from the tower at Jericho to the monumental enclosures at Göbekli Tepe (Celik 2015, Dietrich, Heun, Notroff et al. 2012, Kuijt and Goring-Morris 2002).

It might be thought that the large Neolithic sedentary communities like Abu Hureyra and Çatalhöyük would provide at least indirect evidence for a centralization of power. They pose a puzzle set by the leading archaeologists of Abu Hureyra, one of the early farming sites: “No hunter-gatherers occupying a productive locality with a range of wild foods able to provide for all seasons are likely to have started cultivating their caloric staples willingly. Energy investment per unit of energy return

51 The paucity of evidence in some cases makes it impossible to designate the state or non-state status of a site with any confidence. We have excluded these data from the portions of our data analysis concerning states. We have not classified cases for which there is evidence of state rule but reason to believe that the community under study did not include the elite.
would have been too high” (Moore, Hillman and Legge 2000) p.393. A possible explanation suggested by Acemoglu and Robinson, as we saw in Section 2, is that the transition to farming and the creation of a surplus were accomplished by an “extractive elite.”

If true, their hypothesis that power in such communities was centralized could resolve the puzzle. But on the basis of current archaeological evidence, it seems unlikely that any of the forms of political differentiation at the time of the first farmers early in the Neolithic (ten or so millennia ago) could have included a centralized authority sufficiently powerful to enforce sedentism on and extract surpluses from a heavily armed mobile population. The monumental towers at Göbekli Tepe in Anatolia have been taken by Acemoglu and Robinson (relying on Bender) as evidence of a political elite that might have been able to do this (Acemoglu and Robinson 2009, Bender 1978). But these and other recently discovered monumental structures at “cult sites” in Anatolia (Celik 2015) are not by themselves sufficient evidence of an extractive political elite; nor are such structures present at the known sites of early farming.

The large and apparently nonresidential buildings at Mureybet (in modern-day Syria) and other Early Neolithic sites are now believed by archaeologists to have been used for communal meetings and rituals. Flannery and Marcus refer to them as “men’s houses” or “ritual houses,” not dwellings of members of a political elite, analogizing them to the underground ritual structures (kivas) of the Pueblo sites in the southwestern United States (Flannery and Marcus 2012) p.138, 136.

The best-documented early transition to a more state-like degree of political centralization is from upper Mesopotamia. Gil Stein’s assessment of the Northern Ubaid (5300–4500 BCE) refers to clay tokens, seals, and other “good evidence of administrative technology” at Tepe Gawra during this period but concludes that:

there is no strong evidence for the use of these materials to monitor the mobilization of surpluses by leaders, elites, or an emergent bureaucracy…. At this point we still lack convincing architectural, iconographic, or mortuary evidence for the full development of centralized political leadership [or]

52 See also the evidence on calories acquired per hour of labor exploiting either wild or domesticated species in Bowles (2011).
53 James Woodburn (1982b)p.187 observed that among hunter-gatherers “every man has constantly at hand the weapons needed for killing.”
55 In Wright’s Atlas the earliest listed “primary state” is a bit later, at Sush c 3500 BCE. Wright (2012)
the formalization of ... incipient wealth differences into permanent social structures of stratification and centralization (Stein 2012) p.130.

At about 4500–4200 BCE in upper Mesopotamia we learn that “we still lack mortuary or architectural evidence for formalized leadership and political centralization.... Overall, the LC2 [4200–3850 BCE] period at present provides the earliest evidence in Upper Mesopotamia for the emergence of formalized political leaders” (Stein 2012:132). We consider below this evidence for the emergence of a more centralized system of governance: the archaic proto-state.

7.3. The archaic proto-state in Mesopotamia

While early studies represented Mesopotamian political systems as strongly hierarchical and centralized, research over the second half of the 20th century and since has “tended to see early social organization as a set of three semi-autonomous institutional arrangements: the great estates of temples and palaces and the community” (Yoffee 1995) p.300. I. M. Diakonoff, the ancient historian who reoriented the literature from the 1950s on, writes that a millennium and a half after the first Mesopotamian “state” at Ur, “charioteers and heavy infantry” were the “personal military following of the ruler or belonged to the temples” suggesting the lack of any single entity with monopoly on the use of coercion (Diakonoff 1969) p.187. Though the title of his major work referred to the “despotic state,” Diakonoff’s summary of the system of rule struck a quite different chord:

…the ruler’s authority beyond the territory which he controlled directly (i.e., beyond the temple estate and later crown land) was neither autocratic, despotic nor unlimited…. [T]he state system of ancient Sumer can hardly even be called monarchic. … [T]he authority of a Sumerian ruler beyond his own estate was greatly restricted by … the community—the Council of Elders and to a lesser degree the Popular Assembly…. In the Early Dynastic period and even later, after the Akkadian Dynasty, [third century BCE] the ruler could not exercise his military functions without consulting the community organs… it is probable that the same organs had the right to depose the ruler (p. 184–85)

Diakonoff adds:

---

Owing to patriarchal ties inside the clans or the extended families, the Popular Assembly … could not have been really a democratic organ because the rank-and-file members… were bound by allegiance to their family heads, who constituted an aristocracy in the clan (p. 187).

Recent scholarship is consistent with these formulations. Drawing on the work of Daniel Fleming and others, Stasavage writes that the king at the city-state of Mari “had to negotiate with individual localities in order to raise revenue, and town councils were collectively responsible for these taxes. These councils were probably elite affairs but, in some instances, it is clear that there was broader participation” (Stasavage 2020) p.35. Richard Blanton and Lane Fargher add that “non-palace governing organizations” persisted through the third millennium BCE (Blanton and Fargher 2008) p. 91.

Yoffee and Barjamovic observe that as late as the second millennium BCE, “the character of royal mandate at Assur was reflected in the way rulers styled themselves as ‘stewards’ of the state deity or ‘chairman’ of the Assembly, but never as ‘kings’” (Yoffee and Barjamovic 2018) p. 816. Trevor Watkins writes of the king’s soldiers “that the bulk of their military activity was directed to intercity warfare and skirmishing; there is very little evidence of internal policing, let alone the holding of a subject population” (Watkins 1978) p.485.

Not surprisingly, in light of this lack of an effective monopoly on the use of force, a “characteristic of Mesopotamian history,” according to Nicholas Postgate, is “the alternation of periods of strong centralized political control with periods of turmoil” (Postgate 1992) p. 22. An example is the destruction (some think by its dissident population) of the Eanna district at Uruk III, which had been the location of temples, ceremonial plazas, and possible administrative buildings. There were attempts to construct more centralized, bureaucratic territorial states closer to the Weberian ideal type, for example, by the rulers Sargan at Akkad, Shulgi of the Uruk III dynasty, Hammurabi at Babylon, and their descendants. But without exception, these efforts were short-lived (Yoffee 1995). Stein concluded that

---

57 This tax-collecting polity with substantial autonomy for local elites exemplifies the decentralized pole of Wolf’s continuum of the tributary mode of production, along which he saw polities continually shifting during phases of centralization and collapse Wolf (2010). Stasavage refers to Mari (modern-day Tell Hariri, a third and early second millennium BCE settlement) and similar polities as “early democracies,” drawing on the earlier work of Jacobsen (1943), who used the term “primitive democracy.” But the primary character of early Mesopotamian urban polities appears to have been the bargaining among elites (not subordinate to a powerful central authority) rather than a degree of popular accountability and political equality associated with the term democracy. Jacobsen concedes that “oligarchy” and “democracy …can hardly, at the present state of our knowledge of ancient Mesopotamia, be profitably distinguished” (159). Our reading of the evidence is closer to Fleming’s own view: Mari’s “collective political forms” are certainly not “democratic by any careful definition of the term”; “[t]here was no “primitive democracy,” but there was a lively interplay of individual and group leadership” (Fleming (2004) pp. 235, 237–38).
“when Mesopotamian states pursued maximizing strategies aimed at extracting large consistent surpluses from the countryside, these attempts tended to be short-lived, unstable, and vulnerable to collapse” (Stein 1994) p.13.

This fragility was not confined to the archaic proto-states of Mesopotamia. Teotihuacan, the largest city in Mesoamerica and which dominated its region from about 200 BCE to 550 CE, ended abruptly with the burning of the ceremonial Street of the Dead (and no other part of the city), almost certainly not by external opponents (Mcanany 2019). And in the Inka Empire during the fifteenth century, “people did not always obey the rules of the state, but the rules of their group or community… legitimacy rested primarily with the leaders of local authorities and less so with the higher-ranking state authorities” (Dillehay and Wernke 2019) p.18.

Early in the formation of the Aztec monarchy, a new king had to be confirmed by the men and women over whom he would rule (Trigger 2003). While the nature of their political structure is a matter of debate, the Chaco and Cahokia civilizations in the eleventh and twelfth centuries (in the modern-day United States) both exhibited a degree of concentration of political power, but neither persisted for long (Emerson 2018, Graeber and Wengrow 2021, Mills 2023).

The capacity of multiple groups to resist the concentration of political power, and the resulting imperative of power sharing for the emergence of proto-states, may have constrained the levels of wealth inequality that could have emerged. An example from Mesoamerica is the “republic” of Tlaxcallan, which, though entirely surrounded and frequently attacked by the much larger Aztec Triple Alliance, remained independent. Fargher and Blanton explain:

Residents resisted by building an entirely new sociopolitical order ... one that limited personal income and reduced conspicuous consumption ... of the elite, and ... for commoners [provided] opportunities to advance socially and politically, the distribution of public goods, [and] the fair treatment of commoners and ethnic minorities in judicial matters…. All of this translated into … a powerful and motivated military force. Commoner and noble, Nhua and Otomi [ethnic groups] offered themselves up on the battlefield … for the idea of an independent, just and egalitarian Tlaxcallan. (Fargher and Blanton 2017)p.539–41

It appears that wherever it occurred, the centralization of political power initially had to be negotiated. But these limits on its powers notwithstanding, the emergence of the archaic proto-state was associated with a marked increase in wealth inequality.

7.4 Wealth inequality in archaic proto-states

61
The data in Figure 1 showed that the period after 4200 BCE (during which the archaic proto-state emerged) is associated with the appearance of a few sites with unprecedented levels of wealth inequality and the absence of populations with modest levels of wealth inequality. This impression is confirmed by the distributions shown in Figure 6.

![Graph showing estimated wealth inequality](image)

**Figure 6. The origin of enduring economic inequality: Estimated wealth inequality in four technology/institution configurations before 525 CE.** Source and information on the permutation tests and associated $p$ statistics between groups are in Appendix, Section 1 and 2. More details on the designation of institutions for the cases in our data set in Appendix, Section 3.

By our estimates, the mean wealth inequality for populations ruled by the archaic proto-state ($G = 0.706$), all with land-limited economies, substantially exceeds that for the land-limited non-state populations ($G = 0.577$). The difference in means of this observed magnitude (0.129) and sign is unlikely to have arisen if the null hypothesis, namely that the two groups do not differ with respect to
inequality, were true \(p = 0.003\). As was the case for the increase in wealth inequality associated with the transition to a land-limited economy prior to the emergence of archaic proto-states, these estimates of the differences associated with the development of proto-states and states are not substantially affected by conditioning on a time trend.\(^5^9\)

The dramatic increase in wealth inequality from stateless, labor-limited populations in the Early Neolithic to state-governed, land-limited populations in the Bronze and Iron Ages appears to reflect consecutive changes in both technology and institutions. It is clear from Figures 1 and 6 that across the region a substantial increase in wealth inequality (associated with the shift from labor- to land-limited economies) predates the formation of the first archaic proto-states. Near-modern levels of inequality are observed at both Tel Brak (Level XII) and Çatalhöyük (Level 6a), which in Tel Brak (according to Stein, above) was followed by the apparent formation of an archaic proto-state, and in Çatalhöyük by the subsequent abandonment of the site.

Because all of the archaic proto-state observations are also land-limited, the data in Figure 6 suggest that the total increase in the mean Gini coefficient—from 0.355 to 0.706—could be decomposed into a shift from labor- to land-limited farming among non-state societies—from 0.355 to 0.577—followed by a shift from non-state to proto-state governance among land-limited farmers—from 0.577 to 0.706. But the influence of changes in technology and in institutions cannot really be partitioned in any convincing way. The decomposition above fails to be a clean “change in technology, holding institutions constant” experiment (two different farming practices prior to the proto-state) followed by a “change in institutions, holding technology constant” case (the non-state vs. archaic proto-state comparison for land-limited economies.)

The reason is that, among the non-state observations, labor-limited and land-limited populations almost certainly differed in institutions and social norms, for example, the extent of private storage and de facto, if not de jure, land ownership. As a result, comparing the two means does not provide a clear identification of the stratifying effect of the change in farming technology per se. Similarly, the

\(^5^8\) The mean of our measure of wealth inequality for the forty-five land-limited archaic proto states is somewhat larger than the average of the 182 estimates for the entire data set for state-governed societies after 300 BCE, excepting those in which the enslaved performed a substantial fraction of the labor (namely, from Figure 1, \(G = 0.695\).) For the null hypothesis of the two groups not differing with respect to inequality, we have \(p = 0.404\).

\(^5^9\) From the OLS prediction of wealth inequality in the archaic proto-state and non-state land-limited economies, conditioned on a time trend, the difference in the predicted mean of the two is 63 percent of the predicted mean difference when not conditioned on time, suggesting that the observed mean difference reported above and in Figure 7 is largely independent of time. See Appendix, Section 5, for the details of the regression.
comparison of inequality among land-limited farmers under non-state and archaic proto-state governance probably reflects not only the effects of the change in institutions but also the technological developments of the Bronze Age and Iron Age, notably in the capital intensity of the production of these metals themselves and their use in the production of more effective tools.

7.5. States, slavery, and the emergence of modern levels of wealth inequality

Slavery, however, provides an instance where the effects of institutional differences seem clear. What we define as slave states were based on economies which relied on a substantial amount of enslaved labor; such states conform more closely than do the archaic proto-states to the Weberian ideal. Our estimates for wealth inequality in the slavery-based economies of the Roman Empire are based on land ownership and house size, adjusted to be comparable across the different assets, and for reasons of data availability do not include enslaved people as a form of wealth. Enslaved people are included as members of the population with zero wealth: Their lack of material wealth is treated no different from the poverty of a free landless farm tenant or wage worker.60

Excluding the value of enslaved people in our measure of the owners’ material wealth, however, results in a downward bias in our estimate of material wealth inequality. To see that the resulting underestimate of the Gini coefficient for slave states may be substantial, consider the following hypothetical population. There are 19 enslaved people working for the owner of ten units of land, and 10 (free) households each owning one unit of land. The value of a unit of land is the same as the value of an enslaved person. The Gini coefficient for land, that is, not counting the value of enslaved individuals is $G = 0.78$, matching the mean value for slave states in Figure 6. But counting the value of enslaved persons we have $G = 0.87$, meaning that our method significantly underestimates the degree of wealth inequality (for these hypothetical but not unrealistic numbers) by 11 percent (see Section 6 of the Appendix for the details of the computations).

Even with this systematic downward bias, our estimates show that cases where enslaved people did much of the work and where political power was more effectively concentrated, as in the agrarian economies of the Roman Empire, are distinguished by extraordinary levels of material wealth inequality.

---

60 Due to their unfree status, a case could be made that an enslaved person has negative wealth; but while the bonded nature of the individual surely affects wellbeing, accounting for it as measure of specifically material wealth does not seem consistent. In all cases, the estimates require recalculating the wealth inequality measure after having estimated the number of enslaved people not included in the available measures of wealth holders, as described in Fochesato, et al. (2019).
While enslaved workers were present in the first Mesopotamian cities, their work was “limited both in numbers and in employment to the general domestic service. The productive labor force is represented by the semi-free class of people who constitute the major labor force of the large households involved in agricultural production and derivative processing” (Gelb 1967) p.6. Figure 6 shows that the combination of slavery with a more highly centralized political system is associated with a dramatically greater wealth inequality, already evident in Figure 1. The difference in means between the archaic proto-state and the slave states is 0.122; a difference of this sign and magnitude would be highly unlikely were the null hypothesis that the two groups do not differ with respect to inequality to be true, with \( p = 0.0002 \).

8. Discussion

In closing, we draw on evidence concerning all three of the factors contributing to inequality introduced at the outset—technology/ecology, institutions/politics, and culture/ideology— to propose an explanation of the origin of enduring wealth inequality and to pose some open questions and puzzles.

Before offering our account, however, we want to underline the vast uncertainties necessarily associated with empirical claims about prehistory and ancient times. In a related paper (Fochesato and Bowles 2024b), we sought to explicitly measure the uncertainties in the measurement of wealth inequalities using both historical and archaeological sources. These arise from small and possibly unrepresentative samples, the need to reconcile measures based on wealth held by individuals as opposed to households (or some other wealth pooling unit), to adjust our estimates for the differing assets used to measure wealth (in this case grave goods, house size, storage area and land) and differing population sizes of the communities under study, and to take account of the fact that in many sources of wealth data those without wealth are missing. By our estimates, the uncertainty associated with the Gini coefficient estimates is substantially greater when based on archaeological rather than historical (that is, written) data. In a large data set of 431 estimates of Gini coefficients, each represented as a probability distribution of likely values for the Gini coefficient at the place and date concerned, the mean span of the 95 percent confidence intervals (CIs) for the 241 Gini coefficients measured using archaeological data

---

61 Diakonoff reports that “during the Early Dynastic period all agricultural work including irrigation work both on the temple estates and in the communities was done by freemen” (Diakonoff (1969) p.187. The offspring of enslaved people in Mesopotamia were sometimes absorbed into the semi-free commoner class (Trigger (2003) p.164.
(0.377 Gini points) is more than three times greater than for the corresponding 190 CIs based on historical data (0.101 Gini points).

Aside from the uncertainties intrinsic to the archaeological data, we generally lack the opportunity to exploit the natural experiments that have become such an important foundation for causal inference in historical economics. Prehistory, even more than history, sometimes just does not run the right experiment. We would learn a lot about the relative contributions of differences in technology and institutions, for example, if we had estimates of wealth inequality in labor-limited economies governed by states or archaic proto-states. We lack adequate data, too, on the process by which a labor-limited economy became land limited. One relatively long series of estimates—for Çatalhöyük over three centuries—ends with the abandonment of the town, not a conversion to an enduring land-limited economy. Our two thousand year Thai series records two transitions to a land-limited economy, one ending with a reversion to a more egalitarian labor-limited economy, the other with the emergence of a proto state.

The fact that we have leveraged data on small-scale societies in the recent ethnographic record to fill some of the gaps in the archaeological record imports additional uncertainties, as archaeologists have pointed out.\(^2\) Two findings about the ethnographic data, however, suggest confidence. The first is that the ethnographers’ estimates of the importance to production of labor versus material wealth (in Table 2) correspond remarkably closely to what one would infer from empirically estimated production functions where these can be compared. The second is that these estimates (along with estimates of wealth inequality in recent small-scale societies) provide us with an accurate prediction of the effect of the late Neolithic movement from a labor-limited to a land-limited economy (predicted and observed increase in the Gini coefficient of 0.241 and 0.222 respectively, that we showed in Section 5.3 ), and how this effect differs between western Eurasia and wet rice farming in our Thai robustness check (a predicted and observed increase in the Gini of 0.132 and 0.115 respectively, from Section 5.4).

Bearing the uncertainties in mind, our proposal is that four developments contributed to the origin of substantial and enduring wealth inequality.

First (Section 4), beginning early in the Neolithic, sedentary living, along with harvesting and then cultivating storable cereals and eventually raising large animals, provided opportunities for private ownership of important and readily stored means of subsistence. This allowed the more successful

---

\(^2\) Robert Kelly faults “viewing living foragers as our Paleolithic ancestors” or as “relic populations” (Kelly (1995)p.334. See also Wobst (1978).
households to withdraw from community-based risk-mitigation practices by means of private storage, reducing the extent of egalitarian redistribution. Uninsured wealth shocks also likely increased. But these developments were not associated with enduring modern levels of wealth inequality, due to the persistence (albeit under duress) of “aggressively egalitarian” social norms and cultural practices.

Second (Sections 5 and 6), the subsequent development of capital-intensive, labor-saving methods of production, particularly plow-based farming, raised the shadow price of land and other capital goods relative to labor. This increased the value of forms of wealth—land and draft animals—that, unlike the capacity to perform labor, could be very unequally held, were readily transmitted across generations and were subject to increasing returns and substantial shocks, jointly increasing the inequality of wealth.

Third (Section 7), the gradual centralization of the effective use of coercion represented by the first archaic proto-states, and later by conventional states, mitigated the likelihood of either internal or external challenges to the wealth holdings of the elite, including stores, homes, and animals providing the political conditions for enduring the wealth inequalities fostered by the change in farming technology.

Fourth (Section 7), the enslavement of people converted labor itself to a form of privately owned material wealth that could be very unequally held once sufficient coercive power had been concentrated in elite hands and was both subject to shocks and transmitted over generations. Enslavement of a substantial fraction of the labor force (e.g., in the Roman Empire) thus complemented changes in technology to elevate the importance of material wealth in production and resulted in a large class without material wealth, whose forced labor elevated the returns to their owners’ other forms of wealth.

The second point above (based on joint work with Amy Bogaard) addresses what we consider to be a lacuna in the literature: the lack of a convincing explanation as to how farming populations without marked wealth inequalities persisted through much of the Neolithic and how at least two millennia prior to the emergence of even proto-states, these heavily armed populations that had been committed to egalitarian norms would tolerate the emergence of wealth inequality that would eventually rise to modern levels and would persist over centuries.

The key step, if we are right, was not the emergence of an agricultural surplus or an extractive elite with coercive powers but instead the introduction of the ox-drawn plow. This and similar labor-saving and land-using innovations effectively increased the scarcity of land and the redundancy of labor, providing the conditions for the emergence of a subordinate working class. Our estimates in Section 5.5 suggest that the increase in wealth inequality between the labor-limited and land-limited economies
prior to the emergence of the archaic proto-state is the result of both an increase in both intergenerational transmission of wealth, and (non-insured) wealth shocks.

While the part of this account that can be robustness-checked appears to also hold for prehistoric Thailand, we have not attempted to explore the possibility that similar processes were at work in other parts of the world. We know that the trajectory of wealth inequalities in the Western Hemisphere differs in important respects from the western Eurasian case (Kohler and Smith 2018, Kohler, Smith, Bogaard, et al. 2017). Kohler and his coauthors suggest that the land-limited explanation just summarized may have more general merit: “the best single explanation of these wealth differences is the presence after circa 8000 BC of a variety of large, domesticated animals in the Old World and their general absence in the New World (Kohler, Smith, Bogaard et al. 2018) p.306. But the necessary paleobotanical and other research documenting cultivation practices in the Western Hemisphere to extend the explanation beyond the absence of draft animals there has not yet been done. As in western Eurasia, heightened inequalities may have emerged in the western hemisphere in cases where the value of land relative to labor could be enhanced by improvements such as irrigation.

The contrasting evolution of political and economic disparities in the upper and lower reaches of the Nile Basin from the fourth to the first millennium BCE provides a further check on our proposed explanation, though in this case we lack comparable estimates of wealth inequality. Similar to our account based on Mesopotamian evidence, Wengrow, referring to the upper Nile, writes:

the real puzzle of Nubia’s ancient history and archaeology … is not whether it developed states or empires like those of neighboring pharaonic Egypt. Rather, it is how its population managed to prevent the emergence of similar forms of domination in their own midst, despite the existence of Egyptian models of governance on their doorstep (Wengrow 2023) p.34.

Wengrow suggests that ecology and culture jointly provide an answer. Greater rainfall in Nubia supported a “Green Sahara,” and a highly mobile cattle-based pastoral economy (rather than cultivation) gave subjects of early Nubian kingdoms a ready way to opt out of oppressive rule or taxation. “What guaranteed this freedom,” Wengrow continues, “was the existence of shared norms of hospitality and ritual practice, which cut across local affiliations” (p. 36). Mobility was more limited in the tightly circumscribed Lower Nile Valley of Pharaonic Egypt. There, following the introduction of ox-drawn plows around 3300 BCE, “for the first time, peasant societies took form, whose lives were tightly circumscribed and whose connections to a world beyond were increasingly confined to interactions with landlords priests, merchants [and] tax collectors” (p. 47).
Wengrow’s representation of the Nubian political system and its limited capacity for coercion resemble our archaic proto-state, and his comparison of the Upper and Lower Nile parallels our distinction between labor-limited and land-limited economies and our identification of plow-based farming of cereals as a key contributor to the emergence of durable inequality. Also similar to our account is the role played by culture in resisting elite domination. Wengrow’s “shared norms of hospitality” (p.51) and “acceptance of strangers and non-kin” appear to have placed limits on the concentration of wealth by ensuring an advantageous reservation option via spatial mobility to escape domination in the Upper Nile. Wengrow’s cultural and ecological underpinnings of a relatively egalitarian society are much like the aggressive egalitarianism of Neolithic communities in the Levant, and the existence of an exit option to unoccupied lands, at least prior to plow farming.

Given the importance attributed to cereal production and storage in the emergence of economic inequality, we remain struck by evidence for egalitarian cultural and political and by the persistence of egalitarian farming societies for many millennia following the introduction of readily storable crops (and the domestication of animals). The more than five millennia delay from the first cultivation of storable cereals to the first archaic proto-states poses a further puzzle: What could have sustained the inequalities that did emerge in the course of the Neolithic, long before any substantial political centralization of power?

A possible response is that many highly unequal societies, like Çatalhöyük, did not endure. But to explain those that did persist over long periods—including for over three centuries at Çatalhöyük, for example—the cultural changes and ideological shifts that supported concentrations of power beyond the lineage stressed by Flannery and Markus and Graeber and Wengrow and others must have been important.

Our conjecture is that even in substantial populations, modest differences in privately held wealth, including storage, dwellings, and animals, may have been sustained by the emergence of a novel set of social norms that made respect of others’ property rights a best response, conditional on almost all others doing the same. This could be modeled, for example, by the “bourgeois” strategy in Maynard Smith’s Hawk-Dove-Bourgeois game, adapted for the Late Paleolithic and Neolithic (Bowles and Choi 2013). As long as wealth differences were modest, respect for property rights may have been consistent with an egalitarian ethos.

63 Only by positing extraordinary lags could Mayshar, Moav, and Pascali be correct in writing “when it became possible to appropriate crops, a taxing elite emerged and …this led to the state” Mayshar, et al. (2022) p. 1094.
But this bottom-up and decentralized mechanism for maintaining social order – including the deliberate destruction of some items of wealth, preventing their transmission across generations, and attacks on the dwellings of wealthy households and elite ceremonial centers-- may have been insufficient to sustain the much greater levels of wealth inequality, eventually including enslavement, that emerged following the introduction of the ox-drawn plow.

Turning from sustaining wealth inequalities to “the creation of inequality” (the title of the important book by Flannery and Marcus), however, we are more skeptical about culture as a driver of inequality. Granted, the pivotal role of temples in early wealth concentration in Bronze Age Mesopotamia and the possible introduction of religion among the Chumash islanders prior to the emergence of a wealthy elite are consistent with ideological and symbolic factors being part of the explanation.

But the rulers of early Mesopotamian archaic proto-states were considered ordinary human beings; later attempts to elevate themselves to deity status failed (Trigger 2003) p.84–85. Wright observes of early southwestern Iran in the late fifth millennium BCE, “the residential indications of class segregation and the settlement pattern features indicative of tribute extraction appeared … centuries earlier than the material correlates of the ideology that would justify support of such a political economy” (Wright 1994) p. 81. Unless accompanied by the economic and political changes we have identified, we are not convinced that symbolic (and similar cultural) differences would have been sufficient to sustain the privileges and deprivations including enslavement associated with inequalities of wealth of the magnitude that developed from the early Bronze Age onward.

We have documented a shift from what appear to have been cases of episodic and rare elevated levels of wealth inequality in the Late Paleolithic and Early Neolithic of western Eurasia (that is, prior to about 7 thousand years ago) to sustained high levels of wealth inequality in the Bronze and Iron Ages (that is, beginning about 5 thousand years ago). Our many examples of episodic inequality prior to the emergence of archaic proto states – at Çatalhöyük in Anatolia in the 8th millennium BCE, on the Russian Plain 18-16th millennia BCE, in Northeast Thailand 2000 BCE–600 CE, the salt producers at Provadia near the Black Sea 5400-4200 BCE – suggest a closing conjecture. The increase in material wealth inequality associated with the transition from a labor-limited hoe-based farming to a land-limited plow agriculture might have also proved to be ephemeral had it not been for the Bronze age concentration of coercive power that provided the political precondition for the perpetuation of the emerging economic elite. The changes in both farming technology and political organization may have been jointly necessary and sufficient to explain the origin of enduring wealth inequality.
References


Bogaard, Amy; Michael Charles; Katheryn C. Twiss; Andrew Fairbairn; Nurcan Yalman; Dragana Filipovic; G. Arzu Demirergi; Fusun Ertug; Nerissa Russell and Jennifer Henecke. 2009. "Private pantries and celebrated surplus: Storing and sharing food at neolithic catalhoyuk, central anatolia." Antiquity, 83, 649-68.


Bogaard, Amy; Rebecca Fraser; Tim Heaton; Michael Wallace; Petra Vaiglova; Michael Charles; Glynis Jones; Richard Evershed; Amy Styring; Neils Anderson; Rose-Marie Arbogast; Laszlo Bartoseiwicz; Armelle Gardeisen; Marie Kanstrup; Ursula Maier; Elena Marinova, Ninov, Lazar; Marguerita Schafer and Elisabeth Stephan. 2013. "Crop manuring and intensive land management by europe's first farmers." Proc. Natl. Acad. Sci. USA.


Borgerhoff Mulder, Monique; Samuel Bowles; Tom Hertz; Adrian Bell; Jan Beise; Greg Clark; Ila Fazio; Michael Gurven; Kim Hill; Paul L. Hooper; William Irons; Hillard Kaplan; Donna Leonetti; Bobbi Low; Frank Marlowe; Richard Mcelreath; Suresh Naidu; David Nolin; Patrizio Piraino; Rob Quinlan; Eric Schniter; Rebecca Sear; Mary Shenk; Eric Alden Smith; Christopher Von Rueden and Polly Wiessner. 2009. "Intergenerational wealth transmission and the dynamics of inequality in small-scale societies." *Science*, 326(5953), 682.

Borgerhoff Mulder, Monique; Mary Towner; Ryan Baldini; Bret Beheim; Samuel Bowles; Heidi Colleran; Michael Gurven; Karen Kramer; Siobhan Mattison; David Nolin; Brooke Scelza; Rebecca Sear; Mary Shenk; Eckart Voland and John Ziker. 2019. "Differences between sons and daughters in the intergenerational transmission of wealth." *Philosophical Transactions of the Royal Society (Biological Sciences)*, 374.


Fochesato, Mattia and Samuel Bowles. 2024a. "Accounting for uncertainty and bias in archaeological and historical evidence on wealth inequality." *unpublished manuscript*.


____. 2024b. "Technology, institutions, and wealth inequality: What do we learn from comparable estimates over 12 millennia?" *unpublished manuscript*.


Gurven, Michael; Monique Borgerhoff Mulder; Paul L. Hooper; Hillard Kaplan; Robert Quinlan; Rebecca Sear; Eric Schnitter; Christopher Von Rueden; Samuel Bowles; Tom Hertz and Adrian Bell. 2010. "Domestication alone does not lead to inequality: Intergenerational wealth transmission among horticulturalists." *Current Anthropology*, 51(1), 49-64.

Gurven, Michael; M. Borgerhoff Mulder; Hillard Kaplan; Paul Hooper; Rob Quinlan; Eric Schnitter; Rebecca Sear; Christopher Von Rueden; Samuel Bowles; Thomas Hertz and Adrian Bell. 2009. "Domestication alone does not lead to inequality: Intergenerational wealth transmission among horticulturalists." Current Anthropology.


Kohler, Timothy; Mihael E. Smith; Amy Bogaard; Christian Peterson; Gary Feinman; Rahul Oka; Matthew Pailies; Anna Marie Prentiss; Elizabeth Stone; Alleen Betzenhauser; Timothy Denneby and Laura Ellyson. 2018. "Deep inequality: Summary and conclusions," T. A. Kohler and M. E. Smith, Ten thousand years of inequality: The archaeology of wealth differences. Tucson: University of Arizona Press,


Shenk, Mary K.; Monique Borgerhoff Mulder; Jan Beise; Gregory Clark; William Irons; Donna Leonetti; Bobbi S. Low; Samuel Bowles; Tom Hertz; Adrian Bell and Patrizio Piraino. 2010. "Intergenerational wealth transmission among agriculturalists: Foundations of agrarian inequality." *Current Anthropology*, 51(1), 65-83.


Smith, Eric Alden; Kim Hill; Frank W. Marlowe; David Nolin; Polly Wiessner; Michael Gurven; Samuel Bowles; Monique Borgerhoff Mulder; Tom Hertz and Adrian Bell. 2010. "Wealth transmission and inequality among hunter-gatherers." *Current Anthropology*, 51(1), 19-34.


Twiss, Katheryn; A Bogaard; Doru Bogdan; Tristan Carter; Michael Charles; S Farid; Nerissa Russell; Mirjana Stevanovic; E. Nurcan Yalman and Lisa Yeomans. 2008. "Arson or accident? The burning of a neolithic house at catalhoyuk, turkey." Journal of Field Archaeology, 33(1), 41-57.


Willcox, George and Danuille Storedeur. 2012. "Large scale cereal processing before domestication during the tenth millennium cal bc in northern syria." Antiquity, 86, 99-114.


