The Sparta Game: Violence, Proportionality, Austerity, Collapse

Josiah Ober and Barry R. Weingast
Stanford University

December 2016


Abstract

This paper suggests that contemporary versions of Adam Smith's informal equilibrium theory in Book III of the Wealth of Nations (Smith, 1776) can explain Sparta's regime stability, along with distinctive features of the social system of ancient Sparta, namely the coordinated social uses of systematic violence, the public façade of material equality among the citizen population, the maintenance of a self-enforcing regime of austerity by an extensive body of citizens. In addition, it uses a dynamic element to explain the severe demographic decline that led to Sparta's eventual loss of standing in the Greek world.

Introduction

In The Wealth of Nations, Adam Smith considered the question of why, in early medieval Europe, economic production remained at such a low level for so long. His answer was predicated on the rational choices that were made by economic agents who are assumed to have employed cost-benefit reasoning in light of prevailing sociopolitical conditions:

[T]he occupiers of land in the country were exposed to every sort of violence. But men in this defenceless state naturally content themselves with their necessary subsistence; because to acquire more might only tempt the injustice of their oppressors (Smith 1776: III.iii.12:405).
In a rude state of society there are no great mercantile or manufacturing capitals. The individuals who hoard whatever money they can save, and who conceal their hoard, do so from a distrust of the justice of government, from a fear that if it was known that they had a hoard, and where that hoard was to be found, they would quickly be plundered (Smith 1776: V.iii.9:911).

In these and related passages, in order to explain the social origins and persistence of a low-performing economic order, Smith focuses on the motivations of individuals, who take other’s expected behavior into account in planning their own actions, His explanation is in the form of what we would now call equilibrium theory. Smith’s early medieval economy was in a state of equilibrium because no player in the “medieval economy game” had a move that would improve his own condition, given the moves he knew to be available to other players. The poor “occupiers of the land” did not move to improve their economic standing because they knew that if they attempted to do so the fruits of their efforts would be taken from them by the strong. Those who did have capital hoarded it, rather than investing it in potentially lucrative enterprises, because they distrusted an extortionate government.

The background condition that kept the early medieval economy unproductive, and the game in equilibrium, was what Smith reasonably enough calls injustice. When the basic rule of society (the underlying institutional order) is that “the strong will take what they can,” and when the strong are strong enough to retain what they take against potential challenges, the outcome is a relatively poor but stable social order in which most people live quite near the level of bare subsistence and any available surplus is hoarded rather than profitably invested.

In these two passages, Smith could just as well have been describing ancient Sparta, where helot-slaves were “exposed to every sort of violence,” and the ruling Spartan citizen-elites were famously avaricious, secretive, and ostensibly avoided engagement in any sort of productive industry or trade.

This paper suggests that contemporary versions of Smith’s informal equilibrium theory can explain Sparta’s regime stability, along with distinctive features of the social system of ancient Sparta, namely the coordinated social uses of systematic violence, the public façade of material equality among the citizen population, the maintenance of a self-enforcing regime of austerity by an extensive body of citizens, and the severe demographic decline that led to Sparta’s eventual loss of standing in the Greek world.

Sparta as a limited access order
Sparta was, we suggest, a variant of what North, Wallis, and Weingast (2009) describe as the “limited access order,” a historically common alternative to the

1 Liu and Weingast (2017) explore Smith’s use of equilibrium (and comparative static) logic across his corpus.
2 For a succinct and thoughtful account of these (among other) features of Spartan society, and a discussion of the apparent contradictions and thus fragility of Spartan society, see Hodkinson 1983, Finley 1987: 161-77.
historically rare "open access order" characteristic of a handful of modern, developed states. The primary characteristic of the limited access order is that access to institutions, and the right to create new organizations, is controlled and often monopolized by a small, ruling, rent-creating elite.

In common with other limited access states, the Spartan regime controlled the potential for disruptive violence among powerful elites by refusing access to institutions (especially to property rights, citizenship, and legal protection from violence) to all but those selfsame elites. Limiting access restricted opportunities for gaining valuable rents to the elite, who were then free to use violence, and threats of violence, to dominate others and to expropriate surplus from them. Classical Sparta is an interesting variant on the limited access order because of the relatively large body of citizen-elites (Spartiates) sharing in the distribution of rents, and because of the distinctive social rules that served to coordinate the employment of regime-sustaining violence across that large elite population. The history of Archaic and Classical Sparta sheds light on both the sources of stability and the reasons for the collapse of a particular form of limited access state.

Cox, North, and Weingast (2017) posit that a limited access order will be stable if and only if rents are distributed according to a “Proportionality Principle.” The principle of proportionality holds that each member of the ruling elite must receive a share of rents that is proportionate to his strength: his potential to employ violence in ways that could disrupt the regime. When proportionality is respected, no one with the power to disrupt society has an incentive to do so, and no one with an incentive to disrupt has the power to do so. The result is stability. The Spartan regime doubled down on proportionality. Rents were distributed according to the potential of each Spartiate to use violence against other Spartiates to disrupt the regime – but also according to his active and effective use of violence against helots and external rivals to sustain the regime.

The Spartan social order is interesting from the point of view of proportionality in that a system of military organization, based primarily on the hoplite phalanx, rendered each one of an extensive body of elite citizens effectively equal to each other in his potential to produce regime-threatening and regime-sustaining violence (Cartledge 1977, 1996). Thus, proportionality required equal distribution of rents to each member of this large body of "similars." In reality, property and rents were distributed unequally; some Spartiates were much wealthier than others. Proportionality was, however, respected insofar as the opportunity publicly to consume the benefits of rents was equalized among the members of the citizen elite through social rules that constrained consumption (or at least public consumption) by the wealthy (Holladay 1977; Hodkinson 2000: 76-81, 209-368, 399-445).

In common with other limited access states that respect proportionality, the Spartan system was at once stable and economically low-performing. Stable low economic performance is due to what Cox, North, and Weingast call the "Violence Trap." The stability of the limited access state regime is based on proportional distribution with a fixed referent: “to each according to the strength of his disruptive potential.” This means that rents cannot be redistributed or reduced, in ways that would be more economically productive, without disrupting the existing social
order. Thus the society is not able to move closer to the Pareto optimal situation in which at least someone gains, while none lose, by, for example, lowering rents and encouraging economically valuable forms of capital investment.

Proportionality, along with its relationship to the violence trap, explains why limited access states have been so common and so persistent in history, and why the transition from limited to open access is so difficult and relatively rare. In a paper on how the concept of development might be adapted to fit the conditions of Greek antiquity (Carugati, Ober, and Weingast in progress), we suggest that Sparta’s rival, Athens, did in fact make the transition to a version of open access. This essay presents the other side of the coin of ancient Greek development, explaining why Sparta, as a distinctive variant on the limited access state, did not and indeed could not make that transition, and thus remained relatively speaking under-developed when compared to Athens.

Classical Sparta’s social order

At Sparta’s demographic high point in the late sixth or early fifth century, there were c. 9,000 Spartiates; if we include the families of Spartiates, this was perhaps 35,000 persons – roughly 15% of a total societal population (in the regions of Laconia and Messenia) of something like 250,000. Of that total, perhaps 160,000 or 65% were helots. Helots tilled the land of Laconia and Messenia, retaining enough to live at a level of subsistence, and turning over the rest to their Spartiate masters. A Spartiate was expected to devote his life to preparing for war and to the education of future citizens; he and his family lived (at least in principle) entirely from rents extracted from helots. The remainder of the population, some 55,000 or c. 20%, were perioikoi (“marginals”), free but subordinate residents of some two-dozen dependent towns of Laconia who produced various goods and services and served in the Spartan army. The domination of helots (and to a lesser extent of perioikoi) by Spartiates was expressly through violence and threat of violence, rather than by the ideologies of legitimate authority and subordination (e.g. divine kingship) common in many other pre-modern societies.

The violence potential of each Spartan citizen was increased through standard and intensive military training (in the formal system of education and socialization that came to be known as the agoge – the “upbringing”). So long as the Spartiates were successful in cultivating and coordinating their individually high

---

3 All dates are BCE unless otherwise indicated.
4 Helots were not chattels, in that they could not be bought and sold by those who managed them, and they often remained bound to a specific plot of land. Modern scholars therefore sometimes refer to helots as “serfs,” but classical Greeks used the same terminology for helots as they did for chattel slaves. See Cartledge 2011 and the bibliography cited therein. On helotage, see further Ducat 1990; Luraghi and Alcock 2003; Hodkinson 2008.
5 Population estimates drawn from Ober 2015: 123-56, based on data in Hansen and Nielsen 2004. These estimates are very rough, but are comparable to other recent estimates of Spartan and Peloponnesian populations. See, for example, Figueira 1984; Hodkinson 2000: 134-5, 385-6; Scheidel 2003. On the profound resentment by other residents of Laconia and Messenia of Spartan rule by domination, the locus classicus is Xen. Hell. 3.3.6, on which see Cartledge 2011: 86.
6 On Spartan education, see Ducat 2006.
violence potential, through law, mutual monitoring, and the tactics of the hoplite phalanx, they could effectively protect the regime against threats to the Spartan social order. Threats were internal (helot revolt) as well as external (rival Greek states and foreign invaders). Equality of violence potential was a result of the nature of phalanx warfare and the training it required. The high individual violence potential of each Spartiate, conjoined with the coordination arising from respecting the rules that preserved proportionality, enabled the Spartan elite collectively to extract relatively high rents from a relatively large population of helots. Organized violence was sanctioned by an annual declaration of war against the helots by the Spartan state, and carried out through acts of terrorism. Young Spartiate warriors in the organization known as the krypteia (the Hidden) regularly carried out terror-killings of helots who had in any way stood out from the mass, thereby (in Smith’s words) “tempt[ing] the injustice of their oppressors.”

The Spartiates were known as the Homoioi: the “Equals” or Similars” (Cartledge 1996). As ancient commentators noted (see below) and as recent scholarship on Sparta has emphasized, the ranks of the Spartiates in reality exhibited considerable inequality in respect to distribution of wealth – most obviously in the form of private land holdings. Wealth inequality tracked inequality in access to positions of political influence. But proportionality was preserved because rich Spartiates were limited in the public display and consumption of their wealth by a strict austerity regime.

The austerity regime limited (public) consumption by each Spartiate to a level roughly indexed to the productive capacity of the least-wealthy Spartiate. By consumption, we mean, in the first instance, eating. But consumption is more generally the use of material goods. So consumption in a broader sense includes housing, clothing, luxury goods, and distribution of goods to clients or retainers. As we have seen, each Spartiate, in principle, contributed equally to the employment of violence against the regime’s enemies (internal and external). Each also, in principle, contributed equally to elite consumption through regular contributions to one of a number of regimental “dining associations” (phitidia). The Spartan citizenship regime required that each Spartiate belong to one of these associations.

Each Spartiate’s contribution came from the rents extracted from the lands and the helots he controlled. By taking his meals in his phidition, each Spartiate consumed food provided by those shared contributions on an equitable basis. Because each Spartiate was socially required to eat daily in his phidition, and because the meals were equally provided by all of the members of each phidition, public consumption was standardized across the ranks of the Spartiates. Because contributions and consumption were equal, in principle no Spartiate ate above the level that could be provided by the poorest member of his phidition. Within each phidition and, assuming (perhaps counterfactually) that the poorer Spartiates were

---

7 On the krypteia, see Ducat 1997.
9 Arist. Pol. 1294b26-8. On Spartan phiditia, see Lavrencic 1993. The model proposed here is consonant with the view of the phiditia proposed by Singor (1999), who argues that they were closely linked to military units.
evenly distributed across the *phiditia*, the wealthiest Spartiate in principle consumed at a level determined by the poorest of his fellow Spartiates.

Strong law-like norms restricted display of private wealth in other consumption goods: notably housing and dress. Spartiates were meant to be “similar” in public appearance as well as in public dining (Thuc. 1.6). Thus, despite the fact of unequally distributed wealth, Spartiates could imagine themselves as sharing equally in the rents extracted from the subject population. Imagination could not deviate too far from public practice, because those rents could be extracted only if there were a high level of cooperation in violence among the many elite citizens who were equal in their violence potential. If the distribution of rents, as exemplified by public consumption, was explicitly unfair, such that the poorer Spartiates lost their incentive to cooperate in the provision of regime-sustaining violence or gained an incentive to disrupt the regime in order to increase their rent share, the system would quickly have collapsed. The result, as Smith predicted, was that helots, stuck in a “defenceless state naturally content[ed] themselves with their necessary subsistence,” while wealthy Spartiates hoarded surplus wealth, rather than productively investing it. Their society famously remained, in comparison with other major Greek *poleis*, in a “rude state” with “no great mercantile or manufacturing capitals.”

**Historical emergence of large elite, equality, austerity**

Recent scholarship has placed the origins of the Spartan order, described above, in the seventh or early sixth century. By the eighth century the Greek world had emerged from the relatively impoverished and demographically depressed conditions of the Early Iron Age, and many Greek communities crystallized as city-states. Sparta seems to have been a fairly ordinary, if large, aristocratic proto-*polis*, dominated by a relatively small elite. In the eighth century Sparta may perhaps have already developed exceptionally strong (compared to other mainland Greek *poleis*) status distinctions among its inhabitants. Some native Laconians may already have been locally enslaved (i.e. Laconian helots). There is, however, no reason to believe that there was yet an extensive citizen body, or strong forms of intra-elite equality, or austerity. In the expansion era (later eighth through the seventh century), while some leading Greek *poleis* were establishing colonies in the western Mediterranean and Black Sea regions and intensifying agricultural exploitation of home territories, Sparta focused primarily on a path of military expansion directed at neighboring Messenia – acquiring its highly fertile land and enslaving its large population.¹⁰

The ancient sources make it clear that the conquest of Messenia proved difficult. The conquest was completed by the end of the seventh century and the entire population of Messenia was eventually reduced to helotage. In the early sixth century, after the victory over the Messenians, but perhaps before the final division of Messenian lands among the Spartiates and complete transformation of the Messenian population into helots, Sparta turned its expansionist efforts northwards.

---

In attempting to annex Tegea, a polis in the mountainous region of Arcadia, Sparta appears to have over-reached: The initial campaign against Tegea was a failure, ending in the capture and enslavement of a substantial number of Spartans. Sparta later defeated Tegea, but, rather than reducing the Tegeans to the status of helots, the Spartans made Tegea into an unequal ally of the Spartan state (Herodotus 1.65-8); this was the first step in the eventual creation of a Sparta-dominated Peloponnesian League.

At some point in the course of these events in the seventh or early sixth century, Sparta reached a point of crisis: Although the ancient sources are not very informative, it seems likely that military setbacks led the Spartan elite to fear that Sparta could lose control over Messenia. At this point, we suggest, the ruling coalition (assumed to be at this time a fairly small body of wealthy Spartans) recognized that the only way forward was to increase mobilization rates and build morale among those who were mobilized. In light of the (ex hypothesi) difficult military situation, the elite needed to offer substantial concessions to new recruits in order to persuade them to take the necessary risks. This produced a moment (or an era) of relatively radical franchise inclusiveness, traditionally attributed to the lawgiver Lycurgus, whereby the citizen elite was substantially enlarged. The members of the old elite retained their large land holdings and control over helots in Laconia. But each of those (free, adult, male) Spartans who mobilized against the enemy was offered a substantial share (kleros) of the land taken from the Messenians, and a share in the labor of a Messenian population in the process of being helotized. Those men who took up the offer were mobilized, trained, and armed. Over time they became effective equals of the old elite, and of one another, in their capacity to employ violence. The expanded citizen elite, “the Spartiates” (aka “the Homoioi”), pacified the Messenians and defeated the Tegeans. The c. 9,000 Spartiates who participated and survived were duly rewarded with shares of Messenian land.

The old elite had had two goals in expanding the franchise: survive the crisis and retain their existing wealth – mostly in the form of land in Laconia. The only path that they could devise to achieving both goals was to increase the size of the citizen body and thereby increase collective violence potential in the face of the external (rival Peloponnesian poleis) and internal (helot) threats. The threat was severe enough to force a moment in which access to certain privileges was opened to a broader franchise. Access to citizenship meant a share in some anticipated (Messenian) rents, although not ancestral (Laconian) rents. After this one-time expansion of the franchise, the expanded ruling elite respected (we suggest) the proportionality principle. The old elite was now a wealthy subset of a new and extensive military elite: a large group of citizens with equal and high violence potential. The continued active cooperation in respect to violence of that new elite was essential if the numerous helots (in Messenia and in Laconia) were to be kept

---

11 The treaty that led to alliance with Tegea specified that the Tegeans were to not to give sanctuary to Messenian refugees (Holladay 1977: 125-6). On persistent Spartan fear of helot uprising, see Cartledge 2001: 127-52. On the connection between the dangers represented by helot insurrection and the foundation and structure of the Peloponnesian League, see Baltrusch 2001.
under control and external rivals kept at bay. Because the members of the new citizen elite were equal in their capacity for disruptive and productive (to the Spartan regime) violence, proportionality required that rents be equal. But, if the old elite were to retain their large private holdings in Laconia, “equal rents” would have to be construed in some way other than “equal wealth.”

This, we hypothesize, is the origin of the Spartan austerity regime. According to the mature “Lycurgan” system (i.e. the system that crystallized by the middle of the sixth century, if not before) the consumption benefits from rents were equalized, although wealth remained unequal, through the following mechanism. Each Spartan was a member of a military regiment. During the recent wars, the members of each regiment had (we suppose) eaten together, sharing more or less equitably in what could be secured by plunder or purchase: regiments had, in effect, arranged their own logistics via equitably sharing the goods extracted by the cooperative use of violence. Now, in the post-war era, that system would be retained in another guise: Spartiates would dine in public, by regiments, rather than in private, by families. Each Spartiate would bring to his regimental mess an equal and fixed quantity of agricultural produce that had been extracted as rents from land and helots he controlled. Each member’s land and helots were (at least) sufficient to produce an individual contribution adequate to feed the other members of the regiment for a fixed period of time. All members of the regiment would dine from these equal contributions (i.e. from rents extracted by individuals through the collective violence regime. No Spartiate would be humiliated by differences among the size or lavishness of the individual contributions, because, by the new rules, the contributions would be standardized: the same amount and the same quality of food from each, and to each. By equalizing consumption this system preserved proportionality, while it also allowed for ongoing wealth inequality.12

The Lycurgan equilibrium

Although the old elite had bigger land holdings (their lands in Laconia as well as whatever part they had in the Messenian land distribution), they did not (at least publicly) benefit more than any other Spartan from rents secured by cooperation in violence. The wealthy, like all others, consumed at a standard level set by what the least-advantaged Spartiate (he who owned least land and fewest servants before the new system was put into place) had been able to provide at the time that the system was established and the standard contribution level was set. At the extreme, the least-advantaged Spartiate may be supposed to have provided for this family and his mandatory regimental contributions entirely from whatever share of land and helots was distributed to him after the pacification of Messenia. Proportionality required that the quality and quantity of the standard contribution – i.e. the contribution of the wealthier and poorer members alike, and thus the quantity and quality of food that each Spartiate consumed day in and day out and that the regiment could

12 In reality, there were probably some inequalities introduced in the form of “additional” contributions of hunted game and perhaps other products (Hodkinson 1983: 254). But these were supplementary to the fundamental equal contribution.
publicly redistribute to clients and retainers (helots, according to Figueira 1984: 97) – must be determined by what the least well-off Spartiate was able to provide.

In sum, the poorest Spartiate contributed, both at the level of his rent share and his violence provision, at a level that was equal to that of the wealthiest, and the wealthiest consumed at the level of the poorest. Each Spartiate was dependent on each of his fellows for what he ate, just as he was dependent on each of his fellow infantrymen to perform correctly in the phalanx. The generalized result, when extended to all forms of consumption, was austerity: that is, everyone but the poorest consumed (in public) at a level below what his wealth would otherwise have allowed.\footnote{The reality of Spartan austerity is defended against skeptics by Holladay 1977; see also Hodkinson 1998.}

As soon as a Spartiate was unable to provide for his fellows on equal terms, in regard to material contributions or military service, he was expelled from the ranks of the equals. The expulsion of those who failed in their military service was required by the proportionality principle, which forbids equal distributions of rents to those who are unequal in their violence potential, either to disrupt the system or (as in the Spartan case) to preserve it. A second sort of failure also led to expulsion. Those who failed in their material contributions were expelled because their contributions were equivalent to a regimental system of logistics; failing to supply the regiment with supplies was tantamount to failing in violence provision. The key to proportionality is equalizing rent shares among those who are equal in strength, as measured by violence potential. Failure in either military service or material contribution was evidence of inferiority, and thus required a correspondingly inferior access to rents.

The basis of the Spartan system in proportionality has implications for how we conceptualize the original distribution (or redistribution) of Messenian lands among the Spartiates. The Lycuran system created a condition in which the level of each Spartiate’s consumption was set by the amount and quality of food provided by the poorest Spartiate. The wealthy therefore had an incentive to divide the land of Messenia relatively equitably (or to acquiesce in that equitable distribution). Equitable distribution kept the size of the smallest share large enough to keep each Spartiate’s level of consumption well above subsistence.\footnote{Figura 1984 estimates that a standard contribution was about 5 times the subsistence minimum (based on wheat equivalents). Thus Spartan austerity was far from poverty. But it was also far from luxury. The Spartan standard contribution was not much higher than the mean family income in Athens in the later fourth century (estimated at 4.5 times subsistence, using roughly similar measures). Elite Athenians enjoyed incomes that were 10 or more times subsistence, and had ample opportunity to consume accordingly: Ober 2015:91-97; 2016.}

The Spartan system is a (perverse, in that it is based on violence) version of John Rawls’ (1971) “difference principle” of distributive justice. Rawls’ difference principle holds that, once basic rights (liberties) have been secured for each member of a society, inequality in distribution is justified only to the extent that it benefits the least-advantaged member of society. Admittedly we do not know why the Spartans chose this solution, among various alternative approaches to preserving proportionality and wealth inequality; we can only say that, if the proportionality principle holds, and if we assume that the old elite preferred to keep
their existing wealth, the "Lycurgan" regime of austerity was within the relatively constrained set of solutions available to them.  

We propose that the rest of the Lycurgan system (austerity in dress and housing, intensive focus on military training of citizens, collective responsibility for the education of citizen-youths, social panopticon of mutual monitoring by citizens) coalesced around the equilibrium of proportionality, preserved by wealth-inequality among equal violence-providers being counterbalanced by consumption equality, as expressed in austerity. The Lycurgan system formalized equal consumption, and thus preserved proportionality in violence and rents, while also allowing for private property and considerable inequality of actual wealth. The system ensured that each Spartiate was freed from the need to engage in ordinary forms of productive labor. It also had the effect of eliminating the possibility that accumulated capital would be productively invested, thereby dampening the potential for economic growth.

The advantage of this rule in the short-term is that it freed each Spartiate to concentrate his efforts on developing skills as an effective warrior-\textit{Homoios}. Because the system was reasonably stable, each Spartiate could rationally invest deeply in the very specialized forms of human capital that made for excellence in hoplite warfare (Cartledge 1977). Sparta, as a society, was therefore an active participant in the trend of local and individual specialization, based on identification of comparative advantages, that, elsewhere, drove economic growth across the ecology of Greek states in the Archaic and Classical eras (Ober 2015).

For a social system to be stable in the absence of a third-party coercive authority, it must be self-enforcing. That is, the participants in the system must, themselves, act effectively, and when necessary collectively, against deviations from it: they must both obey the rules and participate in enforcing them. In order to maintain the Lycurgan system over time, each Spartiate had to do his part in educating the next generation of warriors by ensuring that each young Spartiate invested deeply and uniquely in the sorts of human capital necessary for productive and equal violence provision. Each Spartiate also had to take part in monitoring all others. Monitoring was essential in order to avoid free-riding – defection from the cooperative regime by failing to contribute to it while still reaping its benefits. A cascade of free riding would have led to a “race to the bottom” that would quickly have ended in inadequate violence provision, and thus the collapse of the social order based on the master-slave, Spartiate-helot hierarchy.

Because of persistent internal and external threats, each Spartiate faced the likelihood of catastrophic loss if the system were to collapse – i.e. if the citizens were unable to coordinate the right kind of violent responses to threats by helots and external rivals. Thus each had an adequate incentive to sustain the system, even though austerity meant that the wealthier Spartiates were consuming at a level below that which would otherwise have been available to them in light of their privately-held wealth. The mechanism for coordination against defection was

\footnote{15 Other approaches – e.g. progressive taxation, and using tax revenues to provision the regimental messes; differential contributions to the messes collected and then divided equally – might seem either to require more bureaucratic apparatus (which the early \textit{polis} was not prepared to provide) or would make the inequality of rents more obvious.}
provided by common knowledge of the rules – the famous (in Greek antiquity) Spartan adherence to their common master: the laws.\textsuperscript{16}

Mutual monitoring ensured, for example, that each Spartiate treated his helots equally badly, engaged in the appropriate level of military training, shared in training of the youth, did not cheat too blatantly by indulging in private luxury, did his part in violence provision in the phalanx and against the helots, and provided rations for his \textit{phidition} from the rents extracted from his land and helots. He who failed in any way could expect to be reported to the authorities (kings, Gerousia, ephors). If he were found to have failed in a substantial way, the defector was demoted to the level of Trembler or Inferior. This collective enforcement of the rules pushed back against the tendency of the rich to consume and display their unequal private wealth in unacceptable ways within the community. Yet it also left them ample opportunity to display wealth in socially acceptable ways (e.g. chariot races in Laconia and at Olympia and Delphi, Hodkinson 2000: 303-34). It also pushed back against the individual utility-maximizing tendency of each Spartiate to under-supply violence, training, education of the youth, while free-riding on the provision by others of those services, which collectively secured the public good of state security. The imminent danger of free riding underpins the Spartan social panopticon.\textsuperscript{17}

Inherent dynamics in the Spartan system:

\textbf{Demographic decline and security failure}

The Lycurgan system proved highly effective in the first several generations. Extracting high rents from the helots allowed the large citizen population of Spartiates to become uniquely specialized in organized violence – and thus able to dominate their neighbors (Cartledge 1977). After the difficult war against Tegea demonstrated the problems of extending the helotage regime into other parts of the Peloponnese, the Spartans shifted their strategy away from overt military expansion and toward creating unequal alliances with other Peloponnesian states. The elites of those states rationally agreed to acquiesce in Spartan hegemony, and thus to support Spartan foreign policy, so long as Sparta’s violence specialists were potentially available to defend local oligarchies against the threats of tyranny and democracy and did not interfere in local affairs.\textsuperscript{18} This exchange formed the basis of Sparta’s Peloponnesian League. The system was tested in the Persian Wars, in mid-fifth-century conflicts with Athens, and then the Peloponnesian War.

Despite some setbacks, the Spartan system had proved to be robust to challenges. But at the Battle of Leuctra in 371 there were only 1,300 Spartiates in the phalanx. This number, although augmented by \textit{perioikoi} and freed helots serving

\begin{itemize}
  \item \textsuperscript{16} On incentive compatibility and legal mechanisms as foundation for a self-reinforcing order predicated on coordinated collective action against defection, see Weingast 1997; Ober 2012. On Spartan devotion to their laws, see Hdt. 7.104; Xen. \textit{Luc.} 8.1.
  \item \textsuperscript{17} On Spartan mutual monitoring and public authorities, see Cartledge 2001. Security is a public good because it is non-rival and non-excludable.
  \item \textsuperscript{18} On the difficulties faced by Greek oligarchs in maintaining their hold on a given polis after the rise of democracy as an alternative form of social order, see Simonton 2012. On the Peloponnesian League, see Cawkwell 1993 and Wolff 2010.
\end{itemize}
as hoplites, proved to be inadequate to stand against well-trained opponents employing new phalanx tactics. After the severe defeat at Leuctra, Sparta lost control of Messenia and quickly declined to the level of minor regional power.\textsuperscript{19} Writing in the later fourth century, Aristotle noted that, "some Spartans have come to have far too many possessions, while others very few indeed; as a result, the land has fallen into the hands of a small number... although the land was sufficient to support 1500 cavalry and 30,000 hoplites, the number [of Spartans] fell to below 1000."\textsuperscript{20} Aristotle explains the demographic decline in the number of Spartiates by reference to socio-economic rather than biological (e.g. infertility) causes.

How and why did the population of Spartan citizen elites decline so catastrophically by 371? It is worth noting that in the fourth century the overall population of the Greek world was growing. Indeed, the fourth century was the era in which the overall Greek population appears to have reached its pre-modern apex. (Ober 2015: 71-100). To address the puzzle of Spartiate demographic collapse, just as the Greek world was experiencing peak demographic (as well as per capita consumption) growth, we begin with the observation that Sparta’s citizen population decline seems to reflect demotions, generation after generation, of a significant number of Spartiates from the ranks of the citizens. Meanwhile, by the limited access logic that governed the Lycurgan regime, relatively few new full citizen-equals were naturalized. The basic logic of the limited access order explains constrained naturalization. The question is why the demotions were so persistent.

We suggest that Aristotle was correct in identifying material inequality as the underlying cause. Two primary, intertwined social mechanisms were at play in this decline: the rules of partible inheritance that allowed women to inherit real estate and encouraged the consolidation of estates in the hands of a relatively few families, and the continuous pressure of downward mobility conjoined with limits on upward mobility. The role of the inheritance system, which enabled wealthy Spartan women to be heirs, and in turn to marry wealthy men, has been explored in detail by other recent scholarship.\textsuperscript{21} Here we focus on the role of downward mobility; that is, the tendency of the poorest of the Spartan elite, those on the losing side of an inheritance system that consolidated a limited pool (land and labor) into fewer hands, to be demoted from the ranks of the Spartiates.

Individual Spartiates (along with their families) regularly dropped out of ranks of the elite (thus lost their citizenship and its privileges) due to a failure to fulfill the duties incumbent upon each Spartiate. As we have seen, those duties included effective violence provision; a Spartiate could be expelled from the elite and demoted to the ranks of the inferior status group called “the Tremblers for failing in his duty in war. But, following Aristotle, it is probable that a more common reason for demotion was a failure to provide the assigned share of provisions to one’s phidition. The status group known as “the Inferiors” was probably composed of former Spartiates who had been unable to maintain their contributions, due to

\textsuperscript{19} On the battle of Leuctra and its aftermath, see Hanson 1988; Buckler 2003:232-350.
\textsuperscript{21} See Hodkinson 2000: 94-103, 402-6; Fleck and Hanssen 2009, Bresson 2016,
poverty. As noted above, the failure to supply one’s regiment with food was tantamount to a failure in military conduct.

The inheritance system precipitated a dynamic, downward spiral for two reasons. First, the rich married rich, implying that the poor married poor, so the distribution of income became increasingly unequal. Second, and relatedly, the division of an estate among multiple children necessarily implied falling per capita wealth. If a Spartiate had three children and each inherited one-third of the father’s estate, their per capita wealth would be one-third of their father’s. Because poor married poor, the consequences of partible inheritance were not counteracted by “marrying up.” This demographic dynamic meant that, over time, the poorest Spartiates were unable to meet their obligations to provision their regiment and, therefore, were liable to be demoted.22 In principle, economic growth might have offset this dynamic; but as we have observed, Spartan rules discouraged investment and dampened economic growth.

The losses from demotion were not offset by non-elites rising into the ranks of elite violence providers. Although some helots were recruited into the army and were subsequently liberated so as to become, in effect, perioikoi, the sharp cutting edge of Sparta’s infantry remained the citizens who had trained together in the agoge.23 Provisions for bringing new members into the ranks of the Spartan citizen elite were clearly inadequate to make up for those who were demoted from it. As a result, eventually dwindling numbers implied too few elite citizens/violence specialists to defeat external rivals in battle; that in turn led to exogenous shocks – defeat by a coalition led by Thebes and the establishment of Messenia as an independent polis. The result was the collapse of Sparta as a major player on the Greek inter-state scene.

Again we see an interesting variant on the natural state and its “violence trap.” While stable for a long time due to proportionality, the reality of inequality created a slow-motion demographic dynamic that ultimately doomed the system. Notably, the violence trap retained its hold to the end, even as the ongoing demographic decline reduced Sparta’s effectiveness against internal and external threats. There was no obvious way for the regime to redistribute rents in a way that would preserve the consumption-equality/wealth-inequality equilibrium. Replacing partible

---

22 Adam Smith argued that this logic led to a peculiar form of property rights in land during the European feudal era; viz., primogeniture and entails, which, respectively, restricted inheritance of land to the eldest son; and prevented a land owner from dividing his land at other times. Economically, this prevented land from moving from lower valued users to higher valued users, restricting economic growth. But militarily, it helped maintain the ability to fund large armies. “When land was considered as the means, not of subsistence merely, but of power and protection, it was thought better that it should descend undivided... The security of a landed estate, therefore, the protection which its owner could afford to those who dwelt on it, depended upon its greatness. To divide it was to ruin it, and to expose every part of it to be oppressed and swallowed up by the incursions of its neighbours” (Smith 1776: III.i.3:382-83).

inheritance with primogeniture would have replaced one cause of demographic decline with another: Primogeniture would reduce the numbers of Spartiates, insofar as no household would produce more than a single (*kleros*-holding, *phitidion*-contributing) Spartan in a given generation, and some households would naturally fail produce a surviving son. Thus, while the Spartan system is a unique variant on the limited access state, it was subject to the same constraints that impede other limited access states from transitioning into a more economically productive (or in this case, long-term viable) social order. The Spartan case also demonstrates one way that a limited access order may be overtaken in a competitive environment, when rival states begin to transition to the “doorstep conditions” leading towards more open access.

As Hodkinson (1983), among others, has noted, a central contradiction in the Lycurgan system was the tension between the official focus of the regime on the primacy of public interests, and the reality of a persistent attention by individual Spartiates to their private (family) interests. The overriding public interest was in stability and the security of the regime. Stability meant respecting proportionality. Security meant that proportionality must pertain among a large body of elite citizens, each highly adept at the provision of productive violence. Proportionality is scalable; that is, as the numbers of Spartiates shrunk, proportionality was respected so long as those expelled from the ranks lost access to rents in proportion to their loss of power to disrupt (or sustain) the system. On the other hand, security was not indefinitely downwardly scalable. As the loss of Sparta’s standing in the Greek world after the battle of Leuctra demonstrated, when the number of Spartiates dropped to a certain point, security could no longer be guaranteed. Aristotle identified this deadly dynamic readily enough, *ex post.* Even if the Spartan leadership did recognize it *ex ante,* however, the proportionality principle pushed back against making regime adjustments profound enough to reverse the deadly demographic trend.

But the puzzle remains. Presumably the Spartan elite had, collectively, recognized the need to expand the franchise at the moment of crisis that precipitated the Lycurgan reforms. Why did a similar recognition of the danger to collective security of ever-falling numbers of citizen soldiers willing and able to defend the regime not lead to rebooting the system by producing a new distributive order that would respect proportionality among an expanded citizen body, and thereby averting the crash?

One possible answer is provided by the logic of the standard commons tragedy, in which very modest cheating on the margin (the addition of one extra sheep by each shepherd to the jointly-owned grazing ground in the standard example) by sharers in some common pool resource ends in fatally degrading an the resource. But, as we have seen, the mutual monitoring of the social panopticon was

---

24 In fact the “inferior” status classes were a potential source of disruption, if (and only if) they could overcome collective action problems – as at least two conspiracies to overthrow the regime in the late fifth and fourth century demonstrated. For the details, see Cartledge 1987: 93-8, 160-79. Sparta’s solution to this problem was to foment collective action problems by seeking to prevent the emergence of common knowledge about preferences for regime change among the sub-elite. As Simonton 2012 has shown, this was a standard approach by Greek oligarchs in other city-states throughout the Classical era.
designed to prevent visible kinds of cheating, in the form of free riding on the system. Since, over all the monitoring system seems to have worked, the degradation of the jointly-owned violence-resource cannot readily be attributed to individual cheating. What then went wrong, such that the public good of collective security was incrementally undercut by the expulsion of one after another Spartiate from the ranks of the citizens?25

At least part of the answer lies in a conflict between the public goods of security (sustaining a population of Spartiates adequate to dominate helots and rival states) and public order (fulfilling the expectation that Spartiates will obey and enforce the rules). The individual Spartiate’s decision to collaborate in the expulsion from the ranks of the citizen elite of a given citizen, who had indeed failed in the provision of agreed-upon contributions (of goods in the way of the standard contribution to his phidition or services in the way of military service) was in conformity with Spartan law. To demote those who failed was to act in obedience to the law. Obeying the law provided an essential public good insofar as public order was predicated on predictable rule-following and rule-enforcement by all concerned parties. But this may seem to beg the question: Why could the law itself not be changed?

The answer to that question may be sought in the fact that, as Aristotle pointed out and Hodkinson (1983, 2000) has show in detail, Sparta was not a true society of equals. Holdings in Laconia had always been unequal and the initial Messenian distribution of land was probably not perfectly equal either. The Lycurgan system had been based on a compromise: The rich gave up some of the anticipated Messenian rents to the poor, but sharing those rents enabled the rich to keep what they had in Laconia, and to share in Messenian rents as well. Meanwhile, and over time, larger and smaller, richer and poorer families were differentially affected by conditions of partible inheritance and marriage. Through processes of acquisition and inheritance, some Spartan families became very wealthy, others quite poor. If the poor Spartiates threatened to expropriate the lands of the rich, or to require rich women to marry poor men, the rich could turn their disruptive violence potential against them. But it never came to that, because the institutions of the Lycurgan system systematically favored the rich. Wealthy Spartans dominated leadership positions (kings, ephors, Gerousia). As a class, the wealthy could act within the constitutional order as a veto player over any move by the less well off to change the system.

Meanwhile, the wealthy had an incentive to allow demotions, insofar as they benefited (albeit in ways other than public consumption) from those demotions. Although the exact mechanisms involved are obscure, we can presume that the landholdings and helots controlled by those who were demoted from the ranks of the Homoioi ultimately came into the control of other Spartiate families. The takeovers were evidently private, and not regulated by any principle of equitable redistribution.

The “pie” of rents based on land held in Laconia and Messenia and on the labor of helots was essentially fixed. For the reasons identified by Adam Smith, in the passages quoted at the beginning of this paper, there was little endogenous economic growth in Laconia and Messenia – unlike other regions of the Greek world. This “stable-sized rent pie” was divided unequally among an ever-shrinking citizen elite. Thus, as some slices of the pie shrunk to the point that the owner of the slice was dropped from the ranks of the citizens, other slices grew accordingly. Each demotion will have marginally improved the material situation of whichever of the remaining members of the elite came to control his land and helots. There were, therefore, both public (rule following) and private (potential individual benefit) reasons for each Spartan to collaborate in the expulsion of the least successful member of the community. As there were fewer Spartiates to share in the pie, the size of the average (although probably not the median) slice must have increased, resulting in the phenomenon observed and commented upon by Aristotle.

At the moment of decision regarding demotion of a citizen, the combined weight of the public good of legal obedience, the private good of potential individual benefit, along with a potential club good, at the level of the phitidion, of improving the average strength (and thus increasingly the prestige) of its membership was likely to over-balance any concerns about the marginal degradation of the aggregate capacity of the citizen elite to ensure collective security: Like the cheating individual shepherd who reasons that “one more sheep will not destroy our pasture” the law-abiding Spartans could reason that “one fewer Spartan will not lose our next battle.” The aggregate weight of public, private, and club goods created what was, from the perspective of the public good of security, a perverse incentive to expel weaker members from the elite group. The perverse incentive may have strengthened in the late fifth and fourth centuries, for reasons having to do with the Peloponnesian War and its aftermath. The general point we seek to make is that the perverse incentive was, in light of the proportionality principle, strong enough to block the formation of a collective will to reform the system in any very fundamental way. Reform would need to be very substantial if it were to guarantee a constant number of elite violence providers adequate to maintain Spartan security of. Eventually, in the later Hellenistic period, the Kings Agis IV and Cleomenes III did attempt fundamental constitutional changes aimed essentially at rebooting the Lycurgan system. But meanwhile, In the game played by the citizens of Classical-era Sparta, substantial reform was kept off the table by the demands of proportionality and by the perverse incentives created by the proportionality-respecting Lycurgan system.

In conclusion, while some expulsions of Spartans from the ranks of the citizens may have been badly motivated (rivalry, enmity, etc.) there need be no malicious intent. The system collapsed because each Spartan played the game according to the established rules. Each did his job in maintaining the laws of Lycurgus – as his ancestors had done before him. Those most vulnerable, because poorest, were most likely to be put into the position of violating the rules, and the expulsion of the most vulnerable was most likely to benefit the rest. By 371 the effect of maintaining the rules of the game rendered Sparta fatally vulnerable to the combined pressure of the external shock of enemy attack and the internal threat of endemic helot discontent.
Works cited


Fleck, Robert K and F. Andrew Hanssen. 2009. “‘Rulers ruled by women’: An economic analysis of the rise and fall of women’s rights in ancient Sparta.” *Economics of Governance*. 10:221-245.

Hanson, Victor Davis. 1988. "Epameinondas, the Battle of Leuktra (371 B.C.) and the Revolution in Greek Tactics." *Classical Antiquity* 7: 190-207.


